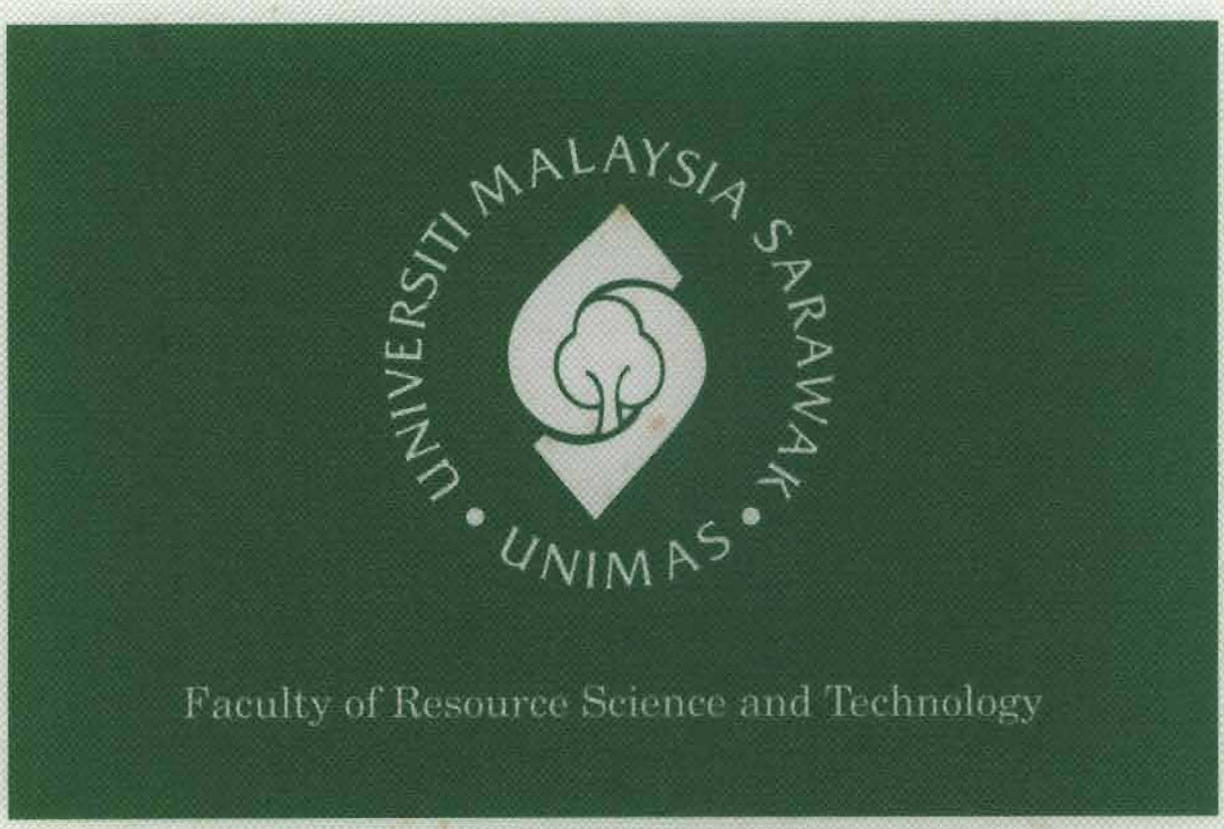

CLONING AND MOLECULAR CHARACTERIZATION OF DNA FRAGMENTS OBTAINED FROM 3' - RACE OF STARCH BRANCHING ENZYME FROM *M. sagu*

Tan Whei Shien

Bachelor of Science with Honours
(Resource Biotechnology)
2009

Cloning and Molecular Characterization of DNA Fragments Obtained from 3'-RACE of Starch Branching Enzyme from *M. sagu*

Tan Whei Shien

This project is submitted in partial fulfillment of the requirements for the degree of
Bachelor of Science with Honours
(Resource Biotechnology)

Department of Molecular Biology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

2009

ACKNOWLEDGEMENT

First, I would like to express my greatest gratitude to my supervisor, Dr. Mohd. Hasnain bin Md. Hussain for his guidance, advice and dedication throughout the project. Besides, I would like to thank Mr. Muhammad Suhaid Mat Hussin (MSc. student), Ms. Anastasia Shera Edward (MSc. student) and Ms. Wee Ching Ching (MSc. student) for their guidance and advice in completing this project. In addition, my special thanks to my lab mates, especially Nikson, Kavitha, Akidah and Amirah for their support and encouragement throughout this project. Last but not least, I would like to thank my family and friends for their moral support and encouragement towards the completion of this project.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

cDNA	Complementary deoxyribonucleic acid
CTAB	Cetyltrimethyiammonium bromide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediamine tetraacetic acid
GSP	Gene specific primer
IPTG	Isopropyl-beta-D-thiogalactopyranoside
KAc	Potassium acetate
LB broth	Luria-bertani broth
LiCl	Lithium chloride
M-MuLVRT	Moloney Murine Leukimia Virus reverse transcriptase
mRNA	Messenger ribonucleic acid
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
RACE-PCR	Rapid amplification of cDNA ends-polymerase chain reaction
RNA	Ribonucleic acid
RT-PCR	Reverse-transcriptase polymerase chain reaction
SBE	Starch branching enzyme
SDS	Sodium dodecyl sulphate
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranosid

LIST OF FIGURES

LIST OF TABLES

Cloning and Molecular Characterization of DNA Fragments Obtained from 3'-RACE of Starch Branching Enzyme from *M. sagu*

Tan Whei Shien

Resource Biotechnology
Department of Molecular Biology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

The complete genome sequence of starch branching enzyme (SBE) in *Metroxylon sagu* is not yet established. The main aim of this study is to amplify the flanking 3'-region of partial coding sequence of SBE I in *M. sagu* using 3'-RACE PCR. RACE PCR utilizes gene specific primer (GSP) and oligo(dT) primer.. The RNA from *M. sagu* leaf was extracted and reverse-transcribed into cDNA using M-MuLV reverse transcriptase. The second strand was then amplified using 3'-RACE PCR method. PCR product was purified using Vivantis GF-1 PCR Clean-up kit. Cloning of PCR product was carried out using *E. coli* XL 1 Blue and pGEM-T Easy Vector Systems. The plasmid was extracted using mini-prep isolation method and Vivantis GF-1 Plasmid Extraction kit and digested by *EcoR1*. Colony PCR was performed for indication of inserted gene. In this study, the transformation and cloning of PCR product failed as the transformants were contaminated with *Bacillus* and the sequence blasted showed the presence of xylanase instead of SBE.

Keywords: 3'-RACE PCR, *Metroxylon sagu*, cloning

ABSTRAK

Susunan keseluruhan genetik 'starch branching enzyme' (SBE) dalam Metroxylon sagu masih tidak ditemui. Tujuan utama kajian ini adalah untuk melengkapkan susunan pada kawasan 3'-kod genetik sebahagian pokok sagu dengan menggunakan kaedah 3'-RACE PCR. RACE PCR melibatkan penggunaan dua primer iaitu, 'gene specific primer' (GSP) dan oligo(dT) primer. RNA daripada daun M. Sagu diekstrak dan diterjemahkan kepada cDNA menggunakan M-MuLV reverse transcriptase. Penjanaan bebenang kedua dilakukan dengan menggunakan teknik 3'RACE PCR. Produk PCR dirawat dengan Vivantis GF-1 PCR Clean-up kit. Pengklonan produk PCR dijalankan dengan penggunaan E. coli XL 1 Blue and pGEM-T Easy Vector Systems. Pengekstrakan plasmid kemudian diteruskan dengan kaedah 'mini-prep' dan Vivantis GF-1 Plasmid Extraction kit dan plasmid dicerna menerusi EcoR1. Koloni PCR dijalankan untuk pemeriksaan kehadiran gen 'insert'. Dalam kajian ini, bahagian pengklonan dan transformasi gagal disebabkan kontaminasi oleh Bacillus dan keputusan 'blast' menunjukan kehadiran xylanase dan bukannya SBE.

Kata kunci: 3'-RACE PCR, *Metroxylon sagu*, pengklonan

INTRODUCTION

The sago palm *Metroxylon sagu* is humankind's oldest food plant. The genus *Metroxylon* is found from 17°S to 15-16°N latitude ranging from Thailand, peninsular Malaysia, and Indonesia to Micronesia, Fiji and Somoa (McClatchey *et al.*, 2006). The species of *Metroxylon* include *M. amicarum*, *M. paulcoxii*, *M. warburgii*, *M. vitiense*, *M. salomonense* and *M. sagu*. These various species of *Metroxylon* are distributed all over the world by aboriginal peoples.

According to Nozaki *et al.* (2004), the sago palm may be a good alternative starch-producing plant. Sago starch accumulates in the pitch core of the stem of the sago palm (Cecil *et al.*, 1982). The starch produced by sago palm has long been a staple food for humans in South-East Asia and now *M. sagu* been produced commercially in Singapore, Malaysia and Philippines. Besides, Stanton (1993) also reported that the advantages of the crop are that it is (1) economically acceptable; (2) relatively sustainable; (3) environmentally stable; (4) uniquely versatile; (5) vigorous; and (6) promotes socially stable agroforestry systems.

In Malaysia, the largest sago-growing areas are outside the Peninsula, in the state of Sarawak, which is now the world's biggest exporter of sago, exporting annually about 25,000-40,000 tons of sago products (Singhal *et al.*, 2007).

Sago is the powdery starch made from the processed pith found inside the trunk of *M. sagu.* and it is nearly carbohydrate and has very little protein, vitamins or mineral. The main forms of starch are consists of the linear polymer, amylose and the branching polymer, amylopectin. According to McCue *et al.* (2002), branching of amylopectin is due to the balanced activities of starch-branching enzyme (SBE) and starch-debranching

enzyme (SDE). SBE can be found in other starch-producing plants such as wheat, pea and maize. The study of 3'-RACE DNA fragment encoding SBE is important in order to complete the whole sequence of SBE from *M. sagu* and thus further study on the gene expression can be done. This is essential to understand the function of this gene completely.

The main goal of this study is to extract RNA encoding starch branching enzyme from sago palm (*M. sagu*) and subsequently reverse transcribe the RNA to cDNA and amplify the DNA fragment by 3'-RACE (Rapid Amplification of cDNA Ends) PCR. The PCR product was then cloned, digested with restriction enzyme and sent for sequencing.

LITERATURE REVIEW

2.1 Selection of Species Studied

The *Metroxylon sagu* (sago palm) is a species of the genus *Metroxylon* belonging to the Palmae family. It is a crop par excellence for sustainable agriculture. The word 'sago' is originally Javanese, meaning starch-containing palm pitch. The scientific name is derived from 'metra', meaning pith or parenchyma and 'xylon' meaning xylem (Singhal *et al.*, 2007). *M. sagu* is commonly known as tha-gu-bin (Myanmar); sagu (Vietnamese); sakhu (Thai); lumbiya (Philiphines); rumbia (Malaysia) and sago palm (English).

The origin of sago palm is believed to be the area extending from Moluccas of Indonesia to New Guinea. At present, sago palm widely grows in Southeast Asia and Oceania, where its natural vegetation takes place in not only low lying wetlands but also dry lands of the tropics (The Society of Sago Palm Studies, 1992). As reported by Flach & Schuilling (1989) and Hisajima (1994), *M. sagu* is an extremely hardy plant, thriving in swampy, acidic peat soils, submerged and saline soils where few other crops survive, growing more slowly in peat soil than in mineral soil.

The habitat of sago palm is at tropical lowland forest and also freshwater swamps, usually found near sea level but can be found 1-700m with rainfall of 2000-5000mm (McClatchey *et al.*, 2006). The temperatures above 25°C and relative air humidity of 70% are favorable for *M. sagu* growth. In addition, incidental light above 800k/cm^2 per day and salinity not exceed 10S/m, which is equivalent to one-eight of the salt concentration of sea water (Singhal *et al.*, 2007) is also optimal for rapid growth.

The size of *M. sagu* can reach 15m (49 ft) in height with bole diameter of 35-60cm (14-24 in). Sago palm is monoecious, having both male and female flowers on the same plant. It is also monocarpic, whereby it fruits and flowers once and then dies. Besides, *M. sagu* has an erect crown of large, pinnate arching leaves. There are two types of fruits produced by *M. sagu*, namely pollinated (seeded) and parthenocarpic (non-pollinated) fruit. Seeded fruits contain a stony, white endosperm and brown testa; whereas parthenocarpic fruits are smaller and contain a spongy mesocarp. The fruits take about 24 months to mature (McClatchey *et al.*, 2006).

M. sagu is economically important and has lots of usage. As reported by McClatchey *et al.* (2006), it is a staple food in the Sepik and Gulf provinces of lowland Papua New Guinea. The leaves of sago are used for the roof thatch and wall siding in house construction, and the wood is also used for floorboards and rafters. Besides, the decaying trunks of the sago palm are a source of sago palm beetle grubs (*Rhynchophorous ferrugineus/bilineatus)*, which is an excellent source of protein.

Figure 2.1: Picture of *M. sagu*. Taken from http://www.cojeco.cz/index

2.2 Starch Biosynthesis

Starch is a complex carbohydrate which consists of glucose polymers and can be found mainly in sago palm, tapioca, maize, wheat and pea. The main forms of starch are amylose, linear chains of glucose monomers linked by 1,4-glycosidic bonds and amylopectin, in which the chains are branched by the addition of 1,6-glycosidic bonds. The starch composition in *M. sagu* is consists of 27% of amylose and 73% of amylopectin (Ito *et al.*, 1979). Starch biosynthesis occurs through the action of four enzymes: ADP-Glc pyrophosphorylase (EC 2.7.7.27), starch synthase (EC2.4.1.21), starch-branching enzyme (EC 2.4.1.18), and starch debranching enzyme (EC 2.4.1.41) (Smith *et al.*, 1997; Myers *et al.*, 2000).

2.3 Starch Branching Enzyme (SBE)

The mRNA encoding for starch branching enzyme can be detected in higher plants especially in sago, maize, wheat, pea and potato. The balanced activities of starch branching enzyme and starch debranching enzyme result in the branching of amylopectin in starch (McCue *et al.*, 2002). According to Blauth *et al.* (2001), starch branching enzyme (SBE) break the α-1,4 linkage of starch, re-attaching the chain to a glucan chain by an α-1,6 bond, altering starch structure. Besides, SBEs also facilitate starch accumulation by increasing the number of non-reducing ends on the growing chain.

The SBEs are present in multiple isoforms in all plants examined (Preiss & Sivak, 1996). There are two types of SBE, SBE I and SBE II, and these isoforms are distinguished by their deduced amino acid sequences. The multiple forms of SBE have been group into two distinct families based on primary structure and functional analysis (Burton *et al.*, 1995). Family A consists of rice SBE III, maize SBE II, pea SBE I and *Arabidopsis* SBE

2.1 and 2.2, whereas family B consists of rice SBE I, maize SBE I, and pea SBE II (Fisher *et al.*, 1996). The regulation of SBE isoforms at development stage has been studied in a number of species. For example, in the study conducted by Guan & Preiss (1993), they found that the SBE I isoform in maize has higher activity on amylase, and SBE II isoform has higher activity on amylopectin. In addition, *in vitro* studies also have shown that SBE I introduces longer branches than SBE II of maize (Guan & Preiss, 1993). In maize, rice and wheat, SBE I has been shown to be a polypeptide that is approximately 83 kDa. However, the role of SBE I in endosperm starch synthesis is unclear.

From the previous study, the partial coding sequences of *M. sagu* encode for starch branching enzyme have been found. The partial sequence of SBE I is 1329bp whereas partial sequence of SBE II is 1383bp, however the complete sequences of SBE I and SBE II are not yet established.

2.4 Reverse-transcriptase Polymerase Chain Reaction (RT-PCR)

RT PCR is a highly sensitive tool in the study of gene expression at the RNA level and, in particular, in the quantitation of mRNA or viral RNA levels (Newton & Graham, 1997). In addition, Marone *et al.* (2001) also claimed that RT-PCR is a highly sensitive and specific method useful for the detection of rare transcripts or for the analysis of samples available in limiting amounts. For example, for research on pathogen-host interactions especially, where only limited amounts of bacterial mRNA are available, the use of a PCR-based technique is desirable (Fislage *et al.*, 1997). This RT-PCR is also known as message amplification phenotyping (MAPPing).

2.5 3'-RACE PCR

Rapid amplification of cDNA ends (RACE) is a polymerase chain reaction (PCR)-based technique which was developed to facilitate the cloning of full-length cDNA 5'- and 3'-ends after a partial cDNA sequence has been obtained by other methods (Brian, 1995). Newton & Graham (1997) also reported that RACE PCR or one-sided PCR is a method by which the PCR technique can be used to amplify the 3'- and 5'- ends of a cDNA using a small stretch of known sequence within the gene.

3'-RACE PCR amplify 3' cDNA from a known region using a sense gene specific primer (GSP) and an anti-sense primer complementary to the adaptor sequence. This technique is widely used to isolate the cDNA of unknown 3'- flanking sequences. 3'-RACE includes a simple procedure that uses naturally occurring poly(A) tail of mRNA as the common anchor site.

3'-RACE procedure uses an oligo(dT) adaptor as a primer for the reverse transcription reaction which anneals to the poly(A) tail of the mRNA. Besides, the oligo(dT) primer is also used as the anchor primer in the subsequent amplifications along with a primer complementary to known sequences within the gene. However, this technique is not efficient when unknown coding regions are large, long 3' untranslated regions are present, or when problematic template regions exist (Ozawa *et al.*, 2004).

2.6 Cloning of PCR Products by using pGEM-T Easy Vector System

Cloning refers to the production of multiple copies of a DNA fragments, cells or organism. The essential steps in DNA cloning using plasmids as vectors involve four steps: (i) DNA recombination, (ii) transformation, (iii) selective amplification and (iv) isolation of desired DNA clones.

The pGEM-T Easy Vector Systems provide a convenient method for the direct cloning of PCR fragments without restriction digestion, modification of primers, or purification of amplified DNA (Robles & Doers, 1994). The system includes the pGEM-T vector, T4 DNA ligase, 2X ligation buffer and a positive control. This system facilitates the cloning of PCR products by providing linear vectors that have a single thymidine extension at the 3'-ends. These terminal thymidines are complementary to the nontemplate-derived 3' adenosine residues which are added to double stranded DNA products by many non-proofreading DNA polymerases (Frackman & Kephart, 1999).

2.7 Colony PCR

Colony PCR is designed to quickly screen for plasmid insert directly from *Escherichia coli* colonies. The plasmid should be high copy number such as pUC18 or pBluescript. Even though blue/white screening can be conducted to determine if inserts are present, this technique however can be used for quick determination of insert size and/or orientation in the vector. Alternately, the presence of an insert can be determined by growing each colony in liquid culture and followed by plasmid extraction and digestion.

MATERIALS AND METHODS

3.1 Primer Design

The primer was designed using the database, NCBI (National Centre for Biotechnology Information) and computer softwares, ClustalW and Primer 3. The gene sequence of *M. sagu* (DQ202304) in Fasta format was aligned with other starch-producing plants, pea (X80009), wheat (AF002820) and maize (NM_001111900) using ClustalW (http://www.ncbi.nlm.nih.gov/). After the multiple alignment process, the sago sequences which have the most number of same bases in the comparison were chose. The conserve region was pasted in the Primer 3 programme in order to get the suggested primers. The primers then were blasted at the NCBI BLAST (Basic Local Alignment Search Tool) in order to make sure that the primer was encoded for starch branching enzyme.

The parameter for designing primer in Primer 3 was set up. The size of primer was in the range of 18 to 23bp and the optimum size was 20bp. The minimum melting temperature, Tm was 50°C and the maximum Tm was set to 65°C whereas the optimum temperature of Tm was set at 55°C. For the G-C%, the range was from 45% to 55%, while the optimum G-C% was set at 50%.

3.2 Media Preparation

i. TE Buffer

0.2mL of 0.5M EDTA (pH 8.0) was added with 1mL of 1M Tris-HCl (pH 8.0). The mixture was then added with distilled water to final volume of 100mL. The solution was autoclaved for sterilization.

ii. 8M LiCl

33.91g of LiCl crystal was added with distilled water up to 100mL of final volume. The solution was autoclaved for sterilization.

iii. 3M NaOAc (pH 5.2)

24.61g of NaOAc powder was dissolved in 50mL of distilled water. The pH of the solution was calibrated to pH 5.2 with addition of concentrated HCl. The solution was added with distilled water to final volume of 100mL. The solution was autoclaved for sterilization.

iv. CTAB RNA Extraction Buffer

5g of CTAB (cetyltrimethyiammonium bromide), 25mL of Tris-HCl pH8, 12.5mL of 25mM EDTA and 29.22g of NaCl was added into 150mL of distilled water. The mixture was then added with 5g of PVP. PVP was slowly added and the mixture was stirred to ensure that it was totally dissolved. The solution then was top up with distilled water to final volume of 250mL. β-mercaptoethanol was only added into extraction buffer prior use.

The extraction buffer without β-mercaptoethanol was then been autoclaved and stored at room temperature.

v. 60mM $CaCl_2$ with 15% glycerol

150mL of 100mM $CaCl_2$ was added with 37.5mL glycerol. The solution was then added with sterile distilled water to a final volume of 250mL. The solution was kept at 4°C prior use.

vi. Solution I

50mM glucose and 1.8M formic acid were added into 25mM Tris-HCl, pH8 in a bijou bottle and top up with sterile distilled water to 10mL of final volume. The solution was kept at 4°C prior use.

vii. Solution II

0.2N of NaOH was mixed with 1% SDS and was added with sterile distilled water in a bijou bottle to a final volume of 10mL. The solution was kept at 4°C prior use. However, fresh solution was recommended.

viii. Solution III

3M KAc was mixed with 10mM EDTA, pH8 in a bijou bottle and was added with sterile distilled water to 10mL of final volume. The solution was kept at 4°C prior use.

3.3 Plate Preparation

i. LB Agar Plate

2.5g of LB powder was added with 1.5g of agar in 100mL of distilled water. The solution was mixed and autoclaved. After been autoclaved, the solution was cooled down until around 50°C and been poured to several petri dishes in laminar flow hood. The LB agar was left in laminar flow to harden and kept at 4°C prior use.

ii. LB/ampicilin/IPTG/X-gal (LAIX) Plate

100µL of 25mg/mL ampicilin was spread onto LB plate using spreader. It was air-dried in laminar flow and followed by spreading of 100µL of IPTG onto the plate. The plate was air-dried again and finally 40µL of X-Gal was spread onto the same plate. The plate was air-dried and kept at 4°C prior use. LAIX plates were recommended to be used within 1 week.

3.4 Plant Materials

The matured *M. sagu* leaves which were used for RNA isolation was obtained from green house of Faculty of Resource Science and Technology (FRST) Universiti Malaysia Sarawak. The leaves were washed under running tab water, distilled water and were wiped with 70% ethanol and was kept at -80°C prior use.

3.5 RNA Extraction

The total RNA of *M. sagu* leaf was extracted using the CTAB (cetyltrimethyiammonium bromide) extraction method by Gasic *et al.* (2004) with slightly modification.

10mL of extraction buffer was added with 200µL β-mercaptoethanol. This step was conducted in a fume hood. The extraction buffer was then pre-warmed at 65°C in a water bath. Around 5g of sago leaves were grind in liquid nitrogen by using mortar and pestle. The leaves powder was then transferred into 50mL polypropylene tubes which contained the pre-warmed extraction buffer by using a chilled autoclaved spatula. The leaves powder was incubated at 65°C for 15 minutes. After that, it was vortex briefly to ensure that all leaves have been mixed. The incubation and vortex step were repeated. After three times of incubation at 65°C, 10mL of C:I:A (chloroform:isoamyl:alcohol) was added into the polypropylene tubes and was immediately vortex for 2 minutes.

The sample then was centrifuged at 7000g for 15 minutes at 4°C. After centrifugation, the supernatant was transferred to a clean polypropylene tube and only the aqueous phase was pipetted. The supernatant was re-extract with equal volume of C:I:A and being centrifuged again at 7000g for 15 minutes at 4°C. After the centrifugation, the supernatant obtained was divided into several 2mL microcentrifuge tubes. 1.5mL of supernatant was added with 500µL 8M LiCl. They were mix by inversion and were stored at 4°C overnight.

After overnight incubation, the samples in the microcentrifuge tubes were centrifuged at 13,000rpm for 30 minutes at 4°C. The supernatant then was discarded and the samples were washed with 500µL 70% ethanol. The pellet was dissolved by flicking the tubes. Then, the samples were centrifuged again at 13,000rpm for 30 minutes at 4°C. The supernatant was discarded and the sample was air-dried. The RNA pellet then was

dissolved in 200μL of distilled water and was added with 20 μL of 3M NaOAc and 500μL of 70% ethanol. The mixture was stored at -20°C overnight.

On the third day of extraction, the samples were centrifuged at 13,000rpm for 30 minutes at 4°C. The supernatant then was discarded and the samples were washed with 500μL 70% ethanol. Then the mixtures were centrifuged again at 13,000rpm for 15 minutes at 4°C. The supernatant was discarded and the RNA pellet was air-dried. Finally, the RNA pellet was dissolved in 40-50μL sterile distilled water and been stored at -80°C prior use.

3.6 Analysis of Sago Palm Total RNA

3.6.1 Qualitative Estimation of RNA by Agarose Gel Electrophoresis

1% of agarose gel was prepared by mixing of 0.25g of agarose powder with 25mL of 1 X TAE buffer. The solution was melted in microwave oven for approximately 2 minutes. The gel was stained with 1μL ethidium bromide and poured into the tray. The gel running tank was filled with 500mL 1X TAE buffer. The gel was ensured to be totally covered by the running buffer. Five micro liters of extracted RNA was mixed with 1μL 6X loading dye solution. The gel was run at 105V for about 35 minutes and finally visualized under a UV transilluminator.

3.6.2 Quantification of RNA by UV Spectrophotometric Analysis

The quantification of RNA by UV Spectrophotometer was done by 1:100 dilutions, whereby 5μL of total RNA was diluted with 495 μL of sterile distilled water. The