CHLOROPLAST DNA DIVERSIFICATION OF MALAYSIAN PINEAPPLE

NORFADILAH BINTI HAMDAN

A dissertation is submitted in partial fulfillment of the requirements for the award of the degree of Master of Science (Biotechnology)

Faculty of Biosciences and Medical Engineering
Universiti Teknologi Malaysia

JANUARY 2013
To my Parent,

Sisters,

and Friends,

Thanks for your pray, attention and spiritual support...
ACKNOWLEDGEMENT

Alhamdulillah, I praise to the almighty Allah for giving me the strength and patience to complete my research. This project would not have been possible without the support of many people and I deeply indebted to many people who helped in the completion of this thesis. I would like to begin by thanking my sole supervisor Dr. Azman Abd. Samad and co-supervisor Dr. Topik Hidayat for their guidance, advice and support. My deep appreciation to both of them for giving me the advice, guidance and suggestions I needed to complete my research. I would also like to express my sincere gratitude to Dr. Faezah Salleh for her tips and consultation in completion and improvement of this thesis. Last but not least, words are inadequate to express my thanks to my parents, family members, lab mate Ms Shahkila Mohd Arif, and numerous friends who endured this long process with me, always offering valuable advice, support and love.
ABSTRACT

There are several genes are normally used for the phylogenetic study such as Mitochondria DNA (mtDNA), Chloroplast DNA (cpDNA) and Nuclear DNA (nDNA). Among these three genomes, cpDNA is commonly being used in phylogenetic study because it is easy to amplify via PCR, evolves at a conservation rate and it has appropriate length and base substitution rate for inferring phylogeny at higher levels. In this study, phylogenetic tree of eight Malaysian pineapple (*Ananas comosus*) cultivars were constructed using sequences of large subunit of the ribulose-bisphosphate carboxylase (rbcL) gene. A rbcL gene was isolated from genomic DNA, amplified and sequenced. Phylogenetic analysis was carried out using maximum parsimony method. Results revealed that rbcL gene of *Ananas comosus* is about 1250 bps. Based on the tree, eight Malaysian pineapple cultivars were classified into two groups. The first group consisted of Yankee and Gandul cultivars while second group consisted of Moris, Moris Bentanggur, Moris Gajah, N36, Josaphine and Sarawak. Bootstrap value in some branches were low this is reflected by the small number of informative characters (981 were conserved and 85 were potentially informative) to build the tree. Formation of several group or subclades is due to its similar genetic pattern, thus support this system classification. The study suggested that rbcL gene could be used to determine the phylogenetic relationship distinguish the pineapple cultivars.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>ii</td>
<td></td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
<td></td>
</tr>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>iv</td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>vi</td>
<td></td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
<td></td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
<td></td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS/ SYMBOLS</td>
<td>xiv</td>
<td></td>
</tr>
</tbody>
</table>

## CHAPTER 1 INTRODUCTION

1.1 Study background 1
1.2 Problem statement 3
1.3 Objectives of the study 4
1.4 Scope of the study 4
1.5 Significance of the study 5
CHAPTER 2  LITERATURE REVIEW

2.1  History and distribution of pineapple (Ananas comosus)  6
2.2  Ananas comosus taxonomy  7
2.3  Plant morphology  9
2.4  Pineapple cultivars  12
2.5  Importances of pineapple  15
2.6  Molecular phylogenetic study on Ananas  16
2.7  Application of rbcL gene in molecular phylogenetic analysis  17
2.8  Method of analysis in phylogenetic studies  19

CHAPTER 3  MATERIALS AND METHODS

3.1  Experimental design and procedure  21
3.2  Plant materials  24
3.3  Genomic DNA extraction  24
3.4  DNA quantification
   3.4.1  Spectrophotometer  25
   3.4.2  Agarose gel electrophoresis  25
3.5  Cloning of rbcL gene
   3.5.1  Amplification of rbcL gene by PCR  26
   3.5.2  Purification of PCR product  28
   3.5.3  Ligation  28
   3.5.4  Transformation  29
   3.5.5  Blue white screening  29
3.6 Plasmid isolation
   3.6.1 Plasmid extraction and glycerol stock preparation
   3.6.2 Restriction Enzyme (RE) Digestion
3.7 Bioinformatics analysis

CHAPTER 4 RESULTS AND DISCUSSION

4.1 Genomic DNA extraction
4.2 Amplification of rbcL gene by PCR
4.3 Selection for transformant
4.4 Plasmid isolation
4.5 Reconfirmation by Restriction enzyme (RE)
4.6 DNA sequencing, homology and similarity search for rbcL gene of Ananas comosus
4.7 Complete sequences alignment and sequences editing using ClustalX and Bioedit software
4.8 Phylogenetic tree construction and analysis

CHAPTER 5 CONCLUSION AND FUTURE WORK

5.1 Conclusion
5.2 Future work

REFERENCES 61
APPENDIX A-Q 71
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Morphology of Malaysian pineapple cultivars adapted from Hidayat et al., (unpublished)</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Primers used in this study</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Reagent for rbcL amplification</td>
<td>27</td>
</tr>
<tr>
<td>3.3</td>
<td>PCR amplification programmes</td>
<td>27</td>
</tr>
<tr>
<td>3.4</td>
<td>Ligation mixture preparation</td>
<td>28</td>
</tr>
<tr>
<td>3.5</td>
<td>Digestion mixture preparation</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>Purity of DNA extracted from eight Malaysian pineapple cultivars</td>
<td>34</td>
</tr>
<tr>
<td>4.2</td>
<td>Concentration of DNA extracted from eight Malaysian pineapple cultivars</td>
<td>35</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Structure of pineapple plant (<em>Ananas comosus</em> var. <em>comosus</em>) adapted from Coppens d’Eeckenbrugge <em>et al.</em>, 2011</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>Matured pineapple leaf (Malaysian Pineapple Industry Board, 2012)</td>
<td>11</td>
</tr>
<tr>
<td>2.3(a)</td>
<td>Moris Gajah cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(b)</td>
<td>Josaphine cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(c)</td>
<td>MD2 cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(d)</td>
<td>N36 cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(e)</td>
<td>MD2/T cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(f)</td>
<td>Sarawak Green Local cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(g)</td>
<td>Moris Bentanggur cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(h)</td>
<td>Yan Kee cultivar</td>
<td>14</td>
</tr>
<tr>
<td>2.3(i)</td>
<td>Gandul cultivar</td>
<td>14</td>
</tr>
<tr>
<td>3.1(a)</td>
<td>Flow chart of experimental design for the amplification of the rbcL gene from chloroplast genome of pineapple leaf</td>
<td>22</td>
</tr>
<tr>
<td>3.1(b)</td>
<td>Flow chart of experimental design for the cloning of rbcL gene</td>
<td>23</td>
</tr>
<tr>
<td>3.1(c)</td>
<td>Flow chart of experimental design for the construction of rbcL phylogenetic tree</td>
<td>23</td>
</tr>
</tbody>
</table>
4.1 DNA extracted from Malaysian pineapple cultivars. **M**: 1kb DNA ladder (New England BioLabs), **1**: Sarawak, **2**: Josaphine, **3**: Moris Gajah, **4**: Moris Bentanggur, **5**: Moris, **6**: Gandul, **7**: Yan Kee, **8**: Josaphine

4.2 PCR products for rbcL gene amplification **M**: 1kb DNA ladder (New England BioLabs), **1**: Sarawak, **2**: Josaphine, **3**: N36, **4**: Gandul

4.3 PCR products for rbcL gene amplification **M**: 1kb DNA ladder (New England BioLabs), **1**: Moris Bentanggur, **2**: Moris, **3**: Moris Gajah, **4**: Yan Kee

4.4 Result for plasmid extraction, **M**: 1 kb DNA ladder (New England BioLabs), **1**: Blue colony (negative control), **2**: pGEM-T Easy::rbcLJ1 (with inserted gene), **3**: pGEM-T Easy::rbcLJ2 (negative- no inserted gene), **3**: pGEM-T Easy::rbcLS1 (with inserted gene), **4**: pGEM-T Easy::rbcLS2 (with inserted gene), **5**: pGEM-T Easy::rbcLN1 (with inserted gene), **6**: pGEM-T Easy::rbcLN2 (with inserted gene), **7**: pGEM-T Easy::rbcLG1 (with inserted gene), **8**: pGEM-T Easy::rbcLG2 (with inserted gene)

4.5 Result for plasmid extraction, **M**: 1kb DNA ladder (New England BioLabs), **1**: pGEM-T Easy::rbcLM1 (with inserted gene), **2**: pGEM-T Easy::rbcLM2 (with inserted gene), **3**: pGEM-T Easy::rbcMG1 (with inserted gene), **4**: pGEM-T Easy::rbcMG2 (with inserted gene), **5**: pGEM-T Easy::rbcLY1 (with inserted gene), **6**: pGEM-T Easy::rbcLY2 (with inserted gene), **7**: pGEM-T Easy::rbcMB1 (with inserted gene), **8**: pGEM-T Easy::rbcMB2 (with inserted gene), **9**: Blue colony (negative control)

4.6 Results for reconfirmation by Restriction Enzyme (RE) using EcoRI, **M**: 1 kb DNA ladder (New England BioLabs), **1**: Moris Gajah, **2**: Moris Bentanggur, **3**: Yan Kee.

4.7 Results for reconfirmation by Restriction Enzyme
(RE) using EcoRI, **M**: 1 kb DNA ladder (New England BioLabs), **1**: Josaphine, **2**: Sarawak, **3**: N36, **4**: Gandul, **5**: Moris

4.8(a) rbcL gene full sequence of Moris Bentanggur cultivar

4.8(b) rbcL gene full sequence of Moris Gajah cultivar

4.8(c) rbcL gene full sequence of N36 cultivar

4.8(d) rbcL gene full sequence of Sarawak cultivar

4.8(e) rbcL gene full sequence of Gandul cultivar

4.8(f) rbcL gene full sequence of Yan Kee cultivar

4.8(g) rbcL gene full sequence of Josaphine cultivar

4.8(h) rbcL gene full sequence of Moris cultivar

4.9 Complete alignment of the eight sequences of Malaysian pineapple cultivars. The black colour regions in the alignment shows the consensus sequences of the eight pineapple cultivars

4.10 Strict consensus tree from the Maximum Parsimony analysis of the rbcL gene
### LIST OF ABBREVIATIONS/ SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>±</td>
<td>plus minus</td>
</tr>
<tr>
<td>%</td>
<td>percentage</td>
</tr>
<tr>
<td>° C</td>
<td>degree Celcius</td>
</tr>
<tr>
<td>μL</td>
<td>microliter</td>
</tr>
<tr>
<td>bp</td>
<td>basepair</td>
</tr>
<tr>
<td>BS</td>
<td>bootstrap support</td>
</tr>
<tr>
<td>CI</td>
<td>consistency index</td>
</tr>
<tr>
<td>cm</td>
<td>centimetre</td>
</tr>
<tr>
<td>cpDNA</td>
<td>chloroplast DNA</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>h</td>
<td>hour</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>ITS</td>
<td>internal transcribed spacer</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kg</td>
<td>kilogram</td>
</tr>
<tr>
<td>LAIX</td>
<td>LB, ampicillin, IPTG, X-gal</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani Broth</td>
</tr>
<tr>
<td>m</td>
<td>meter</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mL</td>
<td>mililiter</td>
</tr>
<tr>
<td>mM</td>
<td>micromolar</td>
</tr>
<tr>
<td>MP</td>
<td>Maximum Parsimony</td>
</tr>
<tr>
<td>MPIB</td>
<td>Malaysian Pineapple Industry Board</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>mtDNA</td>
<td>mitochondria DNA</td>
</tr>
<tr>
<td>nDNA</td>
<td>nuclear DNA</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>rbcL</td>
<td>large subunit of ribulose- bisphosphate carboxylase gene</td>
</tr>
<tr>
<td>RI</td>
<td>retention index</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>sec</td>
<td>second</td>
</tr>
<tr>
<td>SOC</td>
<td>Super Optimal Broth with Catabolite repression</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>annealing temperature</td>
</tr>
<tr>
<td>U</td>
<td>unit</td>
</tr>
<tr>
<td>X- gal</td>
<td>5- bromo-4- chloro- indolyl- β- D- galactopyranoside</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Study background

Chloroplasts are organelles which present abundantly in the leaf cells which play important role in metabolic activity known as photosynthesis process. The other important roles plays by the chloroplast are producing the starch, give the colour pigments to the flowers, certain amino acids, lipids and vitamins. Chloroplasts have their own genome whereby it possess a full complement of transcriptional and translation machinery in order to express their genetic information.

On the other hand, Pineapple (Ananas comosus) is the third most important tropical fruit crop in the world after bananas and mangoes (Botella and Smith 2008; Carlier et al., 2007). It also has become one of the leading commercial fruit crops of the tropics over the past of 100 years. Other than other than Thailand, Philippines, Indonesia, Hawaii, Ivory Coast, Kenya, Brazil, Taiwan, Australia, India and South Africa countries, it was reported that Malaysia has become one of the world producer of pineapple fruit. Development of pineapple industry in Malaysia is monitored by Malaysian Pineapple Industry Board (MPIB). MPIB was established since the year of 1957. The role of MPIB is to manage and develop Malaysian Pineapple Industry.
Pineapple is a type of tropical plant which is believed to originate from East Area of South America. It was introduced in Malaya in the 16th century by the Portuguese. During a year of 1921, pineapple started to be planted in Singapore, Johor and Selangor as cash crop. Now, pineapple plantation is continued to be expended in peat soil area especially in Johor (Malaysian Pineapple Industry Board, 2012). There are many cultivars and ‘Smooth Cayenne’ is the most commonly grown worldwide. In Malaysia, there are about three common cultivars widely planted in Malaysian which are Spanish (also known as Maspine, Josapine and Hybrid pineapple), Smooth Cayenne (Sarawak pineapple) and Queen (Morris pineapple) (Malaysian Pineapple Board, 2012).

Recently, molecular systematics in plants has progressed rapidly with DNA amplification or known as polymerase chain reaction, PCR is mediated by the direct sequencing methods (Schulte et al., 2008; Sheng-Guo et al., 2008; Osaloo and Kawano, 1999). Molecular approach has offered effective method in addressing many phylogenetic questions which cannot be solved using morphology characters. Previous studies (Chase et al., 1993; Spreitzer et al., 2002) claimed that large subunit of the ribulose- bisphosphate carboxylase (rbcL) gene is suitable for inference phylogenetic relationship at higher taxonomic levels. The rbcL gene is usually up to 1250 bp in size and the used of this gene in phylogenetic analysis has been reviewed in many studies (Chase et al., 1993; Clegg, 1993; Spreitzer et al., 2002). This is due to its advantages where it is easy to amplify using PCR, have appropriate length and base substitution rate for inferring phylogeny at higher taxonomy levels and it evolves at a conservation rate which make it suitable to study plant phylogeny (Clegg, 1993).

Pineapple is well known among user and it is highly consumed due to its pleasant taste and medical values. Though it is has become popular, very little is known about the molecular genetics of pineapple. It is due to limitation on available data of Ananas genetic diversity and most of it is based on morphology character only. They apparently arose due to spontaneous mutation, followed by natural selection and cross hybrid with unknown ancestor (Ruas et al., 1996). Hence, in this study, phylogenetic analysis of Malaysian pineapple or Ananas comosus was
conducted by using sequence data from the chloroplast gene which is known as large subunit of the ribulose- bisphosphate carboxylase (rbcL). The information about genetic variability at the molecular level is said to be useful to identify and characterize the unique germplasm that compliments the existing cultivars.

1.2 Problem statement

Pineapple is the third most important tropical fruit in the world after banana and mangoes (Botella and Smith, 2008). However, there is limitation on available data at molecular level for Malaysian pineapple. Previously, most of the classification of pineapple cultivars is based on morphology character. The drawback of classification based on morphological characteristics is inconsistency that arose due to disagreements among morphologist who applied different classification for interpretation of the characteristics.

Hence, in this study, phylogenetic analysis of Malaysian pineapple or Ananas comosus was conducted using sequence data from the chloroplast gene, rbcL to reconstruct a more detailed phylogenetic frame work of Malaysian pineapple and investigate the evolutionary relationships among these cultivars. The information about genetic variability at the molecular level is said to be useful to identify and characterize the Malaysian pineapple cultivars.
1.3 Objectives of the study

The objectives of the study are:

1. To amplify rbcL gene from eight Malaysian pineapple cultivars.
2. To clone the rbcL into pGEM-T Easy Vector System (Promega) and sequence the rbcL gene from eight Malaysian pineapple cultivars.
3. To construct a phylogenetic tree from isolated sequence of rbcL.

1.4 Scope of study

The scope of the study covered the construction of phylogenetic tree from eight commercial Malaysian pineapple cultivars based on rbcL gene. In order to achieve this, genomic DNA was isolated from leaf samples of eight commercial Malaysian pineapple cultivars; Moris, Moris Gajah, Moris Bentanggur, Yan Kee, Sarawak, Gandul and N36. PCR was carried out using rbcL primers and purified PCR product was subjected to cloning procedure comprising ligation of rbcL gene into the pGEM- T Easy Vector System (Promega) followed by transformation into NEB-5α competent cell (New England BioLabs). Positive transformants were sent for sequencing and sequences obtained were aligned using ClustalX software. Phylogenetic tree were constructed using MEGA 5 software and analysis were carried out using Maximum Parsimony method.
1.5 Significance of the study

Chloroplast DNA (cpDNA) has been used widely to infer plant systematic at different taxonomy levels (Clegg, 1993; Gielly and Taberlet, 1994). In this study, gene of the large subunit of ribulose-1,5-bisphosphate carboxylase (rbcL) was used as an alternative approach for morphology identification and to study the evolutionary status and relationship among the eight cultivars of Malaysian pineapple. The important of this study was to provide additional information on the relationship pattern among the cultivars which in future can be used as source of knowledge and information for successful interbreeding on creating new cultivars.
REFERENCES


