Faculty of Resource Science and Technology

**Sequence Polymorphism of *Sucrose Synthase* Gene in Kelampayan (*Neolamarckia cadamba*)**

**Amber Chin Yen Siew (20640)**

**Bachelor of Science with Honours**

**(Resource Biotechnology Programme)**

**2011**

# ACKNOWLEDEMENTS

Above all, I would like to thank God for his blessings and guidance in the completion of this project. My heartfelt gratitude goes to my supervisor as well, Dr. Ho Wei Seng for his precious guidance, advice and encouragement. Apart from that, I would also like to dedicate my sincere appreciation towards my co-supervisor, Dr. Pang Shek Ling for her constant guidance in my experimental process. My special thanks also go to Mr. Liew Kit Siong, Ms. Tchin Boon Ling and all the beloved lab members for their guidance, advice and assistance. Not forgetting my family and friends for their support and encouragement throughout the completion of my final year project.

# DECLARATION

With this, I hereby declare that this thesis is my original work except for quotations and citations, all of which have been duly acknowledged. Apart from that, I would also like to declare that it has not been previously or concurrently submitted for any other degree at UNIMAS or any other institutions.

_____

Amber Chin Yen Siew

Resource Biotechnology Programme

Department of Molecular Biology

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak

# TABLE OF CONTENT

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **ATP** | Adenosine triphosphate |
| **CAPS** | Cleaved amplified polymorphic sequence |
| **cDNA** | Complementary DNA |
| **CESA** | Cellulose synthase |
| **CIA** | Chloroform-Isoamyl Alcohol |
| **CTAB** | Cetyltrimethylammonium Bromide |
| **DNA** | Deoxyribonucleic acid |
| **EDTA** | Ethylenediamine tetraacetic acid |
| **InDel** | Insertion-Deletion |
| **KOR** | Korrigan cellulase |
| **LHD** | Light Hardwoods |
| **LB** | Luria both |
| **MAS** | Marker Assisted Selection |
| **NaCl** | Sodium Chloride |
| **PCR** | Polymerase Chain Reaction |
| **RFLP** | Restriction Fragment Length Polymorphism |

| | |
|---|---|
| **SNP** | Single nucleotide polymorphism |
| **SuSy** | Sucrose synthase |
| $\mathbf{T_a}$ | Annealing temperature |
| $\mathbf{T_m}$ | Melting temperature |
| **UDP** | Uridine diphosphate |
| **UDPG** | Uridine diphosphate glucose |
| **UV** | Ultraviolet |

# LIST OF TABLES

# LIST OF FIGURES

# SEQUENCE POLYMORPHISM OF SUCROSE SYNTHASE GENE OF KELAMPAYAN (*Neolamarckia cadamba*)

## AMBER CHIN YEN SIEW

Resource Biotechnology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

## ABSTARCT

*Neolamarckia cadamba* (Roxb.) Bosser or locally known as Kelampayan possesses great economic and commercial value as its timber is often used for production of plywood, veneer, furniture and hardboard. As conventional plant selection for breeding is time consuming and costly, new approach such as single nucleotide polymorphism (SNP) has often been used as marker for molecular breeding purposes. The main objective of this study was to identify the DNA sequence variation caused by single nucleotide substitution in the sucrose synthase (*SuSy*) gene of Kelampayan. In order to do this, DNA extracted from six Kelampayan trees were firstly subjected to polymerase chain reaction to obtain the desired *SuSy* sequence. BLASTn analysis was then performed to the ~800 bp SuSy amplicons to search for sequence homology against non-redundant nucleotide database available in NCBI. This was followed by sequence alignment using CLC Free Workbench 6.0 for detection of SNPs. Consensus sequence of SuSy was later subjected to *in silico* restriction analysis. A total number of 54 SNP had been detected in the partial sucrose synthase sequence. 46 SNPs are located in the predicted coding region while 8 SNPs are positioned in the predicted non-coding region. Six restriction enzymes which include *EaeI*, *HpyCH4III*, *BsaBI*, *Bpu10I*, *HincII* and *HinP1I* were detected for six SNP sites in partial sucrose synthase sequence as well. The effectiveness of such restriction enzyme can later be used in the development of useful genetic marker.

Key words: *Neolamarckia cadamba*, Sucrose synthase gene (SuSy), Single nucleotide polymorphism (SNP).

## ABSTRAK

*Neolamarckia cadamba (Roxb.) Bosser atau lebih dikenali sebagai Kelampayan memiliki nilai ekonomi dan komersil yang tinggi disebabkan kayunya yang mempunyai pelbagai kegunaan. Memandangkan cara konvensional pemilihan tumbuhan bagi tujuan pembiakbaikkan genetik adalah mahal dan lama, polimorfisme nukleotida tunggal (SNP) sering digunakan sebagai penanda molekul untuk mengatasi masalah ini. Tujuan kajian ini adalah untuk mengenal pasti SNP di gen sukrosa sintase di Kelampayan. Kajian ini bermula dengan pengekstrakkan DNA daripada enam pokok Kelampayan. Tindak balas berantai polimerase kemudiannya digunakkan untuk memperolehi urutan SuSy yang dikehendaki. Selepas itu, ~ 800 bp SuSy amplikon yang diperolehi dianalisis dengan BLASTn dengan tujuan mencari homologi urutan nukleotida yang terdapat di NCBI. Seterusnya, pengesanan SNP dilaksanakan melalui penjajaran urutan dengan menggunakan CLC Free Workbench 6.0. Urutan konsensus SuSy juga dianalisis dengan penyekatan in silico. Hasil kajian ini menunjukkan 54 SNP di sepanjang urutan sukrosa sintase gen. 46 SNP terletak di exon manakala 8 SNP terletak di intron. Selain daripada itu, enam enzim penyekatan (EaeI, HpyCH4III, BsaBI, Bpu10I, HincII dan HinP1I) juga dikesan di enam posisi SNP di urutan sukrosa sintase. Keberkesanan enzim penyekatan tersebut boleh diaplikasikan dalam penghasilan penanda molekul.*

Kata kunci: *Neolamarckia cadamba*, Gen sukrosa sintase (SuSy), Polimorfisme nukleotida tunggal (SNP).

**SECTION 1.0**

**INTRODUCTION**

*Neolamarckia cadamba* (Roxb.) Bosser or locally known as Kelampayan is a fast-growing, medium to large deciduous tree. Kelampayan, which belong to the family Rubiaceae, are widely distributed from India through Southeast Asia to New Guinea. This tree can reach 45 m tall, with diameter up to 100 cm but normally less (Peter, 2007). It is characterized by its orange, small, in dense and globose head flowers. Kelampayan possesses great commercial values as its timber is frequently used for the production of plywood, veneer, furniture and hardboard.

Sucrose is essential in all plants as it supplies carbon and energy needed for actively growing sink tissues such as roots and elongating stems (Yao *et al*., 2009). Apart from that, sucrose is also one of the sources that produce uridine diphosphate glucose (UDP-glucose), which is crucial for wood formation. But sucrose however cannot be utilized directly by plants. It must be hydrolyzed by sucrose synthase (SuSy)**,** which is a homotetrameric enzyme that catalyses the reversible UDP-dependent cleavage of sucrose **i**nto UDP-glucose and fructose (Silvente *et al*., 2003). The products of cleavage, which are glucose and fructose, will then be utilized to facilitate plant growth and development. Therefore, sucrose synthase gene, which encodes for sucrose synthase enzyme plays an important role in both sucrose degradation and UDP-glucose synthesis.

As conventional plant selection for breeding is time consuming and costly, new approach has been utilized to overcome these problems. Single nucleotide polymorphism

(SNP), which is a single base substitution in homologous DNA fragment, can be a promising genetic marker for plant breeding. Although SNPs are biallelic, but it is very informative compared to other markers such as Restriction Fragment Length Polymorphism (RFLP), as large numbers of SNPs can be combined to form haplotypes (Esser *et al*., 2004). SNP marker which is associated with traits of interest is very beneficial to marker-assisted selection (MAS) in plant breeding as desired traits can be selected in a short time.

The main objective of this study was to identify sequence variation caused by single nucleotide substitution in sucrose synthase (SuSy) gene of five Kelampayan trees. Apart from that, synonymous and non-synonymous mutations caused by single nucleotide polymorphism (SNP) were examined as well.

**SECTION 2.0**

**LITERATURE REVIEW**

**2.1**    *Neolamarckia cadamba* **(Roxb.) Bosser**

*Neolamarckia cadamba* (Roxb.) Bosser, which is also known as Kelampayan belongs to the family Rubiaceae. It is widely distributed throughout Bangladesh, Nepal, India, Myanmar, Sri Lanka, the Philippines, Indonesia, and Papua New Guinea (Alam *et al*., 2008). Growth sites of Kelampayan include areas which are adjacent to streams or rivers, open sites and deep moist alluvial soils. According to Timber Technology Center (1999), Kelampayan can be found in lowland to mountain forests at about 1000 m. Kelampayan is characterized as a medium sized to large tree as it can reach a height of 40 m. The bole of Kelampayan is straight and buttresses that sometimes reach the height of two meter can be observed in certain Kelampayan (Timber Technology Center, 1999). As for Kelampayan flower, it is orange in colour, small in size and has a sphere shape, whereas the fruit consists of small capsules that are packed closely together (Peter, 2007).

Kelampayan wood is significant in the economic sector. The timber of Kelampayan is characterized as soft and light. The heartwood is white in colour with a yellow tinge and darkens to creamy yellow colour on exposure (Timber Technology Center, 1999). Timber of Kelampayan is under the classification of Light Hardwoods (LHW), where it has a density below 720 kg m$^{-3}$ and is not naturally durable in exposed condition (Malaysian Grading Rules, 1984 cited in Timber Technology Center, 1999). Since Kelampayan woods are relatively easy

to resaw and cross-cut, the timber has been used for the production of plywood, pulp and paper, boxes and crates, dug-out canoes, furniture components and light construction.

Apart from the timber industry, various parts from Kelampayan have medicinal uses as well. It was reported that the barks and leaves of Kelampayan possess astringent anti-hepatotoxiz, antidiuretic, wound healing, antiseptic and anthelmintic properties (Patel and Kumar, 2008). Alam *et al*. (2008) also established the antidiarrhoeal property of the hydroethanolic extract from the flowers tops of Kelampayan. Presence of indole alkaloids, secoiridoids, triterpenes and saponins in Kelampayan may lead to such antidiarrhoeal property (Alam *et al*., 2008). Besides that, Mondal *et al*. (2009) who studied the diuretic and laxative properties of different extraction from Kelampayan barks also suggested that the methanol extract from Kelampayan barks significantly increased the urinary output as well as urinary electrolyte concentration whereas the chloroform extract from Kelampayan barks produced significant laxative activity. Diuretic and laxative activities in other plants were found to be induced by phytoconstituents such as flavonoids, terpenoids, saponins and presence of these constituents in different extracts of Kelampayan may be responsible for its diuretic and laxative properties (Mondal *et al*., 2009).

## 2.2     Sucrose Synthase (SuSy) Gene

Sucrose synthase (*SuSy*) gene is of great importance to all plants as it encodes for sucrose synthase, which is the key enzyme for sucrose synthesis and breakdown. SuSy enzyme catalyzes the uridine diphosphate (UDP) dependent cleavage of sucrose into UDP-glucose and fructose. Such reaction is reversible.

Sucrose + UDP ⟷ UDP-glucose + Fructose

At pH 8-8.8, *SuSy* enzyme is capable of catalyzing the synthesis of sucrose from UDP-glucose and fructose. Hypoxic or anoxic condition on the other hand, will cause the falling of pH in cells, leading to degradation of sucrose by *SuSy* enzyme into UDP-glucose and fructose (Plaxton and McManus, 2006). Hence, under reduced oxygen condition, plants are still able to survive as *SuSy* enzyme will increase degradation of sucrose. According to Plaxton and McManus (2006), *SuSy* gene is regulated by the level of its own enzyme products. It was reported that sucrose, glucose and D-mannose are able to up-regulate *Sus1*, which is a major *SuSy* gene in *Arabidopsis thaliana* (Ciereszko and Kleczkowski, 2002). Mannose was found to be more effective in the induction of *Sus1* compared to glucose or sucrose.

Plants depend greatly on *SuSy* enzyme as it provides energy and carbon needed for plant growth and development through sucrose catabolism. Glucose and fructose formed from sucrose degradation may be phosphorylated by hexose kinase to produce hexose-6-phosphate

or phosphatases may catalysis hydrolyzation of glucose and fructose to produce free hexoses (Stewart *et al*., 2010). Hence, newly synthesized hexoses and hexoses-6-phosphate can be utilized in glycolytic or pentose phosphate pathway, leading to generation of energy in plants. UDP-glucose produced from breakdown of sucrose can also be used for plant cell wall biosynthesis.

*SuSy* enzyme is involved in starch formation as well. Starch is of great importance as it is used for energy storage in plants (Bettelheim *et al*., 2010). High activities of *SuSy* enzyme can be found in cytoplasm of starch storage tissue. UDP-glucose, which is formed from hydrolysis of sucrose, is converted to glucose-1-phosphate by UDP-glucose pyrophosphorylase (Atwell *et al*., 1999). Glucose-1-phosphate provides glucose moieties for starch synthesis. After glucose-1-phosphate is synthesized, it will be transported to amyloplasts, where synthesis and storage of starch granules happen.

Subbaiah *et al*. (2007) also suggested that *SuSy* enzyme may possess novel and noncatalytic biological function. Their study revealed that two *SuSy* isoforms from maize are partly localized to mitochondria and nuclei, compartments which are not related to sucrose metabolism. Interaction between *SuSy* enzyme and voltage-dependent anion channel (VDAC), which is the major outer mitochondrial membrane protein, may have a role in inter-compartmental signaling under anoxic stress (Subbaiah *et al*., 2007). Prolong anoxia condition will causes de-oligomerization of VDAC. This will result in the release of *SuSy* enzyme from the mitochondria, followed by migration of *SuSy* enzyme to the nucleus. The nuclear accumulation of SuSy then signals the induction of cell death pathway (Subbaiah *et al*., 2007).
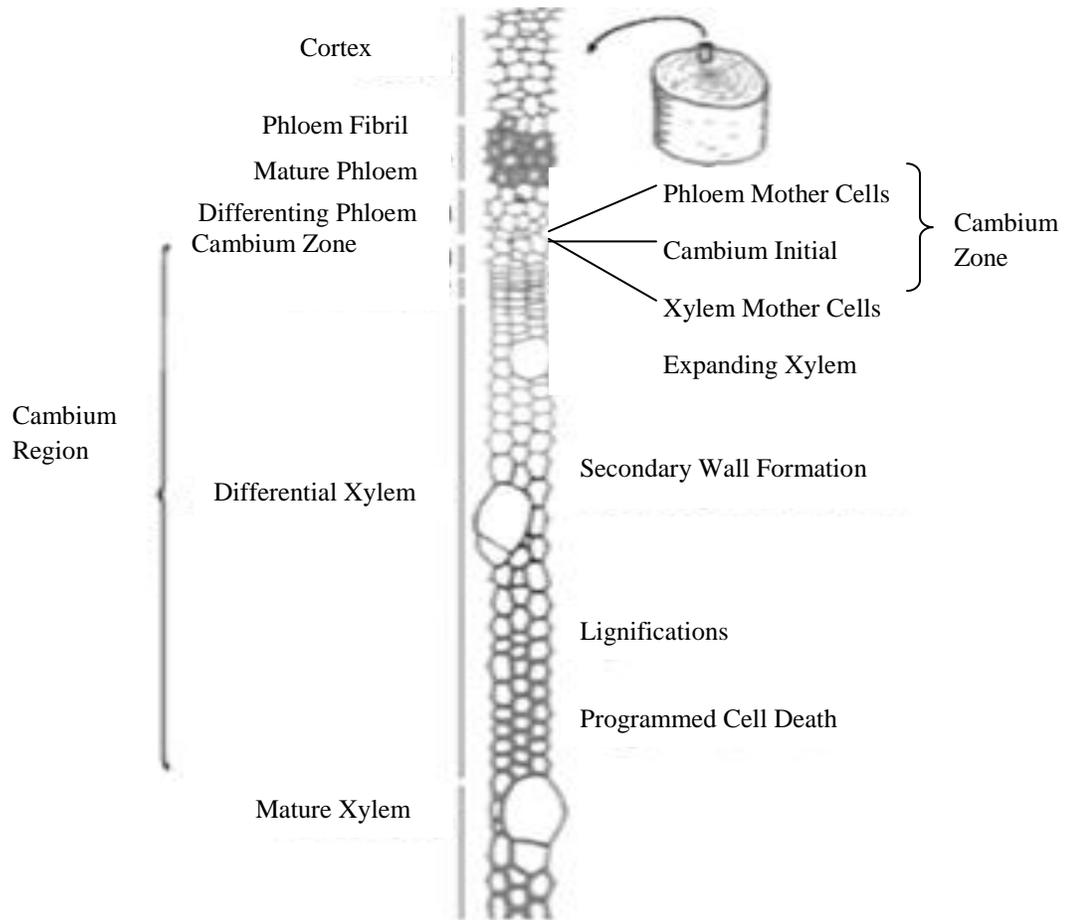
## 2.3     Wood Formation

Proper wood formation is crucial as timber can contribute to both economic and commercial sectors. Wood is formed from cellulose microfibrils, which are associated with hemicellulose, protein, and lignin (Brown and Saxena, 2007).

Wood biosynthesis occurs when there is successive addition of the secondary xylem (Kumar and Fladung, 2004). The processes for wood formation include division, expansion, maturation and cell death (Figure 2.1). First, cells which are located on the xylem side of the cambium will undergo division in order to produce more cells. Then the derivative cells will pass through an expansion zone, in which the cell volume is increased by either extension or elongation (Stokke and Groom, 2006). This is followed by maturation, where lignifications and secondary cell wall thickening happens. The final process in wood formation involved cell death, where all the cellular processes are ceased.

Plants possess primary and secondary cell walls. According to Morohoshi and Komamine (2001), it is the secondary cell walls that give rise to the mechanical strength of plants. Even though formation of both primary and secondary cell walls utilize cellulose, there are differences between the cell wall structure and process of cell wall formation. Primary cell walls are synthesized while the cells are still expanding whereas secondary cell walls are laid down once the cell has attained its final shape (Morohoshi and Komamine, 2001). Apart from that, orientation and arrangement of cellulose microfibrils is random or longitudinal in primary cell walls while secondary cell walls are composed of dense array of helical and almost transverse cellulose microfibrils (Kumar and Fladung, 2004). Morohoshi and Komamine (2001) also suggested that the cellulose-xyloglucan network, which is present in cells that

undergo primary cell wall formation, controls the cell expansion whereas direction of cell expansion is regulated by the orientation of cellulose microfibrils within the walls.

Cortex

Phloem Fibril

Mature Phloem

Differenting Phloem
Cambium Zone

Phloem Mother Cells

Cambium Initial

Cambium
Zone

Xylem Mother Cells

Expanding Xylem

Cambium
Region

Differential Xylem

Secondary Wall Formation

Lignifications

Programmed Cell Death

Mature Xylem

**Figure 2.1** Wood formation process (Source: Retrieved from http://fgilab.com/wp-content/uploads/2010/05/biotechcol3_wsho.pdf).

## 2.4 Sucrose Synthase (*SuSy*) Gene in Cellulose Biosynthesis

Cellulose is the main constituent of wood. It is one of the principal components in both primary and secondary cell walls. According to Sjostrom (1993), 40-45% of the dry substance in most wood species is cellulose, which is located predominantly in the secondary cell walls of plants. Cellulose, which is a homopolysaccharide, is made up of β-D-glucopyranose units that are linked together by (1, 4)-glycosidic bonds (Sjostrom, 1993). Cellulose molecules are linear and are able to form intramolecular and intermolecular hydrogen bonds. Formation of strong hydrogen bonds therefore allows cellulose to exhibit high tensile strength in plants.

Cellulose is synthesized in plants when the glucose residue from UDP-glucose is transferred to the growing 1, 4-β-glucan chain (Rai and Takabe, 2006). Thus, this has indicated the importance of UDP-glucose which serves as a glucosyl precursor in the formation of cellulose. Two different enzymes, which are UDP-glucose pyrophosphorylase and sucrose synthase (*SuSy*), can catalysis for the formation of UDP-glucose in plant cells.

$$\text{UDP-glucose} + (1,4\text{-}\beta\text{- glucan})_n \longrightarrow (1,4\text{-}\beta\text{- glucan})_{n+1} + \text{UDP}$$
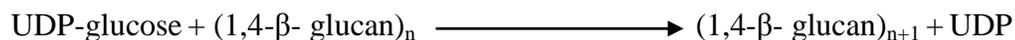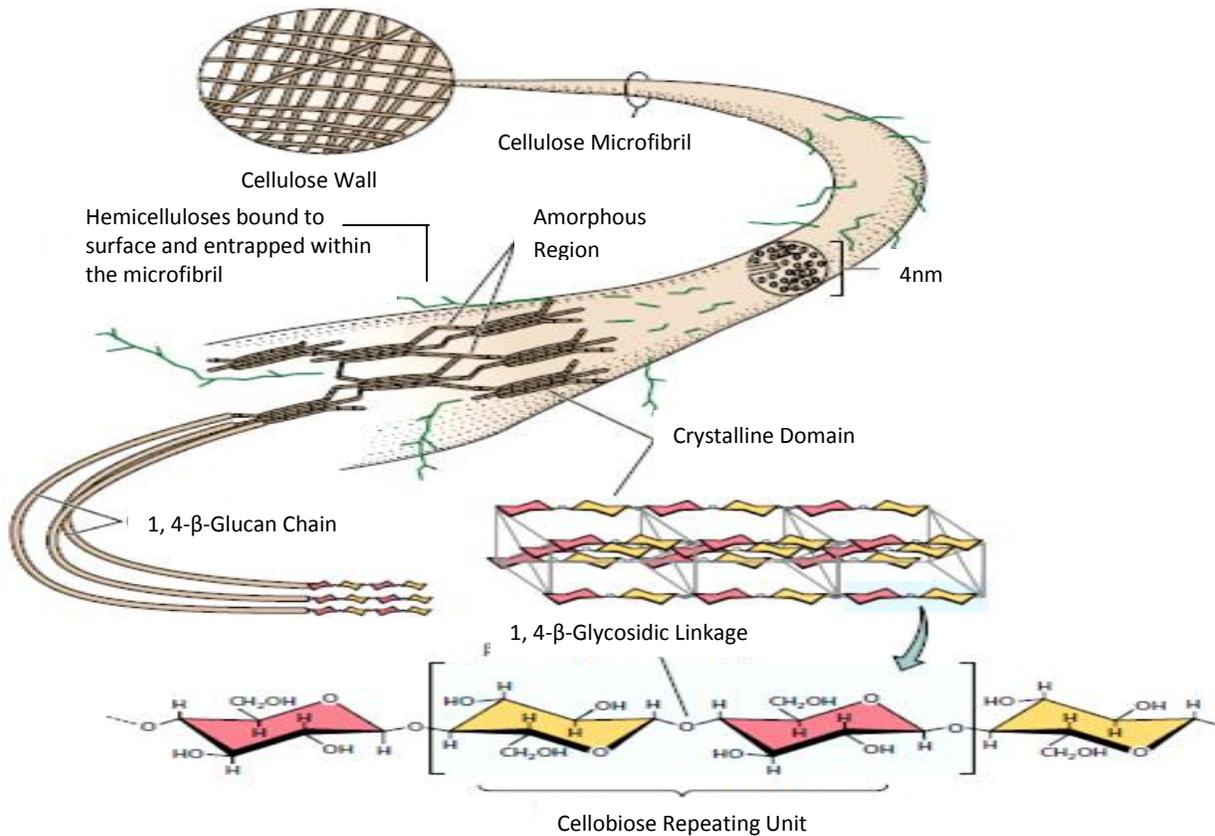
Sjostrom (1993) suggested that *SuSy* enzymes located in the cytoplasm are responsible for the production of UDP-glucose in plant cells. UDP-glucose will then penetrate into the plasma membrane and transfer the glucosyl residue to the growing glucan chain. However, recent studies have discovered that a substantial amount of *SuSy* enzymes are tightly bound to

the plasma membrane of several plant species (Basra, 2006). This plasma membrane associated *SuSy* enzyme facilitates the direct channeling of carbon (UDP-glucose substrate) from sucrose to the cellulose synthesizing machinery. The cellulose synthase complex, which is organized in the form of hexagonal rosettes, will then polymerizes the glucose monomers into glucan chains while recycling UDP back to *SuSy* enzyme (Joshi *et al*., 2004). Another membrane associated cellulase, known as KORRIGAN (KOR) will be monitoring and editing the conversion of glucan chain to cellulose microfibril (Figure 2.2). Microfibrils will further build up fibrils and finally cellulose fibers (Figure 2.3) (Sjostrom, 1993).

Utilization of *SuSy* enzyme in cellulose synthesis is beneficial to plants. This is because *SuSy* enzyme catalyzed an energy conservative reaction with respect to ATP (Rai and Takabe, 2006). On the other hand, two molars of ATP are needed for the synthesis of one molar UDP-glucose by enzyme UDP-glucose pyrophosphorylase. Apart from that, *SuSy* enzyme can prevent accumulation of UDP as well. According to Basra (2006), UDP is a compound that can inhibits the reaction catalyzed by cellulose synthase. Hence, UDP release from cellulose synthesis will be recycled by *SuSy* enzyme to product UDP-glucose (Rai and Takabe, 2006).
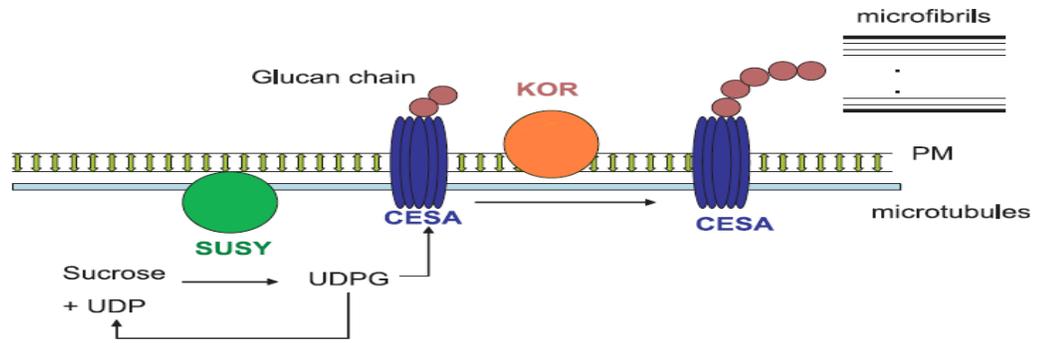
**Figure 2.2** Cellulose biosynthesis in plant. Plasma membrane (PM) associated form of sucrose synthase (*SuSy*) directly channels Uridine diphosphate glucose (UDPG) substrate to cellulose synthase (CESA) rosette complex that aid in glucan chain formation. Glucan chains self assemble into microfibrils and KORRIGAN cellulase (KOR) acts as an editor/monitor of this process (Joshi *et al*., 2004).



**Figure 2.3** Structural model of a cellulose microfibril. (Source: Retrieved from http://www.biol.unlp.edu.ar/biologiavegetal/materialdidactico01-modulo01.pdf).

11

## 2.5     Single Nucleotide Polymorphism (SNP)

Single nucleotide polymorphism (SNP) refers to DNA sequence variation that occurs when a single nucleotide in the genome sequence differs between two individual DNA samples (Cullis, 2004). As SNPs are highly abundant in the plant DNA, they can serve as valuable markers for studying agronomic traits in plant species.

In order for a variation to be considered as a SNP, it must occur in at least 1% of the population (Xu, 2010). Based on the nucleotide substitution, SNPs can be categorized into two primary types, which are transition substitution and transversion substitution. Transition substitution occurs when one purine is replaced by another purine or when one pyrimidine is replaced by another pyrimidine (C/T or G/A). On the other hand, transversion substitution refers to substitution of purine with pyrimidine and vice versa (C/G, A/T, C/A or T/G). As two types of transitions and four types of transversions are observed, ratio of transition to transversion should be 0.5. But estimated rates show that this ratio is biased towards transition (Pratik, 2007). It was reported that C/T transitions constitute about 67% of the SNP in plants (Edward *et al*., 2007a cited in Xu, 2010) as two of every three SNPs involve substitution of cytosine with thymine.

SNPs can occur in the coding and the non coding region of the genome. Nucleotide substitution in the coding region can be further divided into synonymous and non-synonymous. SNP is termed as synonymous when the same amino acid sequence is formed after nucleotide substitution. Such condition is due to the redundancy in genetic code (Xu, 2010). Non-synonymous on the other hand, refers to nucleotide substitution that leads to production of a different polypeptide sequence. Most of non-synonymous mutations are

deleterious (Beaumont *et al*., 2003) and is the cause of genetic disease (Swynghedauw, 1995). Hence, natural selection will eliminate non–synonymous substitution. As for SNP that falls into the non coding region, it may affect gene splicing, transcription factor binding or the sequence of non coding RNA as well (Xu, 2010).

# SECTION 3

# MATERIAL AND METHODS

## 3.1 Plant Materials

Fresh leaves were collected from six selected Kelampayan tree in Kelampayan Trial Plot, Landeh Nature Reserve, Semengoh, Sarawak.

## 3.2 DNA Extraction

### 3.2.1 Chemicals and Reagents

Reagents that were used for DNA isolation include liquid nitrogen, CTAB extraction buffer [100 mM Tris-Cl pH 8.0; 1.4 M NaCl; 20 mM EDTA; 2% CTAB; 1% polyvinylpyrrolidone (PVP); 2% (v/v) β-mercaptoethanol], chloroform/isoamyl alcohol (24:1 v/v), isopropanol, 70% ethanol, TE buffer, and sterile distilled deionized water (ddH$_2$O).

### 3.2.2 DNA Isolation Protocol and DNA Purification

Modified CTAB method from Doyle and Doyle (1990) were used to extract total genomic DNA from Kelampayan whereas Wizard Genomic DNA Purification Kit (Promega, USA) was used to purify the isolated DNA.