

การคัดเลือกลำดับดีเอ็นเอของคลอโรพลาสต์ที่แปรผันได้สูงเพื่อศึกษาความสัมพันธ์
ทางพันธุกรรมของพันธุ์ยาสูบในประเทศไทย



นางสาวมัชฌิมา นันทรัตน์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2552
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

SELECTION OF HIGHLY VARIABLE CHLOROPLAST DNA SEQUENCES FOR
GENETIC RELATIONSHIP STUDY OF TOBACCO CULTIVARS IN THAILAND

Miss Matchima Nantharat

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Biotechnology

Faculty of Science

Chulalongkorn University

Academic Year 2009

Copyright of Chulalongkorn University

521110

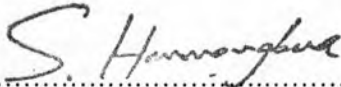
Thesis title SELECTION OF HIGHLY VARIABLE CHLOROPLAST
 DNA SEQUENCES FOR GENETIC RELATIONSHIP
 STUDY OF TOBACCO CULTIVARS IN THAILAND

By Miss Matchima Nantharat

Field of study Biotechnology

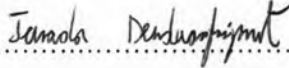
Thesis Advisor Assistant Professor Jessada Denduangboripant, Ph.D.

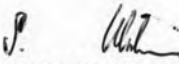
Accepted by the Faculty of Science, Chulalongkorn University in Partial
 Fulfillment of the Requirements for the Master's Degree

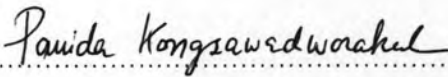

Dean of the Faculty of Science
 (Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE


Chairman
 (Associate Professor Kumthorn Thirakhupt, Ph.D.)


Thesis Advisor
 (Assistant Professor Jessada Denduangboripant, Ph.D.)


 Examiner
 (Sirawut Klinbunga, Ph.D.)


 External Examiner
 (Panida Kongsawadworakul, Ph.D.)

มัชฌิมา นันทรัตน์ : การคัดเลือกลำดับดีเอ็นเอของคลอโรพลาสต์ที่แปรผันได้สูงเพื่อศึกษาความสัมพันธ์ทางพันธุกรรมของพันธุ์ยาสูบในประเทศไทย (SELECTION OF HIGHLY VARIABLE CHLOROPLAST DNA SEQUENCES FOR GENETIC RELATIONSHIP STUDY OF TOBACCO CULTIVARS IN THAILAND) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ.ดร. เจษฎา เด่นดวงบริพันธ์, 166 หน้า.

ประเทศไทยมีอัตราการจัดเก็บภาษีระหว่างยาสูบสายพันธุ์พื้นเมือง และสายพันธุ์นำเข้าที่แตกต่างกัน อย่างไรก็ตาม ยังมีปัญหาใหญ่ทั้งทางด้านกฎหมายและด้านเทคนิคในการแยกแยะสายพันธุ์ยาสูบทั้งสองกลุ่มนี้ออกจากกัน ในการศึกษาครั้งนี้ได้นำบริเวณที่ไม่ถอดรหัสซึ่งมีความแปรผันสูงจากดีเอ็นเอในคลอโรพลาสต์มาใช้ศึกษาความสัมพันธ์ทางพันธุกรรมระหว่างยาสูบพันธุ์พื้นเมืองและพันธุ์นำเข้า (กลุ่มสายพันธุ์เวอร์จิเนีย เตอร์กิซ และเบอร์เลย์) รวม 51 สายพันธุ์ที่ปลูกในประเทศไทย ผลการเพิ่มปริมาณและการอ่านลำดับดีเอ็นเอพบว่า ใน 9 บริเวณนั้นมี 6 บริเวณ (*rpl32-trnL*, *ndhC-trnV*, *ndhF-rpl32*, *psbD-trnT*, *psbJ-petA* และ *atpI-atpH*) ที่แสดงความแตกต่างของลำดับดีเอ็นเอทั้งความยาวของลำดับเบสและการแทนที่ของเบส ยาสูบพันธุ์เวอร์จิเนียเกือบทั้งหมดมีลำดับดีเอ็นเอที่แตกต่างจากพันธุ์อื่นๆ บริเวณ *rpl32-trnL* ยังสามารถแยกแยะพันธุ์พื้นเมือง 7 สายพันธุ์ และอีก 1 ผลิตภัณฑ์ยาเส้นมวนเอง ด้วยดีเอ็นเอชิ้นแทรกขนาดใหญ่ 66 คู่เบส ความแตกต่างของลำดับดีเอ็นเอบริเวณ *rpl32-trnL* มีค่าสูงสุดคือ 10.47% ของความแปรผันและ 75 จำนวนลักษณะที่สามารถให้ข้อมูลได้ เทคนิคมัลติเพล็กซ์พีซีอาร์ได้ถูกพัฒนาขึ้นซึ่งสามารถแยกแยะยาสูบพันธุ์เวอร์จิเนีย และพันธุ์พื้นเมืองพิเศษ 7 สายพันธุ์ ออกจากพันธุ์อื่นๆ ได้ เครื่องหมายโมเลกุลนี้สามารถใช้ในการแก้ปัญหาระหว่างยาสูบสายพันธุ์พื้นเมืองและนำเข้าได้

สาขาวิชา.....เทคโนโลยีชีวภาพ..... ลายมือชื่อนิสิต มัชฌิมา นันทรัตน์.....
ปีการศึกษา.....2552..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก × 100% โฉมงามพร

5072423023 : MAJOR BIOTECHNOLOGY

KEYWORDS: CHLOROPLAST DNA / GENETIC RELATIONSHIP / HIGHLY VARIABLE SEQUENCES / THAILAND / AND TOBACCO

MATCHIMA NANTHARAT: SELECTION OF HIGHLY VARIABLE CHLOROPLAST DNA SEQUENCES FOR GENETIC RELATIONSHIP STUDY OF TOBACCO CULTIVARS IN THAILAND. THESIS ADVISOR: ASST. PROF. JESSADA DENDUANGBORIPANT, Ph.D., 166 pp.

In Thailand, imported and local tobacco cultivars have different regulations in tariff collection. However, there are major legal and technical problems on how to distinguish between the two groups of tobacco cultivars. In this study, highly-variable noncoding regions of chloroplast DNA was introduced to study genetic relationships among totally 51 local and imported tobacco cultivars (Virginia, Turkish, and Burley cultivar groups) grown in Thailand. The amplification and sequencing results revealed that six out of nine regions (*rpl32-trnL*, *ndhC-trnV*, *ndhF-rpl32*, *psbD-trnT*, *psbJ-petA* and *atpI-atpH*) showed polymorphic sequence characteristics, both in sequence lengths and the amounts of base substitutions. Almost all of Virginia cultivars were different in nucleotide sequence from the other cultivars. *rpl32-trnL* region could also distinguish seven local cultivars and one roll-your-own tobacco product with a large 66 bp insertion. The sequence polymorphism of *rpl32-trnL* region was in the highest value with 10.47% variability and 75 potentially informative characters. A multiplex PCR technique was successfully developed to distinguish Virginia cultivars and the seven special local cultivars from the others. This marker can be used as a solution for the "local" and "imported" cultivar problem.

Field of Study :Biotechnology.....

Academic Year :2009.....

Student's Signature Matchima Nantharat

Advisor's Signature x Jessada Denduangboripant

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Assistant Professor Dr. Jessada Denduangboripant for his encouraging guidance, supervision, stimulating suggestions, invaluable help in making the research run smoothly and all support throughout the course of this thesis.

I take this opportunity to thank Associate Professor Dr. Kumthorn Thirakhupt for serving as the chairman of my thesis committee and his suggestions. I would like also to thank the other members of my MSc committee who monitored my work and took effort in reading and proving my thesis with valuable comments: Dr. Sirawut Klinbunga and Dr. Panida Kongsawadworakul.

Special thanks to Mr. Chiwin Kunalai and Mrs. Wilasinee Suwanprasart, the officers of Tobacco Product Analysis Subdivision, Thailand Tobacco Monopoly, for their great help in my collecting of tobacco samples. Furthermore, I would like to thank all of the regional officers and local tobacco farmers in Chiang Mai, Chiang Rai, Phrae, Phayao, Nan, Sukhothai, Phetchabun, Nhonkhai, Nakhon Phanom, Kanchanaburi, Lop Buri and Suphan Buri provinces for their wonderful help, generous hospitality on the sample collection and giving information about tobacco cultivars.

I especially thank the head of Molecular Biology Research Laboratory (MB-RL), Department of Biology, Faculty of Science, Chulalongkorn University for providing laboratory facilities. I wish to send my sincere thanks to all of my good friends in MB-RL and Biotechnology Programme, especially Tianrat Piteekan, Manlika Kilaso and Jirattikarn Kaewmuangmoon for helping me get through the difficult times, comradely, helpfulness, kind assistance and for all of the emotional support.

I would like to express my deep gratitude also to the financial support for laboratory equipments and chemicals from the 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund), Thailand Research Fund (TRF) – Master Research Grant (no. MRG-OSMEP505S010) and Thailand Tobacco Monopoly, Ministry of Finance.

Finally, and most importantly, I would like to express my deepest thank to my family and my relatives for their love, understanding, encouragement, support, my inspiration, care and all support throughout my life.

CONTENTS

	Page
THAI ABSTRACT	iv
ENGLISH ABSTRACT.....	v
ACKNOWLEDGEMENTS.....	vi
CONTENTS.....	vii
LIST OF FIGURES.....	x
LIST OF TABLES.....	xiv
LIST OF ABBREVIATIONS.....	xv
CHAPTER I INTRODUCTION.....	1
CHAPTER II BACKGROUND.....	3
2.1 Morphological characteristics of tobacco.....	3
2.2 History of tobacco.....	5
2.3 Tobacco cultivation.....	8
2.4 Types of tobacco.....	9
2.5 Tobacco products.....	10
2.6 Other non-smoking uses for tobacco.....	13
2.7 Tobacco cultivars growing in Thailand.....	15
2.7.1 Imported tobacco cultivars.....	16
2.7.1.1 Virginia cultivar group (or flue-cured tobacco).....	16
2.7.1.2 Burley cultivar group (or light air-cured tobacco).....	18
2.7.1.3 Turkish cultivar group (or oriental tobacco).....	19
2.7.2 Local cultivar group (or air-and-sun cured tobacco).....	21
2.8 Molecular markers in plant.....	22
2.8.1 DNA fragment markers.....	23
2.8.2 DNA sequencing technique.....	30
2.9 Molecular marker studies in tobacco.....	32

	Page
2.10 Highly-variable noncoding regions of chloroplast DNA and their use in tobacco genetic relationship studies.....	36
2.11 Phylogenetic tree reconstruction.....	42
2.11.1 DNA data matrix preparation.....	43
2.11.2 Genetic distance analysis using neighbour-joining (NJ).....	45
2.11.3 Tree evaluation.....	46
CHAPTER III MATERIALS AND METHODS.....	48
3.1 Materials.....	48
3.1.1 Plant materials.....	48
3.1.2 Equipments.....	54
3.1.3 Chemicals.....	54
3.1.4 Oligonucleotide primers.....	55
3.2 Methods.....	55
3.2.1 DNA extraction and agarose gel electrophoresis.....	56
3.2.2 Highly variable chloroplast DNA region amplification.....	58
3.2.2.1 Preliminary selection of suitable primers.....	59
3.2.2.2 Selected primer pairs amplification.....	60
3.2.3 PCR purification and DNA sequencing.....	61
3.2.4 Genetic relationship analysis of tobacco cultivars in Thailand.	63
3.2.5 Preliminary experiment of multiplex PCR technique.....	64
CHAPTER IV RESULTS.....	66
4.1 DNA extraction.....	66
4.2 PCR amplification and sequencing of highly-variable chloroplast DNA regions.....	69
4.2.1 Selection of suitable primers for the extracted DNA from fresh leaf samples.....	69
4.2.2 Amplification of <i>rp132-trnL</i> region from cured leaves and roll- your-own tobacco samples.....	96
4.3 Genetic relationship analysis of tobacco cultivars in Thailand.....	103

	Page
4.4 Preliminary experiment for multiplex PCR.....	106
4.4.1 <i>rpl32F-trnL</i> molecular marker to differentiate Virginia and local cultivars.....	106
4.4.2 Combined <i>rpl32F-trnL</i> and <i>ndhF-rpl32R</i> markers to differentiate Virginia and local cultivars.....	110
CHAPTER V DISCUSSIONS.....	112
5.1 DNA extraction.....	112
5.2 Selection of suitable PCR primer-pairs of highly-variable chloroplast DNA regions.....	116
5.3 The efficiency of the six selected highly-variable regions.....	119
5.4 Genetic relationships of tobacco cultivars in Thailand.....	124
5.5 Preliminary experiment for multiplex PCR.....	127
CHAPTER VI CONCLUSIONS.....	131
REFERENCES.....	132
APPENDICE.....	149
BIOGRAPHY.....	166

LIST OF FIGURES

		Page
Figure 2.1	Morphological characteristics of tobacco.....	4
Figure 2.2	Global consumption of cigarettes has been rising steadily.....	6
Figure 2.3	Growing tobacco plants around the world.....	7
Figure 2.4	The products of smoking tobaccos.....	11
Figure 2.5	The products of smokeless tobaccos.....	13
Figure 2.6	Virginia tobacco cultivars.....	17
Figure 2.7	Burley tobacco cultivars.....	19
Figure 2.8	Turkish tobacco cultivars.....	20
Figure 2.9	Local tobacco cultivars.....	21
Figure 2.10	Restriction Fragment Length Polymorphism (RFLP) technique.....	24
Figure 2.11	The steps of PCR (Polymerase Chain Reaction) reaction.....	25
Figure 2.12	Random Amplified Polymorphic DNA (RAPD) technique.....	26
Figure 2.13	Amplified Fragment Length Polymorphism (AFLP) technique.....	27
Figure 2.14	Simple Sequence Repeat (SSR) technique.....	28
Figure 2.15	An example of multiplex PCR band pattern.....	29
Figure 2.16	DNA sequencing technique.....	31
Figure 2.17	The complete gene map of the chloroplast DNA from tobacco (<i>N. tabacum</i>) reported by Yukawa et al. (2005).....	35
Figure 2.18	Scaled maps of the 21 noncoding cpDNA regions surveyed by Shaw et al. (2005).....	38
Figure 2.19	The average PIC (potentially informative character) value in Shaw et al. (2005).....	38
Figure 2.20	Scaled maps of the 13 unexplored noncoding cpDNA regions reported by Shaw et al. (2007).....	39
Figure 2.21	The average PIC (potentially informative character) value in Shaw et al. (2007) compared with Shaw et al. (2005).....	40

	Page	
Figure 2.22	The normalised PIC value in Shaw et al. (2007) compared with Shaw et al. (2005).....	40
Figure 2.23	Phylogenetic tree based on rRNA data shows the separation of bacteria, archaea and eukaryotes.....	43
Figure 2.24	A cladogram showing the relationship between various plant groups.....	43
Figure 2.25	FASTA format of sequence input files.....	44
Figure 2.26	Multiple sequence alignment by ClustalX.....	44
Figure 2.27	Tree-building from neighbour-joining method.....	45
Figure 2.28	PAUP* (Phylogenetic Analysis Using Parsimony) version 4.0 used in Macintosh.....	46
Figure 2.29	Bootstrap 50% majority-rule consensus tree.....	47
Figure 3.1	Collecting localities of tobacco samples around Thailand.....	52
Figure 3.2	Field expeditions to collect tobacco leaf samples.....	53
Figure 3.3	An approximate map of nine noncoding chloroplast DNA regions examined in this study.....	58
Figure 4.1	Genomic DNA extracted from fresh leaf samples of two local and six imported tobacco cultivars.....	66
Figure 4.2	Genomic DNA extracted from fresh leaf samples of 14 imported and 3 local tobacco cultivars.....	67
Figure 4.3	Genomic DNA extracted from fresh leaf samples of 17 local tobacco cultivars.....	67
Figure 4.4	Genomic DNA extracted from cured leaf samples of 10 imported tobacco cultivars.....	68
Figure 4.5	Genomic DNA extracted from cured leaf samples of 14 local tobacco cultivars.....	68
Figure 4.6	Genomic DNA extracted from three roll-your-own tobaccos.....	69

	Page	
Figure 4.7	PCR products of Hangkai (local), K326 (Virginia) and B1 special (Burley) cultivars amplified with five primer-pairs and using 50°C annealing temperature.....	70
Figure 4.8	PCR products of Hangkai (local), K326 (Virginia) and B1 special (Burley) cultivars amplified with five primer-pairs and using 50°C annealing temperature.....	70
Figure 4.9	PCR products of Hangkai (local), K326 (Virginia) and B1 special (Burley) cultivars amplified with two primer-pairs and using 50°C, 51°C and 52°C annealing temperatures.....	71
Figure 4.10	Electropherogram of <i>rpl32F-trnL</i> sequence of Hangkai cultivar.....	73
Figure 4.11	Electropherogram of <i>ndhC-trnV</i> sequence of Phu cultivar.....	73
Figure 4.12	Electropherogram of <i>petA-psbJ</i> sequence of Chorlare2 cultivar....	73
Figure 4.13	A 753 bp nucleotide data matrix of <i>petA-psbJ</i> region from fresh-leaf samples of total 18 tobacco cultivars.....	75
Figure 4.14	A 685 bp nucleotide data matrix of <i>ndhC-trnV</i> region from fresh-leaf samples of total 23 tobacco cultivars.....	78
Figure 4.15	A 731 bp nucleotide data matrix of <i>psbD-trnT</i> region from fresh-leaf samples of total 18 tobacco cultivars.....	81
Figure 4.16	A 811 bp nucleotide data matrix of <i>atpH-atpI</i> region from fresh-leaf samples of total 18 tobacco cultivars.....	84
Figure 4.17	A 769 bp nucleotide data matrix of <i>ndhF-rpl32R</i> region from fresh-leaf samples of total five tobacco cultivars.....	87
Figure 4.18	A 716 bp nucleotide data matrix of <i>rpl32F-trnL</i> region from fresh-leaf samples of total 23 tobacco cultivars.....	89
Figure 4.19	A 716 bp nucleotide data matrix of <i>rpl32F-trnL</i> region from all fresh-leaf samples of total 43 tobacco cultivars.....	92
Figure 4.20	PCR products of <i>rpl32-trnL</i> region from cured-leaf samples of 11 local tobacco cultivars.....	97

	Page
Figure 4.21 PCR products of <i>rp132-trnL</i> region from three roll-your-own tobaccos and two cured-leaf samples of local tobacco cultivars...	97
Figure 4.22 A 716 bp nucleotide data matrix of <i>rp132F-trnL</i> region from total 51 fresh-leaf, cured-leaf or roll-your-own tobacco samples.....	99
Figure 4.23 NJ tree of <i>rp132-trnL</i> sequence data from 51 tobacco samples.....	105
Figure 4.24 Electropherogram of <i>rp132F-trnL</i> sequences of K326 and Chorlae1 cultivars amplified together within a single reaction.....	106
Figure 4.25 One part of the alignment of 43 <i>rp132F-trnL</i> sequences showed five base substitutions of K326 Virginia cultivar different from Chorlae1 local cultivar.....	106
Figure 4.26 Electropherogram of <i>rp132F-trnL</i> sequence of K326 and Hangkai cultivars amplified together within a single reaction.....	107
Figure 4.27 One part of the alignment of 43 <i>rp132F-trnL</i> sequences showed five base substitutions of K326 Virginia cultivar and a large 66 bp insertion of Hangkai special-local cultivar different from the other.....	107
Figure 4.28 PCR products of <i>rp132-trnL</i> region amplified from mixed genomic DNA of different tobacco cultivars.....	108
Figure 4.29 PCR products of <i>rp132-trnL</i> region of five tobacco cultivars compared with a mixed reaction of local and imported cultivars.....	109
Figure 4.30 PCR products of four different regions amplified from mixed DNA of K326 (Virginia), B1 special (Burley) and Hangkai (special-local) cultivars.....	111
Figure 4.31 PCR products of two different regions amplified from mixed DNA of K326 (Virginia) and Hangkai (special-local).....	111

LIST OF TABLES

	Page
Table 2.1	Types of tobacco are classified by the main of curing methods..... 9
Table 3.1	Imported and local tobacco fresh-leaf samples were collected from many cultivation crops around Thailand and used in this study..... 49
Table 3.2	Unknown tobacco fresh-leaf samples used in this study..... 50
Table 3.3	Imported and local tobacco cured-leaf samples used in this study.... 50
Table 3.4	Roll-your-own tobacco leaf samples used in this study..... 51
Table 3.5	Primer pairs used for PCR amplification and sequencing..... 59
Table 4.1	Sizes of PCR amplified products and alignment lengths of the six DNA regions..... 74
Table 4.2	Degrees of polymorphism in nucleotide sequence characteristics of the six selected chloroplast noncoding regions..... 104

LIST OF ABBREVIATIONS

A	adenine
AFLP	Amplified Fragment Length Polymorphism
bp	basepair
C	cytosine
°C	degree celsius
cm ²	centimetre square
cpDNA	chloroplast DNA
cpSSR	chloroplast Simple Sequence Repeat
µg	microgram
µl	microlitre
µM	micromolar
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease (DNA endonuclease)
dNTP	deoxynucleotide triphosphate
EDTA	ethylene diamine tetraacetic acid (disodium salt)
EtBr	ethidium bromide
FISH	fluorescent <i>in situ</i> hybridisation
FISSR	Fluorescent Inter Simple Sequence Repeat
G	guanine
GISH	genomic <i>in situ</i> hybridisation
hr	hour
ID	indel
indel	insertion-or-deletion
IR	inverted repeat
ISSR	Inter Simple Sequence Repeat
ITS	Internal Transcribed Spacer
IV	inversion
Kb	kilobasepair

L	the aligned sequence length
LSC	large single copy
MgCl ₂	Magnesium chloride
mg	milligram
min	minute
ml	millilitre
mM	millimolar
mmol	millimole
M.W.	molecular weight
<i>N.</i>	<i>Nicotiana</i>
ng	nanogram
NJ	Neighbour-Joining
NS	nucleotide substitution
nrDNA	nuclear ribosomal DNA
OD	optimal density
PAUP*	Phylogenetic Analysis Using Parsimony
PCR	Polymerase Chain Reaction
PIC	Potentially Informative Character
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	ribonucleic acid
RNase	ribonuclease (RNA endonuclease)
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RYO	roll-your-own
sec	second
SSC	small single copy
SSR	Simple Sequence Repeat
T	thymine
<i>Taq</i>	<i>Thermus aquaticus</i>
TBE	Tris-boric-ethylene diamine tetraacetic acid

T_m	melting temperature
Tris	Tris (hydroxyl methyl) aminomethane
TSNA	tobacco-specific nitrosamine
U	unit
UV	ultraviolet
Volt	voltage