ISOLATION OF cDNA FRAGMENT ENCODING STARCH BRANCHING ENZYME (ISOFORM II) GENE FROM Metroxylon sagu BY RT-PCR METHOD.

Mohd Azinuddin Bin Ahmad Mokhtar

Bachelor of Science with Honours (Resource Biotechnology) 2006
Isolation of cDNA Fragment Encoding Starch Branching Enzyme (Isoform II) Gene

from *Metroxylon sagu* by RT-PCR method.

MOHD AZINUDDIN BIN AHMAD MOKHTAR

This report is submitted in partial fulfillment of the requirement for the degree of

Bachelor of Science with Honours

(Resource Biotechnology)

Faculty of Resource Science and Technology

University Malaysia Sarawak

2006
Acknowledgement

I would like to express my highest gratitude to my supervisor, Dr Mohd Hasnain Bin Md Hussain for his kindness to assist me in various aspects until the final chapter of my project.

I would also like to thank to Dr Hairul Azman Bin Roslan for his kind assistance during the operation of equipment and machine in his laboratory.

My inadequate thanks are also to all the lab assistants, master students and all my friends for their generosity and kindness towards the improvement in completing this project.

Not to mention, appreciation to my parents and family for their positive encouragement and endless support throughout the project.
List of Figures

Figures 1: Distribution of main sago palm areas. 5

Figures 2: A proposed pathway of sucrose breakdown and starch synthesis in storage organs. 9

Figures 3: Analysis results of RNA extracted from leaf tissues. 26

Figures 4: Electrophoresis of first strand cDNA synthesis product. 33

Figure 5: Quantification of first strand treated with DNase. 34
List of Tables

Table 1: Parameters for PCR amplification 22
Table 2: Result of total RNA extraction 28
Table 3: Sequence of forward primer 30
Table 4: Sequence of reverse primer 31
Table 5: Backward of reverse primer 31
Table 6: Design primer sequence 31
### Tables of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgement</td>
<td>i</td>
</tr>
<tr>
<td>List of figures</td>
<td>ii</td>
</tr>
<tr>
<td>List of tables</td>
<td>iii</td>
</tr>
<tr>
<td>Tables of contents</td>
<td>iv-v</td>
</tr>
<tr>
<td>Abstract</td>
<td>1</td>
</tr>
<tr>
<td>Abstrak</td>
<td>1</td>
</tr>
<tr>
<td><strong>CHAPTER ONE</strong></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>1.1 Problem Statement</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Objective</td>
<td>4</td>
</tr>
<tr>
<td><strong>CHAPTER TWO</strong></td>
<td></td>
</tr>
<tr>
<td>Literature Review</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Sago palm</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Composition of starch in plants</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Starch biosynthesis</td>
<td>8</td>
</tr>
<tr>
<td>2.4 Starch branching enzyme</td>
<td>9</td>
</tr>
<tr>
<td>2.5 Complementary DNA synthesis</td>
<td>10</td>
</tr>
<tr>
<td>2.6 Reverse transcription polymerase chain reaction (RT-PCR)</td>
<td>11</td>
</tr>
<tr>
<td>2.7 Bioinformatics tools</td>
<td>12</td>
</tr>
<tr>
<td><strong>CHAPTER THREE</strong></td>
<td></td>
</tr>
<tr>
<td>Materials and Methods</td>
<td></td>
</tr>
<tr>
<td>3.1 Sample collection</td>
<td>14</td>
</tr>
<tr>
<td>3.2 Extraction of total RNA from leaf</td>
<td>14</td>
</tr>
<tr>
<td>3.3 Treatment with DNase</td>
<td>15</td>
</tr>
</tbody>
</table>
3.4 Gel electrophoresis
3.5 Quantifications and determination of quality of total RNA
3.6 Construction of primers
3.7 Synthesis of first strand cDNA
3.9 Synthesis of second strand cDNA

CHAPTER FOUR Results and Discussions
4.1 RNA extraction
4.2 Primer design
4.2 First strand synthesis
4.3 Second strand synthesis

CHAPTER FIVE Conclusion and Recommendation

REFERENCES

APPENDICES
Isolation of cDNA Fragment Encoding Starch Branching Enzyme (Isoform II) Gene from *Metroxylon sagu* by RT-PCR method.

Mohd Azinuddin Bin Ahmad Mokhtar

Resource Biotechnology Programme
Faculty of Resources Science and Technology
University Malaysia Sarawak

ABSTRACT

*Metroxylon sagu* is a starch producing plant that has great potential to be commercialized especially in the field of food industry. Due to economic potential, great emphasis in research toward these plants was done. Researches involve study on molecular aspect especially on gene for starch synthesis. In this project, total RNA from leaves was extracted and the highest concentration extract was used. Amplification of specific gene (isoform II) was done by designing specific primer to be used in reverse transcriptase polymerase chain reaction (RT-PCR). Two steps RT-PCR was conducted thought usage of M-MLV reverse transcriptase enzyme. Synthesis of second strand was also by RNase and DNA polymerase I as comparison. However, due to contamination and optimization problem, synthesis of second stranded cDNA to be sequences was unable to complete. These studies only accomplished halfway and follow up research should be conducted.

Key words: *Metroxylon sagu*, isoform II, cDNA, reverse transcriptase polymerase chain reaction RT-PCR, M-MLV reverse transcriptase.

ABSTRAK


Kata kunci: *Metroxylon sagu*, isoform II, cDNA, tindak balas berantai polimerase enzim reverse transcriptase RT-PCR, enzim reverse transcriptase M-MLV.
CHAPTER ONE

Introduction

Sago (*Metroxylon sagu*) is a starch crop that has great potential for commercial plantation. At present, 80% of cereal world production is starch that is about 1,360 million tons (Flach, 1984). But sago palm starch contributes only 1.5% of total world production of starch (Flach, 1984). According to Stanton, 1993 (cited in Flach, 1997) this plant have number of advantages whereby they are adequate economically, fairly sustainable, ecological friendly, exclusively versatile and the most essential things are they offer established agroforestry systems. Example of sago starch product that is produce commercially was in food industry for production of monosodium glutamate, glucose and custard powder (Danjaji et al., 2001). Even though commercialized of sago starch was new, but this starch has been a staple food for human is South-East Asia for centuries (Flach, 1997).

Sago grows well in the peat-land delta or riverine areas of Southeast Asia especially in Thailand, Indonesia, Papua New Guinea, Philippines and Malaysia. The total acreage for this crop has been estimated to be more than 3.75 million hectares (Flach, 1997). In Malaysia, *Metroxylon sagu* is widely distributed in the area of east-Malaysia such as Sarawak and Sabah and certain part of peninsular Malaysia (Flach, 1984).

There is a great research interest in sago plant in the world. Presently, up to 12th symposium on sago have been held worldwide. But not many information on molecular aspect has been discovered. One of the information is about its DNA sequences and the
function of each gene sequence in its genome. In order to study these, scientist usually will generate complementary DNA (cDNA) and analyzed it.

cDNA is a genetic material that is produce from messenger RNA (mRNA). It was done by using reverse transcriptase enzyme that only present in retroviruses. cDNA was chosen because of it did not have any 'junk' gene or intron that did not encode any protein. By doing this, specific gene for specific protein or enzyme can be obtained. Because sago was a potential starch crop, these studies will focus on enzymes that are involved in starch production.

Enzyme that involve in starch biosynthesis are ADP-glucose pyrophosphorylase (ADPGPase; EC 2.7.7.23), starch synthase (SS; EC 2.4.1.21), and starch branching enzyme (SBE) (SBE; EC 2.4.1.28). There are many enzymes under SBE families such as Isoform I in maize (Guan and Preiss, 1993), Isoform Ila and Isoform IIb in barley (Sun et al., 1997). To make cDNA of gene sequence that encode for SBE Isoform II, there is a need to use other starch crop gene sequence for SBE Isoform II to help in constructing the gene sequences for primer of SBE Isoform II of sago. This was done by using ClustalW computer software to identify conserved region sequences. Reverse transcription polymerase chain reaction (RT-PCR) method was used to produce cDNA and multiply the cDNA produced. The cDNA sequence was analyzed with bioinformatics computer software.
1.1 Problem statement

Research on molecular aspect on starch branching enzyme was not yet been studies. However, this is a continuity of studies before on first strand synthesis of cDNA library from leaf. This study was a start on studying gene encoding the for starch branching enzyme of *Metroxylon sagu*. By generating the cDNA fragment of this gene, further studies on characterization of full length gene can be conducted due to basic knowledge obtained by these studies. Further studies on the expression of this enzyme can also be done from these studies.

1.2 Objective

The objective of this study is to gain knowledge about SBE Isoform II in *Metroxylon sagu*. This study is done to characterize the gene of SBE Isoform II. It also compared the gene with other starch crop that also producing SBE Isoform II for example like rice, maize and potato by using bioinformatics software. In order to achieve these objectives, total RNA of sago needs to be isolated successfully. The productions of cDNA must be successful and should be a correct sequence for SBE isoform II.

Apart from that, it is hoped that by knowing the sequence and comparison, the information can be used in further study in manipulating other starch crop in producing better in quantity and quality of starch. It also can be used to compare the gene sequence if there is mutation happen in sago palm.
CHAPTER TWO

Literature Review

2.1 Sago palm

The sago palm belongs to the Lepidoceylroid subfamily of the Arecaceae (Palmeae). The sago palm is mainly found between longitude 90° to 180° east and between latitude 10° north and south up to an altitude of 700 meters above sea level (Flach, 1984). The economically most important species in the genus *Metroxylon* are doubtfully distinct as they are found mixed in wild stands together with appear to be intermediate. These species are *M. sagu* Rottb. and *M. rumphii* Mart. *M. rumphii* has spines on petioles, spathes and even leaflets (Tan, 1982).

![Figure 1: Distribution of main sago palm areas (Flach, 1984).](image)
Sago palms are monoecious and monocarpic palm (Flach, 1984) that their trunk may reach 10 meters to 12 meter in height and 30 centimeter to 60 centimeter in diameters with basal suckers for suckering (soboliferous) perennial. The leaf was in simple arrangement of pinnate around their trunk and the length is 5 meter to 6 meter in average (Saidin, 1993) with about 50 pairs of leaflets 60 centimeter to 180 centimeter long and approximately 5 centimeters wide (Flach, 1984). The sago palm is a once-flowering (hepaxanthic) that are borne spirally in pairs on the tertiary axis (Flach, 1984). Flowers are bisexual and later will become sago seeds. Flowers were protandrous type where male flowers become mature before female flowers. This is where cross pollination happen. Pollination take place for two years after that the plant dies.

Flower initiation of sago palm named *M. sago* was occur at about leaf scar 54 with a trunk age of 4.5 years, if a rate of leaves form is one per month. Sometimes larger number of leaf scars on their trunks up to 80 scars, corresponding with a trunk age of 6.75 years (Flach, 1984). Reports on the length of the life cycle of sago palm are ranging from 8 to 17 years.

In Malaysia, sago cultivation was mainly concentrated in the state of Sarawak where there are about 1.4 million hectares of peat soil suitable for cultivation of this crop (Jennifer, 2001). The cultivation may reach up to 25 tones of starch per hectares (Danjaji et al., 2001). Sago was also known starch crop that have great potential then other crop. The yield of sago towards any other crops such as cassava and maize are 2000-3000kg/HaYr, 1000kg/HaYr in maize (Stantin, 1992) and 2000kg/HaYr in cassava (Ahmad et al., 1999)
2.2 Composition of starch in plants

Starch is the dominant carbohydrate reserve material of higher plants, being found in leaf chloroplasts and in the amyloplasts of storage organs such as seeds and tubers (Wang et al., 1998). In sago palm, starch was store at its trunks (Flach, 1984). Starch occurs as dense, water-insoluble granules, ranging in size from less than 1μm to over 100μm (Smith and Martin, 1993). The starch can be chemically divided into two type of glucan polymer that is amylose and amyllopectin. Amylose is a linear chain of α(1,4)glycosidic linkage of α-D-glucose unit and amyllopectin is a branching form of amylose that cross-linked by α(1,6)glycosidic bond. The ratio of amylose to amyllopectin varies between species and between cultivars within a species (Shannon and Garwood, 1984). Other than genetic factors, the composition of starch is also affected by developmental stages. For example, the amylose content of starch increases from 9% to 27% between 8 and 28 days post-anthesis in maize endosperm (Tsai et al., 1970). In general, most starches contain 67–70% amyllopectin (Preiss, 1991).
2.3 Starch biosynthesis

Starch is synthesized in leaves during the day from photosynthetically fixed carbon and is mobilized at night. It is also synthesized transiently in other organs, such as meristems and root cap cells, but its major site of accumulation is in storage organs, including seeds, fruits, tubers, storage roots and trunks (Martin and Smith, 1995; Flach, 1984).

Figure 2 illustrates a simplified pathway of sucrose breakdown and starch synthesis in starch storage organs. Sucrose that entering the plant cells in cytosol can be broken down to UDP-glucose and fructose by sucrose synthase. Then, breakdown by UDP-glucose pyrophosphorylase to produce glucose-1-phosphate. There are three major group of enzyme are involve in starch biosynthesis that is ADP-glucose pyrophosphorylase (AGPase), starch synthase (SS) and starch branching enzyme (SBE). AGPase synthesizes ADP-glucose from glucose-1-phosphate were ADP-glucose is the substrate for SS. All SS transfer glucosyl units from ADP-glucose to the nonreducing ends of growing α(1,4) glucans. There end product naturally is amylose and some in amylopectin. SBE involve in synthesis of amylopectin where create the branch point by hydrolysis of an α(1,4) linkage and replace it with an α(1,6) glycosidic bond. These enzymes only present in plastid of plant cell.
Figure 2: A proposed pathway of sucrose breakdown and starch synthesis in storage organs (Andersson, 2001).

2.4 Starch branching enzyme

Starch branching enzymes (SBEs) is involved in synthesis of amylopectin where it create the branch chain by hydrolyze of an α(1,4)glycosidic linkage and subsequent it with an α(1,6)glycosidic bond between the cleaved chain and a C6 hydroxyl group of a α(1,4)glucan. Multiple isoforms of SBE have been reported in various plant species, such as maize, pea, rice, potato, wheat, barley, Arabidopsis, and sorghum (Mutisya, 2004;
Boyer and Preiss, 1987; Smith, 1988; Mizuno et al., 1992; Larsson et al., 1996; Morell et al., 1997; Sun et al., 1997; Fisher et al., 1996; Matisya et al., 2003). The SBE isoforms can be divided into two distinct classes, A and B (Andersson, 2001; Burton et al., 1995). A and B isoforms have highly similar amino acid sequences but usually differ by an N-terminal extension of the B form and a C-terminal extension of the A form (Rydberg et al., 2001). The isoforms nomenclature that usually use is either SBE A and B, or I and II. SBE that was member in A families is potato SBE II, maize SBE II, pea SBE I, rice SBE III and barley SBE II, whereas B family is potato SBE I, maize SBE I, pea SBE II, rice SBE I and barley SBE I.

2.5 Complementary DNA synthesis

In producing complementary DNA (cDNA), extracted total RNA was needed. Total RNA is a single stranded molecule that comprises of various complex of structure and function of RNA e.g. messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA), depend on their molecule’s origin (Strachan & Read, 1999). cDNA was synthesis from mRNA template by using reverse transcriptase enzyme (RT). This enzyme is like any other DNA synthesizing enzyme in that cannot initiate DNA synthesis without primer (Weaver and Hedrick, 1992).

There are three ways to synthesized cDNA that is primed with random primers, oligo (dT), or a gene-specific primer (GSP). Single stranded DNA was produces after the mRNA was copied. This single stranded DNA also known as first strand. The first strand was in form of DNA/RNA hybrid with the template. Then, RNA is removing by using
alkali or ribonuclease H (RNase H). The enzyme hydrolyzed the phosphodiester bond on RNA part of an RNA/DNA hybrid for removing RNA from first strand cDNA (Weaver and Hedrick, 1992). The enzyme was digesting certain part of RNA strand. The next step is by using DNA polymerase I to build the second strand. DNA polymerase used the unresolved region of RNA as primer to syntheses second strand of cDNA. Apart from that, 3’ end of single stranded cDNA (hairpin loop) molecule also can be used as primer to initiate the synthesis (Strachan & Read, 1999).

Another way of synthesis of second strand is by polymerase chain reaction (PCR) of first strand used syntheses before as template in this reaction. For built the second strand by using oligo (dT), first strand of cDNA that been produce before were use as template. Primer still needed but it not the same as first strand primer. In the second strand synthesis, oligo (dC) was making at the 3’-end of the first strand by using enzyme terminal transferase and substrate dCTP. Then, hybridize a short oligo (dG) that function as primer for second strand. RT was using again to make second stand.

2.6 Reverse transcription polymerase chain reaction (RT-PCR)

RT-PCR is the most sensitive technique for mRNA detection and quantitation currently available (Newton, 1994). This technique also known as message amplification phenotyping (MAPPing). It is a method where cDNA was synthesis by using reverse transcriptase enzyme (RT enzyme), permits the simultaneous analysis of large number of mRNAs from small number of cells (Newton, 1994). It can reverse the process of transcription that is the product is RNA to become cDNA. If compared to the two other
commonly used techniques for quantifying mRNA levels, that is Northern Blot analysis and RNase protection assay, RT-PCR is an extraordinarily sensitive method to detect as few as 1–100 copies of a specific RNA (Shuldiner, 1993). Reverse transcriptase is usually used to synthesize first strand cDNA from RNA. Reverse transcriptase can be purified from several sources, e.g., Avian Myeloblastosis Virus (AMV) and Moloney Murine Leukemia Virus (MMLV) (Newton, 1994). AMV reverse transcriptase is RNA-dependent DNA polymerase that use single stranded of RNA as template in producing cDNA in 5' to 3' direction if a primer present. It is also exhibits ribonuclease H activity (Newton, 1994). For the MMLV reverse transcriptase, it acts in the same way as AMV reverse transcriptase, but it lack DNA endonuclease activity and has lower RNase H activity.

2.7 Bioinformatics tools

Bioinformatics is an important tool for complete DNA and protein analysis, secondary structure predictions, primer design, molecular modeling, development of cloning strategies, plasmid drawing or restriction enzyme analyses (Abd-Elsalam, 2003). Bioinformatics software was make primer design easier. In order to design primer, information about the sequences to be amplified is needed. This information was available at National Centre of Biotechnology Information (NCBI) database. Information needed in the form of nucleotide and protein sequence. The sequences need to align and to identify its conserved region. These can be done by using either pairwise alignment in Basic Local Alignment Search Tools (BLAST) program or multiple alignments such as
ClustalW program. BLAST is a program that compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches that can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families (Anonymous, 2001). Whereas, ClustalW is a fully automatic program for progressive global multiple alignment. ClustalW consider the sequence redundancy with some adjustable parameters with reasonable defaults (Lloyd, 1997).
CHAPTER THREE

Materials and Method

3.1 Sample collection

Sample of *Metroxylon sagu* leaves was collected from UNIMAS green house. The leaves were preserved in liquid nitrogen and stored at -80°C in laboratory.

3.2 Extraction of total RNA from leaf

Total RNA was extracted using the method used by Hussain (2002). 1.5g of frozen leaf sample was grounded using mortar and pestle till it becomes powdered. 15ml of RNA extraction buffer (150mM LiCl, 50mM Tris pH 9.0, 5mM EDTA, and 5% w/v SDS) was added to the sample and followed by vortex. After vortex, 15ml of phenol/chloroform/isoamylalcohol (25:24:1) was added to the sample and vortex. Sample was centrifuged at 6,000rpm for 1 minute to allow it precipitate. The aqueous phase of supernatant was transferred carefully into new tube. The phenol/chloroform/isoamylalcohol and centrifuged steps was performed three times. Then, 15ml of chloroform was added to the aqueous solution in the new tube, vortex and followed by centrifuged at 8,000rpm for 1 minutes. The aqueous phase of supernatant was transferred carefully into new tube and this step was done again twice. The aqueous solution (containing the RNA) was precipitated again by adding 2.5ml of 8M LiCl to give final concentration of 2M LiCl for precipitation of RNA. The extract was incubate
overnight at 4°C and followed by centrifuged at 8,000rpm for 30 minutes. Then, pellet was resuspended in 400μl of sterile distilled water. 1ml of 100% (v/v) ethanol and 40μl of 3M Sodium acetate was added to allow the solution to precipitate and was put on ice for 20 minutes. Then, solution was centrifuged at 14,000rpm for 30 minutes followed by washed twice with 70% (v/v) ethanol. The pellet was allowed to air dried and followed by addition of 100μl TE buffer. The extracted total RNA was stored at -80°C.

3.3 Treatment with DNase

DNase treatment of RNA sample was using Promega, Madison, WI, USA kits (catalog #M6101). The following DNase digestion reaction was prepared in a tube on ice:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA in TE buffer</td>
<td>5μl</td>
</tr>
<tr>
<td>RQ1 RNase-Free DNase 10X reaction buffer</td>
<td>1μl</td>
</tr>
<tr>
<td>RQ1 RNase-Free DNase (1u/μl)</td>
<td>1μl</td>
</tr>
<tr>
<td>Nuclease-free water</td>
<td>2μl</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>10μl</strong></td>
</tr>
</tbody>
</table>

Then, the digestion reaction was incubated at 37°C for 30 minutes. 1μl of RQ1 DNase stop solution was added to terminate the reaction and followed by incubate at 65°C for 10 minutes to inactivate the DNase. The reaction mixture was stored at -20°C.
3.4 Gel electrophoresis

Gels were prepared by dissolve 0.8g of agarose in 100ml of 1X TBE buffer. The mixture was boil using microwave oven to completely dissolve. Next, the agar was cooled under running tap water to 50°C to 60°C. Agarose was pour to the tray that has been position with the well-forming comb and was allow being set. Then, the comb was removed carefully and the gel was place into electrophoresis tank that fill with 1X TBE buffer. The sample in a mixture of 5:1 (5μl sample: 1μl of 1X loading dye) was run horizontally at 100mA for one and half hour. Gel then, was stained with ethidium bromide for 30 to 40 second and destained for 2 to 3 minutes. The gel then was visualized under the presence of UV light.

3.5 Quantification and determination of quality of total RNA

Concentration of total RNA was determined by quantified using spectrophotometer to get sufficient RNA concentration. Absorbance of diluted RNA sample at the wavelengths 260nm (A_{260}) and 280nm (A_{280}) was measured by using Ultrospec® 1100 Pro machine with appropriate RNA programme, 5μl of RNA sample was diluted in 500μl of sterile water for 100X dilution was placed in a quartz cuvex. Concentration of RNA sample was calculated using specified equation that gives result in unit μg/μl.

**RNA concentration: 40 x A_{260} reading x 100X dilution factor**
3.6 Construction of primers

Five different starch producing plant has been used is bread wheat (*Triticum aestivum*) (accession no. Y11282), Mungbean (*Vigna radiata*) (accession no. AY622199), Potato (*Solanum tuberosum*) (accession no. AJ000004), Pea (*Pisum sativum*) (accession no. X80010) and Corn (*Zea mays*) (accession no. L08063). These sequences are available in National Centre of Biotechnology Information (NCBI) database. The protein sequences alignments were performed with ClustalW version 1.82 program using the ClustalW www Service at the European Bioinformatics Institute (EBI) (http://www.ebi.ac.uk/clustalw). Conserved region from the alignment were compare with nucleic acid sequence of each plant to determine it nucleic acid sequence. Every selected nucleic acid sequences were analyzed with DNAclub software to determine its melting temperature (52°C), percentage of GC content (less then 50%) and length of nucleotide sequences (range 18 to 24 nucleotide). Designed primer was ordered from the commercial laboratory.