AsPac J. Mol. Biol. Biotechnol. 2011
Vol. 19 (2) : 83-93

Isolation and Characterization of the GA 20-Oxidase cDNA from Sago Palm (*Metroxylon sagu* Rottb.)

Bala Jamel[1]*, Mohd. Hasnain Hussain[2], Mohd. Azib Salleh[2], Noraini Busri[1]

[1]*CRAUN Research Sdn. Bhd. Lot 3147, Block 14, Jalan Sultan Tengah 93055 Kuching, Sarawak, Malaysia*
[2]*Department of Molecular Biology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak.*

Received May 2011 / Accepted June 2011

Abstract. GA 20-oxidase is involved in controlling stem elongation, maturity and flowering. Based on published conserved amino acid sequences of plant GA 20-oxidase cDNA clones, oligonucleotide primers were constructed and used to amplify partial sequence of the GA 20-oxidase gene of *Metroxylon sagu*. A 500 bp PCR product was obtained. BLAST analysis showed that the PCR product was homologous with the GA 20-oxidase gene from other plant species. The PCR product was labelled using the Digoxenin (DIG)-labelling system and used as a probe to obtain the full-length GA 20-oxidase gene sequence of a genomic fragment using the genome walking method. The genomic sequence was used for primer construction which were subsequently used to amplify a full-length cDNA copy of the GA 20-oxidase gene. The cDNA obtained was cloned into a pPCR-Script Amp SK (+) vector for sequence confirmation. The 1161 bp sequence obtained from the cDNA copy was compared with the genomic sequence of GA 20-oxidase. Comparison between genomic and cDNA fragments indicated that the GA 20-oxidase gene from sago palm is comprised of two introns and three exons. Heterologous expression of the GA 20-oxidase cDNA in *Escherichia coli* showed a similar expression pattern as the endogenous GA 20-oxidase of sago palm.

Keywords: DNA sequencing, GA 20-Oxidase, Genome Walking, Sago palm, Southern hybridization, Western hybridization.

INTRODUCTION

Sago palm (*Metroxylon sagu*) is one of the few tropical crops that can tolerate wet growing conditions such as peat swamps. In Sarawak, sago is grown as a starch crop by rural communities living along the coastal areas of certain districts. The total acreage of sago in Sarawak is about 65,000 hectares (ha) of which 45,000 ha are held by smallholders and 20,000 ha consist of plantations. About 75% of the sago planting growing area is located in the Mukah, Igan and Oya-Dalat districts of Sibu Division and Balingian (Tie *et al.*, 1991). There is also a substantial acreage of sago in the Pusa and Saratok districts of Sarawak. The total export of Sarawak sago starch in the year 2010 was 44,192 tons.

Research on sago has begun to gain momentum since the 1970's (Stanton, 1972). Initially R&D focused on agronomic practices to improve growth and yield but later emphasis was placed on downstream applications of sago starch. One of the key issues for sago which needs to be critically addressed is that the palm takes about 10-12 years to reach maturity depending on soil type, while other starch producing crops, potato and cassava, take only 3 and 6 months, respectively (Chulavatnatol, 2002). This reduces the competitiveness of sago as compared to other starch producing crops.

In most quickly maturing plant species, conventional plant breeding techniques are very successful in generating new elite varieties. This technique is unsuitable however for slowly maturing plants such as sago palm. Sago is a hapaxanthic palm flowering only once at the end of its lifetime. Moreover quality palms are harvested prior to flowering for maximum starch yield. Therefore, an alternative breeding approach needs to be explored and developed.

Molecular breeding techniques appear to be the best option to obtain new improved varieties of sago palm. This technique allows researchers to identify and manipulate the potential genes encoding for desired traits, and transform it into the host plant within a shorter time frame. A number of genes involved in the starch biosynthetic pathway have been studied in sago palm (Salleh *et al.*, 2000; Salleh & Lau, 2003; Salleh *et al.*, 2004), however no reports have been published to date concerning the GA 20-oxidase gene of sago. GA 20-oxidase plays an important role in the biosynthetic pathway of growth regulators that control various aspects of plant development, such as seed germination, stem elongation, flower formation and fruit production (Hooley *et al.*, 1994; Swain & Olszewski, 1996; Weiss *et al.*, 1992). This gene was found to be expressed at a high level in leaves compared to expression in the internodes (Garcia-Martinez *et al.*, 1997). A study of GA 20-oxidase in the rice variety IR8 demonstrated that the mutant form of this gene (*sd*-1)

* Author for correspondence:
Bala Jamel. Mailing address: CRAUN Research Sdn. Bhd. Lot 3147, Block 14, Jalan Sultan Tengah 93055 Kuching, Sarawak, MALAYSIA. Tel: +06 082 446489; Fax: +06 082 447385; Email: jbalacraun@yahoo.com.

resulted in greater harvesting index (Jennings, 1964; Walcott & Laing, 1976). This has subsequently led to an increase in the yield production of the crop throughout Asian countries.

Over-expression of this gene has been shown to stimulate growth and flowering (Coles *et al.*, 1999; Huang *et al.*, 1998) and also increased seed and fruit production (Curtis *et al.*, 2000). Another study has shown that the antisense expression of the GA 20-oxidase gene have caused earlier tuberization in potato (*Solanum tuberosum*). The antisense line also produced greater yield. The over-expression line however tuberized only after 30 or more days and the yield was also reduced as compared to control plants (Carrera *et al.*, 2000). The GA 20-oxidase gene is currently used in plant improvement programs in a wide range of species, particularly in crop plants. Thus, genetic manipulation of this gene could be a useful avenue leading towards the generation of a new elite sago variety. For this purpose, the main objectives of this study were to isolate and characterize cDNA coding for GA 20-oxidase in sago palm.

MATERIALS AND METHODS

Plant Material Preparation Leaf (young and mature leaves), trunk and root tissues used in this study were collected from palms in CRAUN Research Experimental Plot, Paya Paloh Research Station, Kota Samarahan, Sarawak. To maintain freshness, tissues were kept cool on ice during the collection process until they were processed in the laboratory. In the laboratory, leaf samples were briefly rinsed with distilled water and dried using clean paper towels. The leaf midriff was then removed. Root samples, however, required thorough cleaning by washing with a large volume of water prior to rinsing with distilled water. After washing, the tissues were blotted dry using clean paper towels. For young leaf base, soft shoot base and trunk tissues, no washing step was required. To facilitate the grinding process, all tissues were cut to approximately 1 cm^2 in size. Samples were then wrapped in an aluminium foil, chilled in liquid nitrogen and stored at -80°C until required. The samples were then used directly for analysis of protein, RNA and DNA.

Construction of Oligonucleotide Primers Prior to generating primers for amplification of the specific target within the sago genome, all available information from the previously published papers regarding GA 20-oxidase genes was studied in detail (Kang *et al.*, 1999; Kusaba *et al.*, 1998; Kusaba *et al.*, 2001; Spielmeyer *et al.*, 2002; Wu *et al.*, 1996). The deduced amino acid sequences of GA 20-oxidase from different plant species were compared. Based on this information, four forward (FAS1, FAS2, FAS3 and FAS8) and two reverse (RS1 and RS2) primers were designed. The primers were constructed commercially after the sequences were designed from published data on the conserved regions of GA 20-oxidase genes (Table 1).

Table 1. Oligonucleotide primers used to amplify the GA 20-oxidase gene from the sago genomic DNA.

Primers	Sequences
FAS1	5'-AACTACTACCCGCCATGC-3'
FAS2	5'-GGCACGGGCCCGCACTGCGAC-3'
FAS3	5'-AACATCGGCGACACCTTC-3'
FAS8	5'-CTCCCATGGAAGGAGACC-3'
RS1	5'-GAAGGTGTCGCCGATGTT-5'
RS2	5'-CGGGCACAGGAAGAACGCCAG-3'

Genomic DNA Isolation, PCR and Probe preparation Genomic DNA was isolated according to the procedure described by Jamel *et al.* (2001). Prior to use, stock solutions of genomic DNA and primers were diluted to a final concentration of 5 ng/µl and 10 µM, respectively. The solutions were stored at -20°C until required. The primers, as described in Table 1, were used to amplify the GA 20-oxidase gene from the sago genomic DNA. Six primer combinations, T1 (FAS1 + RS1), T2 (FAS2 + RS1), T3 (FAS3 + RS2), T4 (FAS1 + RS2), T10 (FAS8 + RS1) and T11 (FAS8 + RS2), were used to amplify the GA 20-oxidase gene from the genomic template. The PCR reaction was performed according to manufacturers' instruction (Qiagen™). The PCR reaction comprised of 10 µl of 5x HotStar HiFidelity PCR Buffer (containing dNTPs), 1 µl forward primer (10mM), 1 µl reverse primer (10 mM) and 1 µl HotStar HiFidelity DNA Polymerase (2.5 units/µl) and 1 µl (5ng) DNA template. Sterile distilled water was added to a final volume of 50 µl. DNA amplification was then carried out in a GeneAmp PCR System 9700, Applied Biosystem, program as follows: one initial denaturation cycle for 4 minutes at 94°C followed by 35 cycles of 94°C for 1 minutes, 60°C for 1 minute and 72°C for 2 minutes. The final step in the last cycle at 72°C was extended for 7 minutes. The PCR products obtained were then separated on a 1.5% agarose gel. DNA sequence was determined according to the procedure described in the following section. PCR product that is homologous with GA 20-oxidase from other plant species was labelled using DIG-labelling system (Roche) and subsequently used as a probe for fishing out the full-length GA 20-oxidase gene from the sago genomic template.

Genome Walking and Cloning of the genomic sequences Purified genomic DNA was digested with several restriction enzymes, *Dra*1, *EcoR*V, *Pvu*II and *Stu*I, provided in the Universal GenomeWalker kit (Clontech, Heidelberg, Germany). Ligation of the genome walker adaptor to the genomic DNA fragment and amplification of the gene was performed according to the manufacturers' instructions. The PCR product carrying the GA 20-oxidase gene fragment was screened using a southern hybridization technique using DIG-labelled probe. The PCR products that showed strong hybridization with this probe was selected for sequence determination.

DNA Sequence Analysis The DNA sequence was determined using a DNA sequencing kit (Big Terminator Cycle Sequencing Ready Reaction Kit, PE-Applied Biosystems) with a DNA sequencer (Model ABI 310, PER-Applied Biosystems). For confirmation of the full length gene sequence, DNA sequencing was also outsourced to First Base Laboratories Sdn Bhd (http//www.base-asia.com/). Homology analysis was carried out using the BLAST program (http://www.ncbi.nlm.nih.gov/blast/). Alignment of amino acid sequences was performed using the ClustalW ver1.81 program (http://www.clustalw. genome.ad.jp).

Total RNA Isolation and cDNA Synthesis Total RNA was isolated from young red leaf tissue according to the procedure described by Jamel *et al*. (2006). After the DNase treatment, 15 to 20 µg of total RNA was used for cDNA synthesis. First strand was prepared using CapFishing Full-Length cDNA Premix kit (Seegene). The entire step was performed in a PCR Thermalcycler (Perkin Elmer).

cDNA Amplification After comparing the candidate genomic DNA sequence with the database, two forward (cDF1 and cDF2) and two reverse (cDR1 and cDR3) primers were synthesized for cDNA amplification (Table 2). The initial reaction was performed using 5'RACE primers included in a reaction mixture comprising 5 µl diluted first-strand cDNAs, 25 µl SeeAmpTM *Taq* Plus Master Mix, 1 µl 10 µM 5'-RACE primer (Forward), 1 µl 10 µM and 3' target primer (Reverse: cDR3) in a 200 µl PCR tube. The volume was made up to 50 µl by the addition of 18 µl sterile distilled water. PCR was carried out for 1 cycle at 94°C for 3 minutes, followed by 35 cycles at 94°C for 40 seconds, 60°C for 40 cycles and 72°C for 1 minute. A final elongation stage was performed at 72°C for 5 minutes. The PCR product obtained was used as template in a second PCR whereby both cDR1 and cDF1 were used as primers.

Table 2. Oligonucleotide sequences used for cDNA amplification.

Primers	Sequences
cDF1	5'-GTAGGTCCCCAACGAGACATG-3'
cDF2	5'-CCATGGTACTCTGCTCTCTTGCT-3'
cDR1	5'-CCAGCCATGTCATCACTGTGGC-3'
cDR3	5'-ACAAGATAAAAGAAATCCCAG-3'
5'RACE	5'-GTCTACCAGGCATTCGCTTCAT-3'

Southern Hybridization Analysis About fifteen micrograms of genomic DNA was digested overnight with several restriction enzymes *BamH*1, *EcoR*1 and *Hind*III. The digested DNA was fractionated on a 1% agarose gel and transferred onto a positively charged nylon membrane (Roche) overnight. Detection was performed using a DIG-detection system according to manufacturers' instruction (Roche). The full-length cDNA fragment that encodes GA 20-oxidase was used as a detection probe.

***Expression of cDNA clone in* E. coli** In order to study gene expression, the full-length GA 20-oxidase gene was cloned in the sense direction into the pPCR-Script Amp SK (+) vector and in the correct reading frame. Gene expression was induced by the addition of Isopropyl-β -D-thiogalactopyranoside (IPTG) to the culture. Protein extraction was performed according to the protocol described by Wu *et al*. (1996) with minor modification. Prior to extraction, the bacterial culture was grown overnight in 2 ml LB-broth containing 50 mg/L ampicillin. The culture was incubated at 37°C overnight. The overnight culture was transferred into a 40 ml LB-broth containing 50 mg/L ampicillin with vigorous shaking at 200 rpm. To induce the expression of the fusion protein, IPTG was added to a final concentration of 5 mM.

The addition of IPTG was conducted when the culture optical density at 600 nm reached 0.5. Then the bacterial culture was incubated for another 2 hours with agitation. The culture was centrifuged at 6000 rpm for 10 minutes. The liquid phase was removed and the remaining bacterial pellet washed with 25 ml LB-broth, followed by centrifugation at 14000 rpm for 10 minutes. The cell pellet was resuspended in 800 µl lysis buffer (100 mM Tris-HCI, pH 8.0, 3 mM DTT, 2.5 mg/ml lysozyme). The mixture was placed at room temperature for 10 minutes, then submerged in liquid nitrogen for 5 minutes. The sample was thawed in an ice bath for 15 minutes. The lysates were centrifuged at 14,000 rpm for 15 minutes and the supernatant transferred into a new sterile microcentrifuge tube, and could then be used directly for an enzyme assay or stored at -30°C (or at -80°C for long term storage) for further analysis.

Production of Antibodies Prior to antibody production, the antigen which in this study, known as residue 173-184, was prepared based on the short peptide sequence of sago GA 20-oxidase gene (C-VHDYFVRTLGEDF). The peptide was used to induce polyclonal antibody production in rabbits. The production of antigen and antibody was outsourced to First Base Laboratories Sdn Bhd. The antibody obtained was stored at -30°C prior to use.

Protein extraction from sago tissue Four different type of tissues, young red leaf, young shoot base, root and trunk, were used for protein extraction. Protein extraction was performed according to the protocol described by Ghesquiere *et al*. (1987) with minor modification. Briefly, 4 g of each tissue type was ground in liquid nitrogen using a pestle and mortar. The powder was transferred into a new sterile mortar which had been pre-cooled using a small amount of liquid nitrogen. Four ml of extraction buffer (2% PVPP, 100 mM DTT, 100 mM L-Cystine and 100 mM Potassium phosphate pH 7.0, 50 mM Herpes pH 7.0) was poured into the powder and mixed using the pestle until even. The homogenate was filtered using a miracloth layer by squeezing into a 40 ml polypropylene tube. The slurry obtained was then transferred into a 2 ml microcentrifuge tube and centrifuged at 14,000 rpm for 20 minutes at 4°C. The supernatant was

transferred into a new sterile 2 ml microcentrifuge tube and stored at -30°C or used directly for SDS-PAGE.

Western hybridization and Immunological detection Two sets of protein samples were prepared and then separated on 15% SDS-PAGE minigel. One set of protein samples was stained with coomassie brilliant blue. Another set was used directly for Western transfer and did not require the staining step. Protein transfer onto the PVDF membrane was performed using Mini Trans-Blot (Bio-Rad) overnight. Immunological detection was carried out according to manufacturer's instruction (Bio-Rad).

RESULTS

DNA Amplification and Probe Preparation Six primer combinations, T1, T2, T3, T4, T10 and T11, were used to amplify the GA 20-oxidase gene from the genomic template. Good PCR products were obtained from each primer combination tested. DNA amplification using primer combinations of T1, T2, T3, T10 and T11 produced a clear single band with estimated sizes of 200, 800, 550, 500 and 550 bp, respectively, while DNA amplification using T4 combination produced three bands with sizes of 290, 650 and 800bp (Figure 1). To determine which PCR product carries the DNA sequence that encodes for GA 20-oxidase, the PCR products obtained were purified using PCR purification kit (Stratagene). Out of the six primer combination tested, T1 and T10 gave PCR products close to the expected size i.e., 200 and 500 bp, respectively. To confirm, the sequence for these fragments needed to be determined. Prior to sequence determination, the PCR product was first purified and subsequently cloned into pPCR-Script Amp SK (+) vector after which the recombinant plasmid was transformed into *E. coli*, DH5α strain. The recombinant plasmid DNA was then isolated from each clone and used as template for DNA sequencing.

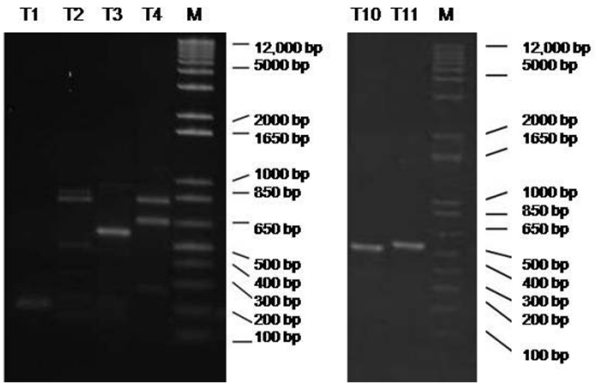

Figure 1. PCR products obtained using primer combinations T1, T2, T3, T4, T10 and T11. Genomic DNA obtained from young red leaf was used as template. PCR products obtained were named as follows: T1 = 200 bp; T2 = 850 bp; T3 = 550 bp; T4 = 290, 650, 800 bp; T10 = 500 bp and T11 = 600 bp.

Sequence Determination of Probe The DNA sequence obtained was used for similarity analysis using BLAST program mentioned earlier. BLAST result showed that, both clones T1 and T10 were homologous to GA 20-oxidase gene from other plant species. For the isolation of full-length gene, the T10 fragment, which is 500 bp long, was selected for use as a detection probe. The probe was labelled with DIG-labelling system according to manufacturers' instruction (Roche) and used for detection of GA 20-oxidase fragment obtained from a genome walking trial.

Isolation of GA 20-oxidase fragment from genomic template Based on sequence information of the GA 20-oxidase gene from other plants species, several sets of oligonucleotide primers were designed and synthesized for cloning the upstream and downstream sequences adjacent to the 500 bp fragment obtained previously. PCR amplification, using oligonucleotide primers obtained above, produced good bands. Southern blot was carried out to identify the fragment that encoded for GA 20-oxidase. This led to the detection of a single fragment with a size of 1150 bp. The DNA fragment was cloned into a pPCR-Script Amp SK (+) vector followed by cloning into *E. coli*, DH5α. The plasmid was isolated and sent to First Base Laboratories for sequence determination. BLAST analysis using the sequence obtained showed that this fragment encoded for GA 20-oxidase. The DNA sequence obtained was used for the construction of primers and for cloning the fragment adjacent to the 5' end of the 1150 bp fragment. As a result, two other fragments with estimated sizes about 700 and 500 bp were obtained. Both fragments showed homology with the GA 20-oxidase gene from other plant species. Full-length sequence for this gene was finally obtained by combining the sequences of all three fragments (Figure 2). Sequence alignment with deduced amino acids of GA 20-oxidases from other plant species showed that the gene has highly conserved regions. This region was finally used to design primers for cloning of full-length cDNA from sago.

Isolation of a full-length GA 20-oxidase cDNA clone The nucleotide sequence of the GA 20-oxidase gene obtained from the genomic template was used to construct primers for the isolation of a full-length gene from the cDNA template. Based on the genomic sequence obtained, two forward (5'-GTAGGTCCCCAACGAGACATG-3', cDF1, and 5'-CCATGGTACTCTGCTCTCTTGCT-3', cDF2) and two reverse (cDR1: 5'-CCAGCCATGTCATCACTGTGGC-3', and cDR 3: 5'-ACAAGATAAAAGAAATCCCAG-3') primers were constructed. A 5'RACE (5'-GTCTACCAGGCATTCGCTTCAT-3') 0primer from the manufacturer were also included for the amplification of the full-length gene from the cDNA template. In this study, the first cDNA strand was generated from total RNA sample isolated from young red leaf tissue using CapFishing Full-Length cDNA Premix Kit obtained from Seegene. To isolate the full-length cDNA encoded GA 20-oxidase, a direct PCR reaction was conducted, whereby 5'RACE (forward)

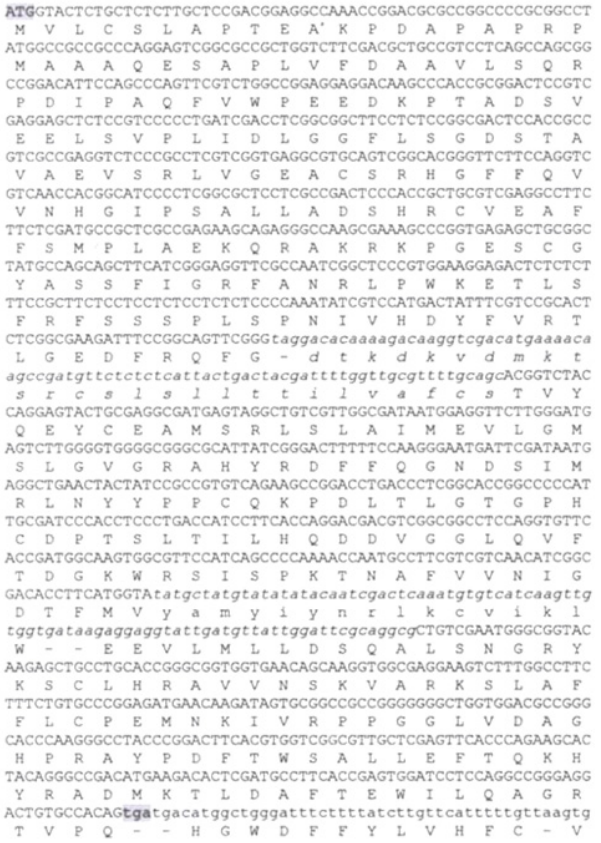

Figure 2. Nucleotide and deduced amino acid sequence for complete genomic sequence of GA 20-oxidase cDNA after combining fragments 1, 2 and 3. The start codon "ATG" and the stop codon "tga" are shaded in black. Complete nucleotide sequence for cDNA of this gene is shown in Figure 4.

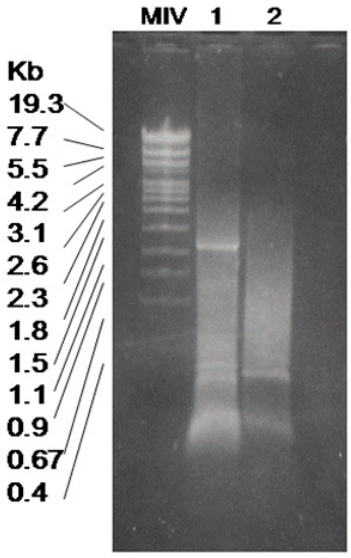

Figure 3. cDNA profile after PCR amplification. The first cDNA strand was generated using Seegene kit. cDNA was then amplified with GSP-cDF1 and cDR3 primers. The size of the PCR product was estimated to be around 1200 bp. MIV: Marker IV; Lane 1: cDNA after amplification using GSP-cDF1 and cDR3; Lane 2: First strand cDNA.

Table 3. Comparison of sago *Ms20ox* with GA 20-oxidase from other plant species.

Species	Common Name	Homology (%)	Identity (%)	Genebank Accession No.
Triticum aestivum	Wheat	61	76	Y14008
Zea mays	Maize	66	67	BT038900
Solanum tuberosum	Potato	55	62	AJ291454
Oryza sativa	Rice	63	52	AB077025
Lolium perenne	Perennial ryegrass	62	74	DQ071620
Lactuca sativa	Garden lettuce	50	63	AB012204
Beta vulgaris	Common beet	44	63	AJ422049
Arabidopsis thaliana	*Arabidopsis*	60	62	1581592

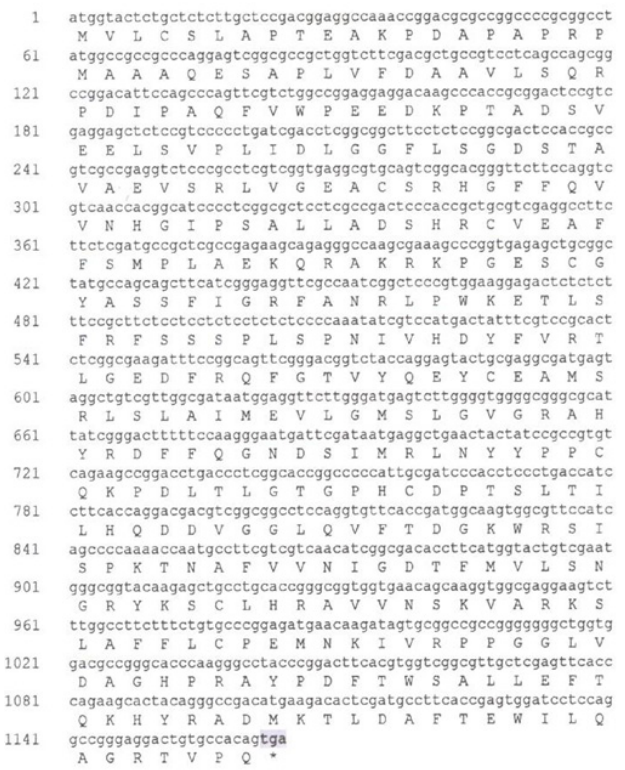

Figure 4. Nucleotide and deduced amino acid sequence for *Ms20ox*. Stop codon "tga" is shaded.

and GSP-cDR3 (reverse) primers were used. To obtain a specific PCR product a second PCR reaction was conducted using another set of GSP primers: reverse cDR1 and forward cDF1 primers. The PCR product obtained earlier was used as a template. One single band was obtained when the PCR was conducted using cDF1 and cDR1 primers. The size of the PCR product was estimated to be between 1100 and

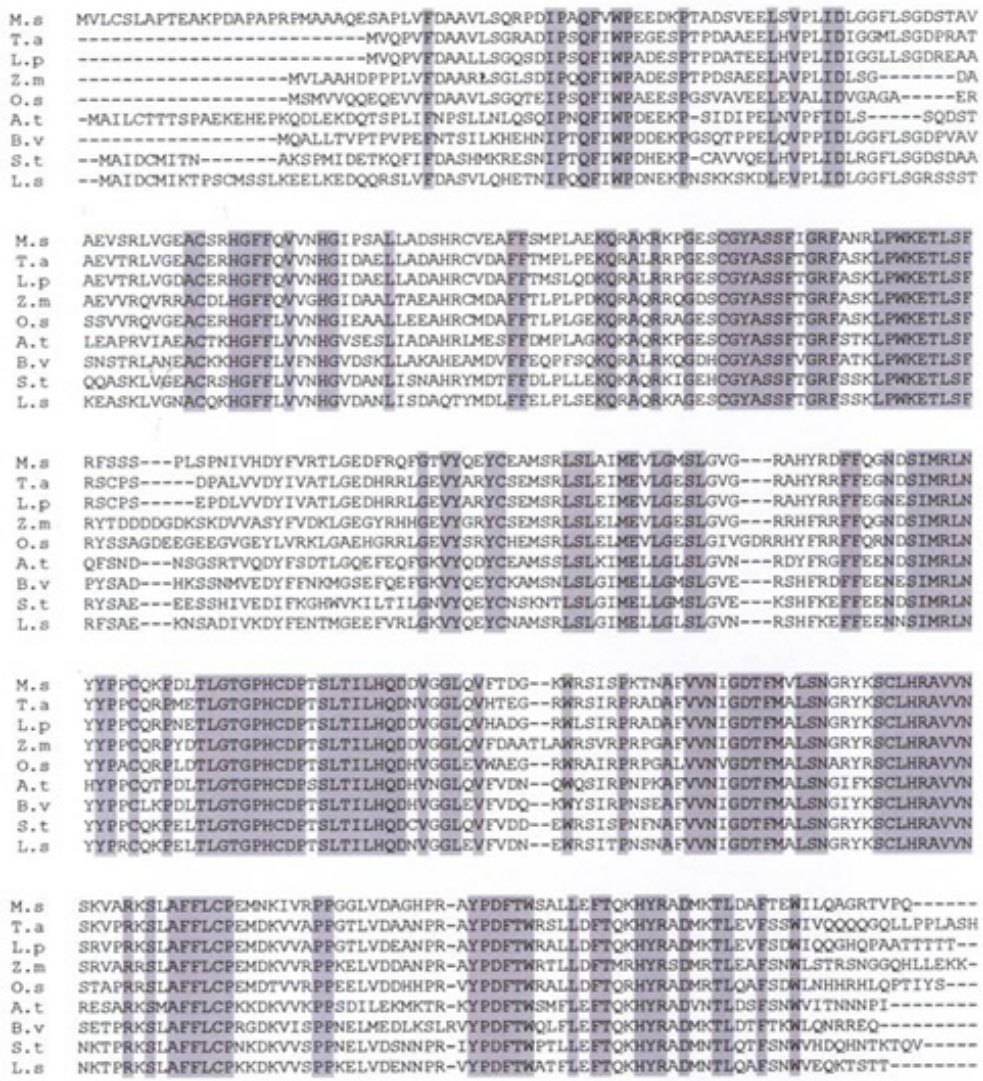

Figure 5. Alignment of deduced amino acid sequence for sago *Ms20ox* cDNA with GA 20-Oxidases from other plant species. Regions identical to all GA 20-Oxidases are shaded. Dashes were introduced for maximum sequence homology. M.s (*Ms20ox*), T.a. (*Triticum aestivum*), L. p (*Lolium perenne*), Z. m (*Zea mays*), O. s (*Oryza sativa*), A. t (*Arabidopsis thaliana*), B. v (*Beta vulgaris*), S. t (*Solanum tuberosum*) and L. s (*Lactuca sativa*).

1300 Kb (Figure 3).

The PCR product obtained was subsequently analyzed with southern hybridization whereby a T10 probe was used as a detection probe. Southern hybridization data showed that the PCR product obtained exhibited strong hybridization with a T10 probe (Data not shown). This observation suggested that the fragment obtained carries a DNA fragment that encodes for GA 20-oxidase gene. The PCR product was cloned into pPCR-Script Amp SK (+) vector and transformed into *E. coli* DH5α. The plasmid was sequenced by First Base Laboratories Sdn Bhd for sequence determination (Figure 4). The sequence obtained was compared with genomic sequence of the GA 20-oxidase gene using ClustalW ver1.81. This analysis showed that the GA 20-oxidase gene, *MsGenom20ox*, comprised 3 exons and 2 introns. The 3 exons are 1161 bp long and encode for 387 amino acid residues with a Mr of 42 kDa. The size of the first, second and third exons were 567, 324 and 270 bp, respectively. Meanwhile, the first intron occurs at base pair 568 bp from the start codon and is 84 bp long. The second intron is located at 975 bp and is 87 bp long. This result was in accordance with GA 20-oxidase identified from another species, *Arabidopsis thaliana* (Xu *et al.*, 1995).

Homology analysis of full-length GA 20-oxidase cDNA

The nucleotide sequence for *Ms20ox* was translated into a deduced amino acid sequence. The amino acid sequence obtained was compared with GA 20-oxidase from other plant species, *Triticum aestivum* (GenBank Acc: Y14008); *Lolium perenne* (DQ071620); *Zea mays* (BT038900); *Oryza sativa* (AB077025); *Arabidopsis thaliana* (1581592); *Beta vulgaris* (AJ422049); *Solanum tuberosum* (AJ291454) and *Lactuca sativa* (AB012204). The sequence was analyzed using ClustalW ver1.81 program, which is available online. The sequence alignment data showed that the *Ms20ox* has homology to the GA 20-oxidase gene from other plant species.

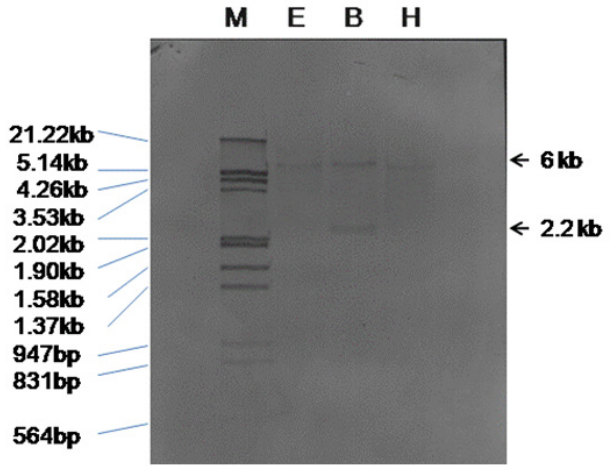

Figure 6. Southern hybridization result for sago genomic DNA digested with three different enzymes, *EcoR*I (E), *BamH*I (B) and *Hind*III (H). M: DIG-labelled Marker III and MIV: unlabelled DNA Marker IV. DIG-labelled full-length cDNA (M.s) of GA 20-oxidase was used as probe.

The nucleotide sequence homology and amino acid identity between sago *Ms20ox* and GA 20-oxidase from other plant species ranged from 44 to 66% and 52 to 76% respectively (Table 3). Thus, *Ms20ox* may encode for GA 20-oxidase in sago palm. It was observed that the sequence Leu-Pro-Trp-Lys-Glu-Thr (LPWKETLSF), at residue 152-160, is highly conserved in all GA 20-oxidase genes cloned so far (Figure 5). This sequence, however, did not show high homology with other 2-oxoglutarate-dependent dioxygenases. A related enzyme, GA 3β -hydroxylase, does not contain this motif. It was proposed that this motif may be involved in binding of the GA substrate (Xu *et al.*, 1995). Other regions that were highly conserved were HGFF, SIMRLN, TLGTGPH-CDP, SLTILHQ, CGYASSF, SCLHRAVVN, AFFLCP and YPDFTW (Figure 5).

Southern analysis of the GA-oxidase gene DNA-blot analysis was conducted to identify the copy number of the *Ms20ox* gene within the sago genome. The blot was hybridized with a cDNA probe against the genomic template digested with *EcoR*I, *BamH*I and *Hind*III. One band of estimated size around 6kb was obtained for samples digested with *EcoR*I and *Hind*III. Meanwhile, samples digested with *BamH*I produced two bands of estimated size around 2.2 and 6 kb. This indicated that there are one to two copies of this gene in sago genome. This result was similar to the work by Kang *et al.* (1999).

Another finding from the Southern blot result was that, when digested with *BamH*I, the gene was cut at one site and divided into two fragments, 2.2 and 6 kb in length, as supported by nucleotide sequence analysis which showed that there is only one cut site, at position 1303 bp, for *BamH*I in this gene. Digestion with *EcoR*I and *Hind*III, however, only produced one band each, as shown in southern blot result (Figure 6). This suggested that both of these enzymes have recognition sites only outside the gene region.

Protein extraction from sago palm tissue In this study, various types of sago tissue, from young red leaf, mature leaf, trunk and root, were initially tested for protein extraction. Various protocols described by Ainsworth *et al.* (1995) and Ghesquiere *et al.* (1987) failed to produce useful protein products for these tissues. This study demonstrated that the selected sago tissues contained an extremely low quantity of protein. Modification was made for each trial but the method still failed to produce good protein profiles even after various changes were made. Buffer volume was reduced from 6 to 4 ml for 4 g of tissue. The extraction was strictly conducted at low temperature. Extraction buffer was pre-cooled on ice prior to use, while the pestle was pre-cooled using liquid nitrogen. However the extraction was still unsuccessful.

A different tissue, the soft shoot base, was finally tested for protein extraction. Unlike the other types of sago tissue, high concentration of protein and a good protein profile was obtained from this tissue (Figure 7). The concentration of protein obtained was determined using the Bradford pro tein assay (Bradford, 1976). This technique demonstrated that the amount of crude protein obtained from soft shoot base tissue could reach as high as 571 µg per gram. This is much higher then the amount of crude protein obtained from trunk, leaf and root tissues, which produced about 60, 49 and 41 µg per gram, respectively. This study suggested that the amount of crude protein obtained from soft shoot base tissue was about 10 times higher compared to either leaf (young red leaf or mature), trunk or root tissues. The protein obtained was separated on SDS-PAGE and used for Western hybridization.

Western Hybridization Western blotting showed that protein samples obtained from sago tissue and its sense clone have produced one band with a molecular weight estimated at around 32 kDa. The protein obtained from the antisense clone and negative control (without gene) however did not show the presence of a 32 kDa band. The presence of a 32 kDa band in both samples, sago soft shoot base tissue (band A) and fusion protein (Band B) suggests that, the *Ms20ox*

clone has been successfully expressed in pPCR Script-Amp SK (+) Vector. The presence of bands A (Figure 7a: lanes 1, 2 and 3) and B (Figure 7a: lane 4) could be seen in the SDS-PAGE result. The protein band for sample *Ms20ox* in lane number 4 showed higher intensity compared to samples from the antisense clone (Figure 7a: lane 5) and without insert (Figure 7a: lane 6). The size obtained was close to 42 kDa, estimated earlier using ExPasy Protparam software (http://au.expasy.org/tools/protparam.html). This finding was consistent over several repetitions. The three samples which were obtained from young shoot base tissue produced 32 kDa bands when hybridized with a polyclonal antibody probe. A few bands of more than 50 kDa were observed in each sample extracted from *E. coli* (both sense and antisense clones). The cause is unknown. These bands were still present even after several repetitions using varying antibody concentration (antibody:buffer ratios; 1: 200, 1: 400, 1:

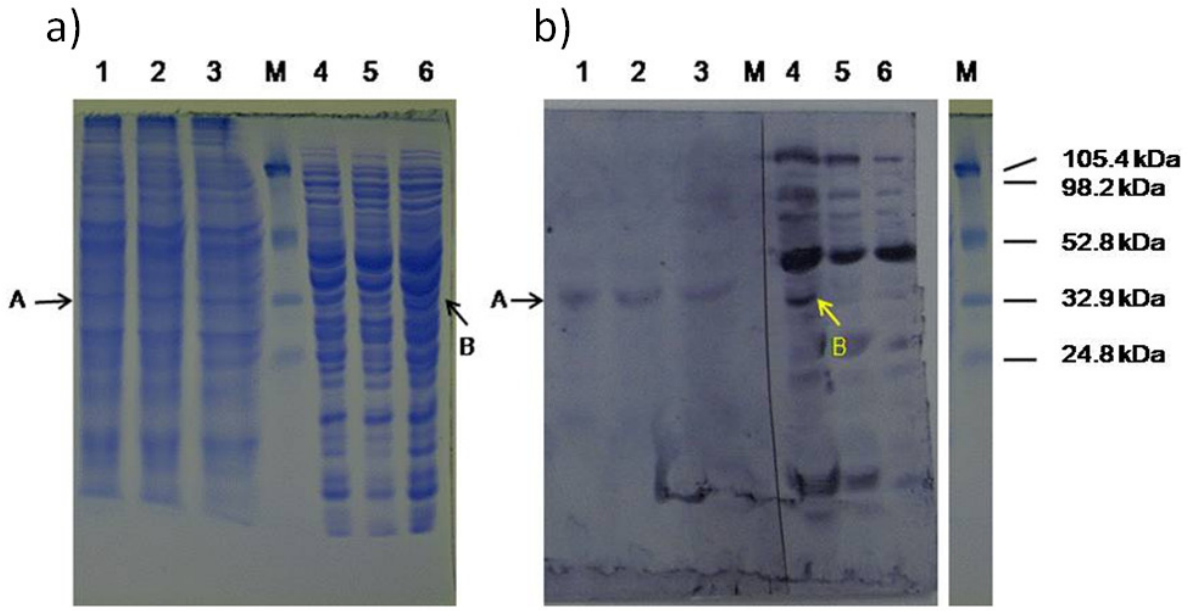

Figure 7. a) Protein profile on 15% SDS-PAGE. Samples 1, 2 and 3 were for protein extracted from young shoot base tissue. Lane 4 is protein extracted from *E. coli* that carries sense strand of GA 20-oxidase gene. Lane 5: protein extracted from *E. coli* in antisense direction and Lane 6: *E. coli* without insert. **b)** The protein was transferred onto the PVDF membrane and hybridized with polyclonal antibody obtained from rabbit. Bands A and B were detected for protein extracted from sago tissue and *E. coli* that carries sense direction of GA 20-oxidase gene. No 32 kDa band was observed for both antisense (Sample 5) and without insert colonies (Sample 6).

1000, and 1: 2000). In this study an antibody dilution of 1: 400 produced the best band resolution. The advantage of using polyclonal antibody is that it allows the partial or-full gene to be detected. The use of a polyclonal antibody would enable the detection of gene expression in a bacterial host. Thus, clones which contain a partial or full GA 20-oxidase gene can be detected. This is because polyclonal antiserum is usually able to react with several epitopes, which might be encoded by different regions of the gene of interest (Sambrook *et al.*, 1989). On the other hand, monoclonal antibodies are able to react with only one epitope, which might enable the detection of only a particular subset of recombinants expressing the gene of interest (Sambrook *et al.*, 1989).

DISCUSSION

DNA amplification using six primer combinations have produced good PCR products of which the 500 bp fragment was obtained. The nucleotide sequence for this fragment has been shown to carry the gene encoding GA 20-oxidase. Based on the nucleotide sequence obtained several sets of Gene Specific Primer (GSP) were constructed, which were subsequently used for the amplification of the entire fragment of the GA 20-oxidase gene. By using a Genome walking technique, three partial fragments of this gene, with sizes estimated to be 1150 bp, 700 bp and 500 bp, were obtained. From the DNA sequence obtained, it was observed that these fragments were adjacent to each other. The size of the full-length gene obtained from the genomic template was determined by combining the nucleotide sequences obtained from the three fragments. BLAST analysis using the full length gene sequence showed high homology with GA 20-oxidases from other plant species. The deduced amino acid sequence for the genomic GA 20-oxidase gene was determined and used for multiple sequence alignment analysis. The multiple sequence alignment result clearly indicated the intron-exon region. This information was used as a guide in the construction of another set of GSP primers for the amplification of the full-length cDNA (*Ms20ox*) fragment.

This study revealed that, the sago GA 20-oxidase is encoded by a relatively small size gene. The size of the sago GA 20-oxidase gene, inclusive of introns, is 1332 bp, while the cDNA is 1161 bp long. This is in line with an earlier study which showed that GA 20-oxidase is encoded by a small multigene family (Garcia-Martinez *et al.*, 1997; Phillips *et al.*, 1995; Rebers *et al.*, 1999; Xu *et al.*, 1995; Wu *et al.*, 1996). This study also showed that the genomic GA 20-oxidase (*MsGenom20ox*) gene comprised three exons and two intronic regions. This was similar to the gene structure for the GA 20-oxidase gene obtained from other plant species such as *Arabidopsis thaliana* (Xu *et al.*, 1995), barley (Jia *et al.*, 2009) and sunflower (*Helianthus annuus*) (Carzoli *et al.*, 2009). The three exons are 1161 bp long in total, and encode for a product which is 387 amino acid residues long with a M_r of 42 kDa.

GA 20-oxidase genes that possess one intron, and containing no introns have both been reported: *OsGA20ox3* (and *OsGA20ox7*) and *OsGA20ox1*, respectively (Han & Zhu, 2011). *OsGA20ox1*, which has no introns, illustrated

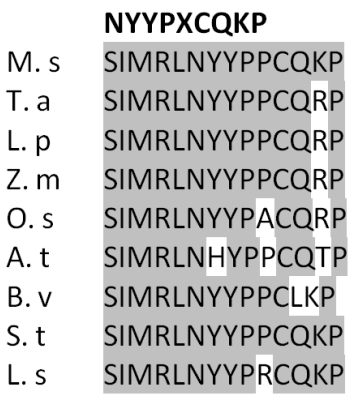

Figure 8. The consensus sequence NYYPXCQKP that has been postulated to be involved in binding the 2-oxoglutarate cofactor. M.s (*Ms20ox*), T.a (*Triticum aestivum*), L. p (*Lolium perenne*), Z. m (*Zea mays*), O. s (*Oryza sativa*), A. t (*Arabidopsis thaliana*), B. v (*Beta vulgaris*), S. t (*Solanum tuberosum*) and L. s (*Lactuca sativa*).

that this gene probably arose from retrotransposon-based random insertions. Alternatively, intron losses may have occurred over the course of intron evolution (Han & Zhu, 2011). GA 20-oxidase genes without introns have so far not been found in soybean or *Arabidopsis* (Han & Zhu, 2011). BLAST analysis showed that the full-length *Ms20ox* cDNA shares homology with GA 20-oxidase from wheat and maize. This strongly suggests that *Ms20ox* is a GA 20-oxidase cDNA from sago palm. Moreover, western blot results showed that the protein samples, obtained from *E. coli* cells that carry the *Ms20ox* gene, exhibit strong hybridization when rabbit antiserum was used as probe. The 32 kDa band was observed for protein obtained from *E. coli* strain DH5α that carries a sense clone of the *Ms20ox* gene. Although the size is lower than the 42 kDa predicted earlier, this might be due to degradation (Curtis *et al.*, 2000) or proteolysis reactions (Fagoaga *et al.*, 2007) of the major band, because the sample had been frozen and thawed several times. This band was not observed in the antisense and negative control (without *Ms20ox*) samples. The presence of a 32 kDa band in both samples, sago soft shoot base tissue (band A) and fusion protein (Band B) suggests that the *Ms20ox* clone had been successfully expressed in pPCR Script-Amp SK (+) Vector.

However, evidence on the functional activity of the protein is still needed, which requires the use of HPLC-GCMS facility that is equipped with a radioactive detection component (Coles *et al.*, 1999; Fagoaga *et al.*, 2007; Hedden & Wu *et al.*, 1996; Kamiya, 1997; Kang *et al.*, 1999; Niki *et al.*, 2001; Xu *et al.*, 2002). This method could not be carried out due to lack of a facility equipped for radioactive work.

Meanwhile, from the multiple sequence alignment data, it was observed that, the LPWKETLSF motif, which is located within residues 152 to 160, is highly conserved in all GA 20-oxidase clones obtained so far. Xu and co-workers (1995) proposed that this motif may be involved in the binding of the GA substrate (Xu *et al.*, 1995). On the other hand, motif NYYPXCQKP which was previously postulated to be involved in binding the 2-oxoglutarate cofactor in GA biosynthetic pathway (Roach *et al.*, 1995) is also present in sago (Figure 8). Other motifs that were highly conserved are TLGTGPHCDP, SCLHRAVVN, SLTILHQ, CGYASSF, AFFLCP, YPDFTW and HGFF. The specific functions of these motifs are still unknown.

ACKNOWLEDGEMENTS

We wish to thank CRAUN Research Sdn Bhd for funding this project and Dr. Zaliha C. Abdullah, Senior Research Fellow, CRAUN Research, for her interest and encouragement in this work, as well as for advice in manuscript preparation.

REFERENCES

Ainsworth, C., Hosein, F., Tarvis, M., Weir, F., Burrel, M., Devos, K.M. and Gale, M.D. 1995. Adenosine diphosphate glucose pyrophosphorylase genes in wheat: differential expression and gene mapping. *Planta* 197(1): 1-10.

Bradford, M.M. 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. *Anals of Biochemistry* 72: 248-254.

Carzoli, F.G., Michelotti, V., Fambrini, M., Salvini, M. and Pugliesi, C. 2009. Molecular cloning and organ specific expression of two gibberellin 20-oxidase genes of *Helianthus annuus. Plant Molecular Biology Reports* 27:144–152.

Carrera, E., Bou, J., Martinez, J.L.G. and Prat, S. 2000. Changes in GA20-oxidase gene expression strongly affect stem length, tuber induction and tuber yield of potato plants. *The Plant Journal* 22(3): 247-256.

Chulavatnatol, M. 2002. Starch utilization in Asia. In *New Frontiers of Sago Palm Studies* (Ed: Kainuma, K., Okazaki, M., Toyoda, Y. and Cecil, J.E.) p. 9-14. Tokyo Japan, Universal Academy Press Inc.

Coles, J.P., Philips, A.L., Croker, S.J., Garcia-Lepe, R., Lewis, M.J. and Hedden, P. 1999. Modification of gibberellin production and plant development in *Arabidopsis* by sense and antisense expression of gibberellin 20-oxidase gene. *Plant Journal* 17:547-556.

Curtis, I.S., Ward, D.A., Thomas, S.G., Philips, A.L., Davey, M.R., Power, J.B., Lowe, K.C., Croker, S.J., Lewis, M.J. and Magness, S.L. 2000. Induction of dwarfism

in transgenic *Solanum dulcamara* by overexpression of a gibberellin 20-oxidase cDNA from pumpkin. *Plant Journal* 23:329-338.

Fagoaga, C., Tadeo, F.R., Iglesias, D.J., Huerta, L., Vidal, A.M., Talon, M., Navarro, L., Garcia-Martinez, J.L. and Pena, L. 2007. Engineering of gibberellins levels in citrus by sense and antisense overexpression of a GA 20-oxidase gene modifies plant architecture. *Journal of Experimental Botany* 58: 1407-1420.

Garcia-Martinez, J.L., Lopez-Diaz, I., Sanchez-Beltran, M.J., Phiilips, A.L., Wardd, D.A., Gaskin, P. and Hedden, P. 1997. Isolation and transcript analysis of gibberellin 20-oxidase genes in pea and bean in relation to fruit development. *Plant Molecular Biology* 33: 1073-1084.

Ghesquiere, M., Barcelos, E., Santos, D.M. and Amblard, P. 1987. Enzymatic polymorphism in *Elaies guineensis* H. B. K (*E. melanococca*). Analysis of population in the Amazone Basin. *Oleagineux* 42(4): 151-153.

Han, F. and Zhu, B. 2011. Evolutionary analysis of three gibberellin oxidase genes in rice, *Arabidopsis* and soybean. Gene 473(1): 23-35.

Hedden, P. and Kamiya, Y. 1997. Gibberellin biosynthesis: enzymes, genes and their regulation. *Annual Review of Plant Physiology and Plant Molecular Biology* 48: 431-460.

Hooley, R. 1994. Gibberellins: Perception, transduction and response. *Plant Molecular Biology* 26: 1529-1555.

Huang, S., Raman, A.S., Ream, J.E., Fujiwara, H., Cerny, R.E. and Brown S.M. 1998. Over-expression of 20-oxidase confers a gibberellins over-production phenotype in *Arabidopsis*. *Plant Physiology* 118: 773-781.

Jamel B., Morshidi M. and Salleh M.A. 2001. Identification of molecular markers in sago palm (*Metroxylon sagu*) using polymerase chain reaction. *Asia Pacific Journal of Molecular Biology and Biotechnology* 9(1): 71-74.

Jamel, B., Noraini, B., Salleh, M.A. and Hussain, M.H. 2006. Total RNA isolation from sago palm. *CRAUN Sago Research Journal* 2: 165-169.

Jennings, P.R. 1964. Plant type as a rice breeding objective. *Crop Science* 4: 13-15.

Jia, Q., Zhang, J., Westcott, S., Zhang, X.Q., Bellgard, M., Lance, R. and Li, C. 2009. GA-20 oxidase as a candidate for the semidwarf gene *sdw1/denso* in barley. *Functional Integrative Genomics* 9(2): 255-262.

Kang, H.G., Jun, S.H., Kim, J., Kawaide, H., Kamiya, Y. and An, G. 1999. Cloning and molecular analysis of a gibberellin 20-oxidase gene expressed specifically in developing seeds of watermelon. *Plant Physiology* 121: 373-382.

Niki, T., Nishijima, T., Nakayama, M., Hisamatsu, T., Oyama-Okubo, N., Yamazaki, H., Hedden, P., Lange, T., Mander, L.N. and Koshioka, M. 2001. Production of dwarf lettuce by overexpressing a pumpkin gibberellin 20-oxidase gene. *Plant Physiology* 126: 965-972.

Phillips, A.L., Ward, D.A., Ukes, S., Appleford, N.E.J., Lange, T., Huttly, A., Gaskin, P., Graebe, J.E. and Hedden, P. 1995. Isolation and expression of three gibberellin 20-oxidase cDNA clones from *Arabidopsis*. *Plant Physiology* 108: 1049-1057.

Roach, P.L., Clifton, I.J., Fulop, V., Harlos, K., Butron, G.J., Hajdu, J., Andersson, I., Schofield, C.J. and Baldwin, J.E. 1995. Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* 375: 700–704.

Rebers, M., Kaneta, T., Kawaide, H., Yamaguchi, S., Sekimoto, H., Imai, R. and Kamiya, Y. 1999. Regulation of gibberellin 20-oxidase gene during flower and early fruit development in tomato. *Plant Journal* 17: 241-250.

Salleh, M.A., Jamel, B.A., Bong, S.K. and Lau, J.S.K. 2000. Molecular studies on the starch biosynthesis pathway of the sago palm, *Metroxylon sagu*. In Yoshida *et al.* (Eds.) *Biotechnology for Sustainable Utilisation of Biological Resources in the Tropics*. 14: 8-15.

Salleh, M.A. and Lau, J.S.K. 2003. Identification and characterization of a genomic DNA sequence coding for granule-bound starch synthase in sago palm, *Metroxylon sagu*. In *Biotechnology for Sustainable Utilization of Biological Resources in the Tropics* (Ed.: Murooka, Y.) 16: 43-50.

Salleh, M.A., Lau, J.S.K., Jamel, B. and Hwang, S.S. 2004. Characterization of the starch biosynthesis pathway in sago palm: Isolation of genes coding for the starch synthase, ADP-glucose Pyrophosphorylase, starch branching enzyme and starch debranching enzyme. In *Biotechnology for Sustainable Utilization of Biological Resources in the Tropics* (Seki, T, *et al.*, Eds.) 17: 54-59.

Sambrook, J., Maniatis, T. and Fritsch, E.F. 1989. Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory, New York.

Spielmeyer, W., Ellis, M.H. and Chandler, P.M. 2002. Semidwarf (*sd*-1) "green revolution" rice contains a defective

gibberellin 20-oxidase gene. *Proceedings of the National Academy of Sciences of the United States of America* 99(13): 9043-9048.

Stanton, W.R. 1972. Report to the government of Sarawak on a reconnaissance of the sago industry in Sarawak. August 10-25, 1972.

Swain, S.M. and Olszewski, N.E. 1996. Genetic analysis of gibberellins signals transduction. *Plant Physiology* 112: 11-17.

Tie, Y.L., Loi, K.S. and Lim, E.T.K. 1991. The geographical distribution of sago (Metroxylon spp.) and the dominant sago growing soil in Sarawak. In Towards Greater Advancement of the Sago Industry in the 90's. *Proceedings of the Fourth International Sago Symposium* 6-9 August, 1990. Kuching, Sarawak, Malaysia. pp 36-45.

Walcott, J.J. and Laing, D.R. 1976. Some physiological aspects of growth and yield in wheat crops: a comparison of a semidwarf and standard height cultivar. *Australian Journal of Experimental Agriculture and Animal Husbandry* 16: 578-587.

Weiss, D., van Blockland, R., Kooter, J.M., Mol, J.N.M. and van Tunen, A.J. 1992. Gibberelic acid regulates chalcone synthase gene transcription in the corolla of the *Petunia hybrida*. *Plant Physiology* 98: 191-197.

Wu, K., Li, L., Gage, D.A. and Zeevaart, J.A.D. 1996. Molecular cloning and photoperiod-regulated expression of gibberellins 20-oxidase from the long-day spinach. *Plant Physiology* 110: 547-554.

Xu, Y., Li, L., Wu, K., Peeters, A.J.M., Gage, D.A. and Zeevaart, J.A.D. 1995. The GA5 locus of *Arabidopsis thaliana* encodes a multifunctional gibberellin 20-oxidase; molecular cloning and functional expression. *Proceedings of the National Academy of Sciences of the United States of America* 92: 6640-6644.

Xu, J., Lange, T. and Altpeter, F. 2002. Cloning and characterization of a cDNA encoding a multifunctional gibberellin 20-oxidase from perennial ryegrass (*Lolium perenne* L.). *Plant Science* 163: 147-155.