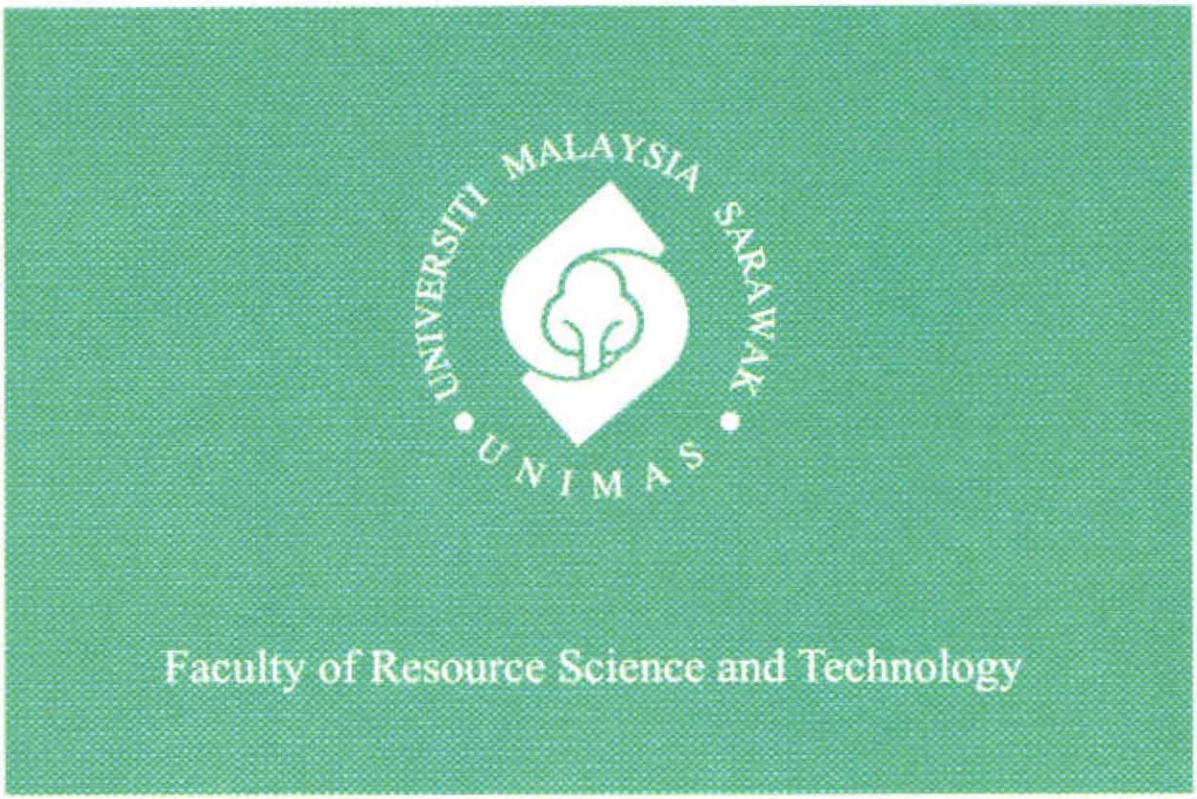

SEQUENCE POLYMORPHISM OF SUCROSE SYNTHASE (SUSY) GENE IN A TROPICAL TIMBER TREE *NEOLAMARCKIA MACROPHYLLA*

Tan Choon Ju

(33215)

Bachelor of Science with Honours
(Resource Biotechnology)
2014

Sequence Polymorphism of Sucrose Synthase (*Susy*) Gene in a Tropical Timber Tree *Neolamarckia macrophylla*

Tan Choon Ju (33215)

A Thesis Submitted in Partial Fulfilment of the Requirement for the Degree of Bachelor of Science with Honours
(Resource Biotechnology)

Supervisor: Dr. Ho Wei Seng

Resource Biotechnology
Department of Molecular Biology

Faculty of Resource Science and Technology
University Malaysia Sarawak
2014

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my supervisor, Dr. Ho Wei Seng for his guidance, suggestion and encouragement. Besides that, I would also like to show my appreciation to my co-supervisor, Dr. Pang Shek Ling for her guidance and advice. I am grateful for all the help and support from my beloved lab members. Last but not least, a special thanks to my family who have been supporting me throughout my final year project.

UNIVERSITI MALAYSIA SARAWAK

Grade: _____

Please tick (√)
Final Year Project Report ☐
Masters ☐
PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the .....day of..... .....2014.

Student's Declaration:

I .... (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled .... is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

1/7/2014
Date submitted

Tan Chun Tu 33215
Name of the student (Matric No.)

Supervisor's Declaration:

I..... (SUPERVISOR'S NAME) hereby certifies that the work entitled .....(TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of ..... (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

Received for examination by: _____ Date:_____
(Name of the supervisor)

I declare that Project/Thesis is classified as (Please tick (√)).

☐ CONFIDENTIAL (Contains confidential information under the Official Secret Act 1972)*
☐ RESTRICTED (Contains restricted information as specified by the organisation where research was done)*
☐ OPEN ACCESS

Validation of Project/Thesis

I therefore duly affirm with free consent and willingly declare that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abiding interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitalise the content for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student itself neither third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Student signature _____ (Date) 1/7/2014

Supervisor signature: _____ (Date)

Current Address:

_____

_____

_____

Notes: * If the Project/Thesis is CONFIDENTIAL or RESTRICTED, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument is duly prepared by The Centre for Academic Information Services]

Table of Content

Acknowledgement	I
Declaration	II
Table of Contents	IV
List of Abbreviation	VI
List of Tables	VII
List of Figures	VIII
Abstract	1
Chapter 1 Introduction	2
Chapter 2 Literature Review	
2.1 *Neolamarckia macrophylla*	4
2.2 Sucrose Synthase *(SuSy)* Gene	6
2.3 Sucrose Synthase (*SuSy*) Gene in Cellulose Biosynthesis	8
2.4 Single Nucleotide Polymorphism (SNP)	11
Chapter 3 Materials and Methods	
3.1 DNA Extractions	13
3.2 Agarose Gel Electrophoresis	13
3.3 DNA Quantification	13
3.4 Polymerase Chain Reaction (PCR)	13
3.5 PCR Product Purification	14
3.6 DNA Sequencing and Data Analysis	14
Chapter 4 Results and Discussion	
4.1 Plant Samples	15

4.2 DNA Extraction and Purification 15

4.3 DNA Quantification 19

4.4 Polymerase Chain Reaction 20

4.5 PCR Product Purification 20

4.6 DNA Sequencing and Data Analysis 22

Chapter 5 Conclusion 32

References 33

List of Abbreviations

ATP	Adenosine triphosphate
CIA	Chloroform-Isoamyl alcohol
CTAB	Cetyltrimethylammonium Bromide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetraacetic acid
KOR	Korrigan cellulose
NaCI	Sodium Chloride
PCR	Polymerase Chain Reaction
RACE	Rapid Amplification of cDNA Ends
SNP	Single nucleotide polymorphism
SuSy	Sucrose synthase
UDP	Uridine diphosphate
UV	Ultraviolet

List of Tables

Tables	Page
4.1 Absorbance reading and absorbance ratio of sample 1 to sample 15	19
4.2 BLASTn analysis for amplified partial *SuSy* DNA	22
4.3 Summary of the sequence variation in exons and introns	30

List of Figures

Figure	Page
2.1 Structure model of cellulose microfibril	8
2.2 Cellulose biosynthesis in plants.	9
4.1 Gel electrophoresis of genomic DNA samples on 0.8% agarose gel	17
4.2 Gel electrophoresis of purified PCR products.	21
4.3 Alignment of 15 partial sucrose synthase using CLC Sequence Viewer 6	25
4.4 Comparison of gene structure between *Populus trichocarpa SuSy 1* genomic DNA (Accession no: GU592809.1), *Populus trichocarpa SuSy* mRNA (Accession no: GU559729.1) and the consensus sequence of the 15 sucrose synthase amplicons.	28
4.5 Alignment of 15 partial sucrose synthase amino acid sequences	29

Sequence Polymorphism of Sucrose Synthase (*SuSy*) Gene in a Tropical Timber Tree *Neolamarckia macrophylla*

Tan Choon Ju

Resource Biotechnology
Faculty of Science and Technology
University Malaysia Sarawak

ABSTRACT

Neolamarckia macrophylla or red kelampayan contains high economic value as the timber can be used to produce plywood, furniture, canoe and paper. However, since traditional plant breeding selection methods are time consuming, single nucleotide polymorphism (SNP) has been used as a molecular marker for molecular breeding. SNPs provide an easier and faster way to select the specific desired plant character. This study was done to determine the nucleotide variation in the sucrose synthase (*SuSy*) gene of red kelampayan and to determine the synonymous and non-synonymous mutations of sucrose synthase gene. Total genomic DNA was first extracted from the 15 red kelampayan samples. Polymerase chain reaction (PCR) was then performed to obtain the desired *SuSy* amplicons and purified amplicons were sent for sequencing. Next, BLASTn analysis was performed on the sequences in order to find sequence homology against the non-redundant nucleotide database on NCBI. CLC Free Sequence Viewer 6 was then used to detect the nucleotide variation in the samples. This study showed that there was total of five SNPs in the 15 partial sucrose synthase (*SuSy*) gene. 4 SNPs were located at the predicted coding region and one SNP located at predicted non-coding region. By determining SNPs in the sucrose synthase gene of red kelampayan, it can be used as a molecular marker for the selection of red kelampayan with the desired phenotypic wood traits for breeding purposes in the future.

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

ABSTRAK

Neolamarkia macrophylla atau kelampayan merah mengandungi nilai ekonomi yang tinggi kerana kayunya boleh digunakan untuk menghasilkan papan lapis, perabot, kanu serta kertas. Disebabkan kaedah pemilihan pembiakan tumbuhan secara tradisional sangat memakan masa, polimorfisme nukleotida tunggal (SNP) telah digunakan sebagai penanda molekul untuk tujuan pembiakan. SNP merupakan cara yang mudah dan cepat dalam pemilihan pembiakan tumbuhan. Kajian ini telah dijalankan untuk mengesan variasi nukleotida dalam gen sukrosa sintase kelampayan merah dan menentukan mutasi sinonim dan bukan sinonim gen sukrosa sintase. Untuk menjalankan kajian ini, DNA telah diekstrak daripada 15 kelampayan merah. Tindak balas berantai polymerase telah digunakan untuk mendapatkan amplikons *SuSy* dan dihantar untuk memperolehi urutan *SuSy* yang dikehendaki. Urutan yang diperolehi akan dianalisis dengan BLASTn untuk mencari homologi urutan terhadap pangkalan data nukleotida di NCBI. CLC Sequence Viewer 6 kemudianya digunakan untuk mengesan variasi nukleotida dalam 15 urutan dengan menggunakan penjajaran urutan. Kajian menunjukkan ada lima SNP dalam 15 urutan gen sukrosa sintase. Empat SNPs terletak di exon dan satu terletak di intron. Dengan menentukan SNP dalam gen sintase sukrosa kelampayan merah, ia boleh digunakan sebagai penanda molekul untuk pemilihan kelampayan dengan sifat yang dikehendaki untuk tujuan pembiakan pada masa hadapan.

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

CHAPTER 1

INTRODUCTION

Neolamarckia macrophylla or commonly known as red kelampayan belongs to the Rubiaceae family. It is a fast growing tropical tree native to Eastern Indonesia and has been cultivated widely in many countries such as in Malaysia, India and Philippines. It is able to reach diameter at breast height of 30 cm within 5 - 7 years (SEAMEO BIOTROP, 2008). The leaves of the red kelampayan are hairy and reddish green in colour and it has hard reddish timber. According to Krisnawati et al. (2011), the timber can be used to make all sorts of stuff such as plywood, furniture, canoe, light construction.

Sucrose is important to plant growth and development. According to Winter and Huber (2000), sucrose is the main photosynthase and major transport carbohydrate in plants. Sucrose synthase (*SuSy*) catalyses the reversible synthesis and degradation of sucrose. Since sucrose cannot be utilized by plant directly, sucrose needs to be cleaved by sucrose synthase (*SuSy*) into fructose and UDP-glucose whereby UDP-glucose serves as the precursor for cellulose biosynthesis. Sucrose synthase gene which encodes for sucrose synthase enzyme is therefore important in wood formation (Winter and Huber, 2000). As the timber of kelampayan contributes mainly to its economic value, study on the SNP of sucrose synthase gene, a gene involved in cellulose biosynthesis can be used as a marker in molecular breeding in order to select the specific desired characters in wood formation.

Traditional plant breeding methods are expensive and time consuming. Single nucleotide polymorphism (SNP) has been used as the simplest form of molecular marker in molecular breeding since single nucleotide base is the smallest unit of inheritance and so it is able to give the maximum markers. It can also be automated easily (Guo, 2013). The

abundance of SNPs in the genomes also provides an easier and faster way to select the specific desired plant characters. Furthermore, SNPs are considered to be stable as the mutation rates are quite low (Jehan and Lakhanpaul, 2006).

The main aim of the study was to determine the DNA sequence variation in the sucrose synthase (*SuSy*) gene of red kelampayan caused by the single nucleotide substitutions. Besides that, this study also aimed to determine the synonymous and non-synonymous mutation of the sucrose synthase gene.

CHAPTER 2

LITERATURE REVIEW

2.1 *Neolamarckia macrophylla*

Neolamarckia macrophylla or *Anthocephalus macrophyllus*, commonly known as red kelampayan or red jabon belongs to the Rubiaceae family. It is a fast growing tropical tree species to Eastern Indonesia and has been cultivated widely in many countries such as in Malaysia, India and Philippines as ornamental tree, in plantation forest as well as community forest. Besides that, it is more resistant towards pests and diseases as compared to white kelampayan due to the presence of tannin. Timber of the tree is also widely used to produce plywood, light construction, furniture as well as moulding (SEAMEO BIOTROP, 2008).

Red kelampayan has straight bole and high free branch. It is tolerant to light and able to grow on different soil type such as in marginal land. Timber of the red kelampayan is also more vigorous and harder than *Neolamarkia cadamba* (white kelampayan). It is a perennial tree and can grow on lowland and lower montane forest in slightly seasonal climate (Sosef, 1897). Red kelampayan can be planted as conservation crop near the river as it has high ability of absorbing or holding water (Yudohartono, 2013). The average height and diameter of red kelampayan is 3.8m and 5cm (Setyaji et al., 2013).

Red kelampayan is multipurpose as different parts of the tree serve for different purposes. According to Krisnawati et al. (2011), timber is the most useful part of the tree and can be used to make all sorts of stuff such as plywood, furniture, canoe, light construction, flooring, beams and rafters, carvings, ceiling boards, boxes and crates, tea-chests, packing cases, wooden shoes, toys, matches, chopsticks and pencils. Pulp of the

tree can also be processed to make papers. It is planted as ornamental trees along roadsides for shading and acts as shelter for crops to avoid excessive sunlight. It is also planted for reforestation to improve the chemical and physical properties of the soil. This is because it has big leaf and non-leaf litter that can increase the amount of organic carbon, nutrients, cation and bases exchange in the soil. Apart from that, the leaves can be used as plates and serviettes or fodder for cattle. The root's bark also contains yellow dye that can be used as dyestuff or tannin (Krisnawati et al., 2011).

2.2 Sucrose Synthase (*SuSy*) Gene

Sucrose is important to the plant growth and development. It is the main photosynthate and serves as the main transport sugar, nutrients and potential signal molecule in plants. Sucrose synthase (*SuSy*) gene encodes for sucrose synthase to catalyse the reversible synthesis and degradation of sucrose. Normally, *SuSy* catalyses the reversible degradation of sucrose and uridine diphosphate (UDP) into UDP-glucose and fructose in plant sink tissues whereby UDP-glucose serves as the precursor for cellulose biosynthesis (Winter and Huber, 2000).

Sucrose + UDP ↔ UDP-glucose + Fructose

SuSy enzyme catalyses the synthesis of sucrose from UDP-glucose and fructose at pH 8 - 8.8. However, when the pH of the cell falls to pH 6 – 6.5 under hypoxic conditions, plants are able to accommodate to the environmental stress by catalysing the degradation of sucrose into UDP-glucose and fructose. *SuSy* enzyme is able to synthesis or degrade sucrose at different environmental condition and the level of the end products regulates the gene itself (Plaxton and McManus, 2006). Degradation of sucrose *in vivo* however is more common in plants than synthesis of sucrose (Harada et al., 2005).

It is proposed that *SuSy* exists in 2 forms: soluble *SuSy* (S-*Susy*) and particulate *SuSy* (P-*SuSy*). Soluble *SuSy* can be found in the cytoplasm while particulate *SuSy* is bound on the plasma membrane (Carlson et al., 2002). S-*SuSy* allocates carbon for metabolic work such as respiration, producing storage and building blocks molecules (Haigler *et al.* 2001). It cleaves sucrose into UDP-glucose and fructose whereby fructose is used in glycolysis and starch formation (Harada et al., 2005). On the other hand, P-*SuSy* supplies

UDP-glucose from the degradation of sucrose for cellulose and callose synthesis. Polysaccharide is produced by using the high energy bond conserved during the process. (Haigler *et al.* 2001). The activation of *SuSy* is thought to be related to adaptation to anaerobic condition. In maize roots, *SuSy* acts as anaerobic proteins (ANPs) under anaerobic conditions to degrade sucrose. It requires less ATP and thus more energy-efficient (Harada et al., 2005). According to Winter and Huber (2000), phosphorylation enabled the conversion of P-*SuSy* into S-*SuSy*.

Apart from catalysing sucrose, *SuSy* is suggested to be present in mitochondrial of maize to regulate nutrient or sugar signalling through the opening of mitochondrial permeability transition pores. In addition, it is also present in chloroplasts whereby it helps in regulating the photosynthetic processes in plastids. Besides that, sucrose synthase is a key regulatory protein and potential abiotic stress biomarker in plants. Under abiotic stress conditions such as osmotic stress or draught, excessive synthesis of sucrose occurs in order to maintain membrane function and structure. However, under energy limiting conditions such as oxygen deprivation, sucrose degradation is favoured due to lower overall energy costs (Jayashree et al., 2008).

2.3 Sucrose Synthase (*SuSy*) Gene in Cellulose Biosynthesis

Cellulose is the main component of wood, located predominantly in the secondary cell walls and makes up to 40-45% of the wood's dry weight. It forms hydrogen bonds and is linked by van der Waals forces (Perez et al., 2002). It combines with hemicellulose and lignin to form woods. It is a homopolymer of β-D-glucopyranose linked by (1, 4)-glycosidic bonds to form long linear unbranched chain that associated to form the elemental fibrils. The elemental fibrils are very stable and resistant to chemical attacks due to high degree of inter and intramolecular hydrogen bonding between chains of native cellulose. Hydrogen bonding between the chains also produces more rigid polymers and prevents flexing of molecules during hydrolytic breaking of glycosidic linkages (Lee, 2013).

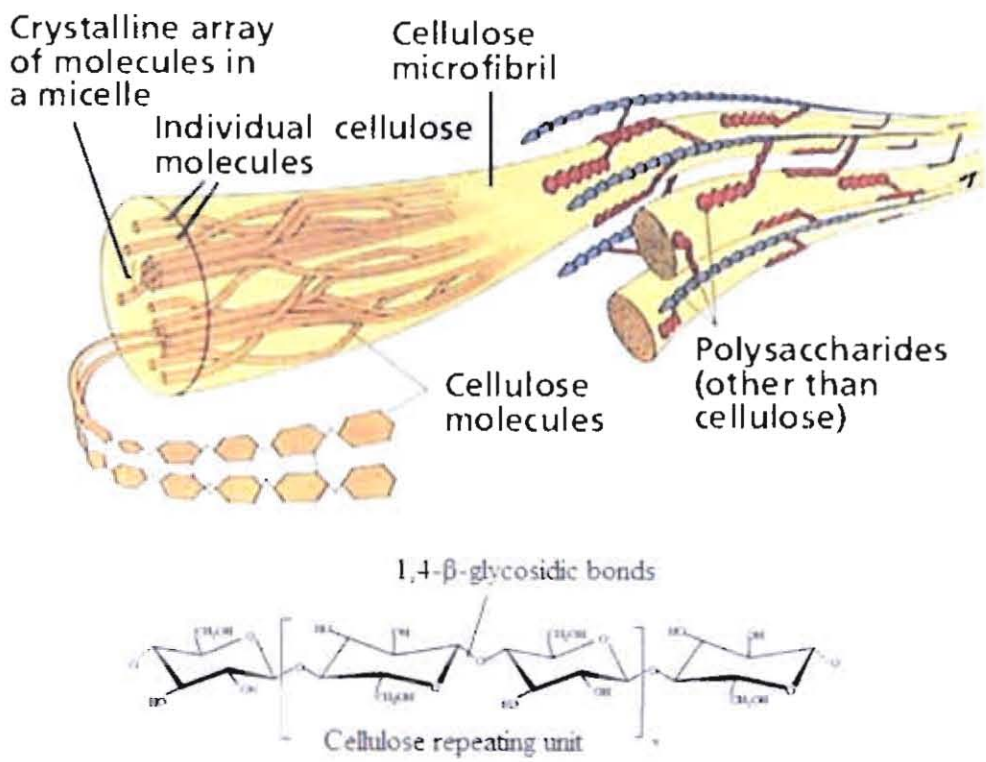

Figure 2.1 Structure model of cellulose microfibril (Source: Retrieved from http://bioserv.fiu.edu/~walterm/human_online/chemistry/water_and_molecules/water_and_organic_molecules.htm).

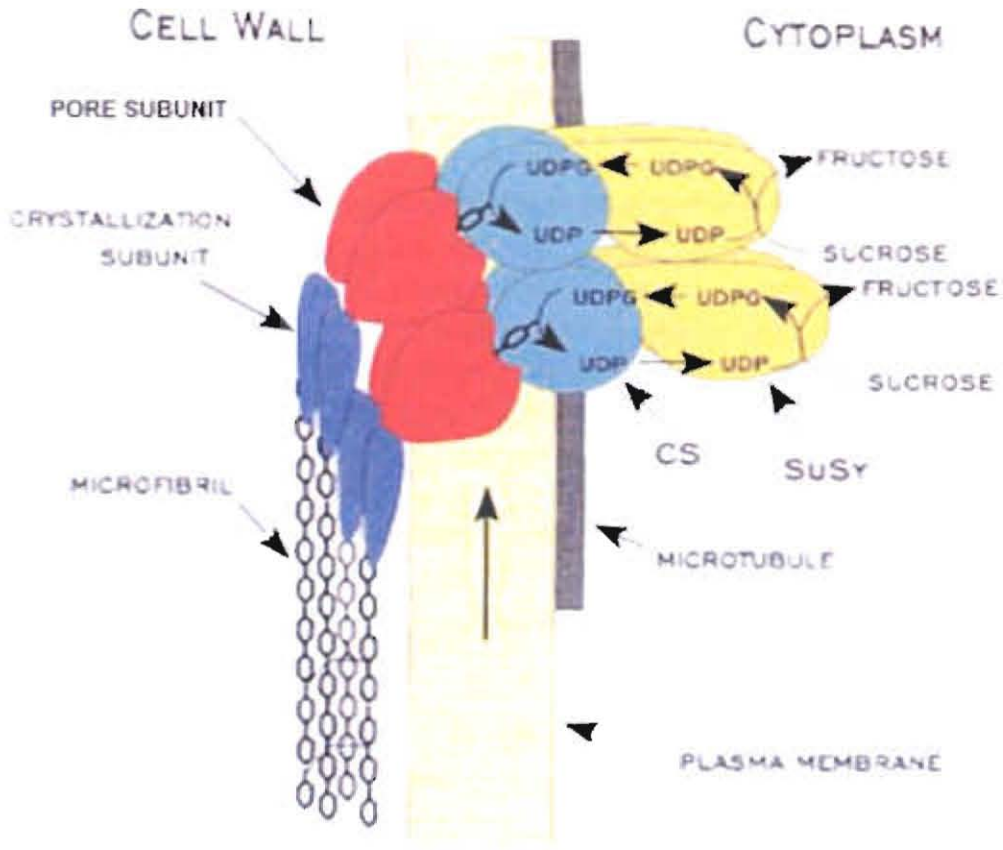

Figure 2.2 Cellulose biosynthesis in plants. Sucrose synthase (*SuSy*) converts sucrose into UDP-glucose that is needed for cellulose synthesis. Cellulose synthase (CS) then catalyzes the β-1, 4-glycosidic bonds to formed cellulose polymers (Delmer and Amor, 1995).

UDP-glucose produced during the degradation of sucrose by S-*SuSy* serve as the precursor for cellulose biosynthesis in the cell wall. According to Rai and Takabe (2006), cellulose is formed when glucose residue from UDP-glucose is transferred to the growing 1, 4-β-glucan chain. Cellulose synthase polymerize the glucose to form 1, 4-β-glucan chain and the UDP is recycled back into the *SuSy* enzymes. P-*SuSy* guides the channeling of UDP-glucose from glucose to the cellulose synthesizing machinery (Dulfresne, 2012). Korrigan, a membrane associated cellulase then controls the production of glucan chain into cellulose microfibril that later associated to form cellulose chain (Sjostrom, 1993).

Cellulose synthesis by *SuSy* enzyme is an energy-saving reaction as it uses less ATP. In addition to that, *SuSy* is able to prevent the accumulation of UDP as it is used in wood formation. It recycles the UDP produced during cellulose synthesis to produce UDP-glucose so that less UDP residue accumulates in the plant and will not inhibit the cellulose synthesis (Rai and Tabake, 2006).

2.4 Single Nucleotide Polymorphism (SNP)

Single nucleotide polymorphisms (SNP) refers to the DNA sequence variation that occurred when a single nucleotide base differs between two individual DNA samples (Guo, 2013) whereby the least frequent allele is at least 1% or greater (Jehan and Lakhanpaul, 2006).

According to Jehan and Lakhanpaul (2006), SNPs is a potent molecular marker since single nucleotide base is the smallest unit of inheritance (Guo, 2013). It is also fairly stable due to the low mutation rates. Besides that, SNPs can be used to construct high-density genetic map and served as basis for superior informative genotyping assay as it is highly abundant in the plants' genome. Functional polymorphisms can be identified from the changing phenotype caused by SNPs in the coding regions. In addition to that, it can be easily automated, detected and time saving due to the non-gel based properties. Since SNPs are biallelic, it makes them less informative per locus. Nevertheless, the abundance of SNPs in the genomes enabled the use of more loci to overcome the problem. However, the start-up cost can be quite expensive. High-quality DNA and high technical equipment have also limit the use of SNPs markers (Jehan and Lakhanpaul, 2006).

Based on the nucleotide substitutions, SNPs can be grouped into transition or transversion. Transition occurs when 1 purine is replaced by another purine or 1 pyrimidine is replaced by another pyrimidine (C/T or G/A) (Guo, 2013). C/T transition is very common as it accounts for 67% of the SNPs (Xu, 2010). On the other hand, transversion occurs when purine is replaced by pyrimidine or vice versa (C/G, A/T, C/A or T/G). In fact, transition is more commonly happened than transversion (Guo, 2013).

SNPs occur on the coding region, non-coding region of a gene or in the intergenic regions between the genes at different frequencies and chromosome regions (Guo, 2013). Redundancy in the genetic code does not necessary altered the amino acid sequence within

a coding sequence. Same polypeptide sequence formed by SNPs in coding region is known as synonymous mutation. It is known as non-synonymous mutation if a different polypeptide sequence is produced in the coding region. Non-synonymous mutation is deleterious and causes genetic diseases. However, it is eliminated by natural selection. In contrast, SNPs in the non-coding regions may affect gene splicing, transcription factor or sequence of the non-coding RNA (Xu, 2010).

CHAPTER 3

MATERIALS AND METHODS

3.1 DNA Extractions

Modified CTAB extraction protocol from Doyle and Doyle (1990) was used to extract total genomic DNA from the plant samples and RNases was used to purify the DNA from RNA.

3.2 Agarose Gel Electrophoresis

Agarose gel of 0.8% was prepared. 3 µl of extracted DNA was mixed with 1 µl of 1X loading dye, 3 µl of λ HindIII (Promega, USA) was used as the DNA marker. The agarose gel electrophoresis was left to run for around 90 minutes with 60V and 80mA. The gel was stained with ethidium bromide for ten seconds. Distilled water was used to de-stain the gel for 30 – 40 minutes. Visualisation of the band was done under UV transilluminator.

3.3 DNA Quantification

NanoDrop ^{TM} 1000 spectrophotometer was used to determine the concentration of DNA. Absorbance reading at wavelength of 230nm, 260nm and 280nm were taken. Absorbance ratio of A_{260}/A_{280} and A_{260}/A_{230} were then used to estimate purify of DNA. Absorbance at wavelength 260nm was also used to estimate the concentration of DNA. By comparing band intensity of HindIII with the DNA band, DNA concentration was estimated as well.

3.4 Polymerase Chain Reaction (PCR)

PCR was performed by using Mastercycler Gradient Thermal Cycler (Eppendorf, Germany). The DNA template was mixed with 1 x PCR buffer, 0.2 mM dNTPs, 1.5 mM $MgCI_2$, 5 pmol of forward and reverse primer, 1 U *Taq* DNA Polymerase (Invitrogen,

USA) and sterile distilled water to a total of 25 µl. The sequence for SuSy forward primer was 5' TTG GAA GAG CAG GCA GAG AT 3' while the sequenee SuSy reverse primer was 3' CCG CAG ACA TCT ACT CAA CAG 5'. The mixture then undergone pre-denaturation for 2 minutes at 95oC, 35 cycles of denaturation for 30 seconds at 94°C, annealing for 30 seconds at 55°C, extension at 72°C for 30 seconds and lastly a final extension of 8 minutes at 72°C. Agarose gel electrophoresis was performed to check on the quality of the PCR product in 1.5% agarose gel.

3.5 PCR Product Purification

PCR product was purified by using Thermo Scientific GeneJet PCR Purification Kit. The purified products were then examined on 1.5% (w/v) agarose gel electrophoresis for 90 minutes at 60 V and 80 mA to check for the concentration and quality.

3.6 DNA Scquencing and Data Analysis

The purified DNA was sent for sequencing by using Applied Biosystems 373xl DNA Analyzer. The homology in the consensus sequence of the samples was checked on BLASTn (http://blast.ncbi.nlm.nih.gov/). The consensus sequences were then aligned by using CLC Sequence Viewer 6 in order to determine the SNP among the samples.