ความผันแปรทางพันธุกรรมของพืชสกุลเปล้า (*Croton* spp.)

นางสาวประสบอร ริ้นทอง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาเภสัชเวท
คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2550

GENETIC VARIATION IN GENUS *CROTON* (*CROTON* SPP.)

Miss Prasoborn Rinthong

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Pharmacognosy
Faculty of Pharmaceutical Sciences
Chulalongkorn University
Academic Year 2007

Thesis Title	GENETIC VARIATION IN GENUS *CROTON* (*CROTON* SPP.)
By	Miss Prasobron Rinthong
Field of Study	Pharmacognosy
Thesis Advisor	Associate Professor Wanchai De-Eknamkul, Ph.D.
Thesis Co-advisor	Assistant Professor Suchart Chanama, Ph.D.

Accepted by the Faculty of Pharmaceutical Sciences, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

..... Dean of the Faculty of Pharmaceutical Sciences
(Associate Professor Pornpen Pramyothin, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Associate Professor Kittisak Likhitwitayawuid, Ph.D.)

..... Thesis Advisor
(Associate Professor Wanchai De-Eknamkul, Ph.D.)

..... Thesis Co-advisor
(Assistant Professor Suchart Chanama, Ph.D.)

..... Member
(Associate Professor Surattana Amnuoypol, Ph.D.)

..... Member
(Assistant Professor Suchada Sukrong, Ph.D.)

..... External Member
(Kongkanda Chayamarit, Ph.D.)

..... External Member
(Assistant Professor Worapan Sitthithaworn, Ph.D.)

ประสบอร รินทอง : ความผันแปรทางพันธุกรรมของพืชสกุลเปล้า (*CROTON* SPP.). (GENETIC VARIATION IN GENUS *CROTON* (*CROTON* SPP.) อ.ที่ปรึกษา : รศ.ดร.วันชัย ดีเอกนามกูล, อ.ที่ปรึกษาร่วม : ผศ.ดร.สุชาติ ชะนะมา, 167 หน้า.

พืชสกุลเปล้า (*Croton* spp.),วงศ์ Euphorbiaceae ถูกใช้ประโยชน์ในการแพทย์พื้นบ้านอย่างแพร่หลายและยังถือเป็นแหล่งผลิตสารออกฤทธิ์ทางชีวภาพ การวิจัยนี้ได้ศึกษาความผันแปรทางพันธุกรรมของพืชสกุลเปล้าที่พบได้ในประเทศไทยจำนวน 15 ชนิด โดยอ้างอิงจากลำดับนิวคลีโอไทด์บริเวณ internal transcribed spacer (ITS) ของนิวเคลียดีเอ็นเอและยีน *trn*L-F ในคลอโรพลาสต์ดีเอ็นเอ การศึกษาพบว่าลำดับนิวคลีโอไทด์บริเวณ ITS ของพืชสกุลเปล้ามีความยาวในช่วง 620-627 คู่เบส, ปริมาณ GC คิดเป็นร้อยละ 52.5-57.1 ความผันแปรของพันธุกรรมบริเวณ ITS โดยส่วนใหญ่เกิดจากกระบวนการแทนที่ของนิวคลีโอไทด์ ส่วนยีน *trn*L-F นั้นมีความยาวในช่วง 1017-1041 คู่เบส, ปริมาณ GC คิดเป็นร้อยละ 31.0-31.9 โดยความผันแปรเป็นผลจากกระบวนการเพิ่มและขาดหายของนิวคลีโอไทด์ นอกจากนี้ผลการวิเคราะห์วงศ์วานวิวัฒนาการโดยอาศัยข้อมูลลำดับนิวคลีโอไทด์บริเวณดังกล่าว พบว่าพืชสกุลเปล้าทั้ง 15 ชนิดมีความสัมพันธ์ทางวงศ์วิวัฒนาการในลักษณะของการอยู่ในกลุ่มเดียวกัน นอกจากนี้ยังได้วิเคราะห์ในเชิงลึกถึงความผันแปรทางพันธุกรรมของ เปล้าน้อย (*C. stellatopilosus*) และเปล้าใหญ่ (*C. roxburghii*) ซึ่งเปล้าทั้งสองชนิดเป็นที่รู้จักอย่างกว้างขวางในด้านการใช้ประโยชน์ทางยา ในกรณีนี้นอกจากการศึกษาลำดับนิวคลีโอไทด์บริเวณ ITS และยีน *trn*L-F แล้ว ยังทำการวิเคราะห์ลำดับนิวคลีโอไทด์เพิ่มเติมบริเวณ *trn*K intron ในคลอโรพลาสต์ดีเอ็นเออีกด้วย ผลของลำดับนิวคลีโอไทด์บริเวณ ITS แสดงให้เห็นว่าเปล้าน้อยมีการผสมข้ามพันธุ์ระหว่างพันธุกรรม Type A และ Type B และลำดับ นิวคลีโอไทด์บริเวณ *trn*K intron บ่งชี้ว่าเปล้าน้อยลูกผสมได้รับการถ่ายทอดพันธุกรรมทางสายพ่อจาก Type A ในขณะที่ถ่ายทอดพันธุกรรมทางสายแม่ได้จาก Type B ส่วนเปล้าใหญ่นั้นพบว่ามีความผันแปรทางพันธุกรรมสูงมาก เปล้าใหญ่ในแต่ละตัวอย่างมีลำดับนิวคลีโอไทด์ในบริเวณ ITS, ยีน *trn*L-F และ *trn*K intron ที่แตกต่างกัน ทำให้แบ่งเปล้าใหญ่ได้เป็น 2 กลุ่ม ผลการวิจัยนี้ทำให้สรุปได้ว่าข้อมูลลำดับนิวคลีโอไทด์สามารถใช้เพื่อการวิเคราะห์และแสดงความความสัมพันธ์ทางวงศ์วิวัฒนาการของพืชสกุลเปล้า อีกทั้งสามารถระบุตำแหน่งของพันธุกรรมที่มีความผันแปรในเปล้าน้อยและเปล้าใหญ่ได้

สาขาวิชา.....เภสัชเวท.....ลายมือชื่อนิสิต.....ประสบอร รินทอง.....

ปีการศึกษา.....2550.....ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4676957933 : MAJOR PHARMACOGNOSY
KEYWORD : *CROTON*/ ITS/ *trn*L-F/ GENETIC VARIATION/ PHYLOGENETIC/ *trn*K/ HYBRIDIZATION/ *CROTON STELLATOPILOSUS*/ *CROTON ROXBURGHII*
PRASOBORN RINTHONG: GENETIC VARIATION IN GENUS *CROTON* (*CROTON* SPP.) THESIS ADVISOR : ASSOC. PROF. WANCHAI DE-EKNAMKUL, Ph.D., THESIS COADVISOR : ASST. PROF. SUCHART CHANAMA, Ph.D., 167 pp.

Plants in the genus *Croton* (Euphorbiaceae) have been widely used in folklore medicine and are considered to be a rich source of bioactive compounds. In this study, genetic variation of fifteen *Croton* species existing in Thailand was evaluated for their phylogenetic relationship. The genes selected for the study included the internal transcribed spacer (ITS) of nuclear DNA and *trn*L-F gene of chloroplast DNA. The length of ITS sequences of various *Croton* species was from 620 to 627 bp with 52.5-57.1 % GC content. Genetic variation in the ITS region was mostly from substitutions. For the *trn*L-F gene, the *Croton* species had their sequences in the range from 1017 to 1041 bp with 31.0-31.9 % GC content. The genetic variation of this gene in *Croton* was caused mainly from insertions and deletions. Phylogenetic analysis revealed that the relationship of all fifteen *Croton* species was monophyletic type. In terms of variation in population, the two well known Thai medicinal plants *C. stellatopilosus* (Plau-noi) and *C. roxburghii* (Plau-yai) were extensively investigated for their genetic variation. Both species were studied not only on the ITS region and *trn*L-F gene, but also on the *trn*K intron of chloroplast DNA. Based on the obtained ITS sequences of *C. stellatopilosus*, two groups of individual sequences, designated as Type A and Type B were found as well as various nucleotide additive sequences which were arised from these two groups and designated as "putative hybrids". These putative hybrids of *C. stellatopilosus* showed their *trn*K intron sequences identical to the Type B but different from the Type A. Thus, it was suggested that Type A was paternal parent and Type B was maternal parent of the hybridization of *C. stellatopilosus*. For *C. roxburghii*, it showed high genetic variation in their ITS, *trn*L-F and *trn*K intron sequences but could be systematically divided into two groups. In conclusion, the technique of direct sequencing was successfully used in this case as a tool for estimating the phylogenetic relationship of *Croton* and identifying the definite sites of genetic variation in the important *Croton* species of *C. stellatopilosus* and *C. roxburghii*.

Field of study.....Pharmacognosy..... Student's signature.Prasoborn Rinthong...
Academic year.....2007.....Advisor's signature..Wanchai De Eknamkul
Coadvisor's signature..Suchart Chanama

ACKNOWLEDGEMENTS

The author wishes to express her deepest gratitude to her advisor, Associate Professor Wanchai De-Eknamkul, Ph.D. of the Department of Pharmacognosy, Faculty of Pharmaceutical Sciences, her co-advisor Assistant Professor Suchart Chanama, Ph.D. of the Department of Biochemistry, Faculty of Science, Chulalongkorn University, for their kindness, guidance, suggestion and encouragement throughout her Ph.D. program.

The author also wishes to express her appreciation to Professor Katsuko Komatsu, Ph.D., Assistant Professor Shu Zhu, Ph.D. and all members of the Department of Pharmacognosy, Institute of Natural Medicine, University of Toyama, Toyama, Japan for their advice and assistance.

The author would like to acknowledge her thanks to Kongkanda Chayamarit, Ph.D. of The Thai Royal Herbarium for her provided *Croton* herbarium specimens and The Institute of Biotechnology and Genetic Engineering, Chulalongkorn University for their provided *C. stellatopilosus* specimens.

The author's appreciation is extended to the thesis committees for their critical review of her thesis; all lecturers, staff members and graduate students of Department of Pharmacognosy, Faculty of Pharmaceutical Sciences, Chulalongkorn University for their kindness and unforgettable friendship.

Finally, the author wishes to express her infinite gratitude to her family and her friends for their love and assistance throughout this graduate study.

CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

Figure **Page**

LIST OF ABBREVIATIONS

5.8s rDNA	5.8s ribosomal RNA gene
18s rDNA	18s ribosomal RNA gene
26 s rDNA	26s ribosomal RNA gene
A, T, G, C	Nucleotides containing the base adenine, thymine, guanine and cytosine, respectively
ATP	Adenosine tri-phosphate
*atp*B gene	Gene encoding the ATP synthase ß-subunit
AUC	Area under the curve
bp	Base pairs
°C	Degree Celcius
cm	Centimeter
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates (dATP, dTTP, dGTP, dCTP)
EDTA	Ethylenediamine tetra acetic acid
HCl	Hydrochloric acid
IGS	Intergenic spacer
ITS	Internal transcribed spacer
ITS1	Internal transcribed spacer 1
ITS2	Internal transcribed spacer 2
IUPAC	International Union Pure and Applied Chemistry
Kb	Kilobase
KCl	Potassium chloride
*mat*K gene	Gene encoding maturase K
mg	Miligram
$MgCl_2$	Magnesium chloride
min	Minutes
ml	Milliliter
mM	Millimolar

*ndh*F gene	Gene encoding NADH dehydrogenase F
ng	Nanogram
nm	Nanometer
PAUP	Phylogenetic analysis using parsimony
PBS	Phosphate buffer solution
PCR	Polymerase chain reaction
*rbc*L gene	Gene encoding the large subunit of the ribulose-bisphosphate carboxylase
Rf	Retention factor
RNA	Ribonucleic acid
*rps*16 gene	Gene encoding ribosomal protein S16
sec	Seconds
TAE	Tris acetate EDTA
TLC	Thin layer chromatography
Tris	Tris (hydroxymethyl) aminomethane
tRNA	Transfer ribonucleic acid
*trn*F gene	Gene encoding $tRNA^{Phe}$
*trn*K gene	Gene encoding $tRNA^{Lys}$
*trn*L gene	Gene encoding $tRNA^{Leu}$
UPGMA	Unweighted pair group method with arithmetic averages
µg	Microgram
µl	Microliter
µM	Micromolar
UV	Ultraviolet

CHAPTER I

INTRODUCTION

Every living organism contains DNA, RNA, and proteins. Closely related organisms generally have their high degree of similarity in the molecular structures of their biomolecules, while non-related organisms usually show patterns of dissimilarity. Studies on the molecular genetics allow us to obtain information on the phylogeny which is the science of estimating and analyzing evolutionary relationships. Molecular phylogeny uses the structural data of biomolecules to build a tree of relationship that shows the probable evolution of various related organisms. Recently, DNA sequencing technique has been used as a tool to determine genetic relationships between different organisms. This is based on the fact that nucleotide sequences are information that retains a record of organism's evolutionary history. The approach of this technique is to compare nucleotide sequences from different organisms and to estimate the evolutionary relationships based on their degree of homology. Number of differences between organisms can then be established and expressed in the form of phylogenetic trees.

Croton is a plant genus belonging to the family Euphorbiaceae sensu stricto. Many of these *Croton* species are widely used in folklore medicines (Gurgel *et al*., 2005 and Salatino *et al.*, 2007) and considered to be the source of bioactive compounds, especially alkaloids and terpenoids (Rizk, 1987). Research works on *Croton* have been on the aspects of phytochemistry (Salatino *et al.*, 2007) and taxonomy (Webster, 1993, 1994; Govaert, 2000). There have been very few reports on the molecular genetics of *Croton* species. Although the taxonomic classification of Thai *Croton* species based on morphological characteristics has been relatively completed (Chayamarit and Welzen, 2005), its genetic data and molecular phylogeny have not yet been well established and thus the subject of this study.

Very recently, molecular surveys on *C. sublyratus* (*C. stellatopilosus*) (Klinbantoom, 2004) and *C. oblongifolius* (*C. roxburghii*) (Chonvanich, 2004) have been carried out to obtain their patterns of DNA fingerprints by using inter simple sequence repeats (ISSRs) technique. The obtained distance based phylogenetic trees have revealed that *C. stellatopilosus* is genetically varied in two genotypes, while *C. roxburghii* is clearly separated in two groups. The results have suggested that both *C. stellatopilosus* and *C. roxburghii* are taxonomical complex species. However, their PCR product profiles obtained from the ISSRs technique could not address the taxonomical status and hypothesize the hybridization of both species. These issues should be investigated in depth to the nucleotide sequence level, which is the main research interest of this work.

Among various DNA sequences in plants, the internal transcribed spacer (ITS) of nuclear ribosomal DNA, which is usually used for plant authentication and phylogeny study (Baldwin *et al.*, 1995), has the potentially to provide novel insights into the plant evolution and hybridization (Campbell *et al.*, 1997). ITS sequences have been shown based on nucleotide direct sequencing, to have nucleotide additive sites that indicate the occurrence of hybridization. This phenomenon has been reported for both the angiosperm (Sang *et al*., 1995; Whittall *et al.*, 2000) and the gymnosperm (Quijada *et al*., 1997). The chloroplast DNA such as *trn*L-F, *mat*K and *trn*K have also been used for the phylogenetic analysis. The combined sequence data between nuclear and chloroplast DNA have well documented for resolving the phylogenetic relationships of the plants in the family Euphorbiaceae and their closely related taxa (Wurdack *et al*., 2005; Berry *et al*., 2005; Benjamin *et al*., 2006). Moreover, these combined data have been subjected to meta-analysis to propose the putative parents of the interspecific hybridization (Padgett *et al*, 1998; Tsukaya, 2004; Zhang *et al.*, 2007; Kondo *et al.*, 2007).

Based on these reports, this dissertation was therefore designed to investigate the genetic information of *Croton* species in Thailand by using the DNA amplification and direct sequencing technique. We focused on the phylogenetic analysis of *Croton* species and on examining the genetic variation of *C. stellatopilosus* and *C. roxburghii*. All *Croton* samples were PCR amplified and investigated for the sequences on the ITS region of nuclear DNA and *trn*L-F gene of chloroplast DNA. All the obtained sequence data were then used to analyze their phylogenetic relationships. In addition, hybridization was observed for *C. stellatopilosus* and *C. roxburghii* as evidence from their ITS nucleotide sequences. Both species were further examined for their nucleotide sequences in *trn*K intron of chloroplast DNA. The nucleotide sequences of ITS region, *trn*L-F gene and *trn*K intron were analyzed to propose their putative parents.

CHAPTER II

LITERATURE REVIEW

1. General Information of Genus *Croton*

The genus *Croton* was established by Carolus Linnaeus in 1737 (Leon, 1945). It is considered one of the largest genera of flowering plants. The number of species is between 1200 and 1300, with herbs, shrubs and trees. *Croton* plants are usually recognized by a suite of characters that include with conspicuous stellate or scalelike trichomes, narrow or condensed inflorescences of unisexual flowers, watery to color sap, frequent petiolar glands and senescent leaves that turn orange before dehiscing. Geographical distribution of *Croton* species is in the tropics and subtropics. The plant genus *Croton* is in horticultural confusion with *Codiaeum* (Han, 1973) whose common name “croton” refers to the worldwide cultivated ornamental varieties of *Codiaeum variegatum* L. For medicinal usage, the genus *Croton* exhibits various well marked medicinal properties as bitter, tonic, stimulant, astringent and diuretic. The best known medicinal plant is probably *C. tiglium*, a native species in Southeast Asia.

For taxonomical studies, the genus *Croton* has traditionally been classified with Euphorbiaceae sensu lato and is retained, subsequent to the recent recognition of four segregate families. So far, this genus belongs to the family Euphorbiaceae sensu stricto that contains only the uniovulated plants. *Croton* is assigned to the subfamily Crotonoideae, tribe Crotonae. This genus has been variously circumscribed by Klotzsch (1841), Muller (1866, 1873), Webster (1993, 1994) and Govaerts (2000). The recent sectional synopsis of *Croton* is recognized as 40 sections (Webster, 1993; Govaerts, 2000). Taxonomical classification of *Croton* is still not well understood. However, the taxonomic classification of the plant family Euphorbiaceae and genus *Croton* in Thailand has been revised by the staff of the Forest Herbarium, Royal Forest Department, Bangkok

and collaborators. The revision has been based on the reliable study of Thai Euphorbiaceae by Airy Shaw (1972). Recently, Chayamarit and Welzen (2005) have reported in the Flora of Thailand that there are 29 *Croton* species and 2 *Croton* sp. existing in Thailand as listed in Table 2.1. In this taxonomical review, *C. kongkandanus* has been reported as a new *Croton* species (Esser and Chayamarit, 2001; Esser, 2002). It differs from all the silvery leaved species of *Croton* in Southeast Asia and Malaysia by the much larger fruits, while its other characters are similar to *C. argyratus*. So far, *C. kongkandanus* has been the only known from Thailand.

Table 2.1 The occurrence of *Croton* species in Thailand (Chayamarit and Welzen, 2005)

No.	Scientific name	Thai name
1.	*Croton acutifolius* Esser	Plao; เปล้า, Plao phae; เปล้าแพะ
2.	*C. argyratus* Blume	Plao ngoen; เปล้าเงิน (Pattani), Mai lot; ไม้ลอด, Plao; เปล้า (Peninsular)
3.	*C. bonplandianus* Baill.	Plao thung; เปล้าทุ่ง (General)
4.	*C. cascarilloides* Raeusch.	Ka-don hin; กะดอนหิน (Loei), Plao nam ngoen; เปล้าน้ำเงิน (Prachuap Khiri Khan), Plao lek; เปล้าเล็ก, Plao ngoen; เปล้าเงิน (Peninsular)
5.	*C. caudatus* Geiseler	Krado hot bai khon; กระดอหดใบขน (Southeastern), Kho khlan; โคคลาน (Chachoengsao), Plao; เปล้า (Peninsular)
6.	*C. columnaris* Airy Shaw	Plao noi; เปล้าน้อย (Lampang), Plao kham; เปล้าคำ (Sukhothai)
7.	*C. crassifolius* Geiseler	Pang khi; พังคี (Northern), Phang khi noi; พังคีน้อย (Eastern)
8.	*C. decalvatus* Esser	

No.	Scientific name	Thai name
9.	*C. delpyi* Gagnep.	Plao yai; เปล้าใหญ่, Nom nam khiao; นมน้ำเขียว (Chon Buri), Plao noi; เปล้าน้อย (Chanthaburi), Plao; เปล้า (Southeastern), Plao phai; เปล้าไผ่ (Surat Thani)
10.	*C. griffithii* Hook.f.	Chik; จิก , Plao; เปล้า (Peninsular)
11.	*C. hirtus* L'Her.	Plao lom luk; เปล้าล้มลุก (Peninsular)
12.	*C. hutchinsonianus* Hosseus in Fedde	Mueat; เหมือด, Plao lueat; เปล้าเลือด (Northern), Plao; เปล้า, Plao phae; เปล้าแพะ (Phetchaburi)
13.	*C. kerrii* Airy Shaw	Plao; เปล้า (General)
14.	*C. kongensis* Gagnep.	Plao ngoen; เปล้าเงิน, Plao nam ngoen; เปล้าน้ำเงิน (General), Plao noi; เปล้าน้อย (Nong Khai)
15.	*C. kongkandanus* Esser	Mi la; มิลา (Sukhothai)
16.	*C. krabas* Gagnep.	Fai nam; ฝ้ายน้ำ (Eastern), Phrik na; พริกนา (Central), Sai khao; ทรายขาว (Northern)

No.	Scientific name	Thai name
17.	*C. lachnocarpus* Benth.	Khi on; ขี้อ้น (Prachuap Khiri Khan)
18.	*C. longissimus* Airy Shaw	Plao noi; เปล้าน้อย (Lamphun)
19.	*C. mekongensis* Gagnep.	Plao nam ngoen; เปล้าน้ำเงิน, Phrik na; พริกนา (Nakhon Sawan)
20.	*C. phuquocensis* Croizat	Mat; มาด
21.	*C. poilanei* Gagnep.	Plao luang; เปล้าหลวง, Plao lueat; เปล้าเลือด (Lampang), Plao; เปล้า, Plao yai; เปล้าใหญ่ (Southeastern)
22.	*C. poomae* Esser	
23.	*C. robustus* Kurz	Plao yai; เปล้าใหญ่ (Prachinburi)
24.	*C.roxburghii* N.P.Balakr.	Plao; เปล้า, Plao yai; เปล้าใหญ่ (General), Plao luang; เปล้าหลวง (Northern)
25.	*C. sepalinus* Airy Shaw	Plao ngoen; เปล้าเงิน (Peninsular)
26.	*C. stellatopilosus* Ohba	Plao noi; เปล้าน้อย (General)
27.	*C. thorelii* Gagnep.	Plao noi; เปล้าน้อย, Plao tawan; เปล้าตะวัน

No.	Scientific name	Thai name
28.	*C. tiglium* L.	Hat sakhuen; หัสคุณ, Salot; สลอด (General), Croton oil plant (English)
29.	*C. wallichii* Müll.Arg.	Plao; เปล้า, Plao na; เปล้านา (Peninsular)
	Croton sp. 1	
	Croton sp. 2	

2. Molecular Phylogeny Methods

Reconstruction of molecular phylogenies can be performed by using either distance based or character based methods as described below:

2.1 Distance Based Methods

Two common algorithms based on pairwise distances are the Unweighted Pair Group Method using Arithmetic averages (UPGMA) (Sneath and Sokal, 1973) and the neighbor joining algorithms (Saitou and Nei, 1987). The first step in these analyses is to compute a matrix of pairwise distances between operational taxonomical units from their sequence differences. To correct for multiple substitutions, it is common to use distances corrected by a model of molecular evolution such as the Jukes-Cantor model (Jukes and Cantor, 1969) and Kimura's two-parameter model (Kimura, 1980).

2.1.1 UPGMA

A simple but popular clustering algorithm for distance data is UPGMA (Michener and Sokal, 1957). This method works by initially having all sequences in separate clusters and continuously joining these. The tree is constructed by considering all initial clusters as leaf nodes in the tree. Each time two clusters are joined. A node is added to the tree as the parent of the two chosen nodes. The clusters to be joined are chosen as those with minimal pairwise distance. Branch lengths are set corresponding to the distance between clusters, which is calculated as the average distance between pairs of sequences in each cluster. The algorithm assumes that the distance data has the so called molecular clock property. This means the leaves of UPGMA trees all line up at the extant sequences and that a root is estimated as part of the procedure.

2.1.2 Neighbor Joining

The neighbor joining algorithm builds a tree where the evolutionary rates are free to differ in different lineages. Some programs always draw trees with roots for practical reasons. But for neighbor joining trees, no particular biological hypothesis is postulated by the placement of the root. The method works very much like UPGMA. The main difference is that, in instead of using pairwise distance, this method subtracts the distance to all other nodes from the pairwise distance. This is done to take care of situations where the two closest nodes are not neighbors in the real tree. The neighbor join algorithm is generally considered to be fairly good and is widely used. Algorithms that improve its cubic time performance exist. The improvement is only significant for quite large datasets (Saitou and Nei, 1987).

2.2 Character Based Methods

Whereas the distance based methods compress all sequence information into a single number, the character based methods attempt to infer the phylogeny based on all the individual characters such as nucleotides or amino acids.

2.2.1 Parsimony

In parsimony based methods, a number of sites are defined which are informative about the topology of the tree. Based on these, the best topology is found by minimizing the number of substitutions needed to explain the informative sites. Parsimony methods are not based on explicit evolutionary models (Hendy and Penny, 1989).

2.2.2 Maximum Likelihood

Maximum likelihood and Bayesian methods are probabilistic methods of inference (Cavalli-Sforza and Edwards, 1967). Both have the pleasing properties of using explicit models of molecular evolution and allowing for rigorous statistical inference. However, both approaches are very computer intensive. A stochastic model of molecular evolution is used to assign probability likelihood to each phylogeny, given the sequence data of the operating taxonomic units. Maximum likelihood inference (Felsenstein, 1981) then consists of finding the tree which assigns the highest probability to the data.

2.2.3 Bayesian Inference

The objective of Bayesian phylogenetic inference is not to infer a single correct phylogeny, but rather to obtain the full posterior probability distribution of all possible phylogenies (Yang *et al.*, 1995). This is obtained by combining the likelihood and the prior probability distribution of evolutionary parameters. The vast number of possible trees means that bayesian phylogenetics must be performed by approximative Monte Carlo based methods (Larget and Simon, 1999).

3. Phylogenetic Analysis

Phylogeny means the explaining observed similarities and differences among organisms. This invites the reconstruction of trees to show the evolutionary (Felsenstein, 1985). Traditionally, phylogenetic trees have been constructed from morphological data. Since the growth of genetic information, it has become common practice to construct phylogenies based on molecular data known as molecular phylogeny (Lio and Goldman, 1998). The data are presented in the form of DNA sequences, protein sequences and DNA fingerprints.

Molecular phylogeny reconstruction consists of four steps. By using DNA sequence as an informative data, the first step includes sequence selection and alignment. It is used for determining site by site homologies and detecting DNA sequence differences. The second step is building a mathematical model to describe the evolution of sequences. A model can be built empirically, using properties calculated through comparisons of observed sequences. Such models permit estimation of the genetic distance between two homologous sequences. It is measured by the expected number of nucleotide substitutions per site that have occurred on the evolutionary lineages between the observed and their common ancestor. Genetic distances are represented as branch lengths in the phylogenetic tree. The extant sequences form the tips of the tree, whereas the ancestral sequences form the internal nodes. The third step is applying an appropriate statistical method to find the tree topology and branch lengths that describe phylogenetic relationships. The final step is interpretation. Nowadays, mathematical models are implemented in numerous software program including MOLPHY (Adachi and Hasegawa, 1995), PHYLIP (Felsenstein, 1995), PASSML (Lio *et al*., 1998) and PAUP* (Swofford, 1998).

The inclusion of outgroup is an important criterion in phylogenetic analysis. The outgroup is assumed to be phylogenetically outside the group of taxa under study and can be any taxa. It is usually chosen as sister taxa that are genealogically closely related to the ingroup, but is not the ancestor. It is used for comparative purposes in determining character polarity and assigning the direction of change of character state transformation (Swofford *et al*., 1996). To use outgroup comparison, one or more species are selected that are relatives of the group being studied but outside of it. These species are equally related to all members of the group being studied. Thus the group under study is the ingroup and the species that are equally related to all ingroup members. The outgroup is also used to determine the root of a phylogenetic tree.

In addition, reliability test on the inferred phylogenetic trees is necessary. The most popular test is bootstrapping analysis. It is introduced to phylogenetic study by Felsenstein (1985). Measure how well a group is reflected by all the data in a sequence alignment, given the data analysis method used. This technique is to resample the alignment columns with replacement. For example, in the resampled alignment, a given column in the original alignment may occur two or more times, while some columns may not be represented in the new alignment at all. The resampled alignment represents an estimate of how a different set of sequences from the same genes and the same species may have evolved on the same tree. If a new tree reconstruction on the resampled alignment results in a tree similar to the original one, this increases the confidence in the original tree. In contrast, if the new tree looks very different, it means that the inferred tree is unreliable. Interpreting the reliability from bootstrapping analysis, if the data is bootstrapped 100 times, a bootstrap score of 100 means that the corresponding branch occurs in all 100 trees made from resampled alignments. Thus, a high bootstrap score is a sign of greater reliability.

4. The Use of Molecular Data in Phylogenetic Studies

Systematics is the process of detecting, describing and explaining diversity in the biological world. In 1758 Linnaeus formulated a hierarchical system of classification prior to the development of theories of evolution. Since the growth of molecular technology and the development of the polymerase chain reaction, PCR (Mullis and Faloona, 1987) resulted in large amounts of data being made available for DNA sequencing and DNA fingerprinting techniques. DNA sequences provide comparable characters that can be used to examine mechanisms of evolution of molecules by using the knowledge of evolutionary history of species. The evolution of molecules can conversely be applied when inferring evolutionary history of taxa. The greater availability of molecular data has caused improvements in the analysis of such data, resulting in the development and expansion of phylogenetic analysis as a whole.

The goals of systematic include phylogenetic reconstruction and the elucidation of the evolutionary processes that generate biological diversity. Many forms of informative characters especially DNA sequences have been applied to the study of plant relationships and evolution. The higher plants are three genomes, nuclear, mitochondrial and chloroplast. Their nucleotide substitutions are not occurring at the equal rate (Kimura, 1980). Nucleotide substitution rate of each genome is estimated (Wolfe *et al.*, 1987) by observing nucleotide substitution in various monocot and dicot. Mitochondrial genes have the lowest substitution rate, averaging (0.2-1.1) $x10^{-9}$ substitutions per site per year. Chloroplast genes are slightly faster around (1.1-2.9) $x10^{-9}$ substitutions per site per year. Nuclear genes, in contrast, has substitution rate up to 150 times as fast as mitochondrial genes (up to $31.5x10^{-9}$ substitutions per site per year). The work on nucleotide substitution rate estimated by using DNA sequence data from rice and maize (Gaut, 1998) show the agreement with Wolfe *et al.*, studies. This plant nucleotide substituted rate observation is fundamentally different from the extremely rapid mitochondrial molecular evolution in animals. In the short history of plant molecular systematics focus has primarily been on the chloroplast genome. However, this has changed as investigators have turned to nuclear gene sequences (Soltis and Soltis, 1998).

4.1 Chloroplast Genome

The chloroplast genome of most common photosynthetic land plants are characterized as a circular molecule. It is the smallest of the three genomes, ranging from 135-160 kb (Judd *et al*., 1999) with two inverted repeat (IR) segments dividing the remaining molecule into a LSC (large single-copy region) and a SSC (small single-copy region) respectively. The chloroplast genome's content, size, structure and linear sequences of genes, have revealed to be highly conservative in evolutionary studies (Palmer, 1991). This conservative mode suggests that any change in structure, arrangement or content may have significant phylogenetic implications. Different portions of the genome evolve at different rates, being fairly slow at the nucleotide sequence level (Downie and Palmer, 1992). Non-coding regions of the chloroplast genome evolve more rapidly than the coding regions (Wolfe and Sharp, 1988). Mutations in chloroplast DNA are either nucleotide substitutions or rearrangements (Palmer *et al*., 1988). Addition or deletion mutations accumulate in non-coding regions at a rate equal to nucleotide substitution and this category of mutation accelerates the divergence of non-coding regions. Most chloroplast genes are essentially single copy while nuclear genes are members of multigene families. The advantage of a conservative rate of nucleotide substitution arises because variation within the chloroplast DNA is appropriate to reconstruct plant phylogenies of evolution from species and genus level to family level and above (Downie and Palmer, 1992; Soltis and Soltis, 1998). The plastid genome is divided into coding and non-coding genes:

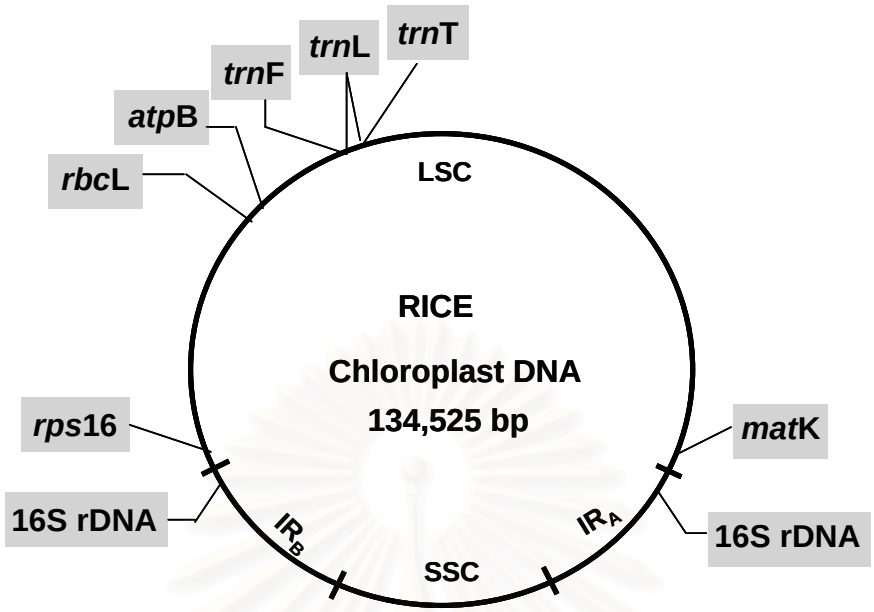


Figure 2.1 Coding and non-coding regions of chloroplast genome.

Coding genes

***rbc*L** gene is located in the LSC region and encodes the large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase (RUBISCO). At the family level and above, *rbc*L is the preferred gene for inferring phylogenies and is among the most frequently sequenced segments of DNA (Goldenburg *et al*., 1990; Soltis and Soltis, 1998). The lower limit of applicability of *rbc*L sequences extends to the generic level and in some groups reaches the specific level.

16S rDNA is a highly conserved chloroplast gene that is located within the inverted repeat and is roughly 1600 bp in length. The gene for the small subunit (SSU) of ribosomal RNA is used to reconstruct phylogeny at deep levels (Wolters and Erdmann, 1986; Mischer *et al*., 1992). Support for the phylogenetic utility of chloroplast 16S rDNA sequences within plants is still preliminary.

***atp*B** gene encodes the ATP synthase ß-subunit (Kuroda & Maliga, 2001). This enzyme couples proton translocation across membranes with the synthesis of ATP. It is located downstream from *rbc*L in the LSC. *atp*B gene is separated from *rbc*L by an intergenic spacer that is roughly 900 bp in length and has phylogenetic utility itself. The length of *atp*B is 1497 bp and no introns or indels are present thus making it easy to align (Soltis and Soltis, 1998). Previous work confirms that *atp*B and *rbc*L have similar rates of evolution in angiosperms (Savolainen *et al*., 1996).

***mat*K** is a rapidly evolving protein coding gene located within the LSC. It is approximately 1500 bp in length, corresponding to 500 amino acids. *mat*K gene encodes a maturase involved in slicing introns from RNA transcript. In examined photosynthetic land plants, *mat*K is located within an intron (2600 bp) positioned between the 5′ and 3′ exons of the tRNA gene *trn*K. Transcription experiments have shown that *trn*K including *mat*K are co transcribed (Chieba *et al*., 1996). A peculiarity of *mat*K are substitution rates in first and second codon positions approaching those in the third (Hilu and Liang, 1997), which contribute to the high overall evolutionary rate of *mat*K in contrast to other chloroplast genes. The *mat*K gene has become one of the most frequently used chloroplast markers in angiosperm phylogenetic studies (Mizukami *et al*., 1998; Katsuko *et al*, 2001). Since *mat*K gene can easily be co-amplified with the flanking non-coding intron parts, the complete *trn*K intron is increasingly used, expanding the dataset to 2400-2700 bp. As a consequence, the utility of this region could be extended to the inter and intraspecies levels (Steele and Vilgalys, 1994; Soltis *et al*., 1996; Shi *et al*., 2002)

Figure 2.2 Schematic diagram of *trn*K gene, *mat*K gene is embedded within *trn*K intron.

Non-coding genes

***trn*L-F gene** include the *trn*L (UAA) intron and the intergenic spacer between the *trn*L (UAA) 3′ exon and the *trn*F (GAA) gene. The *trn*L and *trn*F non-coding regions are used as a source of evidence at the generic level. These regions are located in the LSC, approximately 8 kb downstream of *rbc*L. The *trn*L intron ranges from 350-600 bp and the *trn*L-F intergenic spacer is 120-350 bp in length. These regions are easily amplified and sequenced (Taberlet *et al*., 1991). The *trn*L-F gene is known to be fast rate of evolution, evolving up to three times faster than *rbc*L (Gielly and Taberlet, 1994). Genetic variation of *trn*L-F gene is known to be short sequence repeats of 4 bp and more. Small indels (1-3 bp) are rare and usually confined to poly A/T strings (Borsch, 2003). These genetic variations are less parsimony informative to resolve in closely related taxa (Okaura and Harada, 2002; Fukuda *et al*., 2001). It is usually combined with other sequence data to reconstruct phylogenetic relationships (Hamzeh and Dayanandan, 2004; Kocyan, 2004; Muller *et al*., 2006).

Figure 2.3 Schematic diagram of *trn*L-F gene, *trn*L_{UAA} 5′ exon and *trn*L_{UAA} 3′ exon are separated by *trn*L_{UAA} intron. Whereas *trn*L_{UAA} 3′ exon and *trn*F_{GAA} were separated by intergenic spacer (IGS).

***rps*16** encodes the small subunit (30S) ribosomal protein S16 (Wood *et al*., 2000). It is located within the large single copy region between *trn*Q and *trn*K of the chloroplast genome of most flowering plants. The *rps*16 exons are interrupted by a long (790-967 bp) intron and evolves slower than the nuclear ribosomal ITS region providing a valuable data source for species-level phylogenetic studies (Baker *et al*., 2000; Lee and Hymowitz, 2001). The length of the *rps*16 intron varies from 707-951 bp (Oxelman *et al*., 1997).

4.2 Nuclear sequences

The nuclear genome is the largest in size ($1.1x10^6$ to $110x10^9$ kbp) and includes a variety of genes. Most attempts to infer phylogenies have involved the nuclear ribosomal DNA. The basic structure is a single repeat unit. Each repeat being reiterated thousands of times within most plant genomes. The nuclear genome consists of a transcribed region that comprises an external transcribed spacer (ETS), followed by the small subunit (18S) and large subunit (26S) genes, which are separated by the smaller 5.8S gene. Internal transcribed spacers (ITS1 and ITS2) separate the three genes. Repeats are separated from each other by an intergenic spacer. The approximate lengths of the three coding regions are similar throughout plants: 18S equals 1800 bp, 26S equals 3300 bp and 5.8S equals 160 bp. Contrasting is the length of the intergenic spacer which varies between 1 to 8 kb. The rDNA gene family comprises highly conserved regions such as 18S and 26S genes that can be used to infer phylogeny at higher taxonomic levels. The rapidly evolving segments such as ITS may be best suited for comparing species and closely related genera and even populations.

18S rDNA sequences are used more extensively than 26S rDNA sequences. Although the general taxonomic ranges of application of the two regions appear to be similar, the size of the 26S rDNA gene (over 3000 bp) has deterred investigators, particularly with regard to complete sequencing. In contrast, the size of 18S rDNA (1800 bp), has made it much more amendable to PCR amplification and sequencing. The fact that 18S rDNA topologies are highly congruent with *rbc*L based topologies at a diverse array of taxonomic levels in angiosperms. Despite the potential for reconstructing plant phylogeny, this 18S rDNA has been underused in angiosperm systematics (Soltis and Soltis, 1997).

26S rDNA gene is often noted as a candidate for sequencing as either an alternative or a supplement to 18S rDNA. When compared, the higher-level phylogenetic potential of the entire 26S rDNA gene evolves 1.6 to 2.2 times faster and provides over three times as many parsimony-informative characters than 18S rDNA.

5.8S rDNA gene is easily amplified and sequenced using primers located in the 18S and 26S rDNA genes. It has been rarely used to infer phylogeny because of its highly conserved and small size (164-165 bp). The proportion of potentially informative sites in seed plant sequences is similar to that of 18S rDNA.

ITS (internal transcribed spacer) region, ITS1 separates between 16S rDNA gene and 5.8S rDNA, while ITS2 separates between 5,8S rDNA gene and 26S rDNA. ITS region appears to be universally under 700 bp in flowering plants (Baldwin *et al*., 1995). This small size promotes ITS region easy to amplify even from herbarium materials. The ITS region has highly conserved sequences flanking each of two spacers so many universal primers were described (White *et al*., 1990; Taberlet *et al*., 1991; Urbatsch *et al*., 2000). ITS1 and ITS2 sequences are inherently GC rich making sequencing difficult as these regions are prone to secondary structure (Baldwin *et al*., 1995). Difficulty in sequencing varies from group to group and adding dimethyl sulfoxide (DMSO) or bovine serum albumin (BSA) to PCR and cycle sequencing reactions is an effective aid (Winship, 1989; Bult *et al*., 1992). The alignment of ITS sequences across angiosperm families indicated that plant ITS1 and ITS2 sequences have diverged more than nuclear ribosomal DNA subunits (Yokota *et al*., 1989). These spacers might be sufficiently variable to allow resolution of phylogenetic questions in closely related taxa. Thus, numerous articles presented the success of reconstruct reticulate evolution using ITS sequence data (Ritland *et al*., 1993; Steane *et al*., 1999; Linder, *et al*, 2000; Alvarez and Wendel, 2003; Bellarosa, *et al*., 2005). Moreover, ITS region is parentally inherited allowing it able to detect hybridization (Sang *et al*., 1995; Quijada *et al*., 1997; Whittall *et al*, 2000). This phenomenon is presenting as nucleotide additively.

Figure 2.4 Schematic diagram of the ITS region. The three ribosomal DNA subunits: 18S, 5.8S and 26S are separated by internal transcribed spacers (ITS1 and ITS2).

4.3 Mitochondrial Genome

Molecular study of mitochondrial DNA is a major focus of phylogenetic studies in animals. Rapid changes in the mitochondrial genome structure, size and configuration, results in a genome that is very difficult to analyze. These make the dynamics of mitochondrial DNA in plants different from those of other genomes.

In addition, Soltis (1999) and Qui (1999) suggested that phylogenies based on the combined sequence data representing different genomes are more reliable than phylogenetic trees based on individual genes because gene or genome specific can be rule out. The combining data between nuclear ribosomal DNA and chloroplast DNA could be performed both parent species on the interspecific hybridization (Padgett *et al*, 1998; Tsukaya, 2004; Zhang *et al*., 2007; Kondo *et al*., 2007). Because nuclear ribosomal DNA are parentally inherited data whereas chloroplast DNA are maternally inherited data.

5. Molecular Studies in Family Euphorbiaceae and Genus *Croton*

The family Euphorbiaceae is generally considered taxonomically complex. This problem is due to its extreme morphological diversity and a large number of species and subcosmopolitan distribution. Researchers have used molecular technology as a tool to gain a complete understanding of the family throughout its immense range. Recently, the ITS region of nuclear ribosomal DNA and chloroplast DNA *ndh*F nucleotide sequences have been used to evaluate the phylogenetic relationships of tribe Euphorbieae (Steinmann *et al*., 2002). The obtained parsimony tree has supported its relationship as monophyletic. Wurdack *et al*. (2005) have investigated the phylogenetic relationships of the pantropical family Euphorbiaceae sensu stricto by inferring from chloroplast *rbc*L and *trn*L-F sequence data. They have found that the plants in Euphorbiaceae sensu stricto were recovered as a monophyletic group with no adjustments in circumscription. For the plants in genus *Croton*, Berry *et al*. (2005) have published the first molecular survey within the genus *Croton* using the ITS region and *trn*L-F gene sequence data. Both sequence data have been found to be sufficiently variable to provide high statistic support for many clades in phylogenetic tree and their parsimony tree's topologies are congruent. The combined data set between ITS region and *trn*L-F sequences resolves many clades in the tree with bootstrap supports well above 50 %.

The phylogenetic tree is not only used to study phylogenetic relationships, but also used to determine the taxonomic status of plants. Van *et al*. (2006) have determined the phylogeny and biogeography of *C. alabamensis*, a rare *Croton* species. *C. alabamensis* is first discovered from the limited area in Alabama. After two years, it is surprisingly found in central Texas. The Texas plants are described as a new variety, *C. alabamensis* var. *texensis* Ginzbarg. The taxonomic status of *C. alabamensis* is ambiguously, whether it should be considered a single species with two varieties or two distinct, more localized species. Based on the nucleotide sequences of ITS and *trn*L-F, and the AFLP data, Van *et al*. performed the phylogenetic analysis. The old world *Croton* species, three species each from *Croton* section *Corylocroton* and the plant genus *Moacroton* are included as the next closest sister taxa. The obtained phylogenetic tree

(Fig. 2.5). have shown that *C. alabamensis* emerged alone on a long branch that is sister to *Croton* species section *Corylocroton* and the plant genus *Moacroton.* Moreover, the Alabama and the Texas plants split from their closest relative taxa (Fig 2.5) at 41 million years ago. Van *et al*. have concluded that the taxonomic status of *C. alabamensis* splits should be assigned as two varieties as var. *alabamensis* and var. *texensis*.

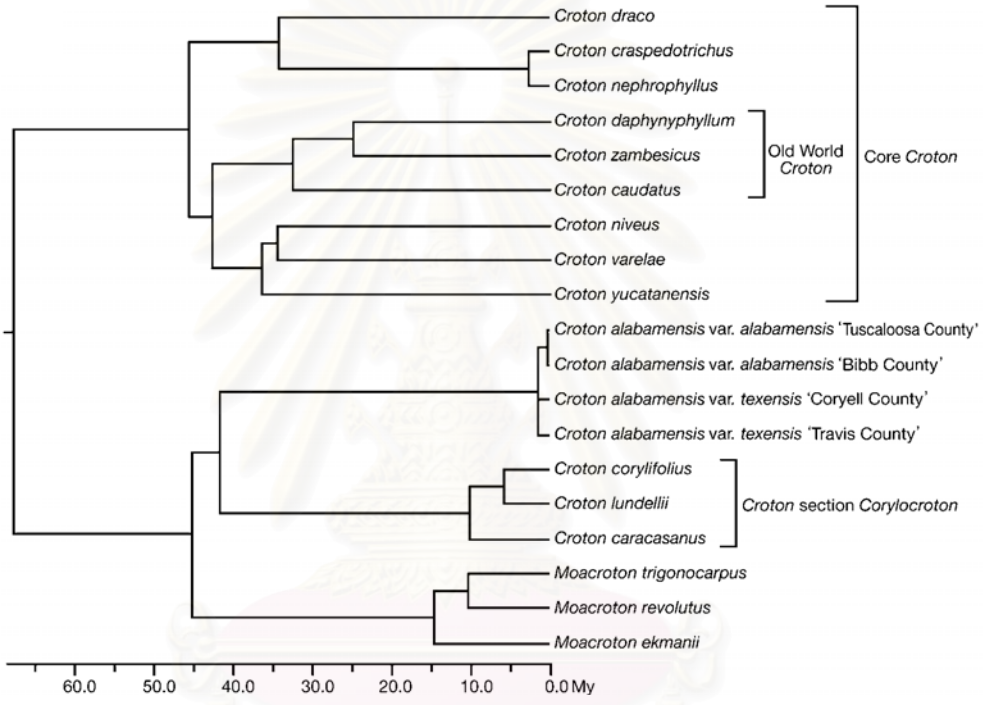


Figure 2.5 Maximum likelihood phylogram of *C. alabamemis*.

6. Molecular Studies on *Croton* Species in Thailand

In Thailand, although several *Croton* plants have been used in traditional medicine and some species are well documented in phytochemical analysis, there have been only a few molecular works reported. Soontornchainaksaeng *et al*. (2003) have studied on chromosome numbers and revealed that *Croton* species have quite unique chromosome number as 2*n*=20, except *C. hirtus* which contains 2*n*=16.

Klinbantoom (2004) has investigated the genetic diversity of Plau-noi, *C. stellatopilosus* that previously known as *C. sublyratus* (Esser and Chayamarit, 2001). This species is a natural source of plaunotol (Ogiso *et al*., 1978 and 1985) a peptic ulcer healing compound. Their plant materials were collected from various parts of Thailand including Nakhon Panom, Prachin Buri and Prachuap Khiri khan. Based on the inter simple sequence repeats (ISSRs) technique and UPGMA parsimony analysis, *C. stellatopilosus* was classified into two groups (Fig. 2.6) with closely genetic relation estimated from their genetic distance. Morphological characters of the leaves were also investigated and showed seven distinctive leaf morphologies (Klinbantoom (2004) as shown in Fig. 2.7. In this study, Klinbantoom (2004) proposed that genetic diversity of *C. stellatopilosus* varied as two genotypes. However, the correct genetic relationship of *C. stellatopilosus* species has not yet been concluded. Since Klinbantoom's work has focused only on the genetic diversity of *C. stellatopilosus* population that lacks the genetic data to relate *C. stellatopilosus* to the other *Croton* species. Moreover, on their UPGMA phylogenetic tree, one of *C. stellatopilosus* groups appears to consist of several sister groups (Fig. 2.6). The results, therefore, imply the taxonomical complexity of *C. stellatopilosus* species.

In addition, Plau-yai, *C. roxburghii* has also been investigated using ISSRs technique (Chonvanich, 2004). Plant materials were sampled throughout the country. Of 115 accessions, neighbor joining analysis (Fig. 2.8) is clearly separated as the major and the minor groups. At the major group, some *C. roxburghii* samples are clustered as a small clade. Based on the results, Chonvanich concluded that *C. roxburghii* is classified into two major groups, but did not indicate the taxonomical classification of both groups as Klinbantoom's conclusion.

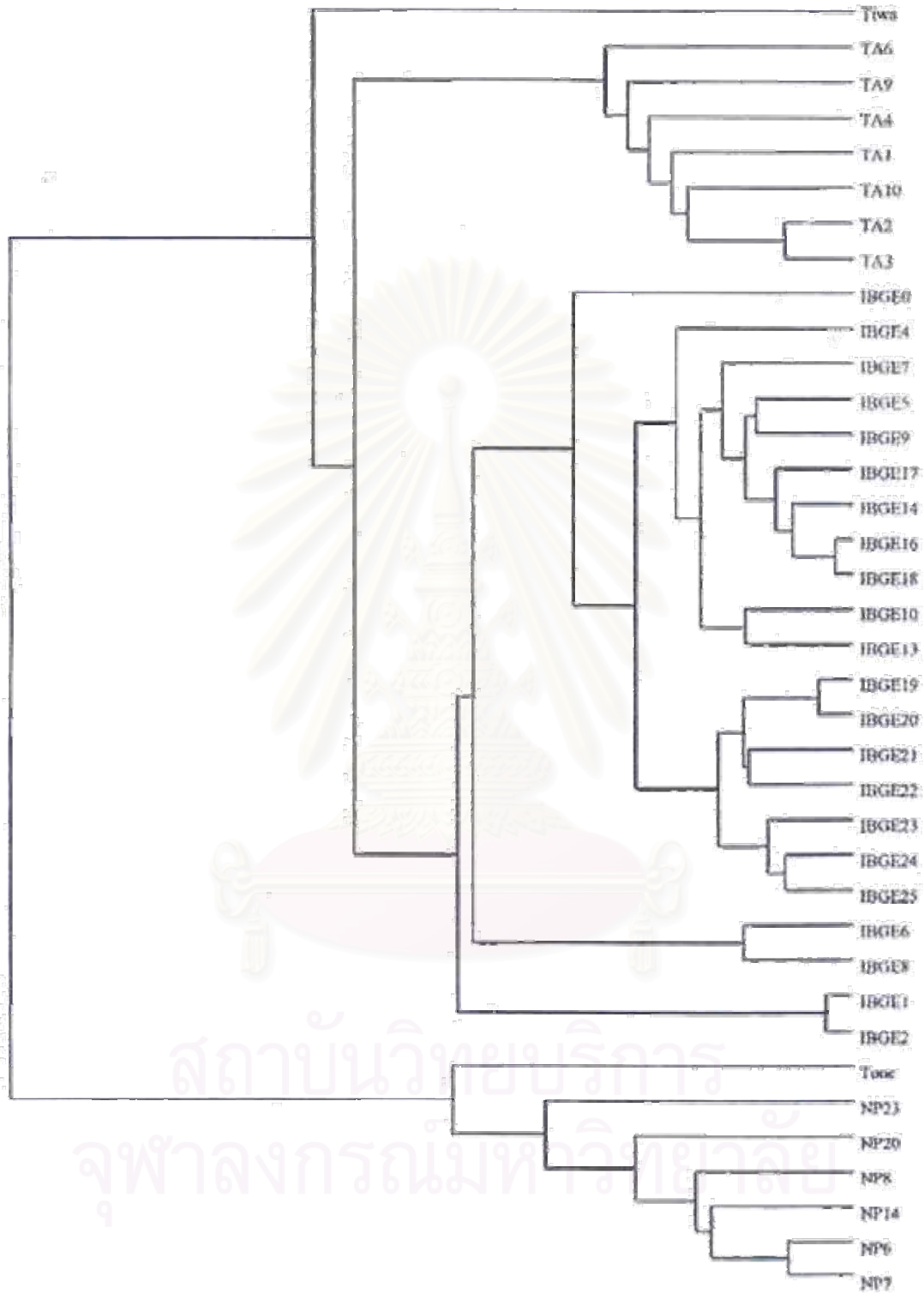


Figure 2.6 UPGMA analysis of *C. stellatopilosus* inferred from ISSRs profile (Klinbantoom, 2004).

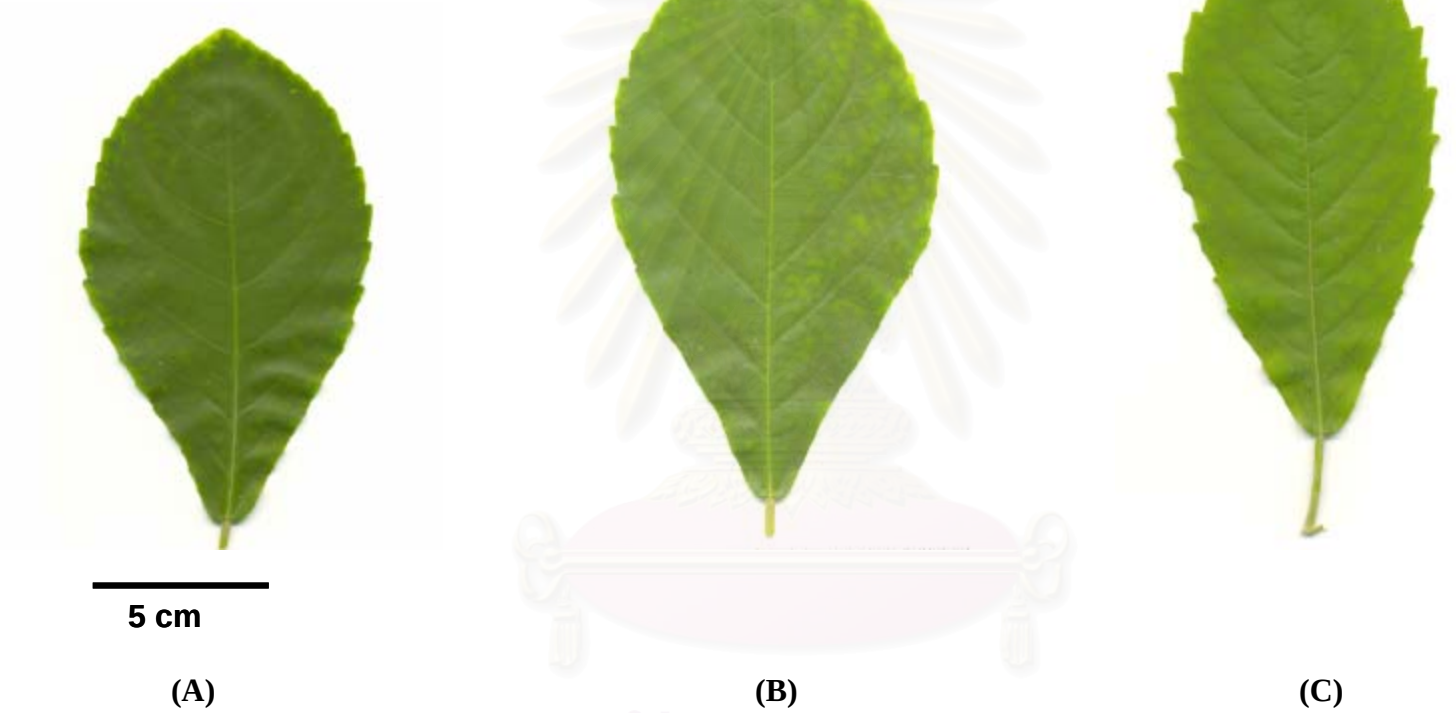


Figure 2.7 Morphological characters of *C. stellatopilosus* leaves (Klinbantoom, 2004).

(A) and (B) were specimens collecting from Muang, Prachaup Khiri Khan; (C) Prachin Buri.

28

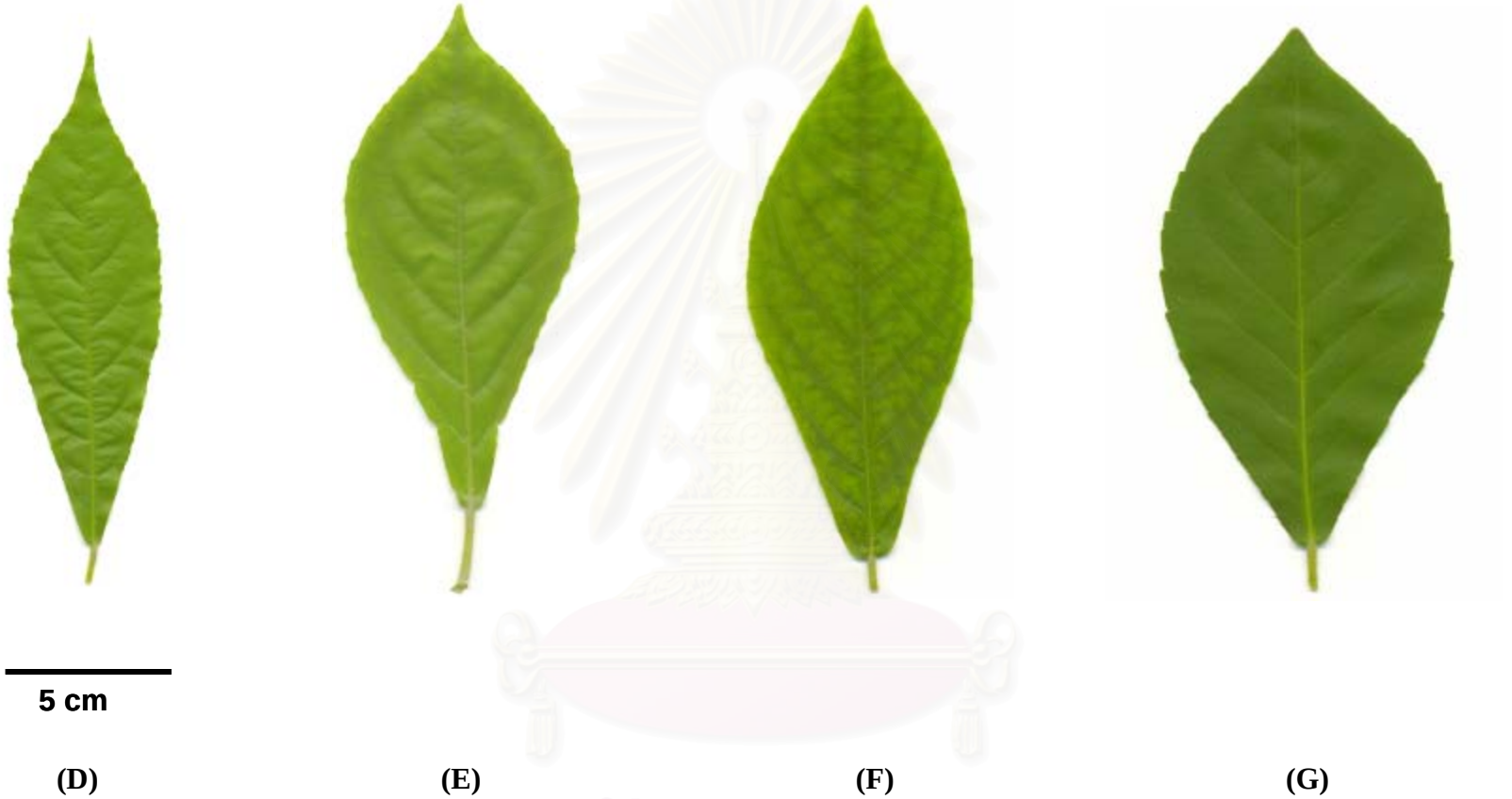
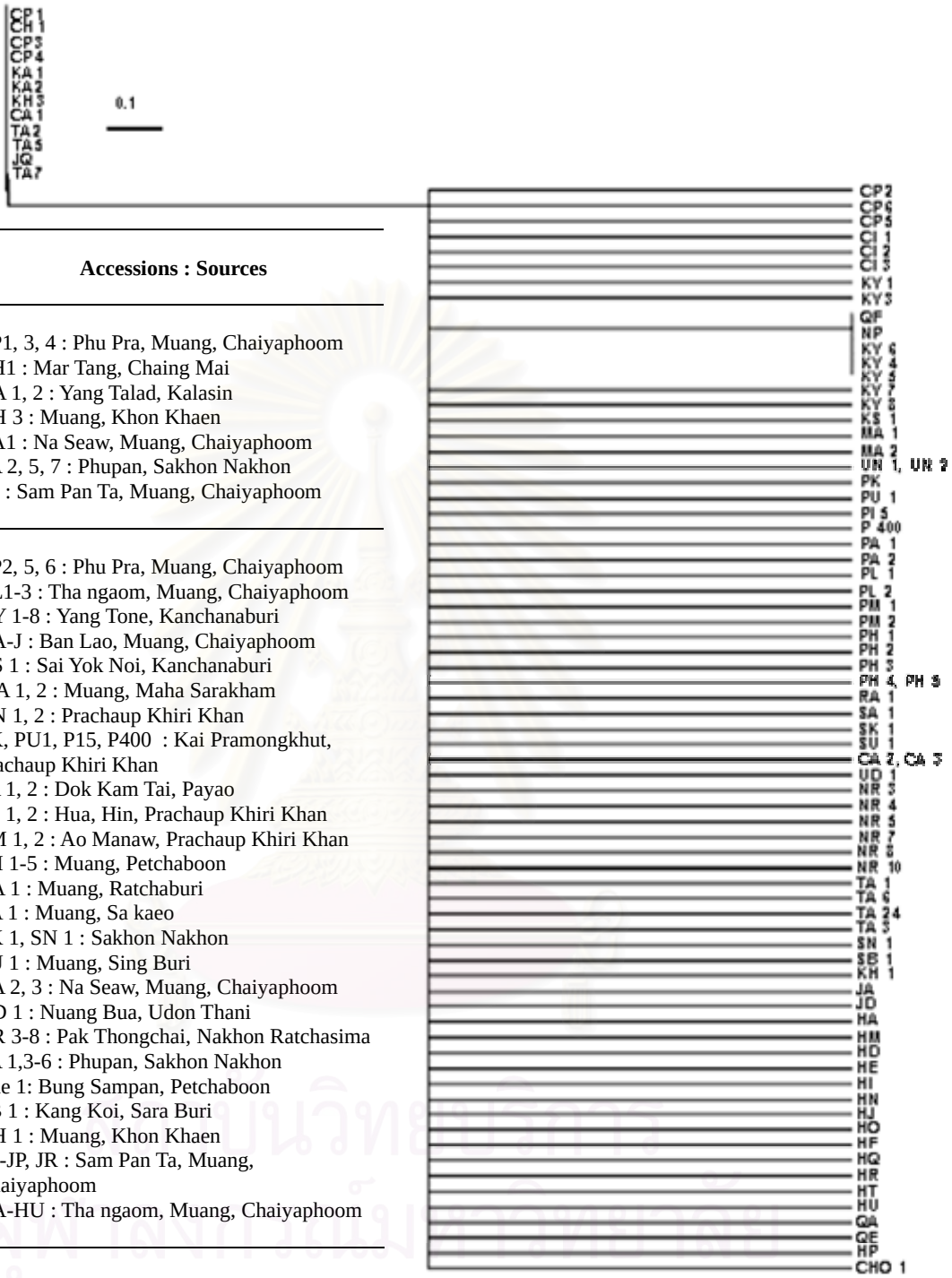


Figure 2.7 Morphological characters of *C. stellatopilosus* leaves (continued).

(D) and (G) were specimens collecting from Muang, Prachaup Khiri Khan; (E) Tha Utane, Nakhon Phanom; (F) Wanakorn, Prachaup Khiri Khan.



Groups	Accessions : Sources
A	CP1, 3, 4 : Phu Pra, Muang, Chaiyaphoom CH1 : Mar Tang, Chaing Mai KA 1, 2 : Yang Talad, Kalasin KH 3 : Muang, Khon Khaen CA1 : Na Seaw, Muang, Chaiyaphoom TA 2, 5, 7 : Phupan, Sakhon Nakhon JQ : Sam Pan Ta, Muang, Chaiyaphoom
B	CP2, 5, 6 : Phu Pra, Muang, Chaiyaphoom CL1-3 : Tha ngaom, Muang, Chaiyaphoom KY 1-8 : Yang Tone, Kanchanaburi QA-J : Ban Lao, Muang, Chaiyaphoom KS 1 : Sai Yok Noi, Kanchanaburi MA 1, 2 : Muang, Maha Sarakham UN 1, 2 : Prachaup Khiri Khan PK, PU1, P15, P400 : Kai Pramongkhut, Prachaup Khiri Khan PA 1, 2 : Dok Kam Tai, Payao PL 1, 2 : Hua, Hin, Prachaup Khiri Khan PM 1, 2 : Ao Manaw, Prachaup Khiri Khan PH 1-5 : Muang, Petchaboon RA 1 : Muang, Ratchaburi SA 1 : Muang, Sa kaeo SK 1, SN 1 : Sakhon Nakhon SU 1 : Muang, Sing Buri CA 2, 3 : Na Seaw, Muang, Chaiyaphoom UD 1 : Nuang Bua, Udon Thani NR 3-8 : Pak Thongchai, Nakhon Ratchasima TA 1,3-6 : Phupan, Sakhon Nakhon Che 1: Bung Sampan, Petchaboon SB 1 : Kang Koi, Sara Buri KH 1 : Muang, Khon Khaen JA-JP, JR : Sam Pan Ta, Muang, Chaiyaphoom HA-HU : Tha ngaom, Muang, Chaiyaphoom

Figure 2.8 Neighbor joining tree of *C. roxburghii* inferred from ISSRs profile (Chonvanich, 2004).

7. Natural Hybridization

Natural hybridization is an important factor in the evolution of many plant taxa. It can create evolutionary novelties that promote adaptive evolution and speciation. The term “hybrid” can be restricted to organisms formed by cross fertilization between individuals of different species, or it can be defined more broadly as the offspring between individuals from populations which are distinguishable on the basis of one or more heritable characters. Gottlieb (1972) has discussed several criteria for testing whether a particular diploid taxon originated through hybridization. These include geographic distribution in the region of parental sympatry, morphological intermediacy in several characters, partial fertility and biochemical additive. The study of interspecific hybridization has been mostly investigated based on phenotypic characters solely. However, it is also essential that these “hybrid” criteria be evaluated carefully because features such as intermediacy and sterility are not invariably associated with hybrids and may result from entirely separate processes (Les and Philbrick, 1993). Morphologically, hybrids typically display a mosaic of parental and intermediate characters, although extreme and novel characters appear quite often in the hybrid phenotype. Morphological characters alone are of limited value when identifying natural hybrids (Marhold *et al*., 2002). A species with morphological characters intermediate between two recognized species has always been considered to be a hybrid. When the polymerase enzyme and direct sequencing technique are developed, such molecular tools have greatly engaged in hybrid studies. Recently, the use of molecular tools has shown that hybridization is even more prevalent than indicated by morphological and cytogenetic evidence (Rieseberg, 1997). Molecular approach has also revealed many historic hybridization events (Rieseberg and Soltis, 1991: Wendel *et al*., 1995; Campbell *et al.*, 1997; Nelson-Jones *et al.*, 2002; Koch *et al.*, 2003; Ritz *et al.*, 2005; Fehrer *et al.*, 2007). Nucleotide sequences of the nuclear and chloroplast genomes have greatly increased the possibility of detecting paternal and maternal genome lineages of hybridization. The hybridizations are found to be unevenly distributed taxonomically. Only 16-34 % of plant families and 6-16 % of genera have one or more reported hybrids (Ellstrand *et al*., 1996).

Thus, contemporary hybridization may not be common, but appears to be concentrated in a small fraction of families and genera.

Tandem repeats such as nuclear ribosomal DNA are ideally suited for studying hybridization events. Because concerted evolution results in sequences conserved within species and divergent between species (Boldwin *et al.*, 1995). Nuclear ribosomal DNA is biparentally inherited thus recent hybrids initially possess both divergent parental genotypes as evidenced by DNA sequence polymorphisms. In the absence of sexual reproduction, concerted evolutionary homogenization of sequences by interchromosomal crossing-over or gene conversion during chromosome pairing at meiosis would not be expected. Consequently, hybrids reproducing strictly vegetatively should retain copies of both divergent sequences for prolonged periods. In such instances, isolation of individual DNA sequences by molecular cloning can reveal the paternal origin of a hybridization event when sequences matching each parental species are recovered. As an adjunct, maternally inherited regions such as chloroplast DNA can distinguish the maternal component of a particular hybrid cross.

In general, most molecular studies of putative hybrid species is the observation from (1) incongruence between phylogenies based on maternally (chloroplast genome) and biparentally (nuclear ribosomal DNA) inherited genes (Wendel *et al*., 1995; Yang *et al*., 2000; Tsukaya *et al*., 2003; Tsukaya, 2004), (2) additivity patterns markers such as nuclear ribosomal DNA (Soltis and Soltis, 1991; Kim and Jansen, 1994; Sang *et al*., 1995; Whittall *et al*., 2000) and (3) significant conflict among morphological, biogeographical, and molecular evidence (Sytsma,1990; Fuertes *et al*., 1999). The ITS region of nuclear ribosomal DNA is considered being the most widely used genetic marker for phylogenetic reconstruction below the family level in angiosperms (Baldwin *et al*., 1995) as the numerous of natural hybridization has been reported (Ritland *et al*., 1993; Sang *et al*., 1995; Whittall *et al*., 2000; Linder, *et al*, 2000; Alvarez and Wendel, 2003; Bellarosa *et al*., 2005).

CHAPTER III

MATERIAL AND METHODS

Plant Materials

Several species of the genus *Croton* were investigated in this study. They were collected from various parts of Thailand. Fresh leaves were collected and then immediately dried with silica gel. All plant samples were identified to their species on the basis of their morphological descriptions in Flora of Thailand (Chayamarit and Welzen, 2005) and by comparison with specimens stored in the Herbarium of the Office of Forest and Plant Conservation Research National Park, Wildlife and Plant Conservation Department, Thailand. Their voucher specimens have been deposited at the Museum of Natural Medicines, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand and the Museum of Materia Medica, Institute of Natural medicines, University of Toyama, Japan. The origins, reference specimens and sample codes are listed in Table 3.1.

In addition, some plant materials were obtained from the voucher specimens that have been stored in the Herbarium of the Office of Forest and Plant Conservation Research National Park. The origin of voucher specimens and sample codes of herbarium samples are listed in Table 3.2. Overall eighty four samples representing twenty six *Croton* species were examined in this study. To differentiate the extremely different characteristics between silica gel dried and herbarium specimen materials, here after, the “fresh sample” was used to represent the fresh sample that was dried with silica gel, while the “herbarium sample” represented the leaf obtained from the herbarium specimens.

Table 3.1 Details of the fresh material.

Species	Origins	reference specimens (BKF number)	Sample codes
Croton argyratus Blume	Hala Bala, Nara Thiwas	131941	ARG 01
C. cascarilloides Raeusch.	Kaeng Krachan, Phetcha Buri	66419	CAS 01
C. columnaris Airy Shaw	Phu Phoi Loom,Udon Thani	140316	COL
C. crassifolius Geiseler	Siri Ruckhacharti Medicinal Plant Garden, Nakhon Pathom	135991	CRA
C. kongensis Gagnep.	Phu Soi Dao, Uttaradit Khun Chae, Chaing Rai	109016	KON 01 KON 02
C. poilanei Gagnep.	Muang, Surat Thani Muang, Surat Thani	103196	POI 01 POI 02
C. roxburghii N.P. Baiakr.	Ang Khang, Chaing Mai Phu Soi Dao, Uttaradit Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok Chao Praya Apai Phubhate Hospital, Prachin Buri Siri Ruckhacharti Medicinal Plant Garden, Nakhon Pathom	2233	ROX 01 ROX 02 ROX 03 ROX 04 ROX 05
C. stellatopilosus Ohba	Huai Sai, Prchuap Khiri Khan Wanakorn, Prchuap Khiri Khan Institute of Biotechnology and Genetic engineering Plant Growing Open Field, Nakhon Pathom Princess Maha Chakri Sirindhorn Herb Garden, Rayong Phu Phoi Loom, Udon Thani Faculty of Science, Chulalongkorn University, Bangkok Faculty of Forestly, Kasetsart University, Bangkok Siri Ruckhacharti Medicinal Plant Garden, Nakhon Pathom Institute of Biotechnology and Genetic engineering Plant Growing Open Field, Nakhon Pathom	21867	STEL 01 STEL 02 STEL 03-04 STEL 05-07 STEL 08-09 STEL 10 STEL 11 STEL 12 STEL 13-30
C. tiglium L.	Pratamnak Pathum, Pathum Thani Bala Bala, Nara Thiwas Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok Siri Ruckhacharti Medicinal Plat Garden, Nakhon Pathom	105937	TIG 01 TIG 02 TIG 03 TIG 04
C. wallichii Mull. Arg.	Kaeng Krachan, Phetcha Buri	71594	WAL 01
Total	**10 *Croton* species, 48 samples**		

Table 3.2 Details of the herbarium material.

Species	Origins of voucher specimens	BKF number	Collected year	Age	Sample code
Croton acutifolius Esser	Doi Saket, Chiang Mai	135991	2002	5	ACUT
C. argyratus Blume	Ton Nga Chang, Hat Yai, Songkhla	131941	2001	6	ARG 02
C. bonplandianus Baill.	Sikhiu, Nakhon Ratchasima	122547	1990	17	BON
C. cascarilloides Raeusch.	Khao Khieo, Chon Buri	130086	2000	7	CAS 02
	Bangsaphan, Prachuap Khiri Khan	133084	2001	6	CAS 03
C. caudatus Geiseler	Ban Na Kum, Muang, Trat	140316	2002	5	CAU 01
	Tharnmayom Fall, Koh Chang, Trat	128510	2000	7	CAU 02
	Khlong Makok, Ko Chang National Park, Trat	132275	2001	6	CAU 03
C. decalvatus Esser	Roipi Botanical Garden, Sa Kaeo	128950	2000	7	DEC 01
	Roipi Botanical Garden, Sa Kaeo	126723	1998	9	DEC 02
	Ban Ang, Makham, Chantha Buri	69911	1977	30	DEC 03
C. delpyi Gagnep.	Taphan Hin	58275	1973	34	DEL 01
C. griffithii Hook. F.	Khao pra-Bang Kram, Krabi	118283	1990	17	GRIF 01
	Pen, Narathiwat, Waeng	52177	1972	35	GRIF 02
C. hirtus L' Her.	Sungai Kolok, Nara Thiwat	111508	1993	14	HIR 01
C. hutchinsonianus Hosseus in Fedde	Thum Pha, Kanchana Buri	58074	1961	47	HUT 01
	Ban Rai, Huai Ka Kaeng, Uthai Thani	65704	1970	38	HUT 02

Table 3.2 Details of the herbarium material (continued).

Species	Origin of voucher specimens	BKF number	Collected year	Age	Sample codes
Croton. kongkandanus Esser	Si Satchanalai Natural Park, Phitsanulok	87372	1983	24	KAN 01
	Pha Wo, Tak	51475	1972	35	KAN 02
C. kerrii Airy Shaw	Phu luang Wildlife Sanctuary, Loei	127357	1998	9	KER 01
	Phu luang Wildlife Sanctuary, Loei	124468	1998	9	KER 02
C. krabas Gagnep.	Suphan Buri	115734	1920	77	KRA 01
	Kantrarom, Sri Sa Ket,	11937	1951	56	KRA 02
	Panampo, Nakhon Sawan	13299	1920	79	KRA 03
C. lachnocarpus Benth.	Erawan National Park, Kanchana Buri	79765	1979	28	LAN 01
C. poilanei Gagnep.	Ko Chang National Park, Trat	132268	2001	6	POI 03
	Koh Chang Island, Chanta Buri	103196	1996	13	POI 04
C. poomae Esser	Phu Wua, Bungkhla, Nong Khai	140308	2002	5	POO 01
C. robustus Kurz	Koh Khram, Satthahip, Chon Buri	128127	1999	8	ROB 01
	Koh Khram, Satthahip, Chon Buri	128123	1999	8	ROB 02
	Kow Kav, Chol Buri	41812	1969	38	ROB 03
C. sepalinus Airy Shaw	Pa Kiampoa, Tachang, Surat Thani	61105	1974	33	SEP 01
	Prachuap Khiri Khan. Bang Saphan	65185	1970	37	SEP 02
C. thorelii Gagnep.	Phu Phan National Park, Nakhon Phanom	84753	1982	25	THO 01
	Ngaw Waterfall Forest Reserve, Ranong	114340	1984	23	THO 02
C. wallichii Mull. Arg.	Sangkhla, Kanchana Buri	43331	1968	39	WAL 02
Total	**20 *Croton* species, 36 samples**				

.

Methods

1. Total DNA Preparation

Approximately 20-40 mg of each fresh sample or 50-100 mg of each herbarium sample was grounded into fine powder with Tissue Lyser (QIAGEN, Germany). The fresh samples were total DNA extracted by DNeasy™ Plant Mini Kit (QIAGEN, Germany) following the manufacturer's protocol. For herbarium samples, total DNA extraction was performed using the modified manufacture's protocol as suggested by Drabkova *et al*. (2002). Total DNA extraction protocols are summarized in Table 3.3. The herbarium total DNA extracts were cleaned by GeneClean II kit (Bio101 systems) according to the manufacturer's protocol. The total DNA extracts were all kept at -20 °C until used.

2. Total DNA Determination

The obtained total DNA extracts were determined qualitatively and quantitatively using gel electrophoresis method. They were electrophoresed through 1 % TAE agarose gel (Bio-Rad Laboratories, U.S.A.). The gel was submerged in an electrophoresis chamber (Bio-Rad Laboratories, U.S.A.) containing 0.5X TAE buffer. Five microliters of the total DNA was mixed with 2 µl of loading dye (Invitrogen Corp., U.S.A.) then loaded into gel. One kb plus DNA ladder (Invitrogen Corp., U.S.A.) was used as a standard molecular marker. Electrophoresis was operated at 100 volts for 40 min. The gel was stained with ethidium bromide solution for 30 min and de-stained in distilled water for 10 min to remove unbound ethidium bromide from the gel. The obtained total DNA was visualized under ultra violet transilluminator and photographed with Gel Doc™ XR System PC/Mac (Bio-Rad Laboratories, U.S.A.). Size and concentration of total DNA extracts were determined by comparing with 1kb plus molecular markers (Invitrogen Corp., U.S.A.).

Table 3.3 The comparison of total DNA extraction protocols.

Step	DNeasy™ Plant Mini Kit total DNA extraction protocols	
	Manufacturer's protocol	**Modified protocol**
1	-	Rinsed the find powder materials with PBS-EDTA solution pH 7.4, 3 times.
2	Add 400 µl buffer AP1 and 4 µl RNase A. Vortex and incubate for 10 min at 65 ℃.	Add **600 µl** buffer AP1 and 4 µl RNase A. Vortex and incubate for **4 hours** at 65 ℃.
3	Add 130 µl buffer AP2. Mix and incubate for 5 min on ice.	Add 130 µl buffer AP2. Mix and incubate for 5 min on ice.
4	Pipet the lysate into a QIAshredder Mini spin column in a 2 ml collection tube. Centrifuge for 2 min at 14,000 rpm.	Pipet the lysate into a QIAshredder Mini spin column in a 2 ml collection tube. Centrifuge for 2 min at 14,000 rpm.
5	Transfer the flow-through fraction into a new tube without disturbing the pellet. Add 1.5 volumes of buffer AP3/E and mix by pipetting.	Transfer the flow-through fraction into a new tube without disturbing the pellet. Add 1.5 volumes of buffer AP3/E and mix by pipetting. **Incubate at -20°C overnight.**
6	Transfer 650 µl of the mixture into a DNeasy Mini spin column in a 2 ml collection tube. Centrifuge for 1 min at ≥8,000 rpm. Discard flow-through. Repeat this step with the remain samples.	Transfer 650 µl of the mixture into a DNeasy Mini spin column in a 2 ml collection tube. Centrifuge for 1 min at ≥8,000 rpm. Discard flow-through. Repeat this step with the remain samples.
7	Place the spin column into a new 2 ml collection tube. Add 500 µl buffer AW and Centrifuge for 1 min at ≥8,000 rpm. Discard flow-through.	Place the spin column into a new 2 ml collection tube. Add 500 µl buffer AW and Centrifuge for 1 min at ≥8,000 rpm. Discard flow-through.
8	Add another 500 µl buffer AW. Centrifuge for 2 min at 14,000 rpm.	Add another 500 µl buffer AW. Centrifuge for 2 min at 14,000 rpm.
9	Transfer the spin column into a new 2 ml microcentrifuge tube and add 100 µl buffer AE for elution. Incubate for 5 min at room temperature. Centrifuge for 1 min ≥8,000 rpm. Repeat this step.	Transfer the spin column into a new 2 ml microcentrifuge tube and add **40 µl** buffer AE for elution. Incubate for **1 hour** at room temperature. Centrifuge for 1 min ≥8,000 rpm.

The bold letters represent the critical steps of the modified total DNA extraction protocol.

3. Design of Oligonucleotide Primers

3.1 Entire ITS Region

To amplify the entire ITS region as overlapping fragments and to perform the nested PCR amplification, four oligonucleotide primers were designed. The primers In 5.8 F, In 5.8 R, 5.8S 77F and 5.8S 77R were designed from the nucleotide sequences of the conserved coding region 5.8S rDNA of rice (Takaiwa *et al*., 1985), tobacco (Kiss *et al*., 1988) and mung bean (Schiebel and Hemleben, 1989). The designed primers were synthesized by Sigma Genosys (Singapore).

3.2 *trn*K Intron

The oligonucleotide primers matK500F and matK500R were designed based on *trn*K and *mat*K nucleotide sequences of species taxonomical related with the genus *Croton*. Four nucleotide sequences published in GenBank, three partial *trn*K nucleotide sequence of *Breynia disticha* (GenBank accession number AY936564), *Margaritaria rhomboidalis* (AY936571) and *Phyllanthus virgatus* (AY936639), family Phyllanthaceae and a partial *mat*K nucleotide sequence of *Croton insularis* (AB233773), family Euphorbiaceae were used in this study. The designed primers were synthesized by Sigma Genosys (Singapore).

4. PCR Amplification

All of the eighty four *Croton* samples representing twenty six *Croton* species were analyzed for their nucleotide sequences on ITS region and *trn*L-F gene. Additionally, thirty samples of *C. stellatopilosus* and five samples of *C. roxburghii* were further examined for their nucleotide sequences on the complete *trn*K intron.

The amplification mixtures of the entire ITS region, *trn*L-F gene and the complete *trn*K intron were carried out with 50-100 ng of total DNA as a template in a total volume of 50 µl reaction mixture. The mixture contained 10mM Tris-HCl (pH 9.0), 50 mM KCl, 0.1 % Triton X-100, 1.5 M $MgCl_2$, 0.2 mM of dNTP, 0.25µM of each primer and 1.5 unit of *Taq* polymerase (Promega, U.S.A.). Some total DNA extracts that were difficult to PCR amplified, 25 mg of non-acetylated Bovine Serum Albumin (Ambion, Japan) was added to enhance the polymerase activity (Paabo *et al*., 1990).

4.1 The Entire ITS Region

The ITS region consisted of 18S, 5.8S and 26S rDNA, or so called the entire ITS region. All total DNA preparations extracted from the fresh samples were PCR amplified using universal primers ITS-1 of Urbatsch, *et al*. (2000) and ITS-4 of White, *et al*. (1990) as forward and reversed primers, respectively (Fig. 3.1a). For the herbarium samples, the entire ITS region was amplified into two overlapping fragments which were designated as the fragment ITS1 and the fragment ITS2. The primer pair ITS-1 (Urbatsch *et al*., 2000) and 5.8S 77R was used to amplify the fragment ITS1, while the primer pair In 5.8S F and ITS-4 (White *et al*., 1990) was used to amplify the fragment ITS2 (Fig. 3.1b).

In addition, for herbarium samples containing degraded total DNA, their entire ITS region was amplified into two overlapping fragments using nested PCR amplification technique. Each target fragment was obtained from two successive PCR amplifications. The first PCR amplification was performed with the outer primer pair. The obtained PCR products were then purified and further used as a template for the second amplification using the inner primer pair. The amplification primers used for nested amplification are summarized in Table 3.4 and Fig. 3.1c. Details of primers used for the entire ITS amplification are listed in Table 3.5. Both standard and nested amplifications were carried out on a Takara PCR Thermal Cycler (Takara, Japan) beginning at 95 °C (5 min), followed by 35 cycles of 95 °C (30 sec), 50 °C (30 sec) and 72 °C (30 sec). After final extension at 72 °C (10 min), the reaction was held at 4 ºC.

Table 3.4 Amplified primers used for the nested amplification of the entire ITS region.

Target PCR fragment	Primer pair	Direction	Primer	Reference
ITS1	Outer	Forward	ITS-5	White *et al*., 1990
		Reverse	In 5.8 R	In this study
	Inner	Forward	ITS-1	Urbatsch *et al*., 2000
		Reverse	5.8S 77R	In this study
ITS2	Outer	Forward	In 5.8 F	In this study
		Reverse	ITS-4	White *et al*., 1990
	Inner	Forward	5.8S 77F	In this study
		Reverse	18S 25S 3'R	Sukrong *et al.*, 2007

Table 3.5 Sequence of amplified primer used for the ITS amplification.

Direction	Primer	Sequence (5′ → 3′)	Length (bp)	Tm (°C)
Forward	ITS-5	GGA AGT AAA AGT CGT AAC AAG G	22	53.9
	ITS-1	GTC CAC TGA ACC TTA TCA TTT AG	23	52.1
	In 5.8 F	TCT CGC ATC GAT GAA GAA CG	20	51.8
	5.8S 77F	AAT GCG ATA CCT GGT GTG AAT TGC	24	57.9
Reverse	In 5.8R	GAC TCG ATG GTT CAC GGG ATT CT	23	57.1
	5.8S 77R	GCA ATT CAC ACC AGG TAT CGC ATT	24	57.0
	ITS-4	TCC TCC GCT TAT TTA TAT GC	20	51.3
	18S 25S 3'R	CCA TGC TTA AAC TCA GCG GGT	21	54.4

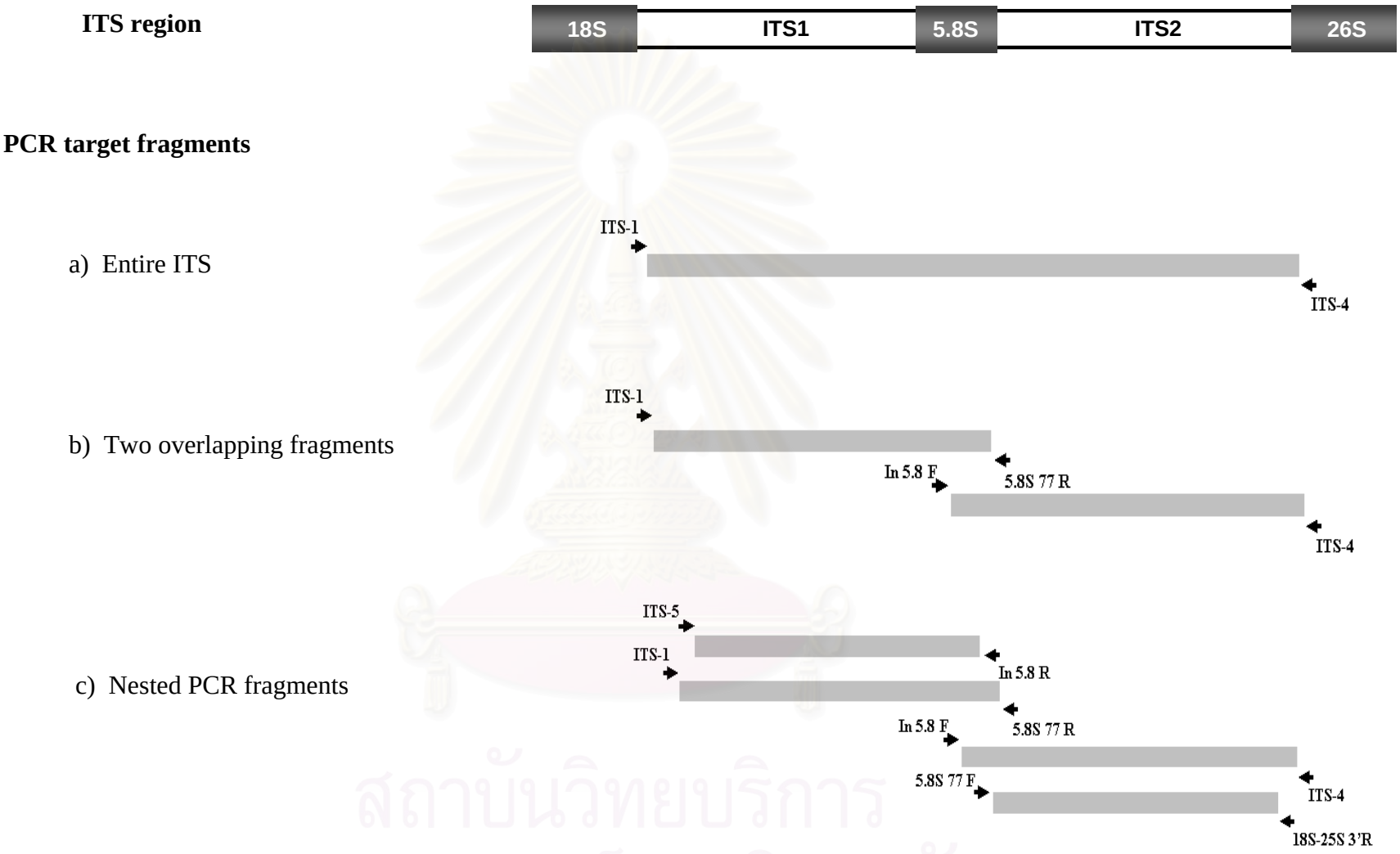


Figure 3.1 The strategies used for the ITS amplification.

4.2 *trn*L-F Gene Amplification

*trn*L-F gene is composed of *trn*L-F intron and intergenic spacer. PCR amplification was performed with four universal reported previously by Taberlet *et al.* (1991). The primer pair *c* and *d* was used to amplify *trn*L intron, while the primer pair *e* and *f* was used to amplify *trn*L-F intergenic spacer. The relative position of primers is presented in Fig. 3.2. PCR amplification was also performed on Takara PCR Thermal Cycler and the amplify reaction was programmed as hot start at 94 °C (5 min), follow by 35 cycles of 95 °C (1 min), 50 °C (1 min), 72 °C (1 min). Final extension was done at 72 °C (10 min), and the reaction was held at 4 °C.

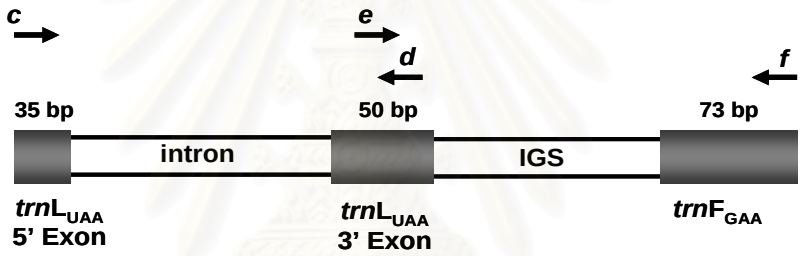


Figure 3.2 The universal primers (Taberlet *et al.*, 1991) used for *trn*L-F amplification.

4.3 The Complete and Partial *trn*K Intron Amplifications

In this study, two types of PCR amplification of *trn*K intron were conducted, the complete and the partial *trn*K intron amplifications. The complete *trn*K intron amplification was carried out on two overlapping fragments (fragment I and II) using two primer pairs. The primers trnK3914F (Johnson and Soltis, 1994) and matK3R (Yang *et al*., 2004) were used to amplify the fragment I, while the primers matK500F and trnK2R (Johnson and Soltis, 1994) were used to amplify the fragment II. For the partail *trn*K intron amplification, it was performed by using primer matK500F and primer trnK2R (Johnson and Soltis, 1994). The relative position of the primers is presented in Fig. 3.3 and the detail of the primers is shown in Table 3.6. Both complete and partial *trn*K intron amplifications were carried out on a Takara PCR Thermal Cycler and programmed as hot start at 95 °C (5 min), followed by 35 cycles of 95 °C (1 min), 50 °C (1 min), 72 °C (1.5 min). Final extension was performed at 72 °C (10 min), and the reaction was held at 4 °C.

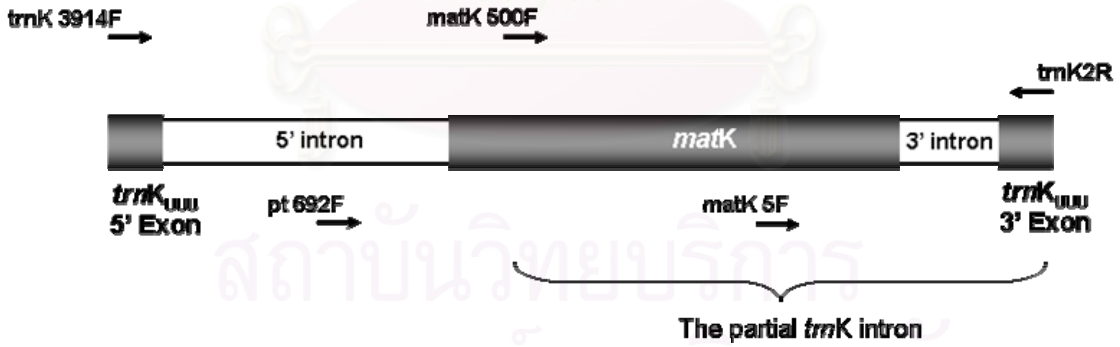


Figure 3.3 The relative position of amplified and sequencing primers used for the complete and partial *trn*K intron amplifications.

Table 3.6 Details of amplified and sequencing primers used for *trn*K gene.

Direction	Primer	Sequence (5' ⟶ 3')	Tm (˚C)	Primer type	Reference
Forward	trnK3914F	TGG GTT GCT AAC TCA ATG G	53.0	Amplified, sequencing	Johnson and Soltis, 1994
	matK500F	GAA AGA TCC CTC GTC TTT GCA T	55.6	Amplified, sequencing	In this study
	pt692F	GAC TGT ATC GCA CTA TGT ATC A	53.9	Sequencing	Sasaki *et al.*, 2002
	matK5F	GTA ACG TAT TGG GGC ATC C	55.2	Sequencing	Yang *et al*., 2004
Reverse	matK3R	CCT TCC CGG TTG AGA CCA C	59.5	Amplified, sequencing	Yang *et al*., 2004
	trnK2R	AAC TAG TCG GAT GGA GTA G	53.0	Amplified, sequencing	Johnson and Soltis, 1994

5. PCR Products Determination and Purification

The PCR products were determined using 1 % agarose gel electrophoresis. Sizes of the target fragments were estimated with 1 kb plus DNA ladder (Invitrogen Corp., U.S.A.). The remaining PCR products were purified with Montage PCR device (U.S.A.) as the manufacturer's protocol.

6. Nucleotide Sequencing Reaction

Purified PCR products (approximately 30 ng) were subjected to nucleotide sequencing reaction using Dye Deoxy Terminator cycle sequencing kit (Applied Biosystems, U.S.A.). The PCR products of the ITS region and *trn*L-F gene were analyzed for their nucleotide sequence with PCR amplification primers. Both strands of PCR products were sequenced. The PCR products of the complete and partial *trn*K intron were analyzed for their nucleotide sequences using the amplified primers. Sequencing primers pt692F and matK5F were also used to determine their nucleotide sequences. Sequencing reactions were carried out on a Takara PCR Thermal Cycler as hot start at 96 °C (2 min), followed by 30 cycles of 95 °C (30 sec), 50 °C (30 sec), 72 °C (1 min). Final extension was done at 72 °C for 5 min and then the reaction was held at 4 ºC.

7. Nucleotide Direct Sequencing Analysis

Initially, the obtained nucleotide sequencing products were removed of the dimer of primer using AutoSeq™ G-50 columns (Amersham Biosciences, UK.) according to the manufacturer's protocol. Then, they were separated on an automated DNA sequencer (ABI Prism 3100-Avant Genetic Analyzer Applied Biosystems, U.S.A.) with POP6 polymer (Applied Biosystems, U.S.A.). The obtained sequences were manually edited by AutoAssemble (version 1.3.0, Applied Biosystems, U.S.A.).

8. Nucleotide Sequence Alignment

For each target PCR fragment, both forward and reverse sequences were assembled using AutoAssemble (Applied Biosystems, U.S.A.) and BioEdit (Hall, 1999) programs. The nucleotide additive sites were confirmed by comparing their intensity ratio of forward and reverse sequencing electropherograms. The nucleotide additive sequences were assigned according to the IUPAC codes. Moreover, some specimens with the ambiguous electropherograms were re-amplified with high fidelity AccuPrime® *Taq* DNA polymerase (Invitrogen Corp., U.S.A.).

The boundaries of ITS1, 5.8S and ITS2 were determined by comparing with the published sequences of rice (Takaiwa *et al.*, 1985), tobacco (Kiss *et al.*, 1988) and mung bean (Schiebel and Hemleben, 1989). The boundaries of *trn*L-F intron and intergenic spacer were compared with the published sequences of tobacco (Yamada *et al*., 1983). For the complete *trn*K intron, the boundary of *trn*K intron and the open reading frame of *mat*K gene were defined by comparison with the sequence of tobacco (Sugita *et al*., 1985). All the obtained nucleotide sequences are deposited in GenBank.

9. Phylogenetic Tree Reconstruction

The obtained entire ITS and *trn*L-F sequences were used to manipulate data matrix, excluding nucleotide sequences that contained nucleotide additive sequences. Phylogenetic analysis was performed on the individual ITS and *trn*L-F sequences data and the combine sequence data of both. Three data matrixes were analyzed for phylogenetic relationships using PAUP* (Version 4.0 beta 10a, Sinauer Assoc. Inc., U.S.A.) program. Phylogenetic trees were reconstructed under the maximum parsimony (MP) criteria, character states were treated as unordered and optimal tree was estimated using a heuristic search strategy. Tree bisection reconnection (TBR) branch swapping was used with multiple parsimonious trees (MULPARS) and minimum branches of zero were collapsed. ACCTRAN option was used for character optimization. One hundred replications were searched. Any gaps were treated as the fifth base. The consensus tree was performed using 50% majority consensus method. Bootstrap analysis was performed with 1000 replicates. The consistency index (CI), rescaled consistency index (RC) and retention index (RI) were calculated for each analysis. *C. alabamensis* var. *alabamensis* (Crotoneae, Euphorbiaceae) was used as outgroup (GenBank accession number AY971177 and AY794692, respectively).

10. Quantitatively Analysis of Plaunotol Content in *C. stellatopilosus*

10.1 *C. stellatopilosus* Sample Preparation

The dried leaves of thirty *C. stellatopilosus* samples (codes STEL 01-30) were ground to fine powder in a grinder connected with a cool water circulator (IKA Labortechnik A10 Blender, Janke & Kunkel Gmbh & Co. Kg, Germany). After passing a sieve no. 40, 200 mg of the ground leaves were extracted with 10 ml 95 % ethanol (Analytical grade, Merck, Germany) under reflux at 80 °C for one hour in a 20x2.5 cm tube connected with a 15 cm condenser. Each ethanolic extract was cooled and used directly for the analysis of plaunotol content using the TLC densitometric method.

10.2 Plaunotol Standard Preparation

Plaunotol standard compound was purified from the light yellow-brown liquid form of Kelnac® (Sankyo Phrama, Japan). The purified plaunotol was dissolved in chloroform (Analytical grade, Merck, Germany) to prepare the standard solution in the concentration range of 25-175 µg/ml.

10.3 TLC Densitometric Analysis

The quantitative analysis of plaunotol content in the leaves of *C. stellatopilosus* was performed following the previous publication by Vongchareonsathit and De-Eknamkul (1998). Five microliter aliquot of each ethanolic extract obtained was spotted on a TLC plate by Linomat® semi-automatic spotter (Camag, Germany). The stationary phase was silica gel 60 F_{254} (Merck, Germany) and mobile phase was chloroform:n-propanol (Analytical grade, Merck, Germany) in the ratio 24:1. The TLC plate was scanned using the wavelength of 220 nm. Plaunotol in each *C. stellatopilosus* sample in the TLC plate was quantitatively analyzed according to the previously established TLC densitometric method and calculated based on its standard curve.

CHAPTER IV

RESULTS

1. Total DNA Preparations

Total DNA was extracted from the leaves of each sample using DNeasy® Plant Mini Kit (QIAGEN, Germany). All total DNA extracts prepared from fresh samples appeared as clear solutions whereas some total DNA extracts from herbarium samples were in brown color. After being purified with the DNA purification kit, the total DNA extracts were obtained as colorless solutions.

All samples of total DNA extracts were determined on 1 % agarose gel electrophoresis. Figure 4.1 shows the fresh sample total DNA extracts separated by eletrophoresis, the size and concentration of total DNA extracts are estimated by comparing with the 1kb plus molecular marker (Lane 1). All forty-eight fresh samples of total DNA extracts have their sizes longer than 12000 bp with high concentration causing the appearance of bright banding as showed in Lane 2-5. In contrast, those thirty-six herbarium samples showed their total DNA extracts on agarose gel as smear bands as shown in Fig. 4.2. Lane 2-4 and Lane 6 are the typical patterns of the herbarium samples of total DNA extracts separated by eletrophoresis. For some old herbarium samples, their total DNA extracts showed low DNA concentration which could not be detected on the gel as shown in Lane 5. The total DNA extracts were stored at -20 °C until used.

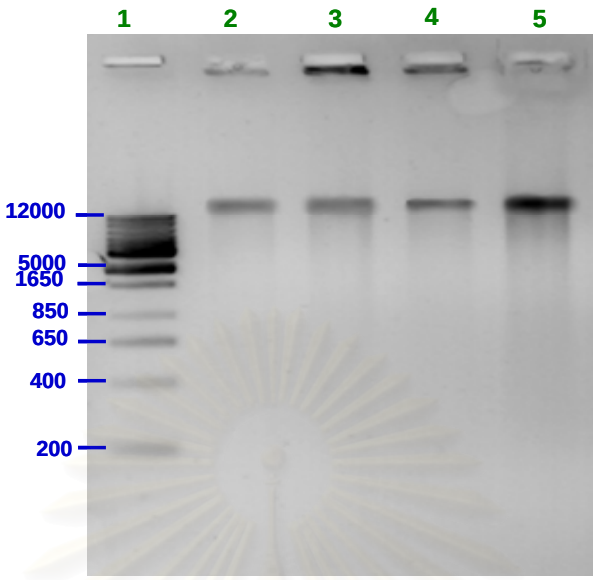


Figure 4.1 The fresh sample total DNA extracts separated by electrophoresis.

Lane 1: 1kb plus molecular marker,

Lane 2-5: typical patterns of fresh samples of total DNA extracts

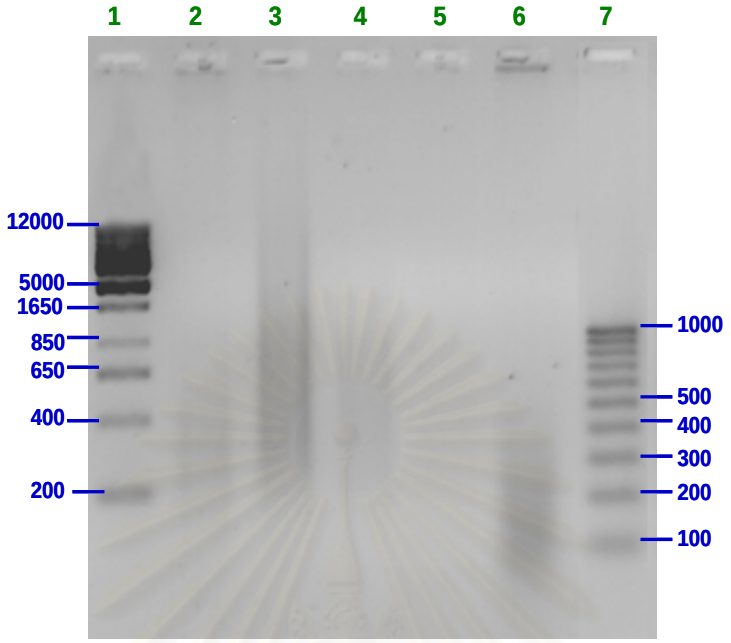


Figure 4.2 The herbarium sample total DNA extracts separated by electrophoresis.

Lane 1: 1kb plus molecular marker,

Lane 2-6: typical patterns of herbarium samples of total DNA extracts,

Lane 7: 100 bp molecular marker

2. PCR Amplification

Using the obtained total DNA extracts as PCR templates, eighty-four specimens representing twenty six *Croton* species were subjected to PCR amplification on the entire ITS region and *trn*L-F gene. In addition, thirty *C. stellatopilosus* samples were further examined for their nucleotide sequences on the complete *trn*K intron and five *C. roxburghii* samples were investigated for nucleotide sequences on the partial *trn*K intron. The results of PCR amplification in the entire ITS region, *trn*L-F and the *trn*K intron are listed in Tables 4.1 and 4.2.

2.1 The Entire ITS Region

Forty eight of total DNA extracts extracted from the fresh samples were PCR amplified in the entire ITS region using two universal primers, ITS-1 (Urbatsch *et al*., 2000) and ITS-4 (White *et al*., 1990). All fresh samples were successfully amplified (Table 4.1). The size of the obtained PCR products was about 700 bp long (Fig 4.3, Lane 2).

For the herbarium samples, most of their total DNA extracts showed low DNA concentration. Therefore, the entire ITS amplifications of their total DNA extracts were amplified into two overlapping fragments as the fragment ITS1 (Fig 4.3, Lane 3) and the fragment ITS2 (Fig 4.3, Lane 5). Both target fragments were similar size of about 350 bp long. Based on the standard PCR amplification method, herbarium sample codes ACUT, BON, CAS 02, CAS 03, DEC 01 and KAN 01 could be amplified on both of the fragment ITS1 and ITS2 (Table 4.2) whereas the other samples failed in this step. Thus, the remaining samples were further amplified by using the nested technique. The results (Table 4.2) showed that the herbarium sample codes ARG 01, DEC 02, DEC 03, ROB 01 and WAL 02 were able amplified on both the fragment ITS1 (Fig 4.3, Lane 4) and ITS2 (Fig 4.3, Lane 6). However, sample codes CAU 01, HIR 01, KER 01, LAN 01 and POI 03 could be amplified only the fragment ITS1. By using both standard and nested amplification, sixteen herbarium samples (30.56 %) could be amplified on both of the

fragment ITS1 and ITS2. However, five samples (13.89 %) could be amplified only the fragment ITS1. Moreover, it was found that most of herbarium samples that could be amplified had their age below 30 years. The samples above this age could not be amplified.

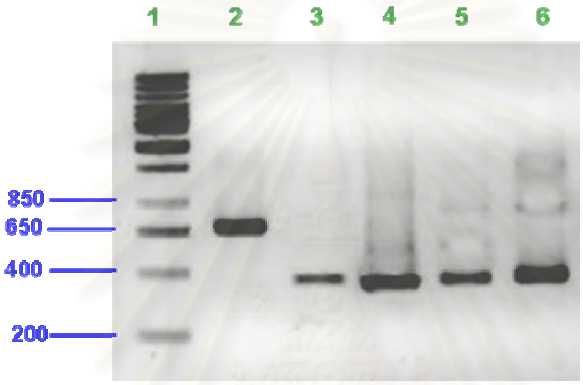


Figure 4.3 The PCR products of the ITS region separated by electrophoresis.

Lane 1: 1kb plus molecular marker,

Lane 2: the entire ITS PCR product,

Lane 3: the ITS1 PCR products with the standard amplification,

Lane 4: the ITS1 PCR products with nested amplification,

Lane 5: the ITS2 PCR products with the standard amplification,

Lane 6: the ITS2 PCR products with nested amplification

2.2 *trn*L-F gene

Amplification of *trn*L-F gene was performed using the universal primers of Taberlet *et al*. (1991). Firstly, the fresh samples of total DNA extracts were subjected to amplification on the entire *trn*L-F gene which include *trn*L intron and *trn*L-F intergenic spacer by using the universal primers *c* and *f* (Taberlet *et al*., 1991). But this was not successful (Table 4.1). Thus, they were further PCR amplified by separating the entire *trn*L-F gene into two overlapping fragments namely the *trn*L intron (Fig. 4.4, Lane 2) and the *trn*L-F intergenic spacer (Fig. 4.4, Lane 3). The primer pair *c* and *d* was used to amplify the *trn*L intron while the primer pair *e* and *f* was used to amplify the *trn*L-F intergenic spacer. Based on this approach, all fresh sample total DNA extracts could be amplified on both target fragments (Table 4.1). The sizes of *trn*L intron and *trn*L-F intergenic spacer fragments were about 650 and 550 bp in length, respectively. For the herbarium samples, they were also amplified on the *trn*L-F gene with two overlapping fragments, the *trn*L intron and the *trn*L-F intergenic spacer using the standard PCR amplification. The results (Table 4.2) showed that only five samples (13.89 %) of BON, CAS 02, CAS 03, DEC 01 and KAN 01 could be amplified on both *trn*L intron and *trn*L-F intergenic spacer target fragments.

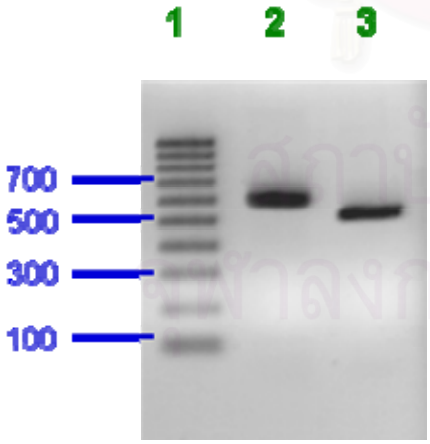


Figure 4.4 The PCR products of the *trn*L-F gene separated by electrophoresis

Lane 1: 100 bp molecular marker,

Lane 2: *trn*L intron PCR product,

Lane 3: *trn*L-F intergenic spacer

2.3 *trn*K intron

Thirty *C. stellatopilosus* samples were PCR amplified and examined for their nucleotide sequences on the complete *trn*K intron. Firstly, the primer pair tk3914F (Johnson and Soltis, 1994) and trnk2R (Johnson and Soltis, 1994) designed from the conserve coding region of *trn*K 3' exon and *trn*K 5' exon was used. The expected PCR product was about 2500 bp long. All *C. stellatopilosus* samples could not be amplified on this target fragment (Table 4.1). Then, they were amplified for the complete *trn*K intron as two separate but overlapping fragments designated as fragments I and II. The primer pair trnK3914F (Johnson and Soltis, 1994) and matK3R (Yang *et al*., 2004) was used to amplify fragment I, (Fig. 4.5, Lane 2) while the primer pair matK500F and trnK2R (Johnson and Soltis, 1994) was used to amplify fragment II (Fig. 4.5, Lane 3). With this approach, all *C. stellatopilosus* samples were successfully amplified on the target regions (Table 4.1). The sizes of both fragments I and II were about 1300 bp long. For *C. roxburghii* species, five samples were analyzed for their nucleotide sequences on the partial *trn*K intron. The PCR amplification was performed using the primer tk3914F (Johnson and Soltis, 1994) and primer matK3R (Yang *et al*., 2004). All *C. stellatopilosus* samples were successfully DNA amplified (Table 4.1). Their sizes of the partial *trn*K intron (Fig. 4.5, Lane 3) was about 1300 bp long.

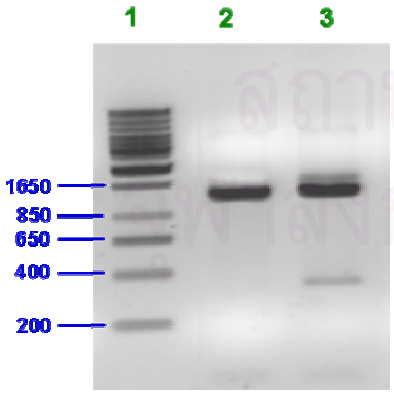


Figure 4.5 The PCR products of the *trn*K intron separated by electrophoresis

Lane 1: 1kb plus molecular marker,

Lane 2: fragment I of the *trn*K intron amplification,

Lane 3: fragment II of the *trn*K intron amplification

Table 4.1 The ITS region, *trn*L-F gene and *trn*K intron amplifications from the fresh samples.

Species	Sample code	ITS region			*trn*L-F		*trn*K intron	
		Entire ITS	ITS1 + ITS2 fragments	Nested technique	Entire gene	Intron + IGS fragments	Entire gene	Fragments I and II
Croton argyratus Blume	ARG 01	+	NA	NA	-	+	NA	NA
C. cascarilloides Raeusch.	CAS 01	+	NA	NA	-	+	NA	NA
C. columnaris Airy Shaw	COL	+	NA	NA	-	+	NA	NA
C. crassifolius Geiseler	CRA	+	NA	NA	-	+	NA	NA
C. kongensis Gagnep.	KON 01	+	NA	NA	-	+	NA	NA
	KON 02	+	NA	NA	-	+	NA	NA
C. poilanei Gagnep.	POI 01	+	NA	NA	-	+	NA	NA
	POI 02	+	NA	NA	-	+	NA	NA
C. roxburghii N.P. Baiakr.	ROX 01	+	NA	NA	-	+	-	+
	ROX 02	+	NA	NA	-	+	-	+
	ROX 03	+	NA	NA	-	+	-	+
	ROX 04	+	NA	NA	-	+	-	+
	ROX 05	+	NA	NA	-	+	-	+
C. stellatopilosus Ohba	STEL 01	+	NA	NA	-	+	-	+
	STEL 02	+	NA	NA	-	+	-	+
	STEL 03	+	NA	NA	-	+	-	+
	STEL 04	+	NA	NA	-	+	-	+
	STEL 05	+	NA	NA	-	+	-	+
	STEL 06	+	NA	NA	-	+	-	+
	STEL 07	+	NA	NA	-	+	-	+
	STEL 08	+	NA	NA	-	+	-	+
	STEL 09	+	NA	NA	-	+	-	+
	STEL 10	+	NA	NA	-	+	-	+
	STEL 11	+	NA	NA	-	+	-	+
	STEL 12	+	NA	NA	-	+	-	+
	STEL 13	+	NA	NA	-	+	-	+
	STEL 14	+	NA	NA	-	+	-	+

Table 4.1 The ITS region, *trn*L-F gene and the *trn*K intron amplifications from the fresh samples (continued).

Species	Sample code	ITS region			*trn*L-F		*trn*K intron	
		Entire ITS	ITS1 + ITS2 fragments	Nested technique	Entire gene	Intron + IGS fragments	Entire gene	Fragments I and II
Croton stellatopilosus Ohba	STEL 15	+	NA	NA	-	+	-	+
	STEL 16	+	NA	NA	-	+	-	+
	STEL 17	+	NA	NA	-	+	-	+
	STEL 18	+	NA	NA	-	+	-	+
	STEL 19	+	NA	NA	-	+	-	+
	STEL 20	+	NA	NA	-	+	-	+
	STEL 21	+	NA	NA	-	+	-	+
	STEL 22	+	NA	NA	-	+	-	+
	STEL 23	+	NA	NA	-	+	-	+
	STEL 24	+	NA	NA	-	+	-	+
	STEL 25	+	NA	NA	-	+	-	+
	STEL 26	+	NA	NA	-	+	-	+
	STEL 27	+	NA	NA	-	+	-	+
	STEL 28	+	NA	NA	-	+	-	+
	STEL 29	+	NA	NA	-	+	-	+
	STEL 30	+	NA	NA	-	+	-	+
C. tiglium L.	TIG 01	+	NA	NA	-	+	NA	NA
	TIG 02	+	NA	NA	-	+	NA	NA
	TIG 03	+	NA	NA	-	+	NA	NA
	TIG 04	+	NA	NA	-	+	NA	NA
C. wallichii Mull. Arg.	WAL 01	+	NA	NA	-	+	NA	NA
Total		**48**	**NA**	**NA**	**0**	**48**	**0**	**30**

NA: not amplified

Table 4.2 The ITS region, *trn*L-F gene and *trn*K intron amplifications from the herbarium samples.

Species	Sample code	ITS region			*trn*L-F		*trn*K intron	
		Entire ITS	ITS1 + ITS2 fragments	Nested technique	Entire gene	Intron + IGS fragments	Entire gene	Fragments I and II
Croton acutifolius Esser	ACUT	NA	+	-	NA	-	NA	NA
C. argyratus Blume*	ARG 02	NA	-	+	NA	-	NA	NA
C. bonplandianus Baill.	BON	NA	+	-	NA	+	NA	NA
C. cascarilloides Raeusch.*	CAS 02	NA	+	-	NA	+	NA	NA
	CAS 03	NA	+	-	NA	+	NA	NA
C. caudatus Geiseler	CAU 01	NA	-	-	NA	-	NA	NA
	CAU 02	NA	-	-	NA	-	NA	NA
	CAU 03	NA	-	-	NA	-	NA	NA
C. decalvatus Esser	DEC 01	NA	+	-	NA	+	NA	NA
	DEC 02	NA	-	+	NA	-	NA	NA
	DEC 03	NA	-	+	NA	-	NA	NA
C. delpyi Gagnep.	DEL 01	NA	-	-	NA	-	NA	NA
C. griffithii Hook. F.	GRIF 01	NA	-	-	NA	-	NA	NA
	GRIF 02	NA	-	-	NA	-	NA	NA
C. hirtus L' Her.	HIR 01	NA	-	-	NA	-	NA	NA
C. hutchinsonianus Hosseus in Fedde	HUT 01	NA	-	-	NA	-	NA	NA
	HUT 02	NA	-	-	NA	-	NA	NA
C. kongkandanus Esser	KAN 01	NA	+	-	NA	+	NA	NA
	KAN 02	NA	-	-	NA	-	NA	NA
C. kerrii Airy Shaw	KER 01	NA	-	-	NA	-	NA	NA
	KER 02	NA	-	-	NA	-	NA	NA
C. krabas Gagnep.	KRA 01	NA	-	-	NA	-	NA	NA
	KRA 02	NA	-	-	NA	-	NA	NA
	KRA 03	NA	-	-	NA	-	NA	NA

Table 4.2 The ITS region, *trn*L-F gene and *trn*K intron amplifications from the herbarium samples (continued).

Species	Sample code	ITS region			*trn*L-F		*trn*K intron	
		Entire ITS	ITS1 + ITS2 fragments	Nested technique	Entire gene	Intron + IGS fragments	Entire gene	Fragments I and II
Croton lachnocarpus Benth.	LAN 01	NA	-	-	NA	-	NA	NA
C. poilanei Gagnep.*	POI 03	NA	-	-	NA	-	NA	NA
	POI 04	NA	-	-	NA	-	NA	NA
C. poomae Esser	POO 01	NA	-	-	NA	-	NA	NA
C. robustus Kurz	ROB 01	NA		+	NA	-	NA	NA
	ROB 02	NA	-	-	NA	-	NA	NA
	ROB 03	NA	-	-	NA	-	NA	NA
C. sepalinus Airy Shaw	SEP 01	NA	-	-	NA	-	NA	NA
	SEP 02	NA	-	-	NA	-	NA	NA
C. thorelii Gagnep.	THO 01	NA	-	-	NA	-	NA	NA
	THO 02	NA	-	-	NA	-	NA	NA
C. wallichii Mull. Arg.*	WAL 02	NA	-	+	NA	-	NA	NA
Total		**NA**	**6**	**5**	**NA**	**5**	**NA**	**NA**

NA: not amplified

3. Genetic Diversity of *Croton* Species

3.1 The Entire ITS Nucleotide Sequences

Among the overall eighty four *Croton* samples, fifty-five samples representing sixteen *Croton* species could be amplified and determined for their nucleotide sequences on ITS region. The obtained entire ITS nucleotide sequences had the length between 620 (*C. columnaris*) and 627 bp (*C. crassifolius* and *C. tiglium*), similar with previous reports (Berry *et al*., 2005 and Xue, *et al*., 2007). The GC contents were in the range of 52.50-57.10 %. Details of the entire ITS nucleotide sequences of *Croton* species are listed in Table 4.3.

The ITS1 regions of the samples were in the range 264 to 269 bp long (Table 4.3). At the 3' terminal of ITS1 region (nucleotide position 256-276, Appendix A), all of *Croton* species presented the single nucleotide repeated unit as the poly T and the poly A. This single nucleotide repeating unit was 4-9 repeats for the poly T and 5-12 repeats for the poly A. However, the number of the poly T and the poly A repeats appeared to be similar among the same *Croton* species.

For 5.8S rDNA coding regions, all *Croton* samples appeared to have similar nucleotide sequences with the same length of 162 bp (Table 4.3). Only *C. bonplanianus* and *C. columnaris* showed single nucleotide substitution (Appendix A) at the nucleotide positions 375(C) and 409(A), respectively. In addition, most *Croton* species showed nucleotide substitutions at nucleotide positions 402, 404 and 405 (Appendix A).

Table 4.3 Details of the ITS nucleotide sequences of *Croton* species.

Croton Species	Sample code	Length (bp)				GC %	No. of nucleotide additive sites	GenBank accession number
		ITS1	5.8S	ITS2	Entire region			
acutifolius	ACUT	269	162	194	625	55.20	6	AB375078
argyratus	ARG 01, 02	264	162	200	626	55.91	1	AB375093
bonplandianus	BON	265	162	194	621	52.50	0	AB375094
cascarilloides	CAS 01, 02	265	162	194	621	54.91	6	AB375087
	CAS 03	265	162	194	621	55.72	1	AB375088
columnaris	COL	264	162	194	620	56.29	0	AB375080
crassifolius	CRASS	265	162	200	627	57.10	0	AB375089
decalvatus	DEC 01, 03	266	162	194	622	55.95	0	AB375075
	DEC 02	266	162	194	622	55.95	1	AB375076
kongkandanus	KAN 01	268	162	194	624	56.41	1	AB375084
kongensis	KON 01	265	162	196	623	55.06	1	AB375090
	KON 02	265	162	196	623	54.90	2	AB375091
poilanei	POI 01	269	162	194	625	56.32	1	AB375086
	POI 02	269	162	194	625	56.48	0	AB375085
robustus	ROB 01	266	162	194	622	56.11	1	AB375077
roxburghii	ROX 01	268	162	194	624	55.45	0	AB428632
	ROX 02	268	162	194	624	55.77	4	AB428633
	ROX 03	268	162	194	624	56.57	3	AB375082
	ROX 04	268	162	194	624	56.41	5	AB375081
	ROX 05	269	162	194	625	56.16	3	AB375083
stellotopilosus	STEL 01	268	162	194	624	54.65	15	AB375070
	STEL 02	267	162	194	623	55.22	12	AB375072
	STEL 03, 15, 26-30	267	162	194	623	55.86	0	AB375068
	STEL 04, 08, 09,10, 13, 16	267	162	194	623	56.34	0	AB375069
	STEL 05-07	268	162	194	624	54.97	15	AB375073
	STEL 11	267	162	194	623	54.74	15	AB375071
	STEL 12	267	162	194	623	55.86	2	AB375074
	STEL 14	267	162	194	623	55.54	7	AB428634
	STEL 17	267	162	194	623	55.54	7	AB428635
	STEL 18	267	162	194	623	55.22	11	AB428636
	STEL 19	267	162	194	623	55.54	7	AB428637
	STEL 20	267	162	194	623	55.54	4	AB428638
	STEL 21	267	162	194	623	54.74	14	AB428639
	STEL 22	267	162	194	623	55.22	12	AB428640
	STEL 23	267	162	194	623	55.06	11	AB428641
	STEL 24	267	162	194	623	55.06	12	AB428642
	STEL 25	267	162	194	623	55.06	13	AB428643
tiglium	TIG 01-04	264	162	201	627	56.37	0	AB375092
wallichii	WAL 01-02	266	162	194	622	56.59	1	AB375079

The ITS2 regions of *Croton* species were found to have their range of 194 -200 bp in length (Table 4.3). At this region, *Croton* samples showed only the insertions. *C. kongensis* and *C. tiglium* were 2 bp insertion at the nucleotide position 578-579. *C. argyratus*, *C. crassifolius* and *C. tiglium* were 6 bp insertion at the nucleotide position 614-619. Details of insertion are shown in Table 4.4 and their relative position is illustrated in Fig. 4.6.

Table 4.4 The nucleotide insertions at ITS2 region of *Croton* species.

Nucleotide position	**Insertion length (bp)**	**Sequence**	***Croton* Species**
578-579	2	AC	*C. kongensis* and *C. tiglium*
614-619	6	GCGTCC	*C. crassifolius* and *C. tiglium*
		ACGTCC	*C. argyratus*

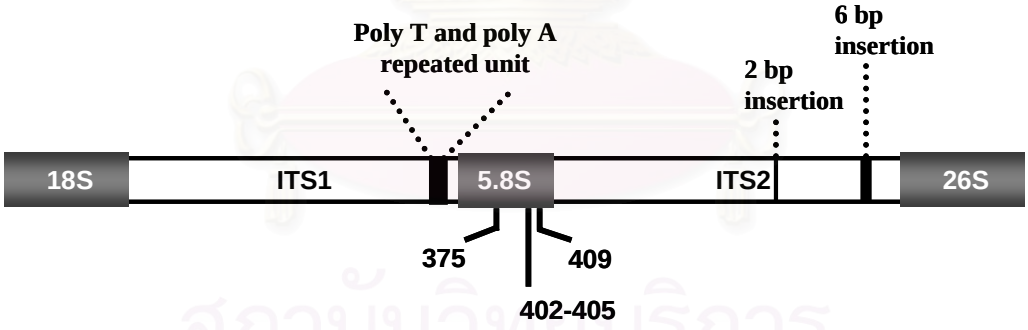


Figure 4.6 The relative position of the poly A and poly T repeated units at ITS1 region, nucleotide substitutions at 5.8S rDNA coding region and the insertion sites at ITS2 region.

The obtained ITS sequences were all aligned as shown in Appendix A. The results revealed that genetic variations of ITS region were mostly caused from nucleotide substitutions. In addition, based on the direct sequencing technique, the electrophoregrams of some samples appeared to have double signaling peaks at one nucleotide position (Fig 4.7). From a single PCR product sequencing, the obtained electrophoregrams showed similarity in both forward and reverse sequencing. This phenomenon has been called nucleotide additive sites (Sang *et al.*, 1995). The nucleotide additive sites probably indicated that more than one ribotype differing at these sites was involved within a single individual (Baldwin *et al.*, 1995; Sang *et al.*, 1995 and Campbell *et al.*, 1997). Figure 4.7 shows the electropherograms of the forward and reverse sequencing of the sample KON 02. It can be seen that each electropherogram had one nucleotide additive site that showed the double signaling peaks. The blue signaling peaks represent C and those with red represent T. The electropherograms indicated that the considerable nucleotide positions of the sample KON 02 contained two types of nucleotide sequences as C and T. These nucleotide additive positions were assigned following the IUPAC code as Y.

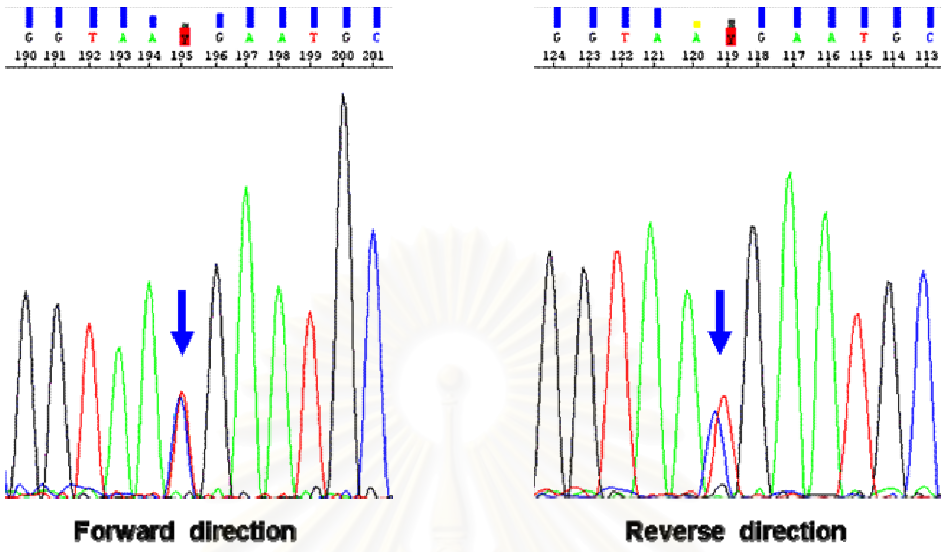


Figure 4.7 The electrophoregrams presenting the nucleotide additive sites. (A) and (B) were electrophoregrams in forward and reverse sequencing directions, respectively. The arrows indicate the nucleotide additive sites.

The obtained entire ITS sequences (Appendix A) indicated that *C. bonplanianus*, *C. columnaris*, *C. crassifolius* and *C. tiglium* appeared to have no nucleotide additive sites. Whereas eleven *Croton* species, including *C. acutifolius*, *C. argyratus*, *C. cascarilloides*, *C. decalvatus*, *C. kongkandanus*, *C. kongensis*, *C. poilanei*, *C. robustus*, *C. roxburghii*, *C. stellatopilosus* and *C. wallichii* possessed the nucleotide additive sites. The number of the nucleotide additives was in the range of 1 to 16 sites (Table 4.3). *C. stellatopilosus* and *C. roxburghii* were found to contain several nucleotide additive sites.

3.2 *trn*L-F Nucleotide Sequences

The *trn*L introns and *trn*L-F intergenic spacers could be amplified from fifty-five samples of thirteen *Croton* species. The obtained *trn*L-F nucleotide sequences had their length from 1017 (*C. cascarilloides*) to1041 bp (*C. argyratus*). The *trn*L introns were from 527 to 550 bp while the intergenic spacers (IGS) were from 411 to 428 bp long. The GC contents of *trn*L-F gene were 31.03-31.86 %. Details of the *trn*L-F nucleotide sequences of *Croton* species are summarized in Table 4.5. In addition, *trn*L-F nucleotide sequences of *Croton* were found to have single nucleotide repeat unit as poly T (nucleotide position 925-939) as shown in Appendix B. The number of poly T repeat units was from 11 to 15 units. For the insertions and deletions (indels), *trn*L introns were 5 insertion sites. The intergenic spacers had one insertion and one deletion. Details of indels and their relative positions are shown in Table 4.6 and Fig. 4.8. The obtained *trn*L-F sequences were all aligned as shown in Appendix D. It was found that genetic variations of the *trn*L-F genes were caused from mainly indels.

Table 4.5 Details of *trn*L-F nucleotide sequence of various *Croton* species.

***Croton* species**	**Sample code**	**Length (bp)**			**GC %**	**Accession number**
		***trn*L intron**	**IGS ***	***trn*L-F gene**		
argyratus	ARG 01	544	424	1041	31.03	AB375095
bonplndianus	BON	550	411	1034	31.43	AB375109
cascarilloides	CAS 01, 03	527	417	1017	31.86	AB375096
	CAS 02	527	417	1017	31.76	AB375097
columnaris	COL	534	423	1030	31.46	AB375099
crassifolius	CRASS	527	420	1020	31.47	AB375100
decalvatus	DEC 01	535	424	1032	31.59	AB375110
kongensis	KON 01, 02	527	422	1022	31.70	AB375101
kongkandanus	KAN 01	528	426	1037	31.34	AB375111
poilanei	POI 01, 02	528	426	1037	31.34	AB375102
roxbughii	ROX 01, 02	534	426	1033	31.46	AB375103
	ROX 03	534	424	1031	31.52	AB375103
	ROX 04	534	423	1030	31.65	AB375104
	ROX 05	534	428	1035	31.11	AB375105
stellatopilosus	STEL 01-30	535	422	1030	31.65	AB375106
tiglium	TIG 01-04	528	422	1023	31.77	AB375107
wallichii	WAL 01	533	425	1031	31.52	AB375108

* IGS: *trn*L-F intergenic spacer

Table 4.6 Indels in the *trn*L-F nucleotide sequences of various *Croton* species.

Region	Indels	Length (bp)	Nucleotide position	Sequence	*Croton* Species
Intron	Insertion (A)*	7	110-116	CCAAAAA	*C. bonplanianus*
	Insertion (B)	6	127-132	CAAATA	*C. columnaris*, *C. roxburghii*, *C. stellatopilosus*, *C. wallichii*
	Insertion (C)*	8	328-335	TTAATTAA	*C. decalvatus*
	Insertion (D)*	14	352-365	ATTTTTTTATCTTT	*C. argyratus*
	Insertion (E)*	16	352-367	ATTTTTTTATCTTTCT	*C. bonplanianus*
IGS	Insertion (F)*	6	826-831	CAAAAA	*C. roxburghii* (sample ROX 05)
	Deletion (G)*	6	1023-1028	GACATA	*C. cascarilloides*

- Nucleotide positions start since the first position of the *trn*L-F gene.
- * The letters indicate the insertions and deletion of *trn*L-F sequence that are shown in Fig. 4.4.

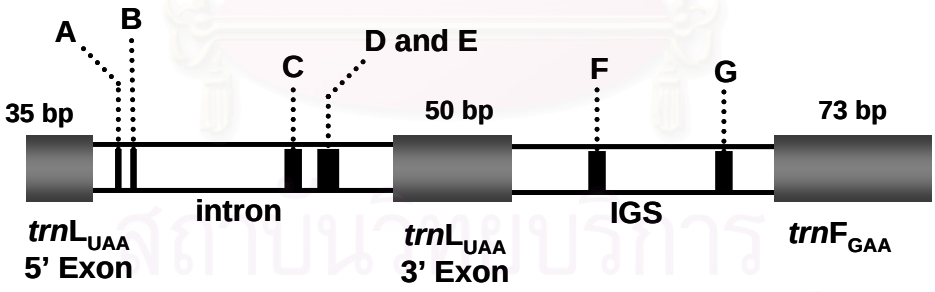


Figure 4.8 The indels relative position of *trn*L-F sequences. The letters A, B, C, D and E represent insertions of *trn*L intron. The letters F and G represent insertion and deletion of the intergenic spacer, respectively.

4. The Maximum Parsimony Trees of *Croton* Species

The obtained sequences of both entire ITS region and *trn*L-F gene were used to manipulate the data matrixes. The nucleotide sequences that contained the additive sites were excluded because of the limited performance of the computer simulation. Those ITS, *trn*L-F and the combined sequence data matrixes were analyzed for the phylogenetic relationships by using PAUP* (Version 4.0 beta 10a, Sinauer Assoc. Inc., U.S.A.) program. All of possible phylogenetic trees were consensus with 50 % majority tree method. Bootstrapping analysis was performed with 1000 replications. The entire ITS and *trn*L-F sequences of *C. alabamensis* var. *alabamensis* (Crotoneae, Euphorbiaceae) was used as the outgroup. Data set and parsimony based tree characteristics of the ITS, *trn*L-F and combine data analyses are summarized in Table 4.7. The topology and bootstrapping analysis of the obtained maximum parsimony trees are described below.

4.1 The Maximum Parsimony Tree Based on The Entire ITS Sequence Data

The entire ITS sequences of nineteen taxa representing fifteen *Croton* species were used to produce the data matrix. The obtained ITS data matrix was 639 in length, 137 variable characters and 76 informative characters. Based on these data matrix, maximum parsimony method could be simulated for possible parsimony trees as forty-one trees. All possible trees were consensused with the 50 % majority rule and gave a consensus parsimony tree as shown in Fig. 4.9. It could be concluded that the phylogenetic relationship of the fifteen *Croton* species was the monophyletic. The consistency index, retention index and rescaled consistency index of the consensuses tree were 0.8144, 0.6417 and 0.5226, respectively. Based on a consensus ITS parsimony tree, all of the *Croton* taxa could be divided into two major clades (major clade I and II). Only the *cascarilloides* 01 and *bonplandianus* taxa were splited as the individual branches (at the node C) with 52 and 100 bootstrapping percentages, respectively. The details of each major clade can be described as follows:

The major clade I: At the node A, the sister group of *kongkandanus* 01 and *poilanei* 01 taxa (94 bootstrapping percentage), and eleven taxa of *columnaris*, *decalvatus* 01, *stellatopilosus* Type A, *acutifolius*, *robustus*, *roxburghii* 01, *roxburghii* 02, *roxburghii* 03, *roxburghii* 04, *roxburghii* 05 and *wallichii* 01 were assigned in the same group (67 bootstrapping percentage). At the node B, the *stellatopilosus* Type B taxon was split as a separate branch from those thirteen taxa with moderate bootstrapping support (86). Focusing the parsimony assignment among *C. roxburghii* taxa, at the node A, the *roxburghii* 01 and *roxburghii* 02 taxa were located at the same level as the other taxa of this major clade. But the *roxburghii* 03, *roxburghii* 04 and *roxburghii* 05 taxa were cluster as a sister group with the high strong bootstrapping support (94). From the node A, the branch lengths of the *roxburghii* 01, *roxburghii* 02, *roxburghii* 03, *roxburghii* 04 and *roxburghii* 05 taxa were 6, 1, 7, 7 and 8, respectively.

The major clade II: The *crassifolius* and *tiglium* 01-04 taxa were assigned as a sister group with moderate bootstrapping supported (82). The *argyratus* and *kongensis* 01 taxa were respectively, joined with this sister group with 82 and 73 bootstrapping percentages.

.

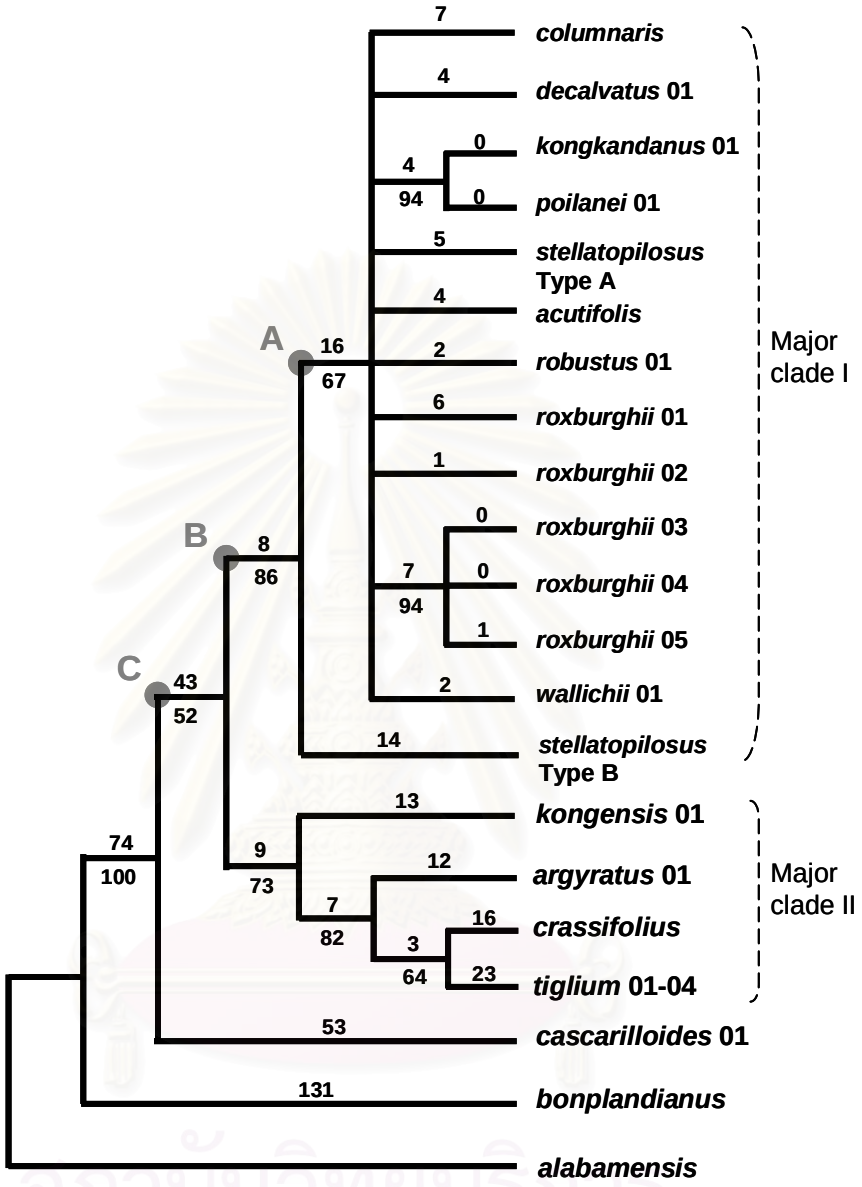


Figure 4.9 The 50 % majority consensus tree resulting from parsimony analysis of the entire ITS sequence data. The numbers under the branches are bootstrapping percentages ≥50 % and the numbers above the branches are branch lengths. *Croton alabamensis* is assigned as the outgroup. The Thai name of the *Croton* samples are as follow; *columnaris* (เปล้าคำ), *decalvatus*, *kongkandanus* (มิลา), *poilanei* (เปล้าเลือด), *stellatopilosus* (เปล้าน้อย), *acutifolius* (เปล้าแพะ), *robustus* (เปล้าใหญ่), *roxburghii* (เปล้าใหญ่), *wallichii* (เปล้านา), *argyratus* (เปล้าเงิน), *crassifolius* (พังคี), *tiglium* (สลอด), *kongensis* (เปล้าน้ำเงิน), *cascarilloides* (เปล้าเล็ก) and *bonplandianus* (เปล้าทุ่ง).

4.2 The Maximum Parsimony Tree Based on The *trn*L-F Sequence Data

The obtained *trn*L-F nucleotide sequences of sixteen taxa were used to produce the data matrix. This data matrix was representative of the thirteen *Croton* species. The *trn*L-F sequence data matrix was 1063 in length. It had 108 variable characters and 58 informative characters. The number of possible maximum parsimony trees was eight. A 50 % majority consensus parsimony tree (Fig. 4.10) showed that the phylogenetic relationship of this thirteen *Croton* species was monophyletic. The consistency index, retention index and rescaled consistency index of the consensus tree were 0.8995, 0.8601 and 0.7737, respectively. Based on the obtained consensus tree (Fig. 4.10), the fifteen taxa were clustered as a major clade. While the *crassifolius* taxon was split as an individual branch with moderate bootstrapping support (72). The major clade consisted of the clade I, the clade II, the sister group A, the sister group B and a branch of *cascarilloides* 01-03 taxon. The details of clades and sister groups of the major clade are described below.

The clade I: The *columnaris*, *roxburghii* 04 and *roxburghii* 05 taxa were assigned as a sister group with moderate bootstrapping support (64). At the node A, the *roxburghii* 01-02, *roxburghii* 03, *stellatopilosus* 01-30 and *wallichii* 01 taxa were joined with this sister group with high bootstrapping support (100).

The clade II: The *kongkandanus* and *poilanei* 01, 02 taxa were assigned as a sister group with high bootstrapping support (99). At the node B, the *decalvatus* 01 taxon was joined with this sister group with high bootstrapping support (97).

The sister groups A: The *kongensis* 01 and *tiglium* 01-04 taxa were assigned as a sister group with the moderate bootstrapping supported (61).

The sister groups B: The *argyratus* 01 and *bonplandianus* taxa were assigned as a sister group with high bootstrapping support (95).

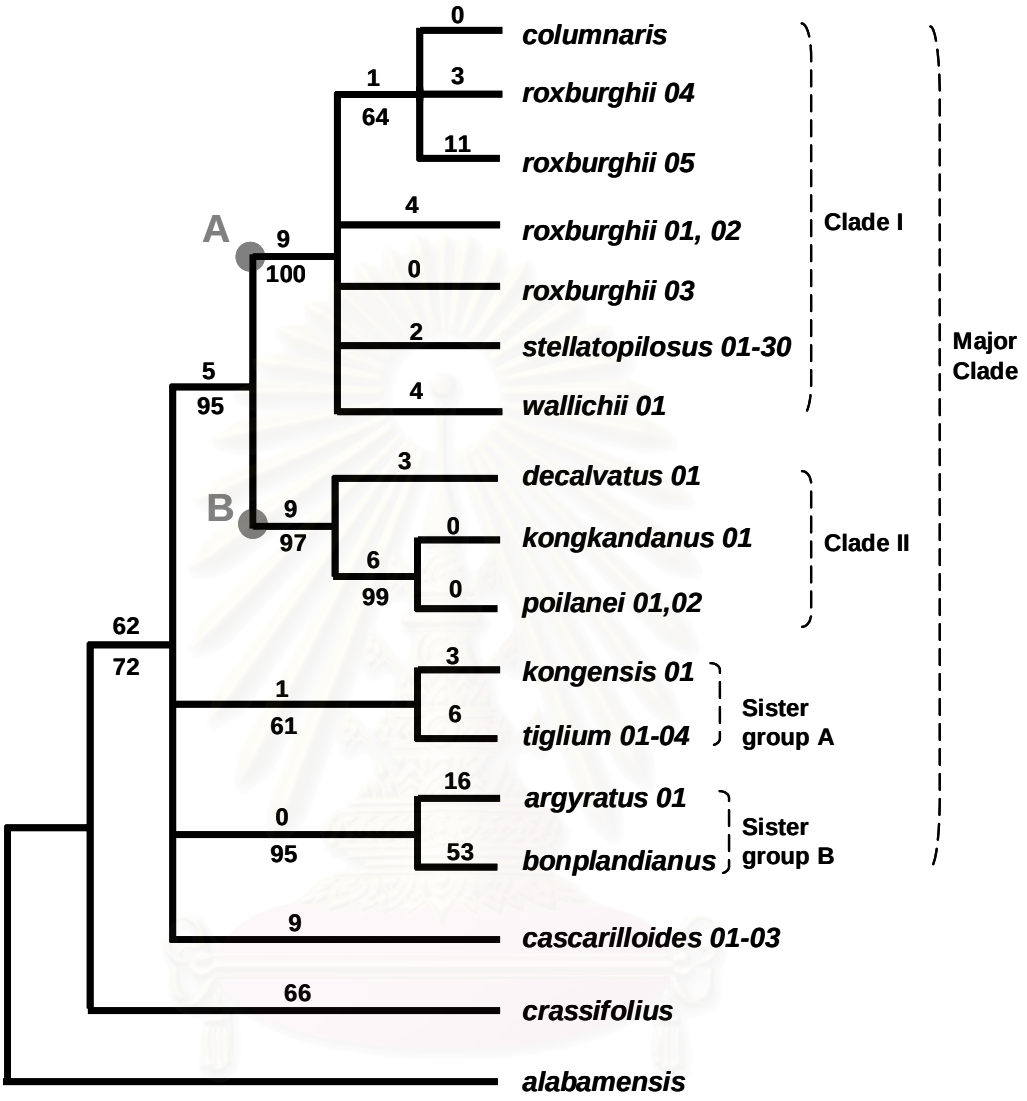


Figure 4.10 The 50 % majority consensus tree resulting from parsimony analysis of the *trn*L-F sequence data. The numbers under the branches are bootstrapping percentages ≥50 % and the numbers above the branches are branch lengths. *Croton alabamensis* is assigned as the outgroup. The Thai name of the *Croton* samples are as follow; *columnaris* (เปล้าคำ), *stellatopilosus* (เปล้าน้อย), *roxburghii* (เปล้าใหญ่), *wallichii* (เปล้านา), *decalvatus*, *kongkandanus* (มิลา), *poilanei* (เปล้าเลือด), *kongensis* (เปล้าน้ำเงิน), *tiglium* (สลอด), *argyratus* (เปล้าเงิน), *bonplandianus* (เปล้าทุ่ง), *cascarilloides* (เปล้าเล็ก) and *crassifolius* (พังคี).

4.3 The Maximum Parsimony Tree Based on The Combined Data

In this approach the entire ITS and *trn*L-F nucleotide sequences were combined. The data from nineteen combined taxa were used to produce the data matrix. This data represented twelve *Croton* species. The combined data matrix was 1714 in length. It contained 259 variable characters and 128 informative characters. The number of possible maximum parsimony trees was three. A 50 % majority consensuses tree (Fig. 4.11) showed monophyletic relationship among these examined *Croton* species. The consistency index, retention index and rescaled consistency index of the consensus tree was 0.8342, 0.7143 and 0.5959, respectively. An obtained consensus parsimony tree showed that almost the examined taxa fell into three major clades (major clade I, II and III). Only the *cascarilloides* 01 and *bonplandianus* taxa were split as individual branches with 51 and 96 bootstrapping percentages, respectively. The details of each major clade are described below.

The major clade I: At the node A, three sister groups (A, B and C) and two individual branches of the *columnaris* and *wallichii* 01 taxa were formed as the major clade I with 95 bootstrapping percentage. At the node B, the *stellatopilosus* Type A and the *stellatopilosus* Type B taxa were clustered as the sister group A with moderate bootstrapping support (68). At the node C, the *roxburghii* 01 and *roxburghii* 02 were assigned as the sister group B with moderate bootstrapping support (71). At the node D, the *roxburghii* 03, *roxburghii* 04 and *roxburghii* 05 were formed the sister group C with strong bootstrapping support (98).

The major clade II: At the node E, the *kongkandanus* and *poilanei* 01 taxa were assigned as the sister group with strong bootstrapping support (100). The *decalvatus* 01 taxon was joined with this sister group at the node F, with 100 bootstrapping percentage.

The major clade III: At the node G, the *crassifolius*, *tiglium* 01-04 and *argyratus* 01 were formed a sister group with 76 bootstrapping percentage. At the node H, the *kongensis* 01 taxon was joined with this sister group with moderate bootstrapping support (77).

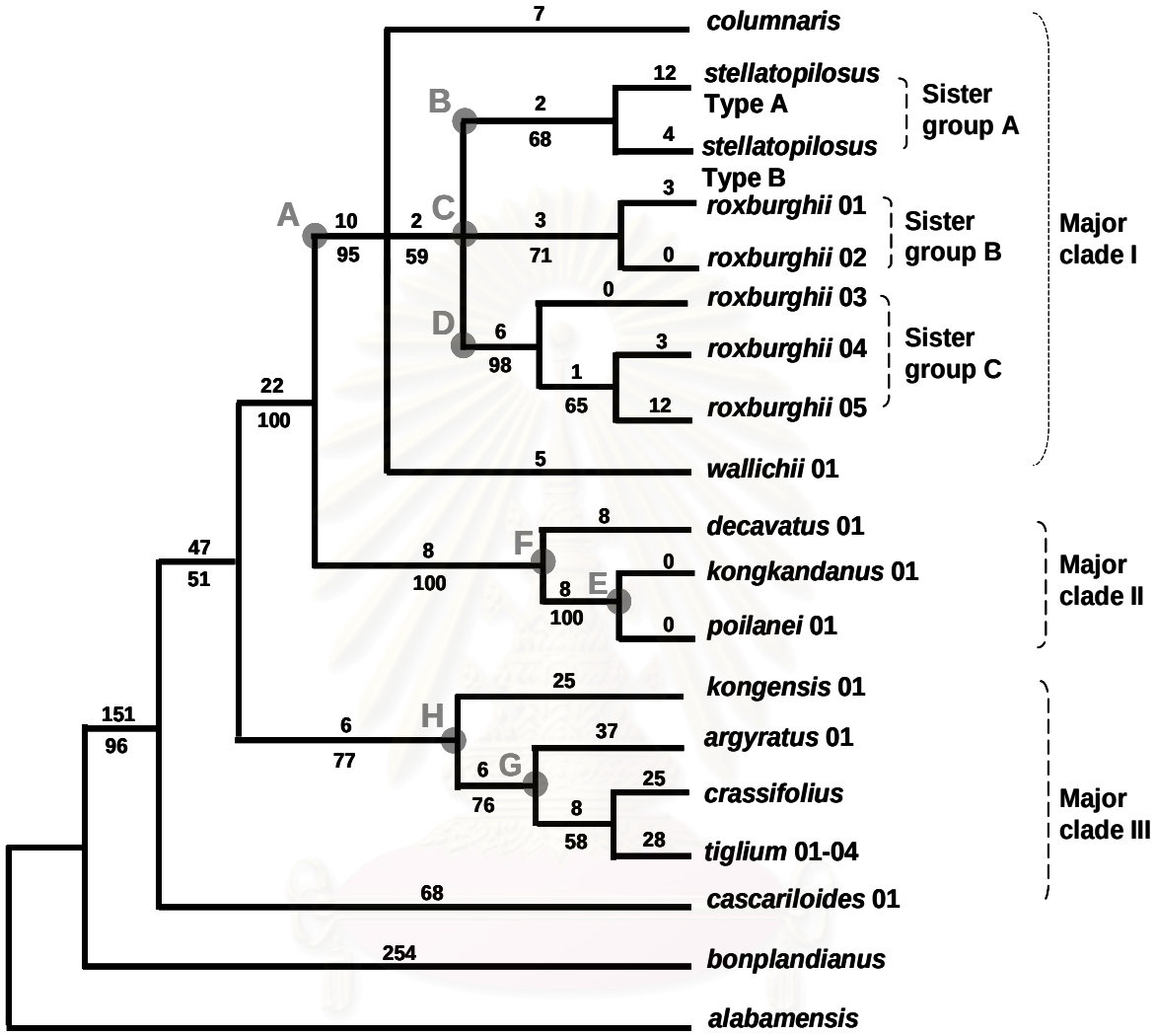


Figure 4.11 The 50 % majority consensus tree resulting from parsimony analysis of the combined sequence data. The numbers under the branches are bootstrapping percentages ≥50 % and the numbers above the branches are branch lengths. *Croton alabamensis* is assigned as the outgroup. The Thai name of the *Croton* samples are as follow; *columnaris* (เปล้าคำ), *stellatopilosus* (เปล้าน้อย), *roxburghii* (เปล้าใหญ่), *wallichii* (เปล้านา), *decalvatus*, *kongkandanus* (มิลา), *poilane* (เปล้าเลือด), *argyratus* (เปล้าเงิน), *crassifolius* (พังคี), *tiglium* (สลอด), *kongensis* (เปล้าน้ำเงิน), *cascarilloides* (เปล้าเล็ก) and *bonplandianus* (เปล้าทุ่ง).

4.4 The Congruence of *Croton* Phylogenetic Trees Reconstructed From ITS, *trn*L-F and Combined Sequence Data

In this study, we reconstructed the phylogenetic trees of the *Croton* species from their ITS, *trn*L-F and the combined sequence data. The obtained parsimony trees from all sequence data appeared to have the consistency index, retention index and rescaled consistency index higher than 0.500 (Table 4.7). This result indicated that the ITS, *trn*L-F and combined sequence data were consistent. Furthermore, all obtained parsimony trees were analyzed for the congruence of tree topologies by comparing their branching patterns. Figure 4.12 shows the layout of the ITS, *trn*L-F and the combined data parsimony trees. It could be found that all parsimony trees had similar branching patterns and they could be divided the *Croton* species into two major groups as indicated by the dash line. It is clear that the topologies of parsimony trees reconstructed from ITS, *trn*L-F and the combined sequence data were congruent. Table 4.7 shows the parsimony based tree characteristics of each data set. Comparing the number of informative characters and number of possible parsimony trees, it was found that the combined sequence data had the number of informative character as 259, which was more than the ITS and *trn*L-F data and it had the least number of possible parsimony trees. This result suggested that the combined sequence data could be used to demonstrate the phylogenetic relationships of *Croton* species better than the other data. In addition, each individual sequence data could be used to resolve the phylogenetic relationships of *Croton* species in the different parts. Figure 4.14 shows that the ITS sequence data could better resolve the relationships among *C. argyratus, C. crassifolius, C. tiglium, C. kongensis* and *C. cascarilloides* (the *Croton* species below the dash line), while the *trn*L-F sequence data could resolve the relationships among *C. columnaris, C. stellatopilosus, C. roxburghii, C. wallichii, C. decalvatus, C. kongkandanus* and *C. poilanei* (the *Croton* species above the dash line). The parsimony trees reconstructed from the combined sequence data could well resolve the phylogenetic relationships for all *Croton* species. Therefore, among three parsimony trees, the parsimony tree reconstructed from the combined data was considered to be the tree that best illustrated the phylogenetic relationships of *Croton* species.

Table 4.7 Data set and parsimony based tree characteristics of entire ITS, *trn*L-F and the combine data analyses.

Parameters	Nucleotide sequence data matrix		
	ITS	***trn*L-F**	**Combined**
Number of accessions (excluded the outgroup)	20	16	19
Data matrix length	639	1063	1714
Variable characters	137	108	259
Informative characters	76	58	128
Number of trees	41	8	3
Consistency index (CI)	0.8144	0.8995	0.8342
Retention index (RI)	0.6417	0.8601	0.7143
Rescaled consistency index (RC)	0.5226	0.7737	0.5959

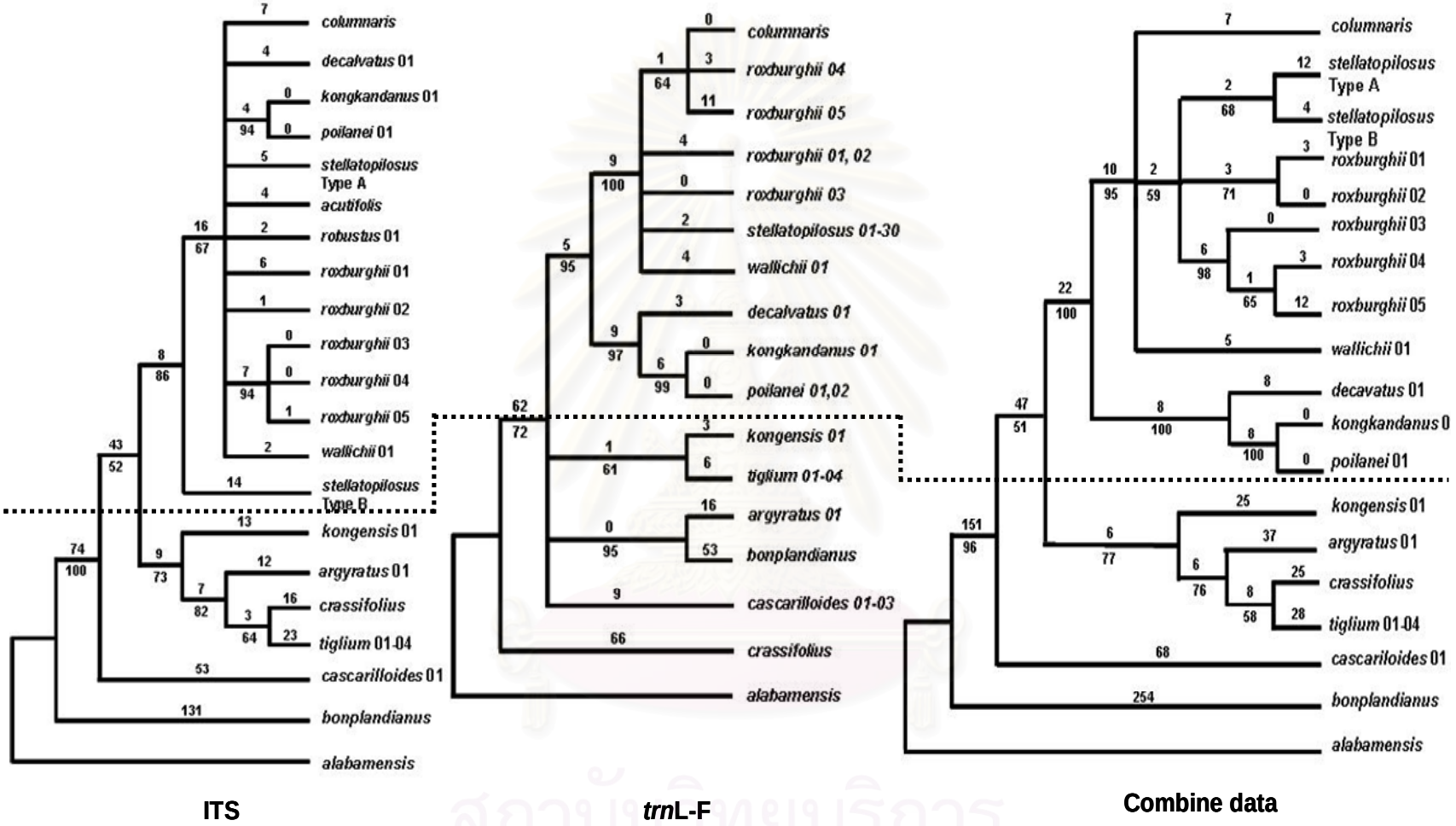


Figure 4.12 The topologies of the consensus parsimony trees of the ITS, *trn*L-F and the combined sequence data. The dash line separates the two major groups of the *Croton* species

.

5. Nuclear and Chloroplast DNA Nucleotide Sequences of *C. stellatopilosus*

In this study, thirty *C. stellatopilosus* samples were examined for their nucleotide sequences of the entire ITS region of the nuclear ribosomal DNA, as well as the *trn*L-F gene and the complete *trn*K intron of chloroplast DNA. The entire ITS nucleotide sequences of all *C. stellatopilosus* samples were found to be in range 623-624 bp in length (Table 4.3). The obtained sequences were all aligned and the results are shown in Table 4.8. All the *C. stellatopilosus* samples appeared to contain 15 polymorphic sites. Their distribution of the polymorphism was at the regions of ITS 1, 5.8S and ITS 2 with the number of 8, 2 and 5 sites, respectively. Moreover, some *C. stellatopilosus* samples showed nucleotide additive sequences at these polymorphic sites.

Based on the alignment of the ITS nucleotide sequences (Table 4.8), thirty *C. stellatopilosus* samples could be divided into three groups. One group with six *C. stellatopilosus* samples (codes STEL 04, STEL 08-10, STEL 13 and STEL 16) had identical nucleotide sequence (GenBank Accession AB375069) and was designated as STEL Type A. This sequence did not contain the additive nucleotide. Another group of seven *C. stellatopilosus* samples (codes STEL 03, STEL 15 and STEL 26-30) had another site of identical sequence and also did not contain the additive nucleotide. This second identical sequence (Accession AB375068) was designated as STEL Type B. Previously, Klinbantoom (2004) has proposed that there are two genotypes of *C. stellatopilosus*. Our results on the individual STEL Type A and STEL Type B nucleotide sequences, therefore, strongly supported Klinbantoom's work. The remaining seventeen samples appeared to have nucleotide additive sites in their ITS nucleotide sequences, ranging from 11 to 15 sites (Table 4.8). These samples were designated as "Putative Hybrid Group". The nucleotide additive patterns of the putative hybrid samples were considered to be the combination between the individual sequences of the STEL Type A and the STEL Type B samples at most nucleotide additive sites. This ITS nucleotide additive sequences in *C. stellatopilosus* was similar to the previous publications reported for the hybridization of *Paeonia* (Sang *et al.*, 1995), *Nuphar rubrodisca* (Donald *et al.*, 1998), *Rhododendron agastum* (Zhang *et al.*, 2000) and

Atractylodes (Shiba *et al.*, 2006). Based on the appearance of two individual and the additive nucleotides of the obtained ITS sequences, we proposed that the hybridization of the two types did occur in *C. stellatopilosus* species.

In addition, for the Putative Hybrid Group, a quantitative aspect could be analyzed to see whether the Type A or Type B was more dominated. This information could be obtained based on the peak intensity of the nucleotide additive sites of the hybrid population. All nucleotide additive sites were classified into three categories of the peak intensity as (the Type A sequence > the Type B), (the Type A sequence ≈ the Type B) and (the Type B sequence > the Type A). For each sample, the number of nucleotide additive sites of each peak intensity category were used to determine the dominate type. The results are shown in Table 4.9. It can be seen that three samples of STEL 18, STEL 02 and STEL 21 had their nucleotide additive sites with peak intensity of the Type A higher than the Type B (the Type A sequence > the Type B). For the samples STEL 01 and STEL 11, it was found that both were dominated rather equally by the Type A and Type B (the Type A sequence ≈ the Type B). Seven samples of STEL 05-07, STEL 25, STEL 23, STEL 22 and STEL 24 showed that Type B was more dominating than Type A (the Type B sequence > the Type A). For the samples of STEL 17, STEL 19, STEL 14, STEL 20 and STEL 12, it was found that Type B was extremely dominating (the Type B sequence >> the Type A). Thus, among seventeen samples of the Putative Hybrid Group, twelve samples were Type B dominant. Based on these results, we proposed that the Type B was the dominating parent in the *C. stellatopilosus* hybridization.

To evaluate the maternal and paternal genotypes of *C. stellatopilosus*, the nucleotide sequences of their chloroplast DNA as *trn*L-F gene and the complete *trn*K intron were investigated. The obtained nucleotide sequences of both regions were all aligned. The results revealed that all the *trn*L-F nucleotide sequences of *C. stellatopilosus* samples were identical (Table 4.5). On the other hand, the complete *trn*K intron sequences appeared to be present in two individuals among the samples (Appendix C). Their alignments are summarized in Table 4.10. One sequence was 2527 bp long

(Fig. 4.13) while the other was 2536 bp long. The complete *trn*K intron and *mat*K gene nucleotide sequences are shown in Fig. 4.13. When the two *trn*K intron nucleotide sequences of *C. stellatopilosus* were aligned together (Table 4.10), six nucleotide substitution sites were found at the nucleotide positions of 579, 620, 1106, 1353, 1858 and 2199. The longer sequence had 9 bp insertion at the nucleotide position 2363-2371.

Since, the nuclear DNA is inherited parentally while the chloroplast DNA is inherited maternally, the incongruence between the nuclear DNA and the chloroplast DNA sequence data should indicate the maternal and paternal of the hybridization (Tsukaya, 2004; Whittall *et al*., 2000; Kondo *et al*., 2007). Therefore, the three groups of *C. stellatopilosus* samples inferred from the entire ITS sequence data were compared with the complete *trn*K intron nucleotide sequences. The results showed that *C. stellatopilosus*, STEL Type A group had the length of 2536 bp of *trn*K intron nucleotide sequence (Accession AB428645). The STEL Type B and the Putative Hybrid groups had the same *trn*K intron nucleotide sequence (Accession AB428644) with the length of 2527 bp. As a result, the STEL Type B group was the maternal genotype whereas the STEL Type A group was the paternal genotype of *C. stellatopilosus* hybridization.

In addition, the leaf morphological characters of each *C. stellatopilosus* group were compared: the maternal group, the Type A sequence > the Type B hybrid group, the Type A sequence ≈ the Type B hybrid group, the Type B sequence > the Type A hybrid group and the Type B sequence >> the Type A hybrid group, with assignments based on the ITS and the complete *trn*K nucleotide sequences were observed. The young, medium and old leaves of each group were compared with regard to their morphological characters as shown in Fig. 4.14 to Fig. 4.17.

Table 4.8 Alignment of the entire ITS nucleotide sequences of *C. stellatopilosus*.

Sample	Nucleotide position															Group
	ITS 1								5.8S		ITS 2					
	65	77	102	115	126	128	139	247	402	405	489	572	597	606	620	
STEL 04, 08, 09, 10, 13, 16	**C**	**C**	**C**	**A**	**A**	**T**	**G**	**C**	**G**	**A**	**C**	**C**	**G**	**A**	**A**	**Type A (n = 6)**
STEL 18	Y	Y	Y	R	W	Y	R	Y	K	M	*	S	*	*	*	
STEL 02	Y	Y	Y	R	W	Y	*	Y	K	M	M	S	*	*	M	
STEL 21	Y	Y	Y	R	W	Y	R	Y	K	M	M	S	K	*	M	
STEL 01, 11	Y	Y	Y	R	W	Y	R	Y	K	M	M	S	K	M	M	
STEL 05-07	Y	Y	Y	R	W	Y	R	Y	K	M	M	G	K	M	M	
STEL 25	Y	Y	Y	R	W	Y	R	Y	K	M	A	G	K	M	C	
STEL 23	Y	Y	Y	*	W	Y	R	Y	K	M	M	G	K	M	C	**Putative hybrid (n = 17)**
STEL 22	Y	Y	Y	R	W	Y	A	T	K	M	M	G	K	M	C	
STEL 24	Y	Y	Y	R	W	Y	A	Y	K	M	M	G	K	M	C	
STEL 17	Y	T	T	G	W	Y	A	Y	K	C	A	G	K	M	C	
STEL 19	Y	T	T	G	W	Y	A	Y	K	C	A	G	K	M	C	
STEL 14	Y	T	T	G	W	Y	A	Y	K	C	A	G	K	M	C	
STEL 20	T	T	T	G	T	C	A	T	K	M	A	G	K	M	C	
STEL 12	Y	T	T	G	T	C	A	Y	T	C	A	G	T	C	C	
STEL 03, 15, 26-30	**T**	**T**	**T**	**G**	**T**	**C**	**A**	**T**	**T**	**C**	**A**	**G**	**T**	**C**	**C**	**Type B (n = 7)**
No. of polymorphic sites				**8**					**2**				**5**			

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.
* Asterisks indicate the nucleotide sequences that are the same as the sequence in the first line.
The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

Table 4.9 The relative peak intensities at the nucleotide additive sites of the individual STEL Type A and Type B sequences.

	Nucleotide position																
	ITS 1								5.8S		ITS 2						
Sample	65	77	102	115	126	128	139	247	402	405	489	572	597	606	620	**Group**	**n**
STEL 04, 08, 09, 10, 13, 16	**C**	**C**	**C**	**A**	**A**	**T**	**G**	**C**	**G**	**A**	**C**	**C**	**G**	**A**	**A**	**Type A**	**6**
																Putative hybrid :	
STEL 18	C>T	C>T	C>T	A>G	A>T	C>T	G>A	C>T	G>T	A=C	*	C>G	*	*	*	Type A>Type B	
STEL 02	C>T	C>T	C>T	A>G	A>T	T>C	*	C>T	G>T	A>C	C>A	C>G	*	*	A>C	Type A>Type B	**3**
STEL 21	C>T	C>T	C>T	G>A	A>T	T>C	G>A	C>T	G>T	A>C	C>A	G>C	G>T	*	A>C	Type A>Type B	
STEL 01	C=T	C=T	C>T	G>A	A=T	T>C	G=A	C=T	G>T	A=C	C>A	G>C	G>T	A=C	A>C	Type A≈Type B	**2**
STEL 11	C>T	C>T	C>T	G>A	A=T	T=C	G>A	C>T	G>T	A=C	C>A	C=G	G=T	A>C	A>C	Type A≈Type B	
STEL 05	T>C	T>C	T>C	G>A	T>A	C>T	G=A	T>C	T>G	C>A	C>A	G	T>G	C>A	C>A	Type B>Type A	
STEL 06	C>T	T>C	T>C	G>A	T>A	C>T	G=A	T>C	T>G	C>A	A>C	G	T>G	A=C	C>A	Type B>Type A	
STEL 07	T>C	T>C	T>C	G>A	T>A	C>T	G=A	T>C	T>G	C>A	A>C	G	T>G	A=C	C>A	Type B>Type A	
STEL 25	T>C	T>C	T>C	G>A	T>A	C>T	A>G	T>C	T>G	C>A	A	G	T>G	C>A	C	Type B>Type A	**7**
STEL 23	T>C	T>C	T>C	*	T>A	C>T	A>G	T>C	T>G	C>A	A>C	G	T>G	C>A	C	Type B>Type A	
STEL 22	T>C	T>C	T>C	G>A	T>A	C>T	A	T	T>G	C>A	A>C	G	T>G	C>A	C	Type B>Type A	
STEL 24	T>C	T>C	T>C	G>A	T>A	C>T	A	T>C	T>G	C>A	A>C	G	T>G	C>A	C	Type B>Type A	
STEL 17	T>C	T	T	G	T>A	C>T	A	T>C	T>G	C	A	G	T>G	C>A	C	Type B»Type A	
STEL 19	T>C	T	T	G	T>A	C>T	A	T>C	T>G	C	A	G	T>G	C>A	C	Type B»Type A	
STEL 14	T>C	T	T	G	T>A	C>T	A	T>C	T>G	C	A	G	T>G	C>A	C	Type B»Type A	**5**
STEL 20	T	T	T	G	T	C	A	T	T>G	C>A	A	G	T>G	C>A	C	Type B»Type A	
STEL 12	T>C	T	T	G	T	C	A	T>C	T	C	A	G	T	C	C	Type B»Type A	
STEL 03, 15, 26-30	**T**	**T**	**T**	**G**	**T**	**C**	**A**	**T**	**T**	**C**	**A**	**G**	**T**	**C**	**C**	**Type B**	**7**

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.
* Asterisks indicate the nucleotide sequences that are the same as the sequence in the first line

Table 4.10 Alignment of the complete *trn*K intron sequences of *C. stellatopilosus.*

Sample	Nucleotide position: The complete *trn* K intron							Group inferred from ITS data
	579	620	1106	1353	1858	2199	2363 - 2371	
STEL 04, 08-10, 13, 16	**A**	**G**	**A**	**T**	**C**	**A**	**ATAATGAAA**	**Type A (n = 6)**
STEL 18	G	A	C	G	T	G	---------	
STEL 02	G	A	C	G	T	G	---------	
STEL 21	G	A	C	G	T	G	---------	
STEL 01, 11	G	A	C	G	T	G	---------	
STEL 05-07	G	A	C	G	T	G	---------	
STEL 25	G	A	C	G	T	G	---------	
STEL 23	G	A	C	G	T	G	---------	**Putative hybrid (n = 17)**
STEL 22	G	A	C	G	T	G	---------	
STEL 24	G	A	C	G	T	G	---------	
STEL 17	G	A	C	G	T	G	---------	
STEL 19	G	A	C	G	T	G	---------	
STEL 14	G	A	C	G	T	G	---------	
STEL 20	G	A	C	G	T	G	---------	
STEL 12	G	A	C	G	T	G	---------	
STEL 03, 15, 26-30	**G**	**A**	**C**	**G**	**T**	**G**	**---------**	**Type B (n = 7)**
No. of polymorphic sites					**7**			

Nucleotide positions of the complete *trn*K nucleotide sequences start from the first position of this region.

Hyphens indicate deletions

```
   1   TAGAGTACTC GGCTTTTAGT GCGGCTAACA TCTTTTACAC GTTTGTATGA     50
  51   AGAAAGAAAT TCGTCCATAC TATCGGTATA GCTTGTAAGA CCACGACTGA    100
 101   TCCTGAAAGT GAAAGGAATG AATGGAAAAA ATAGCATGTC GTATCAATGG    150
 151   AAAATTCTGC AAATATTTCA TTTTTGCCGG ATCGGTCCAA ATTTTTTTGA    200
 201   ATTCTTGATG CGCAACATAA CAAAATGAAT TCAGAGTTGG GTCGAATTAA    250
 251   TAATAAATGG ATGGAGTCCT ACGATTCCAA TTATAGCGAA ACAAAAAAAG    300
 301   CAACGAGCTT ACGTTCTTAA TTTGAATGAT TTTCCGATCT AATTAGACGT    350
 351   TAAAAATAAA TTAGTACCTA ATGCGGGAAA GGTTTCTCTA TAAGCAGATT    400
 401   TTCGATTTTC TTTTAATAAG TCCTAACTAT TAGTTTATTC TACACTATGA    450
 451   AGAGGAGATG AATGTGTAGA AGAAAGAGTA TATTGATAAA GAGATTTTGA    500
 501   AAGAGATTTT TTCCAAAATC AAAAGAGCGA TTGACTTGAA AAAATAAAGG    550
 551   ATTTCAAATC ATCTTTTTAT CCATCTTAAA ATGCACATAA ACAAATTGGA    600
 601   TGGAAAAAAG AAGAGGATGG GGAGTCCATT AATGATAATG AGTTTACCTA    650
 651   TTTCCGAGGT ATCTACTCTT TTCTTATTAT ATTACTTTGG TTTTATTGTA    700
 701   CTGTATCGCA CTATGTATCA CCCAGGACGG TCCTTTATCC TTACTTCAAT    750
 751   CAAATCGAAT AAAAAAAAAT GTGGGAATAC CAGAGATATT TAGAACTAGA    800
 801   TAGATCTCGA AAAAAGGACT TCCTGTACCC ATTTATTTTT CGGGAGTATA    850
 851   TTTATACATT TGCTCATGAT CATAATTTAA ATAGATCTAT TTTGTTGGAA    900
 901   AATGTGGGTT ATGACAATAA ATCTAGTTTT TTAATTGTAA AACGTTTAAT    950
 951   TACTCGAATG TATCAACAGA ATCATTTTAT TTTTTCTGCT AATGATTCTA   1000
1001   ACCAAAATTC ATTTTTTAGG TACAACAAGA ATTTGTATTA TCAAATGATA   1050
1051   TCAGAGAGTT TTGCAATTAT TGTGGAAATT CCATTTTCTC TACAATTAGT   1100
1101   ATCTTATTTG GAAAGGTCGG AGCCAGTAAA ATCTCATAAA TTACGATCAA   1150
1151   TTCATTCAAT ATTTCCTTTT TTAGAGGATA AATTTCCACA TTTAAATTAT   1200
1201   GTGTCAGATG GGGTAATACC TTACCCCATC CATGTAGAAA AATTAGTTCA   1250
1251   AACCCTTCGC TATTGGATGA AAGATCCCTC GTCTTTGCAT TTATTACGAC   1300
1301   TCTTTCTTCA TGAGTATTGG AATAGGAGCA GTCTTTTTAT TCAAAAGAAA   1350
1351   TCTATTTTTA TTTTTACAAA AAGGTAATCC AAGATTTTTC TTGTTCCTAT   1400
1401   ATAATTCTCA TGTATATGAA TACGAATCAA TCCTCTTTTT TCTTCGTAAC   1450
1451   CAATCCTTTC ATTTACGATC AACATTTTCT CGAGTCCTTC TTGAACGAAT   1500
1501   TTTTTTCTAT GGAAAAATAG AACATTTTGC AGAAGTCTTT GCTAATGATT   1550
1551   TTCAGACTAT CCTAGGGTTG GTCAAAGATC CTTTCCTGCA TTATGTTAGA   1600
1601   TATCAAGGAA AATCCATTCT GGCTTCAAAA GATGGGCTTC TTCTGATGAA   1650
1651   AAAATGGAAA TATTACCTTG TCAATTTATG TCAATGTCAT TTTCATGTGT   1700
1701   GGTTTCAACC CAAAAAGATC TATATAAGTT CATTACCCAA GCATTCTCTC   1750
1751   AACCTTTTGG GCTATCTTTC AAATGTACAA TTAAATCCTT TGGTCGTACG   1800
1801   AAGTCAAATG CTAGAAAATT CATTTTTAAT AGAAAAAGAT AATACTATGA   1850
1851   AGAAACTCGA TACAATAGTT CCAATTATTC CTTTAATTGG ATCATTATCA   1900
1901   AAAATGCAAT TTTGTAACAC AGTCGGGCAT CCCATTAGTA AATCGGCCTG   1950
1951   GACTGATTCA TTGGATTCTG ATATATCGAC CGCTTTGTGC GTATATGCAG   2000
2001   AAGCTTTCTC ATTATTATAG TGATCTTCAA AAAAAAAGAG TTTGTATCGA   2050
2051   GTAAGATATA TACTTCGACT TTCGTGTGTT AAAACTTTGG CTCGTAAACA   2100
2101   CAAAAGCACT GTACGCGCTT TTTTGAAAAG ATTAGGTTCG GAATTATTCG   2150
2151   AAGAATTTTT TACGGAAGAG GAACAGATTC TTTCTTTGAT CTTCCCGAAA   2200
2201   GTTTCCTCTA TTTCGCGCAG GTTATATAGA GGGCGGGTTT GGTATTTGGA   2250
2251   TATCATTTCT ATCAATGATT TGGCCAATCA TGAATAATTG GTTATGAAAT   2300
2301   CAAAATTTTC TCTAAATACT CAATTAAATA CTATCAATTA AATACTAAAT   2350
2351   AATACTAAAT AATAAAAAGA ATATTAATAT AATGATTTAA TATAATGATA   2400
2401   ATGAAGGGAT AACCAAAAAT TCATTCATTT CTATTATGAA ATGCTCATAC   2450
2451   ATTATACGAG TAAGGATTGA GCGACTGAGT ATTCAACTTC CTTATAAGTT   2500
2501   TCGTCTAGGA AGGGAACGGA GATTAGA                           2527
```

Figure 4.13 The complete *trn*K intron nucleotide sequence of *C. stellatopilosus* STEL Type B and the putative hybrid groups that length 2527 bp. The shaded area indicates nucleotide sequence of *mat*K gene.

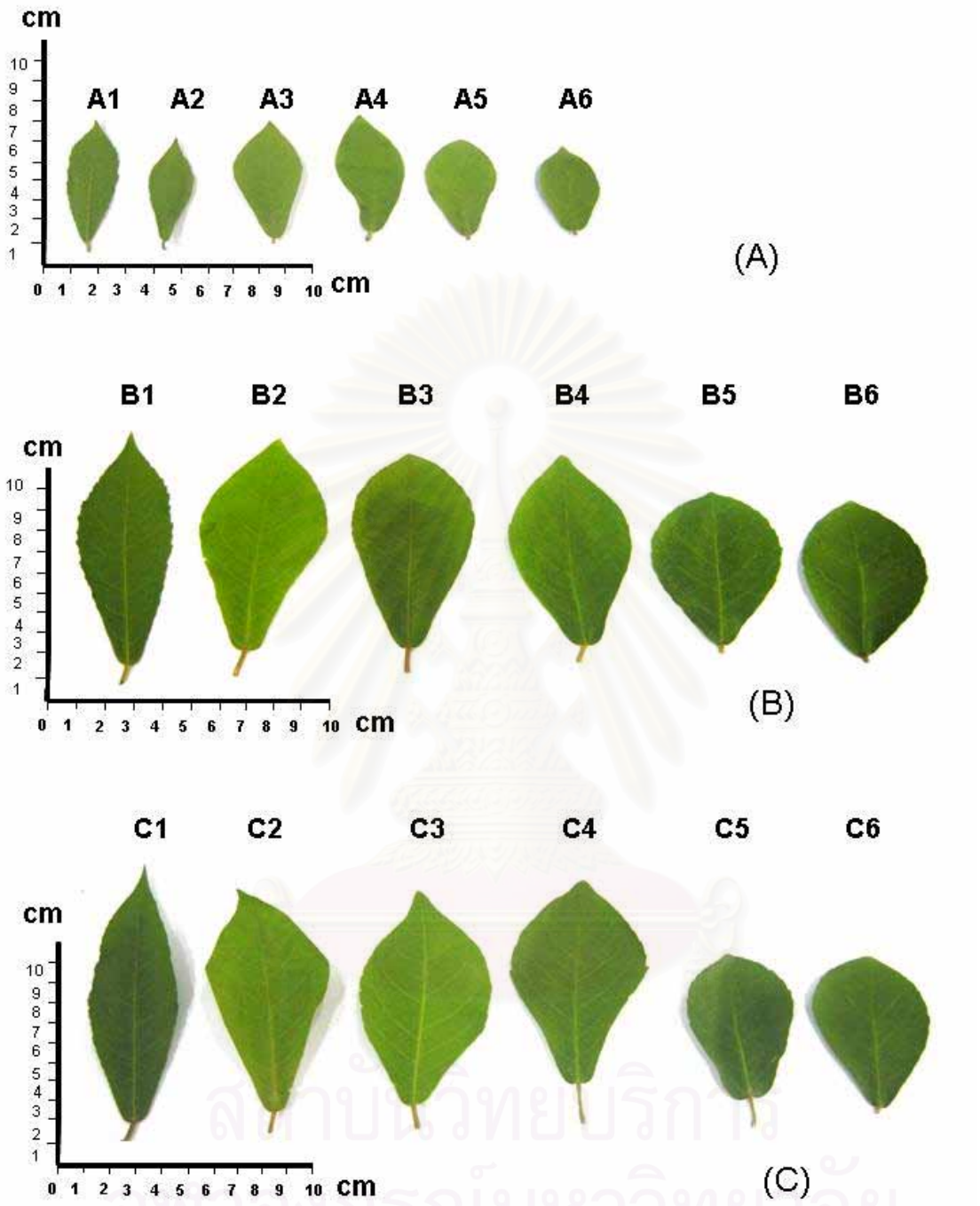


Figure 4.14 Morphological characters of the *C. stellatopilosus* leaves.

(A) young leaves

(B) medium age leaves

(C) mature leaves

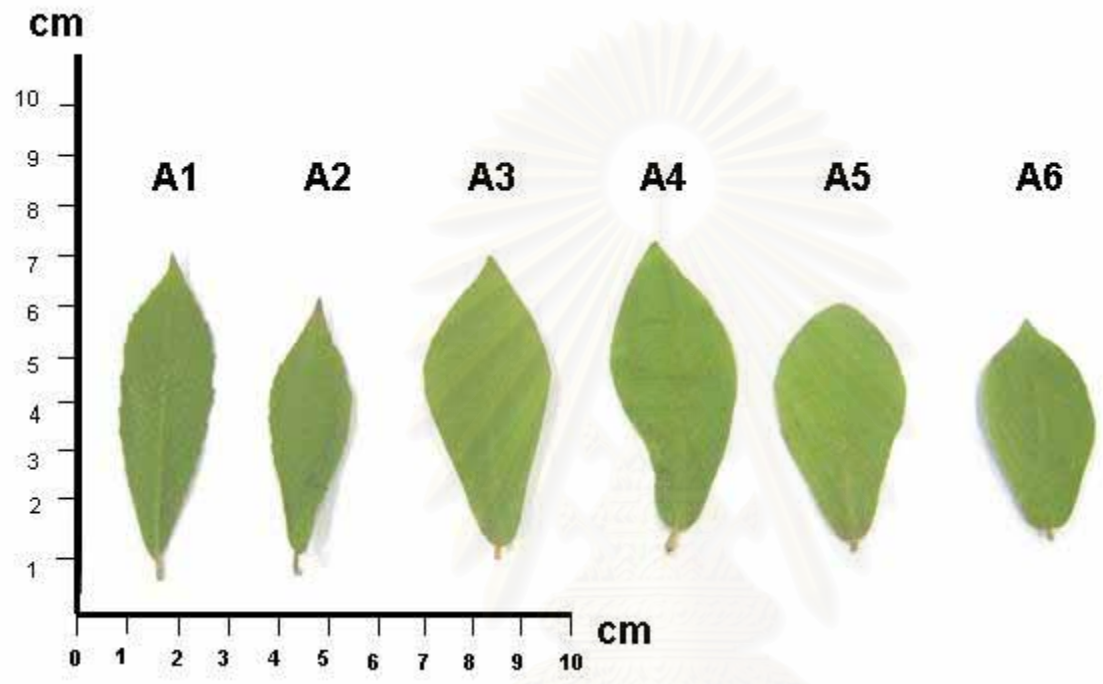


Figure 4.15 Young leaves of each *C. stellatopilosus* groups

A1 : the paternal group (Type A),
A2 : the Type A sequence > the Type B hybrid group,
A3 : the Type A sequence ≈ the Type B hybrid group,
A4 : the Type B sequence > the Type A hybrid group,
A5 : the Type B sequence >> the Type A hybrid group,
A6 : the maternal group (Type B)

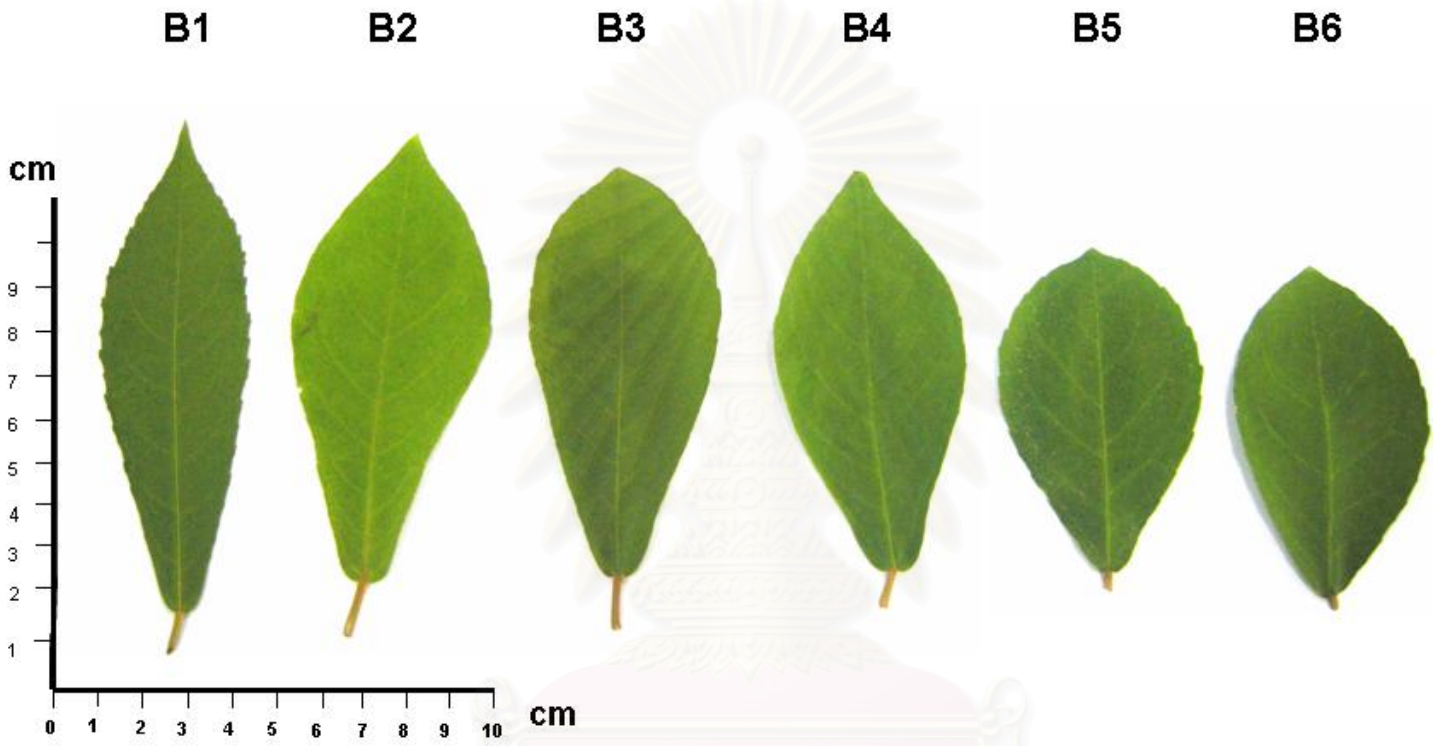


Figure 4.16 Medium age leaves of each *C. stellatopilosus* groups

B1 : the paternal group (Type A),
B2 : the Type A sequence > the Type B hybrid group,
B3 : the Type A sequence ≈ the Type B hybrid group,
B4 : the Type B sequence > the Type A hybrid group,
B5 : the Type B sequence >> the Type A hybrid group,
B6 : the maternal group (Type B)

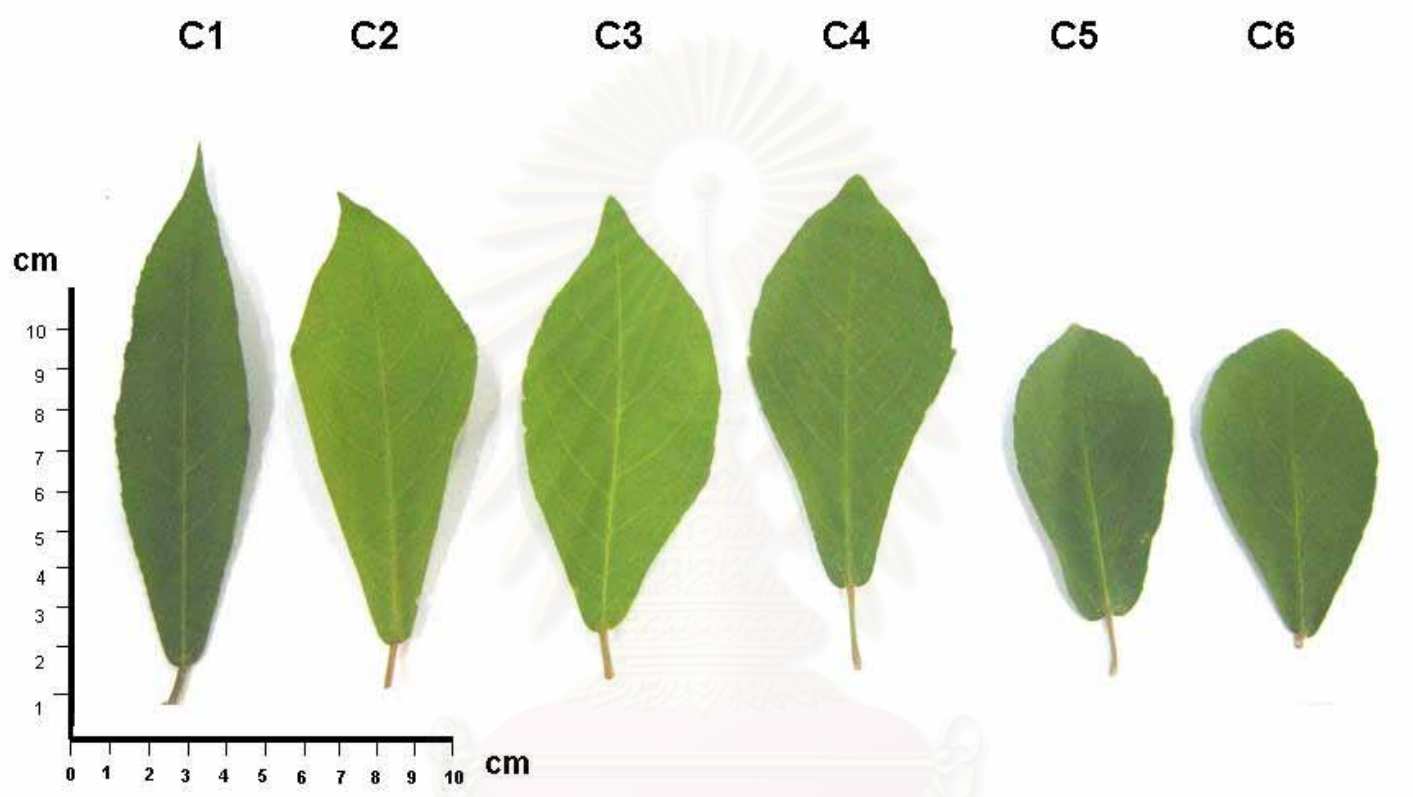


Figure 4.17 Mature leaves of each *C. stellatopilosus* groups

C1 : the paternal group (Type A),
C2 : the Type A sequence > the Type B hybrid group,
C3 : the Type A sequence ≈ the Type B hybrid group,
C4 : the Type B sequence > the Type A hybrid group,
C5 : the Type B sequence >> the Type A hybrid group,
C6 : the maternal group (Type B)

6. Plaunotol Content in the Parental and Hybrid Groups of *C. stellatopilosus*

6.1 Plaunotol Quantitative Analysis Based on the TLC Densitometry

All TLC chromatograms obtained in this study showed that the solvent system of chloroform and n-propanol (24:1) on TLC (silica gel) plate gave a good separation of plaunotol from other components in the ethanolic extracts of *C. stellatopilosus* leaves. The Rf value of plaunotol was found to be 4.3 as shown in Figs 4.18 and 4.19. To confirm that the peak of plaunotol in the TLC densitometric chromatogram was absolutely contributed by pure plaunotol, the spot on the TLC plate corresponded to the Rf value of standard plaunotol was scanned to produce a UV-absorption spectrum. The obtained UV-absorption spectrum was then compared with that of authentic plaunotol. As shown in Fig 4.20, it was clear that plaunotol in the ethanolic extract and standard plaunotol had identical absorption spectrum with their maximum wavelength at 220 nm. Therefore, it was clear that the peak with Rf value of 4.3 on the TLC plate was contributed by only pure plaunotol present in the ethanolic extract of *C. stellatopilosus* leaves.

In addition, the complete separation of plaunotol from the other leaf constituents by the development of TLC plate allowed the compound be quantitived by the method of densitometry using the calibration curve of the authentic plaunotol for calculation. The calibration curve of standard plaunotol obtained by ploting the peak areas against plaunotol concentration is shown in Fig. 4.21. This curve shows the linearity of the relationship between 29.06 to 162.75 µg plaunotol per ml. The linear regression was found to be 0.9902.

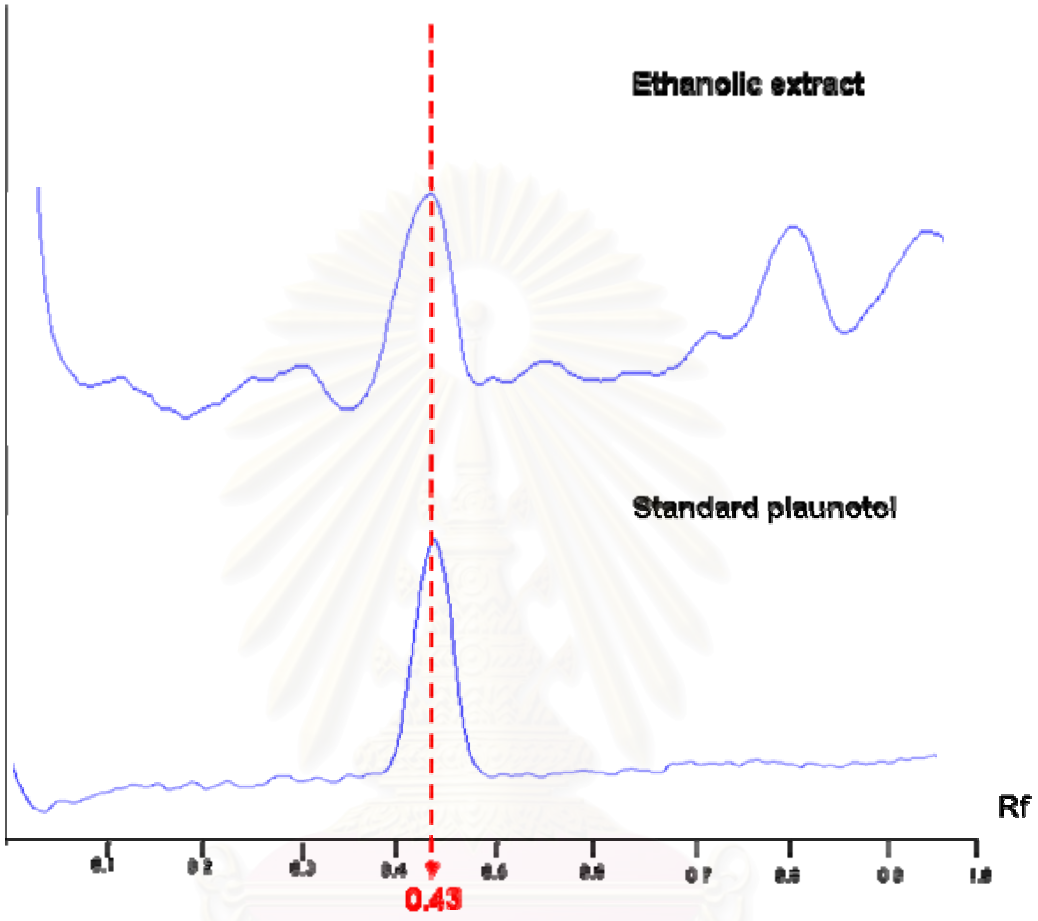


Figure 4.18 TLC chromatogram of the *C. stellatopilosus* ethanolic extracts and authentic plaunotol.

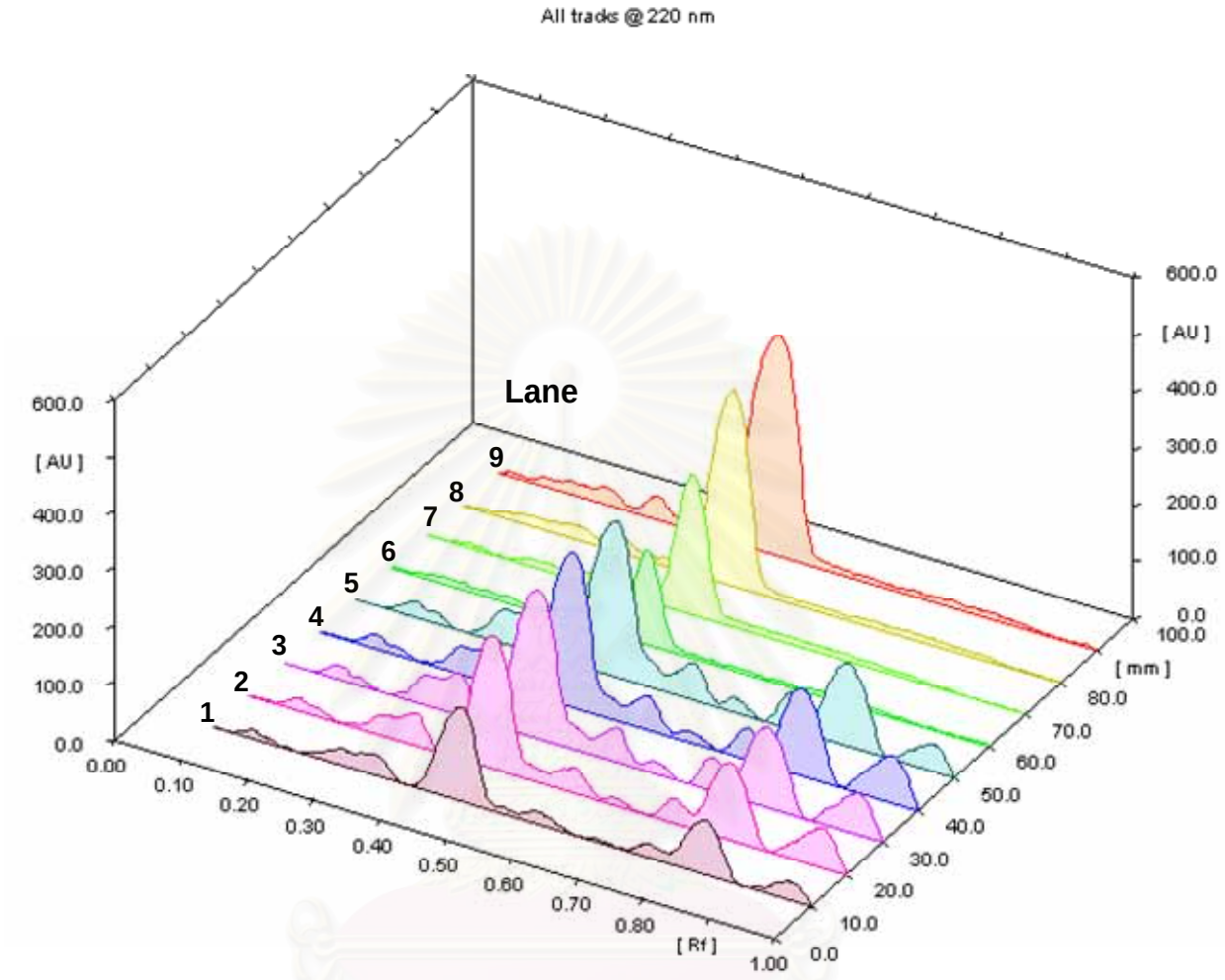


Figure 4.19 Three dimension chromatogram of the ethanolic extracts of *C. stellatopilosus* leaves and the authentic plaunotol.

Lane 1-5 : the ethanolic extracts of *C. stellatopilosus* leaves

Lane 6-9 : the authentic plaunotol

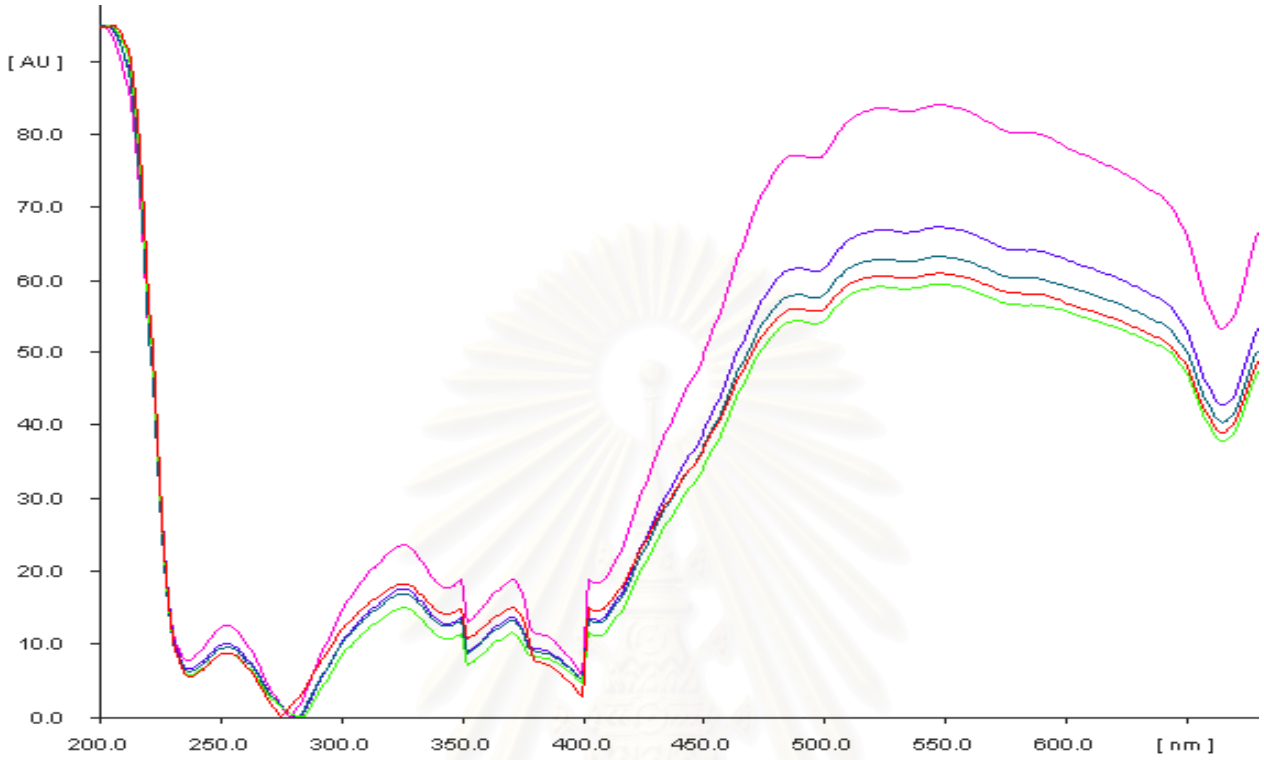


Figure 4.20 The UV-absorption spectrum of the authentic plaunotol and the compound on TLC plate with similar Rf value to plaunotol which was separated from the ethanolic extract of *C. stellatopilosus* leaves.

the authentic plaunotol

the compounds similar Rf value to the authentic plaunotol

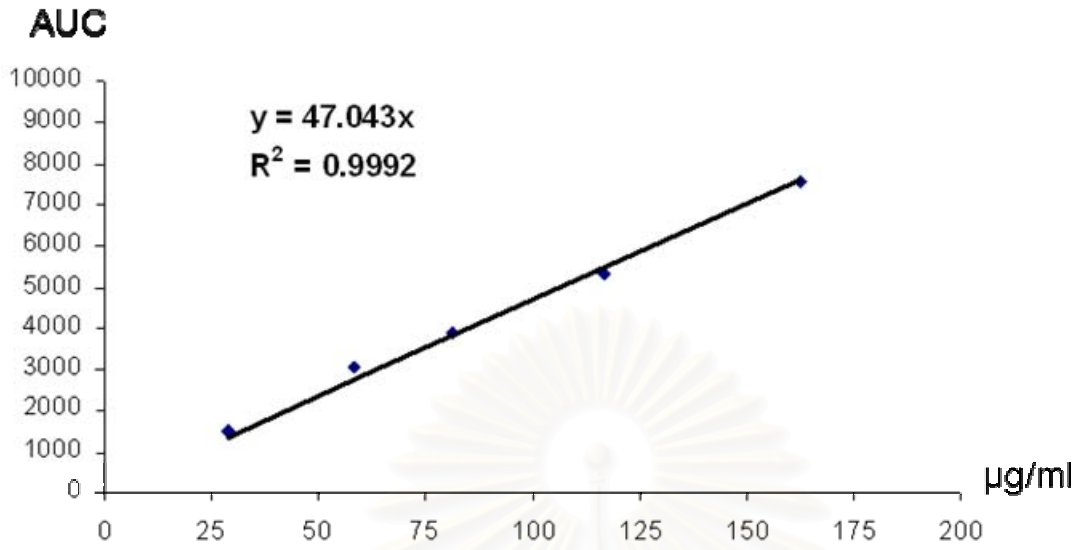


Figure 4.21 Calibration curve of plaunotol by TLC densitometric method

6.2 Plaunotol Content in the Parental and Hybrid Groups of *C. stellatopilosus*

As inferred from the ITS and *trn*K intron sequence data, the thirty samples of *C. stellatopilosus* appeared to have three groups of the maternal group of the STEL Type B, the paternal group of the STEL Type A, and the hybrid group of the two types. Based on the signaling peak intensities of each nucleotide additive site, the *C. stellatopilosus* hybrid group was further divided into four subgroups. Besides the direct sequencing analysis, all *C. stellatopilosus* samples were analyzed for their content of the plaunotol in their leaves using the TLC densitometry. Plaunotol in each *C. stellatopilosus* sample in the TLC plate (Appendix D) was calculated based on its standard curve. The results of the plaunotol content in the parental and the hybrid groups are summarized in Table 4.11. It can be seen that the maternal group (Type A) of *C. stellatopilosus* contained variable plaunotol content ranging from 0.209 to 0.493 % (w/w) dry weight. The hybrid four subgroups of (the Type A sequence > the Type B), (the Type A sequence ≈ the Type B), (the Type B sequence > the Type A) and (the Type B sequence >>the Type A) contained plaunotol content in the range of 0.394-0.662 %, 0.416-0.729 %, 0.391-0.896 % and 0.496-0.848 % (w/w) dry weight, respectively. The paternal group (Type B) contained plaunotol content ranging from 0.442 to 1.00 % (w/w) dry weight.

Table 4.11 Plaunotol content of the paternal, maternal and each hybrid groups of *C. stellatopilosus.*

Group		Sample code	Plaunotol content (% w/w dry weight ± SD)
Paternal (Type A)		STEL 04	0.418±0.0020
		STEL 08	0.289±0.0012
		STEL 09	0.256±0.0005
		STEL 10	0.492±0.0021
		STEL 13	0.448±0.0009
		STEL 16	0.209±0.0014
Hybrid	**Type A sequence > Type B**	STEL 18	0.394±0.0011
		STEL 02	0.662±0.0014
		STEL 21	0.434±0.0011
	Type A sequence ≈ Type B	STEL 01	0.416±0.0050
		STEL 11	0.729±0.0334
	Type B sequence > Type A	STEL 05	0.419±0.0023
		STEL 06	0.503±0.0014
		STEL 07	0.319±0.0028
		STEL 25	0.896±0.0039
		STEL 23	0.698±0.0090
		STEL 22	0.711±0.0044
		STEL 24	0.816±0.0052
	Type B sequence >> Type A	STEL 17	0.711±0.0011
		STEL 19	0.848±0.0004
		STEL 14	0.602±0.0006
		STEL 20	0.496±0.0008
		STEL 12	0.501±0.0014
Maternal (Type B)		STEL 03	0.497±0.0019
		STEL 15	0.442±0.0012
		STEL 26	0.883±0.0042
		STEL 27	0.727±0.0052
		STEL 28	1.000±0.0013
		STEL 29	0.675±0.0277
		STEL 30	0.525±0.0095

7. Nuclear and Chloroplast DNA Nucleotide Sequences of *C. roxburghii*

In addition to *C. stellatopilosus*, samples of another species of *C. roxburghii* were also examined for their nucleotide sequences of the entire ITS region and *trn*L-F gene in a similar manner. The *trn*K intron was also determined only in a partial region. The results showed that the entire ITS nucleotide sequences of all *C. roxburghii* samples were in range of 624-625 bp long (Table 4.3). The obtained sequences were all aligned. It was found that all *C. roxburghii* samples did not show identical sequences (Table 4.12). Instead, they appeared to contain 17 polymorphic sites. Their distribution of the polymorphism was at the ITS1, 5.8S and ITS2 regions with the number of 7, 3 and 7 sites, respectively. Among these, three polymorphic sites of the ITS1 region appeared to have the substitutions. At the nucleotide positions 146, 251 and 258, the nucleotide sequences of the sample codes ROX 01 and ROX 02 were A, T and T, respectively. Whereas, the sample codes ROX 03, ROX 04 and ROX 05 were G, C and C, respectively. Furthermore, four *C. roxburghii* samples (codes ROX 01-04) were found to have the nucleotide additive sequences (at the nucleotide positions 13, 111, 112, 402, 404, 405, 482, 552, 573, 591, 593 and 624) whereas the sample ROX 01 did not show such sequences. Based on the obtained ITS sequences, the nucleotide additive distribution of *C. roxburghii* samples did not have the typical nucleotide positions as found in the ITS nucleotide sequences of *C. stellatopilosus.*

For the *trn*L-F nucleotide sequences, all *C. roxburghii* samples had the nucleotide sequences in the range from 1030 to 1035 bp long (Table 4.5). The obtained sequences were all aligned and the results are shown in Table 4.13. Four *C. roxburghii* samples (codes ROX 01-04) appeared to have similar sequences whereas the sample ROX 05 had the extreme different sequence with 8 nucleotide substitution sites (at nucleotide positions 239, 241, 383, 518, 557, 630, 661 and 1078) and the 6 bp insertion (at nucleotide positions 831-836). On the other hand, the number of single nucleotide repeated unit (Poly T, at nucleotide position 930-944) of five *C. roxburghii* samples was different. *C. roxburghii* samples ROX 01 and ROX 02 had 14 Poly T repeated units. The samples ROX 03, ROX 04 and ROX 05 possessed 12, 11 and 10 repeated units,

respectively. However, Xu *et al.* (2002) have reported that the number of single nucleotide repeated unit in chloroplast DNA varies in the same plant species.

In addition, all the *C. roxburghii* samples were analyzed for their nucleotide sequences of the partial *trn*K intron that included the partial *mat*K gene and the complete *trn*K 3' intron. Their length and GenBank accession number are listed in Table 4.14. The obtained sequences were all aligned (Table 4.14 and Appendix E). All *C. roxburghii* samples did not have identical sequence. The partial *trn*K intron nucleotide sequences appeared to contain 22 and 5 polymorphic sites at the partial *mat*K gene and the complete *trn*K 3' intron, respectively. The polymorphic sites of the partial *mat*K gene were mostly substitutions while those of the complete *trn*K 3' intron were indels. The substitutions at nucleotide positions 592, 656, 752, 764, 787, 817, 834, 1024, 1359 and 1431 of the partial *mat*K gene could be distinguished between the samples ROX 01 and ROX 02 and the samples ROX 03, ROX 04 and ROX 05 (Table 4.14). Furthermore, five substitutions at nucleotide positions 854, 926, 935, 1261 and 1429 of the partial *mat*K gene, the substitution at nucleotide position 1534 and three indels of the complete *trn*K 3' intron could help to differentiate the sample ROX 03 from the *C. roxburghii* samples ROX 04 and ROX 05.

Based on the obtained nucleotide sequences, *C. roxburghii* showed high genetic variation with its ITS, *trn*L-F and *trn*K intron sequences differing from one another among the samples. However, five *C. roxburghii* samples could be separated into two groups. The first group consists of the sample ROX 01 and ROX 02 while the second group was comparing of the samples of ROX 03, ROX 04 and ROX 05. Our results were in agreement with the previous work by Chonvanich (2004).

Table 4.12 Alignment of the ITS sequences *C. roxburghii*.

	Nucleotide position																	
	ITS 1							5.8S			ITS 2							
Samples	13	111	112	146	207	251	258	402	404	405	482	504	552	573	591	593	624	
ROX 01	T	T	T	A	A	T	T	G	C	A	C	A	A	C	G	T	T	
ROX 02	*	K	K	*	C	*	*	K	*	M	*	G	*	*	R	*	*	
ROX 03	Y	*	*	G	C	C	C	T	Y	C	*	G	W	*	*	*	Y	
ROX 04	Y	*	*	G	C	C	C	T	T	C	*	G	W	*	*	*	Y	
ROX 05	*	*	*	G	C	C	C	T	T	C	M	G	*	Y	*	Y	*	
No. polymophic sites				**7**					**3**					**7**				

Table 4.13 Alignment of the *trn*L-F sequences *C. roxburghii*.

	Nucleotide position											
	***trn* L intron**					***trn* L-F intergenic spacer**						
Samples	239	241	383	518	557	630	661	831- 836	930-944	1048	1049	1078
ROX 01	C	G	C	G	A	G	T	------	14T	A	G	T
ROX 02	*	*	*	*	*	*	*	------	14T	*	*	*
ROX 03	*	*	*	*	*	*	A	------	12T	*	*	*
ROX 04	*	*	*	*	*	A	A	------	11T	C	C	C
ROX 05	A	T	T	A	G	A	A	CAAAAA	10T	*	*	T
No. polymophic sites			**5**						**7**			

Nucleotide positions of ITS region and *trn*L-F gene start from the first position of both regions.
* Asterisks indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate deletions.

Table 4.14 Alignment of the partial *trn*K intron sequences of *C. roxburghii*.

Samples	Nucleotide position: Partial *mat*K gene																					
	592	602	656	692	699	752	764	787	817	831	854	926	935	944	1024	1133	1162	1261	1359	1431	1439	1454-1459
ROX 01	T	G	A	C	A	A	A	A	T	A	G	C	T	C	A	G	T	A	C	G	C	TTATAT
ROX 02	*	A	*	A	G	*	*	*	*	*	*	*	*	A	*	A	*	*	*	*	*	******
ROX 03	A	A	G	*	G	G	G	G	C	G	T	T	G	A	G	A	*	C	A	A	A	------
ROX 04	A	A	G	*	G	G	G	G	C	G	*	*	*	A	G	A	C	*	A	A	*	------
ROX 05	A	A	G	*	G	G	G	G	C	G	*	*	*	A	G	A	C	*	A	A	*	------
No. polymorphic sites												**22**										

Samples	Nucleotide position: The complete *trn* K 3' intron					Length (bp)	GenBank Accesion number
	1534	1553 -1564	1573-1579	1618-1630	2459		
ROX 01	A	----------	CAATTAA	TTTAATATAATGA	A	1338	AB428646
ROX 02	*	----------	*******	**********	*	1338	AB428647
ROX 03	C	CAATTAAATACT	-------	----------	G	1328	AB428648
ROX 04	*	----------	-------	**********	G	1329	AB428649
ROX 05	*	----------	-------	**********	G	1329	AB428650
No. polymorphic sites			**5**				

Nucleotide positions start from the first position of the *mat*K gene.
* Asterisks indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate deletion

CHAPTER V

DISCUSSION AND CONCLUSION

1. Genetic Variation of *Croton* Species

This study has revealed the genetic variation of fifteen *Croton* species in Thailand by determining their entire ITS and *trn*L-F nucleotide sequence. Based on all the obtained ITS sequences, the length of ITS1 regions are in range of 264-265 bp while that of ITS2 are in range of 194-201 bp. The ITS1 region of all the species is consistently longer than the ITS2 region. Their 5.8S coding regions are all 162 bp in length. The appearent entire ITS sequences of all fifteen *Croton* species are, therefore, in the range from 620 to 627 bp long. The GC contents of their ITS region are 52.5-57.1 %. In literature, the ITS region of angiosperms has generally been reported to have low intragenomic diversity and their length remains in a relatively narrow range from 565 to 700 bp (Baldwin *et al.*, 1995; Wendel *et al.*, 1995). Therefore, the entire ITS regions of all *Croton* species in this study have the length falling in the range of angiosperms. Moreover, although some polymorphisms have been observed within the individuals and among species, the ITS regions do not differ significantly in length.

For *trn*L-F gene, the sequences of the *trn*L 5' exon, *trn*L 3' exon and *trn*F coding regions are 35, 50 and 73 bp long, respectively. Based on all the obtained *trn*L-F sequences, the *trn*L-F introns shows to have their length in range of 527 to 550 bp while the intergenic spacers between *trn*L and *trn*F coding genes are in range of 417 to 428 bp. The GC contents of their *trn*L-F genes are 31.03-31.86 %. Apparently, the sizes of the entire ITS and *trn*L-F nucleotide sequences are similar to the previously published sequences of the plants in the genus *Croton* (Berry *et al*., 2005; Xue, *et al*., 2007) and their GC contents are also within the in range of angiosperms (Baldwin *et al*., 1995; Gielly and Taberlet, 1994).

Both the entire ITS region and *trn*L-F gene have spacers. The ITS1 and ITS2 of the entire ITS region have the spacers between the transcribed regions while the *trn*L-F gene has an intergenic spacer between the *trn*L and *trn*F coding genes. The results of this study show that mutations of the entire ITS region and *trn*L-F gene are different. The variation of the ITS sequences indicate that they have more nucleotide substitutions than indels, while the *trn*L-F sequences have high frequency of indels. The ITS1 and ITS2 regions appear to function, at least in part, in the maturation of nuclear ribosomal RNAs (Baldwin *et al.*, 1995). In *Saccharomyces cerevisiae*, the deletions of certain regions within ITS1 can inhibit formation of small and large subunits of ribosomal RNAs (Nue *et al.*, 1994), whereas the certain deletions in ITS2 can prevent or reduce the processing of large subunits of ribosomal RNAs (Sande *et al.*, 1992). Therefore, the mutations in ITS sequences are attributable to substitutions rather than indels that cause the length variations. The polymorphic sites of the ITS and *trn*L-F sequences found in this study are 21.5 % and 6.1 %, respectively. Generally nuclear genome has substitution rate faster than chloroplast (Wolfe *et al.*, 1987; Gaut, 1998). The distribution of polymorphism is not random along the nuclear genome nor within a gene. This unevenness is associated with differences in the many factors such as recombination rate and gene density in the genomic region (Zhang and Hewitt, 2003). The genomic regions with low recombination rates generally have reduced levels of polymorphisms (Begun and Aquadro, 1992). Although substitution rate is estimated from the interspecific comparisons, it may not reflect the intraspecific variability of a given DNA region and often it is an underestimate of that variability. The higher polymorphism of the ITS region of nuclear DNA is consistent with its nucleotide substitution rate.

In addition, based on the direct sequencing approach, the electropherograms of ITS region show some nucleotide positions containing double signaling peaks. These can be tentatively interpreted by two nucleotide states at site, implying superimposition of two sequence patterns or nucleotide additively. Persistence of the same apparent site polymorphisms of both complementary DNA strands reduces the likelihood that the double peaks are attributable to sequencing artifact. In this study, the ITS sequences of some *Croton* species especially *C. stellatopilosus* and *C. roxburghii* appear to have the

a high number of nucleotide additive sites. In the plant kingdom, this nucleotide addition tends to be restricted to a small fraction of families and genera, not a random appearance (Ellstrand *et al*., 1996).

2. Phylogenetic Relationships of *Croton* Species

2.1 The Monophyletic Group of *Croton*

At present, no particular method of phylogenetic reconstruction can be claimed to be better than the others under certain conditions. Each of the methods of phylogenetic reconstruction has advantages and disadvantages. Several strategies are available to minimize random and systematic errors in phylogenetic analysis. The best way to minimize random errors is to use large number of data. A tree based on large number of molecular data is almost invariably more reliable than the one based on limited amount of data. As a result, the topologies of parsimony trees that are reconstructed from the ITS, *trn*L-F and combine data are similarly. However, the combined sequence data gives better phylogeny that is well resolving for all the *Croton* species than individual sequence data. This is obviously to be the result of the large informative parsimony characters of the combined sequence data. Therefore, in this study, the parsimony tree reconstructed from the combined data is selected to estimate the phylogenetic relationships of the *Croton* species. In the past years, the complex family Euphorbiaceae was classified to comprise of five subfamilies: the Acalyphoideae, the Crotonoideae, the Euphorbioideae, the Phyllanthoideae and the Oldfieldioideae. The three former are uni-ovulate families while the two latter are bi-ovulate. Recently, the circumscription of the Euphorbiaceae sensu lato has been modified. The molecular studies (Chase *et al.*, 2002; Davis *et al.*, 2005; Wurdack *et al.*, 2005) have splited the Euphorbiaceae sensu lato into five families: Euphorbiaceae sensu stricto, Pandaceae, Phyllanthaceae, Picrodendraceae and Putranjivaceae (Tokuoka, 2007). The molecular phylogenetics within Euphorbiaceae sensu stricto is analyzed by Wurdack *et al.* (2005). They have reported the monophyletic group of all the major clades within this family. Moreover, Berry *et al*. (2005) have analyzed the molecular phylogenic of worldwide *Croton* based on the ITS and *trn*L-F

sequence data. They have also reported the monophyletic group of the genus *Croton*. Based on the parsimony tree obtained in this study, the branching pattern of the tree appears as one group. Therefore, it was suggested that these examined *Croton* species are closely related as a monophyletic group.

2.2 Agreement between the Morphological Classification and Molecular Phylogenetic

In this study, we examined not only the phylogenetic relationship of the genus *Croton*, but also determined the agreement between the morphological classification and the molecular phylogeny of the genus *Croton*. For the traditional classification, the genus *Croton* is classified into the tribe Crotoneae, one of the twelve tribes in the subfamily Crotonoidae. However, the description and relationships of the tribe Crotoneae have been considered to be relatively unstable (Webster 1975, 1994; Radcliffe-Smith, 2001). Webster (1993) has published the sectional synopsis of *Croton* consisting of 40 sections which provide hypotheses concerning groups of related species. However, it contains only some exemplars, not all taxa of the genus *Croton* that are explicitly placed within the specific sections. The most recent attempt at the sectional synopsis of *Croton* has been performed by Govaert (2000). He has assigned the section of *Croton* into 40 sections similar to the assignment by Webster (1993). However, the section assignment by Govaert (2000) is more complete as it includes all the *Croton* species. Table 5.1 shows comparison between the section of *Croton* which assigned by Webster (1993) and by Govaert (2000). To determine the agreement between the proposed morphological characters and our molecular phylogeny, the section assignment of Govaert (2000) was used. Figure 5.1 shows the parsimony tree of the combined sequence data comparing with their taxonomically classification. It can be seen that the examined *Croton* species are fallen into three clades (clade I-III) and all clades are bootstrapping supported higher than 95 %. The *Croton* species in clade I and II are assigned in sections of cascarilla and argyrocroton. Based on the morphological data, the *Croton* species in clade I and II are overlapping into two sections. This is corresponded to the molecular phylogeny that showed the clade I and II closely related with 10 in branch length.

For the species of *C. argyratus*, *C. crassifolius*, *C. tiglium*, *C. kongensis* and *C. cascarilloides* in clade III, they are assigned in different sections. It means that these five *Croton* species have different morphological characters. Their genetic distances are reflected as the branch length that longer than those in clade I and II. Moreover, *C. bonplandianus* is only one species that located far from other *Croton* species with the long branch length. This assignment corresponds with morphological characters of *C. bonplandianus* that mostly differed from the other *Croton* species with its herb and small indumentum appearances.

Table 5.1 The section assignments of *Croton* species.

Assignment in phylogenetic tree	Species	Section	
		Webster (1993)	**Govaert (2000)**
Clade I	*Croton columnaris* Airy Shaw	Cascarilla	Cascarilla
	C. stellatopilosus H. Ohba	ND.	New section 41?
	C. roxburghii N.P.Balakr.	Argyrocroton	Argyrocroton
	C. wallichii Mull.Arg.	ND.	Cascarilla or new section 41?
	C. acutifolus Esser	ND.	Cascarilla?
	C. robustus Kurz	ND.	Argyrocroton?
	C. decalvatus Esser	ND.	Cascarilla?
Clade II	*C. kongkandanus*	ND.	ND.
	C. poilanei Gagnep	ND.	Argyrocroton
	C. kongensis Gagnep.	Cascarilla + Argyrocroton	Cascarilla + Argyrocroton
	C. argyratus Blume	Eutropia	Lamprocroton
Clade III	*C. crassifolius* Geiseler	Croton	Corylocroton
	C. tiglium L.	ND.	Tiglium
	C. cascarilloides Raeusch.	Argyroglossum	Argyroglossum
	C. bonplandianus Baill.	ND.	Cascarilla?

ND. : No data

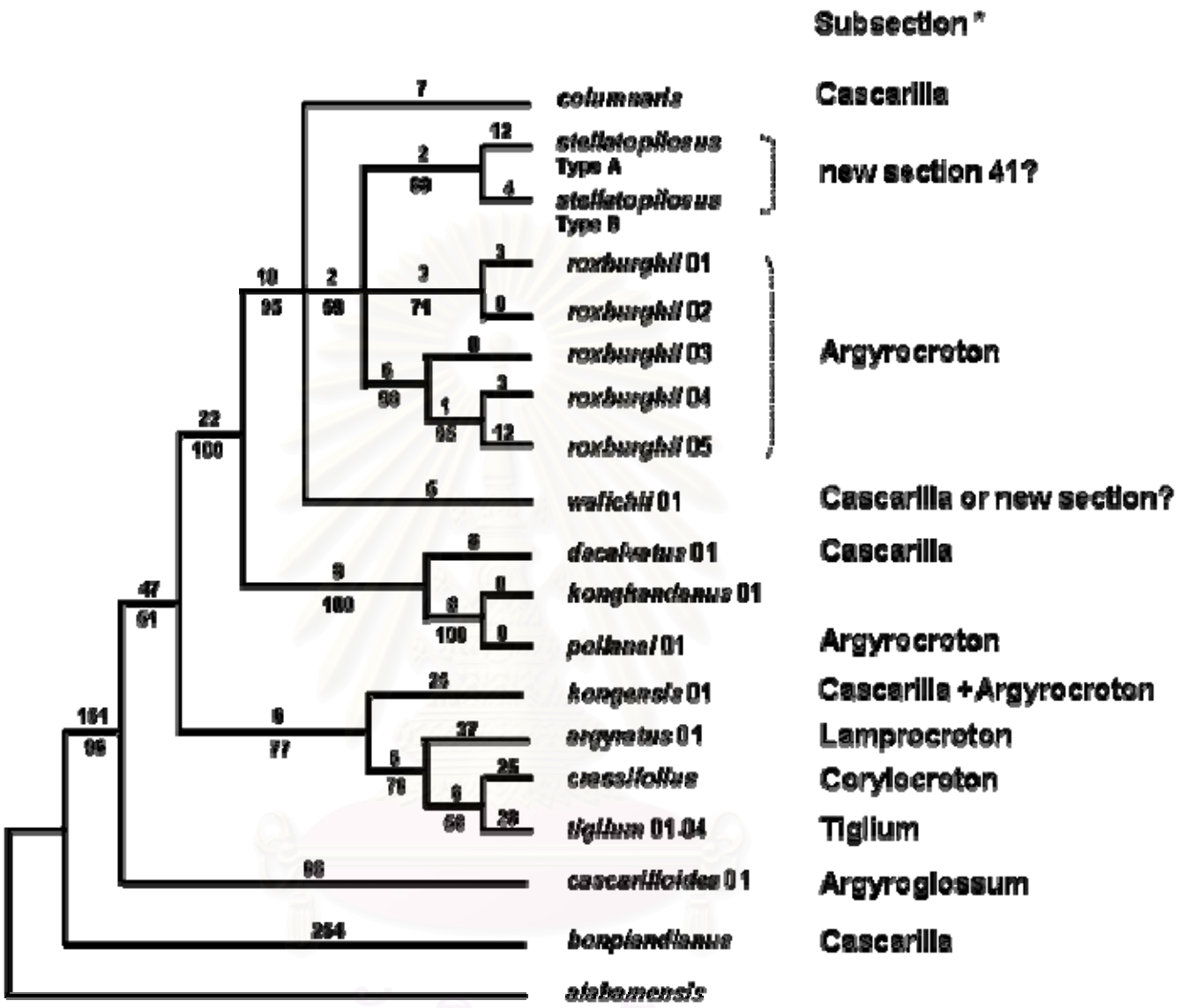
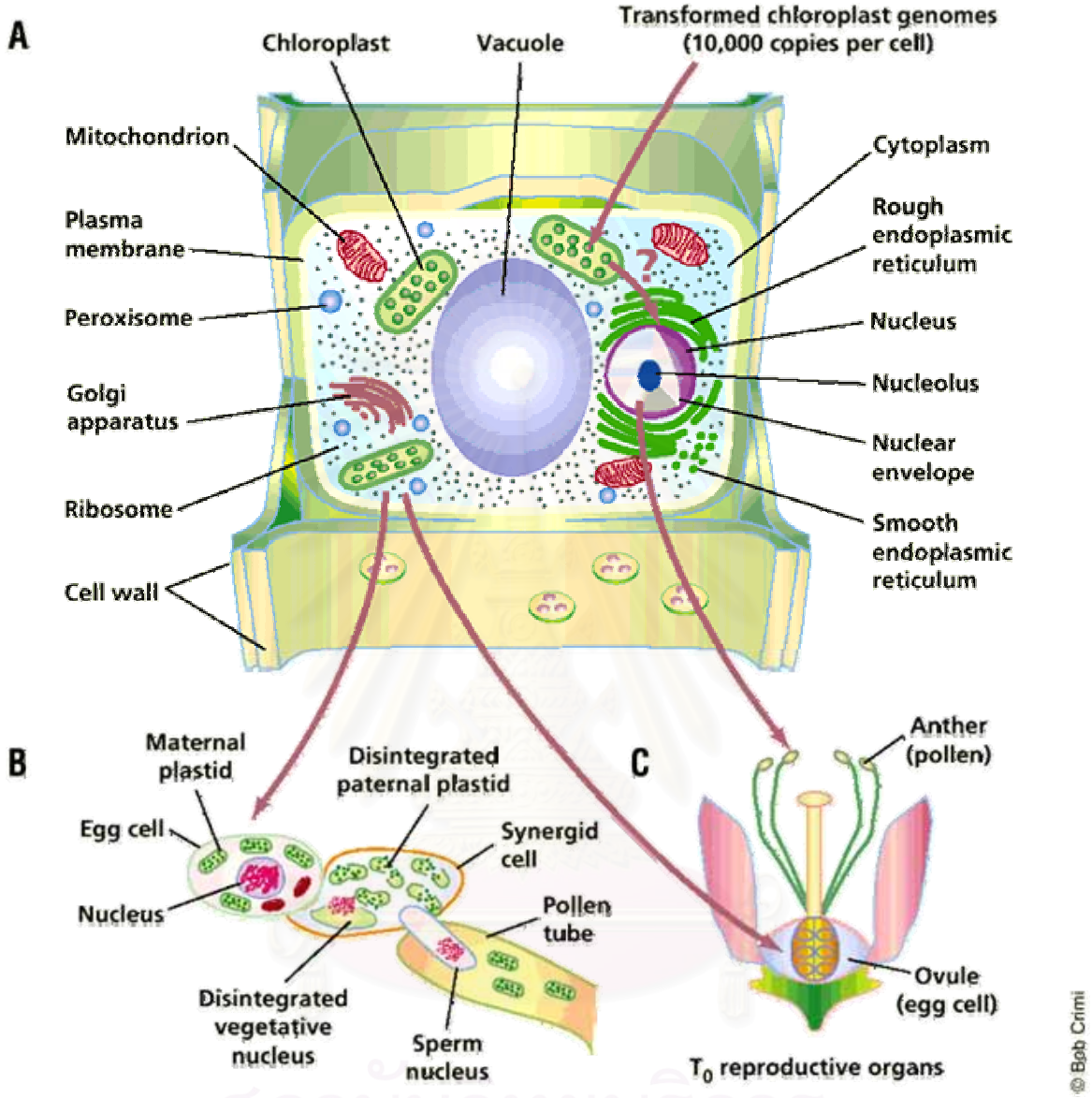


Figure 5.1 The parsimony tree of the combined sequence data comparing with the taxonomic classification of various *Croton* species. The numbers under the branches are bootstrapping percentages ≥50 % and the numbers above the branches are branch length. *Croton alabamensis* is assigned as an outgroup. The right side is the *Croton* section assignment of Govaert, 2000.

3. Hybridization of *C. stellatopilosus*

3.1 Hybridization and Hybrid Origin of *C. stellatopilosus*

Hybridization is a widespread phenomenon that has markedly contributed to diversity and speciation in the plant kingdom. In literature, hybridization is mostly based on the binary lineage (Kapan and Fehrer, 2007). The term “hybrid” can be restricted to organisms formed by cross fertilization between individuals of different species, or it can be defined more broadly as the offspring between individuals from populations which are distinguishable on the basis of one or more heritable characters (Rieseberg, 1997). Thus, this study prefers the broader definition of hybrids as it provides greater flexibility in usage. In general, a plant cell shows three compartments that contain DNA: nuclear, chloroplast, and mitochondrial genomes. The nuclear genome transfers genetic data to the next generation follow as a rule of inheritance (Fig. 5.2). In contrast, both mitochondrial and chloroplast genomes have their own special mode of inheritance. Figure 5.2 explains the maternal inheritance in plants. For the reproductive organs of plants, anthers produce pollen while ovules contain egg cells. During meiosis, haploid egg and sperm cells are formed. Fertilization begins when the pollen tube enters the synergid cell. Once inside the cytoplasm of the synergid cell, the pollen tube ruptures releasing its contents. The paternal chloroplasts are disintegrated. Only the sperm nucleus enters the egg cell and fuses with the egg to form zygote. Thus, the chloroplast DNA generally shows strictly maternal inheritance while the nuclear DNA is paternally inherited (Birky, 1976).



(This image was retrieved from http://www.nature.com)

Figure 5.2 The maternal inheritance in plants.

(A) The three genomes of plants that contain DNA: nuclear, chloroplast, and mitochondrial genomes. (B) Egg and sperm cells fertilization. (C) The reproductive organs, anthers and ovule.

It has been proposed that concerted evolution is the main factor that causes the molecular process of DNA sequence homogenization among different loci within multigene families (Arnheim *et al.*, 1980; Dover, 1982; Avise, 1994). This homogenization is driven by two molecular processes: gene conversion and unequal crossing over (Arnheim, 1983). The ITS region is multiple copies in the order of thousands of copies per cell and it arranges in one to several arrays of tandem repeats. Relative homogeneity of the ITS sequences across multiple copies of rDNA is maintained through concerted evolution which involves such processes as gene conversion and unequal crossing over (Arnheim, 1983). More than one sequence pattern is maintained in a genome (O'Kane *et al.*, 1996; Widmer and Baltisberger, 1999; Hughes *et al.*, 2002; Koch *et al.*, 2003). The occurrence of these polymorphisms indicating intragenomic variation can, therefore, provide evidence for past hybridization events and allow the tracing of the taxa and genomes involved (Sang *et al.*, 1995). Moreover, the two main outcomes that concert evolution produces on ITS sequences are additive patterns from both parental lineages and biased homogenization toward one of the parental lineages. The predominance of either one depends on the intensity and direction of concerted evolution, which is proposed to act through the recombination mechanisms occurring during meiosis (Li, 1997).

This study has found clearly the hybridization evidence of *C. stellatopilosus*. The ITS nucleotide sequences of *C. stellatopilosus* appear to have two groups of distinct non-nucleotide additive sequences (STEL Type A and Type B) and a group of putative hybrid with nucleotide additive sequences. In general, these additive sequences are therefore, cloning and sequencing of the individual PCR clone. They reveal the parental origin of the hybridization event when sequences matching each parental species are recovered. This individual PCR clone sequencing technique is considered ideally to identify the hybrid origin of *C. stellatopilosus*. For each *C. stellatopilosus* sample, the highest number of nucleotide additive sites is fifteen. Thus, the number of possible individual sequences for this sample is as large as 2^{15} patterns. Routine sequencing of a the large number of individual PCR clones from each *C. stellatopilosus* sample is expensive and labor intensive. Alternatively, Sang *et al*. (1995) have detected the

hybridization based on the ITS sequence data. Without sequencing the individual PCR clone, they have successfully proposed the hybrid origin of *Paeonia* by considering the combination patterns of nucleotide additive sites. This hypothesis of *Paeonia* hybrid origin has been further confirmed with the nucleotide sequences of the chloroplast DNA (Sang *et al.*, 1997). Similar to the work by Sang *et al*. (1995) for our case, the nucleotide additive sequences of *C. stellatopilosus* obtained in this study showed the perfect nucleotide addition at most sites that are variable between the two distinct non-nucleotide additive sequences. The additive sequences appear to locate not only at the ITS1, and ITS2, as well as in the 5.8S coding region. Although the 5.8S coding region is always conserved, there are few publications of the nucleotide additive sequences at this coding region (Campbell *et al.*, 1997; Fuertes Aguilar and Nieto Feliner, 2003). Therefore, our ITS sequence results strongly support the hybridization evidence of *C. stellatopilosus*. For the hybrids, different repeat units contributed by parental genomes can be rapidly homogenized towards one of the parental variant (Franzke and Mummenhoff, 1999; Fuertes Aguilar *et al.*, 1999) or produce recombinant ITS sequences (Wendel *et al.*, 1995; Mummenhoff *et al.*, 1997). In principle, there are three different ways in which two different ITS copies are evolved within a single individual (Koch *et al.*, 2003): 1) unidirectional concerted evolution leads to the loss of one copy and fixation of the second, as reported for *Gossypium gossypioides* (Wendel *et al.*,1995*a*) and Chinese *Yinshania* (Koch and Al-Shehbaz, 2000); 2) concerted evolution leads to a new ITS type that represents a mixture of the two original ITS sequences, as reported for *Gossypium* (Wendel *et al.*, 1995*b*), *Microseris* (Van *et al.*,1993), *Thalspi* (Mummenhoff *et al.*, 1997) and *Cardamine* (Franzke and Mummenhoff,1999); 3) both ITS copies are still present in a single individual which may be mostly the case in young hybridogenous taxa (Koch *et al*, 2003), as reported for Krigia (Kim and Jansen, 1994) and *Arabidopsis suecica* (O'Kane *et al.*, 1996). According to our results, the ITS sequences of *C. stellatopilosus* fall into the third possibility. This combination pattern suggested that the hybridization of *C. stellatopilosus* may have occurred recently because of their complete homogenization of both parental species.

In addition, the biased homogenization toward one of the parental lineages is also analyzed based on the peak intensity of the nucleotide additive sites of the hybrid population. Based on the direct sequencing data, some caution in estimating the relative amounts of ITS parental types is needed. However, it can be assumed that polymorphic sites in the putative hybrids result from amplification of the ITS copies from both parents. Similarly, some estimation on the relative copy number can be made. The relative intensity of peak signals was considered following careful analysis of electropherograms with the forward and reverse strand. The possibility of sequencing error cannot be excluded. However, each PCR fragment was read both from direct and reverse primers and the patterns were replicated. Thus, the possibility of sequencing error is unlikely and the data are considered as a reliable estimation. Additionally, it was assumed that the number of ITS copies in the STEL Type B was higher than that of Type A parental lineages. Thus, it is clear that the hybridization of *C. stellatopilosus* is more extreme to the STEL Type B dominated.

In this study, the *trn*L-F gene and the complete *trn*K intron were also determined for their sequences to identify the maternal and paternal lineages of *C. stellatopilosus* hybridization. The complete *trn*K intron nucleotide sequences of the STEL Type B and the hybrid group appear to have the identical sequence, while the nucleotide sequences of the STEL Type A group have a distinct sequence. Although both the *trn*L-F gene and the complete *trn*K intron are part of chloroplast genome, all *C. stellatopilosus* samples are found to have their genetic variation only in the complete *trn*K intron. In principle, the *mat*K gene is denoted as a rapidly evolving gene (Lavin, 2005). Since the complete *trn*K intron includs in the *mat*K gene, their nucleotide sequences of all *C. stellatopilosus* samples show the genetic variation as they appear to have the two distinct sequences. Whereas the *trn*L-F gene shows their identical sequence. The results of the ITS and the complete *trn*K intron sequence data clearly indicate that the STEL Type B and Type A group are respectively, the maternal and paternal lineages of the *C. stellatopilosus* hybridization. The results of the nucleotide sequence analysis thus provide new insights into the taxonomical status of *C. stellatopilosus.* Moreover, the results of hybridization and the hybrid origin postulated in this study could well explain the previous study of

C. stellatopilosus genetic diversity by Klinbantoom (2004) which proposed that *C. stellatopilosus* varied as two genotypes. The STEL Type A and the STEL Type B samples of this study represent the two *C. stellatopilosus* genotypes.

3.2 Plaunotol content in the Population *C. stellatopilosus* Hybrids

The effect of hybridization on plant secondary metabolism is immensely variable (Orians, 2000). Qualitatively, hybrids may express all of secondary metabolites of the parental taxa, or may fail to express certain parental chemicals, or may express novel chemicals that are absent in each parent. Quantitatively, concentration of parental chemicals may vary markedly among hybrids. For *C. stellatopilosus,* plaunotol is a biological active compound. (Ogiso *et al.*, 1978 and 1985). Vongcharoensathit (1997) and Klinbantoom (2004) have reported that *C. stellatopilosus* populations have different plaunotol content. Both previous studies suggested that the genetic background is the major factor effecting the plaunotol expressions. The ages and localities of *C. stellatopilosus* are not correlated with their plaunotol expression (Vongcharoensathit, 1997). In this study, the *C. stellatopilosus* population is investigated for both the genetic variation and the plaunotol producing aspects. It was found that the six *C. stellatopilosus* groups of the maternal, (the Type A sequence > the Type B hybrid, the Type A sequence ≈ the Type B hybrid, the Type B sequence > the Type A hybrid, the Type B sequence >>the Type A hybrid and the paternal groups which are assigned based on their ITS and the complete *trn*K intron sequences) contain different plaunotol contents. Figure 5.3 compares the individual plaunotol contents in each *C. stellatopilosus* group. It can be seen that plaunotol content in the paternal group is lowest. In contrast, the *C. stellatopilosus* samples in the maternal group appear to contain high plaunotol content. In terms of biosynthetic pathway of plaunotol (Fig. 5.4), in principle, plaunotol starts from geranylgeranyl diphosphate (GGPP), a general precursor of diterpene compounds. GGPP phosphatase then cleaves the phosphate group from GGPP yielding geranylgeraniol (GGOH). Finally, GGOH is hydroxylated at C-18 by geranylgeraniol 18-hydroxylase, affording plaunotol (Tansakul and De-Eknamkul, 1998). Although the ITS region and the complete *trn*K intron are not the genes that are involved in the

biosynthetic pathway of plaunotol. The results of this study show that the genetic variations on both regions correspond with the plaunotol contents in the *C. stellatopilosus* leaves.

3.3 Taxonomical Status of *C. stellatopilosus* Type A and Type B

At present, the taxonomical status of *C. stellatopilosus* Type A and Type B is not clear. Based on the phylogenetic tree reconstructed based on the combined data between the ITS and *trn*L-F sequences, both types are clearly related. However, genetic evidence only is not sufficient to evaluate their definite taxonomical status. More pieces of evidence are needed such as morphological characters and geographical distribution. However, as the term of "hybridization" can be defined as the offspring between individuals from populations which are distinguishable on the basis of one or more heritable characters (Rieseberg, 1997), the taxonomical status of STEL Type A and Type B may be explained by two possibilities. First, STEL Type A and Type B are considered to be two different *Croton* species. Second, *C. stellatopilosus* is a single species with two varieties of the STEL Type A and Type B.

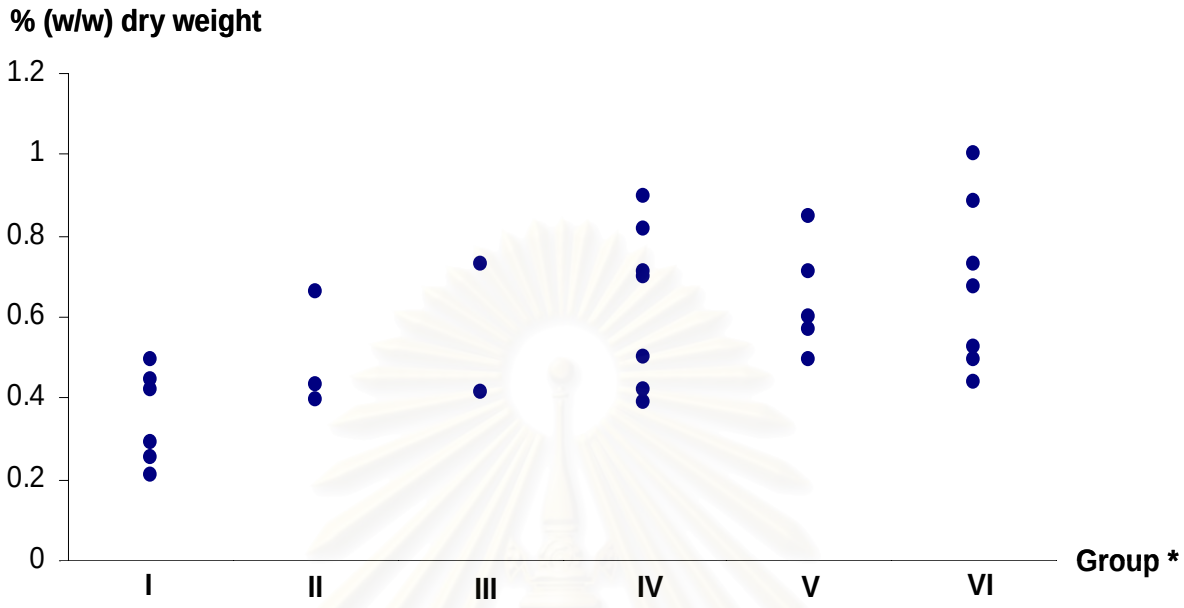


Figure 5.3 Plaunotol content of the maternal, paternal and each hybrid groups of *C. stellatopilosus.*

I : the paternal group (the Type A)

II : the hybrid group that the Type A sequence > the Type B

III : the hybrid group that the Type A sequence ≈ the Type B

IV : the hybrid group that the Type B sequence > the Type A

V : the hybrid group that the Type B sequence >> the Type A

VI : the maternal group (the Type B)

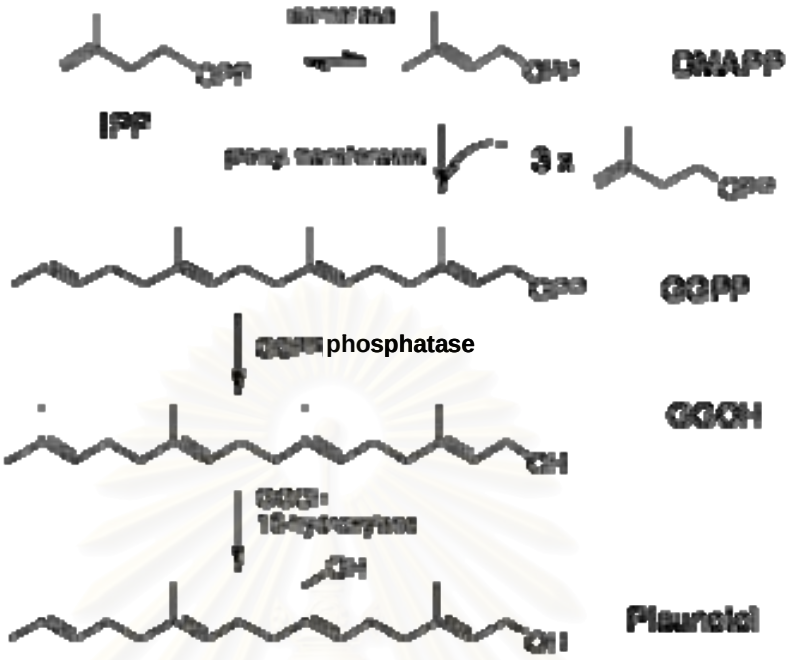


Figure 5.4 Biosynthetic pathway of plaunotol.

4. Genetic Variation of *C. roxburghii*

From the previous genetic study by Chonvanich (2004), *C. roxburghii* is clearly separated into one major and one minor group. For the major group, some *C. roxburghii* samples are clustered as a small clade. Based on the study on the nucleotide sequences of the ITS region, *trn*L-F gene and the partial *trn*K intron, the five *C. roxburghii* samples can be separated into two groups, one with the sample codes ROX 01 and ROX 02 and the others with codes ROX 03, ROX 04 and ROX 05. The results also suggest that the examined *C. roxburghii* are extremely genetic varied. Soontornchainaksaeng *et al*. (2003) have shown that the Thai *Croton* species have quite unique as 2*n*=20, but the chromosome number of *C. roxburghii* is varied. Bedi *et al*. (1980) and Gill *et al.* (1981) have reported the chromosome number of *C. roxburghii* as 2*n*=20, while Sarker *et al.* (1972) has reported to have 2*n*=22. Phytochemical studies have been shown that *C. roxburghii* plants from different localities have different chemical profiles (Roengsumran *et al.*, 1998a, 1998b, 1999a, 1999b, 2001, 2003; Ruangweewat, 2001;

Singtothong, 1999; Siriwat, 2000; Sirimongkhon, 2000; Sriyangnok, 2000). Both the chromosome number and chemical profiles support the genetic variation of *C. roxburghii.* However, to obtain a conclusion, a sufficient number of samples from different localities are requiring for the study so that the variations of the *C. roxburghii* can be analyzed.

5. The Use of Herbarium Specimens in Molecular Studies

In the past decade, molecular studies based on nucleotide sequence comparison have been extended progressively. However, an important difficulty is the collection of living plant samples. Most interesting taxa from the point of view of evolution are often rare and geographically restricted. Thus, herbarium specimens potentially represent an invaluable source of material for molecular studies. In this study, we used thirty six herbarium samples of twenty *Croton* species as plant materials. Among these, we successfully amplified on the ITS region and *trn*L-F gene in sixteen (30.56 %) and five (13.89 %) herbarium samples, respectively. It was found that herbarium samples that could be amplified should have their age below 30 years. The samples above this age could absolutely not be amplified. Although some previous works reported amplified products from the ancient samples, only short target fragments about 200 bp long were obtained. For long target fragments, amplifications were often unsuccessfully. Moreover, the success of PCR amplification of DNA from herbarium samples depends on many factors such as chemical particularities of the species, developmental stage of the collected herbarium specimens, drying method, duration and conditions of herbarium preparation.

Initially, rapid desiccation is the primary factor that limited the extent of senescence processes (Chase and Hills, 1991). Generally, herbarium samples are not dried as rapidly as in silica gel. This could be one of the reasons for the degraded total DNA. In addition, the breaking of cellular compartments during the drying stage may liberate nucleases, producing endogenous hydrolytic damage. The electron microscope study shows a good correlation between the state of degradation of the chloroplast

membranes and the success of the total DNA preparation and amplification (Lindahl, 1993). However, there is little literature about the modifications of DNA of herbarium specimens stored over long periods of time. The herbarium tissues also remain subject to oxidative damages (Eglington and Logan, 1991). Gano and Poinar (1993) observed the poor conservation of DNA extracts from herbarium sheets. This may be the consequence of interactions of some herbarium-specific chemical components co-extracted with DNA. Thus, the oxidative DNA damages occurring during herbarium specimen preparation may be reactivated during the extraction procedures (Paabo, 1989). Some chemicals used in herbarium disinfection or collection could interfere with DNA extraction and amplification (Coradin and Giannasu, 1980). However, the DNA preparation and amplification procedures may be optimized in order to overcome the low quality and/or quantity of the DNA remaining in herbarium samples. Our results show that the yield of DNA extraction and amplification can be increased by extracting total DNA with the modified protocol and then purifying the total DNA extracts. However, such procedures were not successful with all samples. The extremely low amount of total DNA in certain herbarium samples is problematic. In this study, amplification of smaller target fragments has indeed produced positive results as compared with the larger fragment. Moreover, using the nested technique, some of herbarium samples were successful amplified. Based on these results, herbarium specimens could be used in molecular studies, but this is far from being routine.

In addition, the herbarium specimens that have been deposited in herbaria are more accessible than the plant materials in natural habitat, especially the rare species. But those herbarium specimens may lack the molecular materials. If herbarium specimens contain sufficient amount of DNA, they would be great useful. Herbarium specimens containing enough genetic materials would be useful to a wide range of researchers. In general, the taxonomist collects plant specimens for their morphological study not for molecular study. Thus DNA are damaged. In contrast, the molecular biologist collects plant samples by placing the leaf part in silica gel to preserve the genetic materials. However, this collecting method is not suitable for large scale collection because it requires a large amount of silica gel. Recently, Tsukaya (2003)

reported the used of the FTA Cards® (Whatman Company) for collecting DNA without any special equipment. The FTA Cards® are paper cards that immobilize nucleic acids. Leaf samples are simply rubbed or squashed on the sheets and allow them to dry at room temperature. Dried cards can be stored indefinitely at room temperature without special care. The card can be washed, dried and used successfully as PCR templates. With this approach, herbarium specimens can be new source of genetic materials for molecular studies.

CONCLUSION

1. Fifteen *Croton* species in Thailand were examined for their nucleotide sequences on the ITS region of nuclear DNA and *trn*L-F gene of chloroplast DNA. The results showed that genetic variations of the ITS region were mostly from substitutions. The *trn*L-F gene of all *Croton* species was genetic varied mainly from insertions and deletions. Moreover, *C. stellatopilosus* and *C. roxburghii* were further investigated their genetic variations on the *trn*K intron of chloroplast DNA.

2. Based on the obtained ITS sequences, *C. stellatopilosus* had two individual sequences designated as Type A and Type B. Also, various nucleotide additive sequences were found to be obtained from these two types were designated as “putative hybrids”.

3. Both the Type B and Putative hybrids of *C. stellatopilosus* showed identical *trn*K intron sequences whereas the Type A had different sequences. The results suggested that the Type A and Type B were respectively, the paternal and maternal inheritances of *C. stellatopilosus* hybridization.

4. *C. roxburghii* showed genetic variation with its ITS, *trn*L-F and *trn*K intron sequences differing from one another. Therefore, *C. roxburghii* was considered to have highly genetic variation

5. Maximum parsimony tree was reconstructed from ITS, *trn*L-F and the combined sequences data. The results showed that topologies of the obtained parsimony trees were

congruent. The parsimony tree reconstructed from the combined sequence data was the best tree for estimating the phylogenetic of all *Croton* species. .

6. Phylogenetic relationships of all fifteen *Croton* species was shown to be monophyletic type and correspond with the taxonomic classification.

In conclusion, in this study the technique of direct sequencing was successfully used to elucidate the phylogenetic relationships of *Croton* species and to identify the definite sites of genetic variation for the important *Croton* species, *C. stellatopilosus* and *C. roxburghii.*

REFERENCES

Adachi, J. and Hasegawa, M. 1995. MOLPHY: Programs for molecular phylogenetics version 2.3. Institute of Statistical Mathematics:Tokyo.

Airy Shaw, H. K. 1972. The Euphorbiaceae of Siam. Kew Bull. 26: 191-363.

Baldwin, B. G., Sanderson M. J., Porter, J. M., Wojciechowski, M. F., Cambell C. S. and Donoghue, M. J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Missouri. Bot. Gard. 82: 247-277.

Begun, D. J. and Aquadro, C. F. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. Nature 356: 519-520.

Bellarosa, R., Simeone, M. C., Papini, A. and Schirone, B. 2005. Utility of ITS sequence data for phylogenetic reconstruction of Italian *Quercus* spp. Plant Phyl. Evol. 34: 355-370.

Benjamin, W., Van, E. E., Jelinski, N., Berry, P. E. and Hipp, A. L. 2006. Phylogeny and biogeography of *Croton alabamensis* (Euphorbiaceae), a rare shrub from Texas and Alabama using DNA sequence and AFLP data. Mol. Ecol. 15: 2735-2751.

Berry, P. E., Hipp, A. L., Wurdack, K. J., Van, B. E. and Riina, R. 2005. Molecular phylogenetics of the giant genus *Croton* and tribe Crotonae (Euphorbiaceae sensu stricto) using ITS and *trn*L-*trn*F DNA sequence data. Am. J. Bot. 92(9): 1520-1534.

Birky, C. M. 1976. The Inheritance of Genes in Mitochondria and Chloroplasts BioScience 26(1): 26-33.

Borsch, T., Hilu, K. W., Quandt, D., Wilde, V., Neinhuis, C. and Barthlott, W. 2003. Non coding plastid *trn*T-*trn*F sequences reveal a well resolved phylogeny of basal angiosperms. J. Evol. Bio. 16: 558-576.

Bult, C. J., Kallersjo, M. and Suh, Y. 1992. Amplification and sequencing of 16/18S rDNA from gel-purified total plant DNA. Plant Mol. Bio. Rep. 10: 273-284.

Campbell, C. S., Wojciechowski, M. F., Baldwin, B. G., Alice, L. A. and Donoghue, M. J. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* Agamic complex (Rosaceae). Mol. Biol. Evol. 14(1): 81-90.

Cavalli-Sforza, L. L. and Edwards, A. W. F. 1967. Phylogenetic analysis: Models and estimation procedures. Am. J. Hum. Genet. 19: 233-257.

Chase, M. W. and Hills, H. H. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. Taxon 40: 215-220.

Chase, M. W., Zmarzty, S., Lledo, M. D, Wurdack, K. J, Swensen, S. M. and Fay, M. F. 2002. When in doubt, put it in Flacourtiaceae: a molecular phylogenetic analysis based on plastid *rbc*L DNA sequences. Kew Bull. 57:141–181.

Chayamarit, K. and Welzen, P. 2005. Euphorbiaceae genera A-F. Flora of Thailand, Santisuk, T., and Larson, K. (Editor). The Forest Herbarium, National Park Wildlife and Plant Conservation Department, Bangkok. 8: 220-221.

Chieba, T., Harada, T., Goto, S., Ishikawa, R. and Niizeki, M. 1996. Transcription of tRNA genes from large-scale plastid DNA deletion clearly reveals the action of nuclear encoded RNA polymerase in the plastid. J. Plant Physio. 148: 652-656.

Chonvanich, O. 2004. ISSR analysis of *Croton roxburghii* in Thailand. Master's thesis, Department of Biotechnology, Faculty of Science, Chulalongkorn University.

Coradin, L., Giannasi, D. L. 1980. The effects of chemical preservations on plant collections to be used in chemotaxonomic surveys. Taxon 29: 33-40.

Davis, C. C., Webb, C. O., Wurdack, K. J., Jaramillo, C. A., Donoghue, M. J. 2005. Explosive radiation of Malpighiales supports a Mid-Cretaceous origin of modern tropical rain forests. Am. Nat. 165:36–65.

Downie, S. R. and Palmer, J. D. 1992. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In Soltis, P. S., Soltis, D. E. and Doyle, J.J. (eds.), Molecular Systematics of Plants. Chapman & Hall, New York.

Eglington, A., Logan, G. A. 1991. Molecular preservation. Philos. Trans. Roy. Soc.B 333: 315-328.

Ellstrand, N. C., Whitkus, R. and Rieseberg, L. H. 1996. Distribution of spontaneous plant hybrids. Proc. Natl. Acad. Sci. U S A. 93: 5090–9329.

Esser, H. J. and Chayamarit, K. 2001. Two new species and a new name in Thai *Croton* (Euphorbiaceae). Thai For. Bull. (Bot.) 29: 51-57.

Esser, H. J. 2002. Novelties in *Croton* (Euphorbiaceae) from Southeast Asia. Novon 12: 42-46.

Fehere, J., Gemeinholzer, B., Chrtek, J. and Brautigam, S. 2007. Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in *Pilosella* hawkseeds (*Hieracium*, *Cichorieae*, Asteraceae). Mol. Phyl. Evol. 42: 347-361.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17(6): 368-376.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791.

Felsenstein, J. 1995. PHYLIP (Phylogenetic inference package) version 3.57. Department of Genetics, University of Washington:Washington.

Fuertes, A., Rossell, J., Nieto, J. A. and Feliner, G. 1999. nrDNA concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Mol. Ecol. 8: 1341-1346.

Fuertes Aguilar, J. and Nieto Feliner, G. 2003. Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). Mol. Phyl. Evol. 28: 430-447.

Fukuda, T., Yokoyama, J. and Ohashi H. 2001. Phylogeny and biogeography of the genus *Lycium* (Solanaceae): inferences from chloroplast DNA sequences. Mol. Phyl. Evol. 19(2): 246-258.

Gano, R. J., Poinar, H. N. 1993. Rapid isolation of DNA from fossil and museum specimens suitable for PCR. Biotechniques 15: 432-435.

Gaut, B. S. 1998. Molecular clocks and nucleotide substitution rates in higher plants. Evol. Biol. 30: 93-120.

Gielly, L. and Taberlet, P. 1994. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences. Mol. Bio. Evol. 11(5): 769-777.

Goldenberg, E. M., Giannasi, D. E., Clegg, M. T., Smiley, C. J., Durbin, M., Henderson, D. and Zurawski, G. 1990. Chloroplast DNA sequence from a Miocene *Magnolia* species. Nature 344: 656-658.

Gottlieb, L. D. Leaves of confidence in the analysis of hybridization in plants. Ann. Mo. Bot. Gard. 59: 435-446.

Govaerts, R., Frodin, D. G. and Radcliffe-Smith, A. 2000. World checklist and bibliography of Euphorbiaceae (with Pandaceae). Royal Botanic Garden, Kew.

Gurgel, L. A., Sidrim, J., Domingos, T., Martins, V., Filho, C. and Rao, S. 2005. In vitro antifungal activity of dragon's blood from *Croton urucurana* against dermatophytes. J. Ethnopharmaco. 77(2): 409-412.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series. 41: 95-98.

Hamzeh, M. and Dayanandan, Y. 2004. Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast *trn*T-*trn*F region and nuclear rDNA. Am. J. Bot. 91(9): 1398–1408.

Han, A. S. 1973. Chromosomal conspectus of the Euphorbiaceae. Taxon 22: 591-636.

Hendy, M. D. and Penny, D. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38: 297-309.

Hilu, K. W. and Liang, H. 1997. The *mat*K gene: sequence variation and application in plant systematics. Am. J. Bot. 84: 830-839.

Hughes, C. E., Bailey, C. D. and. Harris, S. A. 2002. Divergent and reticulate species relationships in *Leucaena* (Fabaceae) inferred from multiple data sources: insights into polyploid origins and nrDNA polymorphism. Am. J. Bot. 89: 1057–1073.

Kim, K. J. and Jansen, R. K. 1994. Comparisons of phylogenetic hypothesis among different data sets in dwarf dandelions (*Krigia*): Additional information from internal transcribed spacers sequences of nuclear DNA. Plant Syst. Evol. 190: 157–185.

Klinbantoom, R. 2004. ISSR analysis of *Croton sublyratus* Kurz. in Thailand. Master's thesis, Department of Biotechnology, Faculty of Science, Chulalongkorn University.

Koch, M., Dobes, C. and Mitchell-Olds, T. 2003. Multiple hybrid formation in natural populations: concerted evolution of the internal transcribed spacer of nuclear

ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). Mol. Bio. Evol. 20: 338-350.

Johnson, L. A. and Soltis, D. E. 1994. *mat*K DNA sequences and phylogenetic reconstruction in *Saxifragaceae* s.s. Sys. Bot. 19: 143-156.

Judd, W. S., Campbell, C. S., Kellogg, E. A. and Stevens, P. F. 1999. Plant Systematics, A Phylogenetic Approach. Sinauer Associates, Inc. USA.

Jukes, T. and Cantor, C. 1969. Evolution of protein molecules. In Mammalian Protein Metabolism. Academic Press:New York.

Kaplan, Z. and Fehrer, J. 2007. Molecular evidence for a natural primary triple hybrid in plants revealed from direct sequencing. Ann. Bot. 99: 1213-1222.

Katsuko, K., Zhu, S., Fushimi, H., Qui, T. K., Cai, S. and Kodota, S. 2001. Phylogenetic analysis based on 18S rRNA gene and *mat*K gene sequences of *Panax vietnamensis* and five related species. Planta Med. 67: 461-465.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111-120.

Kiss, T., Kis, M., Abel, S. and Solymosy, F. 1988. Nucleotide sequence of the 17S-25S spacer region from tomato rDNA. Nucleic Acids Res. 16: 7179.

Kita, Y. and Ito, M. 2000. Nuclear ribosomal ITS sequences and phylogeny in East Asian *Aconitum* subgenus *Aconitum* (Ranunculaceae), with special reference to extensive polymorphism in individual plants. Plant Syst. Evol. 225: 1-13.

Klotzsch, J. F. 1841. Neue und weniger gekannte sudamerikanische Euphorbiaceen-Gattungen. Archiv fur Naturgeschichte. 7: 175-200.

Kocyan, A. Qiu, Y. L., Endress, P. K. and Conti, E. 2004. A phylogenetic analysis of Apostasioideae (Orchidaceae) based on ITS, *trn*L-F and *mat*K sequences. Plant Syst. Evol. 247: 203-213.

Kondo, K., Shiba, M., Yamaji, H., Morota, T., Zhengmin, C., Huixia, P. and Shiyama, Y. 2007. Species identification of licorice using nrDNA and cpDNA genetic markers. Biol. Pharm. Bull. 30(8): 1497-1502.

Larget, B. and Simon, D. 1999. Markov chain monte carlo algorithms for the bayesian analysis of phylogenetic trees. Mol. Biol. Evol. 16: 750-759.

Lavin, M., Herendeen, P. and Wojciechowski, M. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. Sys. Biol. 54(4): 575-594.

Leon, C. 1945. History and Nomenclature of the Higher Units of Classification. J. Torrey Bot. Soc. 72(1): 52-75.

Li, W. H. 1997. Molecular Evolution. Sinauer, Sunderland, Massachusetts.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. Nature 362: 709-715.

Linder, C. R., Moore, L. A. and Jackson, R. B. 2000. A universal molecular method for identifying under ground plant parts to species. Mol. Ecol. 9: 1549-1559.

Lio, P. and Goldman, N. 1998. Models of Molecular Evolution and Phylogeny. Genome Res. 8: 1233-1244.

Lio, P., Goldman, N., Thorne, J.L. and Jones, D.T. 1998. PASSML: Combining evolutionary inference and protein secondary structure prediction. Bioinformatics 14: 726–733.

Marhold, K., Lihova, J., Perny, M., Grupr, R. and Neuffer, B. 2002. Natural hybridization in *Cardamine* (Brassicaceae) in the Pyrenees: evidence from morphological and molecular data. Bot. J. Linn. Soc. 139: 275-294.

Michener, C. and Sokal, R. 1957. A quantitative approach to a problem in classification. Evolution 11:130-162.

Mischler, B. D., Thrall, P. H., Hopple J. S., Luna, E. and Vilgalys, R. 1992. A molecular approach to the phylogeny of bryophytes: cladistic analysis of chloroplast encoded 16S and 23S ribosomal RNA genes. Bryologist. 95: 172-180.

Mizukami, H., Shimizu, R., Kohjyouma, M., Kohda, H., Kawanishi, F. and Hiraoka, N. 1998. Phylogenetic analysisof Atractylodes plants based on chloroplast *trn*K sequences. Biol. Pharm. Bull. 21: 474-478.

Muller, J. 1866. Euphorbiaceae. In Candolle, A. (ed.), Prodomus sytematis naturalis regni vegatabilis. 15: 189-1261.

Muller, J. 1873. *Croton*. In Martius, C. F. P. (ed.), Flora brasiliensis. 11: 81-274.

Muller, K. F., Borsch, T. and Hilu, K. W. 2006. Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: Contrasting *mat*K, *trn*T-F, and *rbc*L in basal angiosperms. Mol. Phyl. Evol. 41: 99–117.

Mullis, K. B. and Faloona, F. A. 1987. Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. Methods Enzymol. 155: 335-350.

Nelson-Jones, E. B., Briggs, D. and Smith, A. G. 2002. The origin of intermediate species of the genus *Sorbus*. Theo. App. Gene. 105: 953-963.

Nue, R. W., Rientjes, M. J., Sande, C. A. F. M., Zerp, S. F., Sluiter,C., Venema, J., Planta, R. J. and Raue, H. A. 1994. Separate structural elements within internal transcribed spacer 1 of *Saccharomyces cerevisiae* precursor ribosomal RNA direct the formation of 17S and 26S rRNA. Nucleic Acids Res. 22: 912-919.

Ogiso A., Kitazawa E., Mikuriya I. and Promdej C. 1978. Isolation and structure of antipeptic ulcer diterpene from Thai medicinal plant. Chem.Pharm.Bill. 26: 3117-3123.

Ogiso A., Kitazawa E., Kobayashi, S., Komai, T., Matsunuma, N. and Kataumi, S. 1985. Plaunotol (CS-684), a new antipeptic ulcer agent. Sankyo Kenkyusho Nempo. 37: 1-39.

Okaura, T. and Harada, K. 2002. Phylogeographical structure revealed by chloroplast DNA variation in Japanese beech (*Facus crenata* Blume). Heredity. 88: 322-329.

Orians, C. M. 2000. The effects of hybridization in plant on secondary chemistry: implications for the ecology and evolution of plant-herbivore interactions. Am. J. Bot. 87(12): 1749-1756.

Oxelman, B., Liden, M. and Berglund, D. 1997. Chloroplast *rps*16 intron phylogeny of the tribe *Sileneae* (*Caryophyllaceae*). Plant Sys. Evol. 206: 393-410.

Paabo, S. 1990. Amplifying ancient DNA. In Innis, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J.(eds.), PCR protocols: a guide to methods and applications. 315-322. Academic Press:New York.

Padgett, D., Les, D.H. and Crow, G.E. 1993. Evidence for the hybrid origin of *Nuphar X Rubrodisca* (Nymphareaceae). Am. J. Bot. 85:1468-1476.

Palmer, J. D., Jansen, R. K., Michaels, H. J., Chase, M. W. and Manhart, J. R. 1988. Chloroplast DNA variation and plant phylogeny. Ann. Missouri. Bot. Gard. 75: 1180-1206.

Palmer, J. D. 1991. Plastid chromosomes: structure and evolution. In Bogorad, L. and Vasil, I. K. (eds.), Cell Culture and Somatic Cell Genetics in Plants, 7: 5-53. The Molecular Biology of Plastids. Academic Press:New York.

Qiu, Y. L., Lee, J., Bernasconi-Quadroni, F., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A., Chen, Z., Savolainen, V. and Chase, M. W. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature 402: 404-407.

Quijada, A., Liston, A., Robinson, W. and Alvarez, E. 1997. The ribosomal ITS region as a marker to detect hybridization in pines. Mol. Ecol. 6: 995-996.

Rieseberg, L. H. 1997. Hybrid origins of plant species. Annu. Rev. Ecol. Syst. 28: 359-389.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16: 351-358.

Rieseberg, L. H. and Soltis, D. E. 1991. Phylogenetic consequences of cytoplastmic gene flow in plants. Evol. Trend Plant 5: 65-84.

Ritland, C. E., Ritland, K. and Straus, N. A. 1993. Variation in the ribosomal internal transcribed spacers (ITS1 and ITS2) among eight taxa of the *Mimulus guttataus* species complex. Mol. Bio. Evol. 10(6): 1273-1288.

Ritz, C. M., Schmuths, S. and Wissemann, V. 2005. Evolution by reticulation: European dogroses originated by multiple hybridization across the genus Rosa. Heredity 96: 4-14.

Rizk, M. 1987. The chemical constituents and economic plants of the Euphorbiaceae. Bot. J. Linn. Soc. 94: 293-326.

Roengsumran, S., Achayindee, S., Petsom, A., Pudhom, K., Singtothong, P., Surachetapan, C. and Vilaivan, T. 1998a. Two new cembranoids from *Croton oblongifolius* Roxb. J. Nat. Prod. 61: 652-654.

Roengsumran, S., Achayindee, S., Petsom, A., Pudhom, K., Singtothong, P., Surachetapan, C. and Vilaivan, T. 1998b. Two new cembranoids from *Croton oblongifolius* Roxb. J. Nat. Prod. 61: 652-654.

Roengsumran, S., Petsom, A., Sommit, D. and Vilaivan, T. 1999a. Labdane diterpenoids from *Croton oblongifolius* Roxb. Phytochemistry 50(3): 449-453.

Roengsumran, S., Singtothong, P., Pudhom, K., Ngamrochanavanich, N., Petsom, A. and Chaichantipyuth, C. 1999b. Neocrotocembranal from *Croton oblongifolius* Roxb. J. Nat. Prod. 62: 1163-1164.

Roengsumran, S., Petsom, A., Kuptiyanuwat, N., Vilaivan, T., Ngamrojnavich, C., Chaichantipyuth, C. and Phuthong, S. 2001. Cytotoxic labdane diterpenoids from *Croton oblongifolius* Roxb. Phytochemistry 56(1): 103-107.

Roengsumran, S.,Pornpakakul, S., Muangsin, N., Sangvanich, P., Nhujak, T., Sintothong, P., Chaichit, N., Puthong, S. and Petsom, A. 2003. New halimane diterpenoids from *Croton oblongifolius* Roxb. Planta Med. 7: 1-3.

Ruangweewat, N. 2001. Chemical constituents and biological activity from stem bark of *Croton oblongifolius* Roxb. from Amphoe Phurua, Loei Province. Master's degree, Department of Chemistry, Faculty of Science, Chulangkorn University.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4(4): 406-425.

Salatino, A., Maria, L., Salatino; F. and Negri, G. 2007. Traditional uses, chemistry and pharmacology of *Croton* species (Euphorbiaceae). J. Braz. Chem. Soc. 18(1): 11-33.

Sande, C. A. F. M., Kwa, M., Nue, R. W., Heerikhuizen, H. A., Raue, H. A. and Rientjes, Planta, R. J. 1992. Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. J. Mol. Biol. 223: 899-910.

Sang, T., Crawford, D. J. and Stuessy, T. F. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. Evolution 92: 6813-6817.

Sang, T., Crawford, D. J. and Stuessy, T. F. 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84(9): 1120-1136.

Sasaki, Y., Fushimi, H., Cao, H., Cai, S. Q. and Komatsu, K. 2002. Sequence analysis of Chinese and Japanese *Curcuma* drugs on the 18S rRNA gene and *trn*K gene and the application of amplification-refractory mutation system analysis for their authentication. Biol. Pharm. Bull. 25(12): 1593-9.

Savolainen, V., Morton, C. M., Hoot, S. B. and Chase, M. W. 1996. An examination of phylogenetic patterns of plastid *atp*ß gene sequences among eudicots. Am. J. Bot. (suppl.) 83: 541.

Schiebel, K. and Hemleben, V. 1989. Nucleotide sequence of the 18S-25S spacer region from rDNA of mung bean. Nucleic Acids Res. 17: 2852.

Shi, S., Zhong, Y., Huang, Y., Du, Y., Qiu, X. and Chang, H. 2002. Phylogenetic relationships of the Rhizophoraceae in China based on sequences of the chloroplast gene *mat*K and the internal transcribed spacer regions of nuclear ribosomal DNA and combined data set. Biochem. Syst. Eco. 30: 309-319.

Singtothong, P. 1999. Chemical and and biological activity of deterpenoids from *Croton oblongifolius* Roxb. Doctoral dissertation, Department of Chemistry, Faculty of Science, Chulangkorn University.

Siriwat, K. 2000. Chemical constituents and biological activity from stem bark of *Croton oblongifolius* Roxb. from Amphoe Dan Sai, Loei Province. Master's degree, Department of Chemistry, Faculty of Science, Chulangkorn University.

Sirimongkhon, S. 2000. Kuarane diterpened from stem bark of from stem bark of *Croton oblongifolius* Roxb. from Amphoe Kui Buri, Prachuap Khiri Khan Province. Master's degree, Department of Chemistry, Faculty of Science, Chulangkorn University.

Sneath, P. and Sokal, R. 1973. Numerical Taxonomy. Freeman:SanFrancisco.

Sriyangnok, S. 2000. Chemical constituents and biological activity from stem bark of *Croton oblongifolius* Roxb from Amphoe Pran Buri, Prachuap Khiri Khan Province. Master's degree, Department of Chemistry, Faculty of Science, Chulangkorn University.

Soltis, D. E., Kuzoff, R. K., Conti, E., Gornall, R. and Ferguson, K. 1996. *mat*K and *rbc*L gene sequence data indicate that *Saxifraga* (Saxifragaceae) is polyphyletic. Am. J. Bot. 83: 371-381.

Soltis, D. E. and Soltis, P. S. 1997. Phylogenetic relationships among Saxifragaceae sensu lato: a comparison of topologies based on 18S rDNA and *rbc*L sequences. Am. J. Bot. 84: 504-522.

Soltis, D. E. and Soltis, P. S. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In Soltis, D. E., Soltis, P. S. and Doyle, J. J. (eds.), Molecular systematics of Plants II. DNA sequencing. Kluwer Academy Publications:Boston.

Soltis, P. S., Soltis, D. E. and Chase, M. W. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402: 402-404.

Soontornchainaksaeng, P., Chantaranothai, P. and Senakun, C. 2003. Genetic Diversity of *Croton* L. (Euphorbiaceae) in Thailand. Cytologia 68(4): 379-382.

Steane, D. A., McKinnon, G. E., Vaillancourt, R. E. and Potts, B. M. 1999. ITS sequence data resolve higher level relationships among the eucalyptus. Mol. Phyl. Evol. 12(2): 215-223.

Steele, K. P. and R. Vilgalys. 1994. Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene *mat*K. Syst. Bot. 19: 126-142.

Steinmann, V. W. and Porter, J. M. 2002. Phylogenetic relationships in Euphorbiaceae (Euphorbiaceae) based on ITS and *ndh*F sequence data. Ann. Missouri. Bot. Gard. 89: 453-490.

Sukrong, S., Zhu, S., Ruangrungsi, N., Phadungcharoen, T., Palanuvej, C. and Komatsu, K. 2007. Molecular analysis of the genus *Mitragyna* existing in Thailand based on rDNA ITS sequences and its application to identify a narcotic Species : *Mitragyna speciosa*. Biol. Pharm. Bull. 30(7): 1284-1288.

Sugita, M., Shinozaki, K. and Sugiura, M. 1985. Tobacco chloroplast tRNA ^{Lys} (UUU) gene contains a 2.5-kilobase-pair intron: an open reading frame and a conserved boundary sequence in the intron. Proc. Natl. Acad. Sci. U S A. 82: 3557-3561.

Swofford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer, Sunderland:Massachusetts.

Sytsma, K. J. 1990. DNA and morphology: Inference of plant phylogeny. Trends Ecol. Evol. 5: 104–110.

Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Bio. 17: 1105-1109.

Takaiwa, F., Oono, K. and Sugiura, M. 1985. Nucleotide sequence of the 17S-25S spacer region from rice rDNA. Plant Mol. Bio. 4: 355-364.

Tansakul, P. and De-Eknamkul, W. 1998. Geranylgeraniol-18-hydroxylase: The last enzyme on the plaunotol biosynthesis in *Croton sublyratus.* Phytochemistry 47: 1241-6.

Tokuoka, T. 2007. Molecular phylogenetic analysis of Euphorbiaceae sensu stricto based on plastid and nuclear DNA sequences and ovule and seed character evolution. J. Plant Res. 120:511–522.

Tsukaya, H. 2003. A simple method for collecting DNA samples in the field. Newslet. Him. Bot. 32: 15-17.

Tsukaya, H. 2004. Gene flow between *Impatiens redicans* and *I. javernsis* (Balsaminaceae) in Gunung Pangrango, Central Java, Indonesia. Am. J. Bot. 91: 2119-2123.

Tsukaya, H., Fukuda, T. and Yokoyama, J. 2003. Hybridization and introgression between *Callicarpa japonica* and *C. mollis* (Verbenaceae) in central Japan, as inferred from nuclear and chloroplast DNA sequence. Mol. Ecol. 12: 3003-3011.

Urbatsch, L. B., Baldwin, B.G. and Donoghue, M. J. 2000. Phylogeny of the coneflower and relatives (Heliantheae: Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. Sys. Bot. 25: 539-565.

Webster, G. L. 1993. A provisional synopsis of the sections of the genus *Croton* (Euphorbiaceae). Taxon 42: 793-823.

Webster, G. L. 1994. Synopsis of the genera and supergeneric taxa of Euphorbiaceae. Ann. Missouri. Bot. Gard. 81: 33-144.

Wendel, J. F., Schnabel, A. and Seelanan, T. 1995a. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. U S A. 92: 280–284.

Wendel, J. F., Schnabel, A. and Seelanan, T. 1995b. An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic, intergenomic introgression. Mol. Phyl. Evol. 4: 298–313.

White, T. J., Bruns, T., Lee, S. and Taylor, J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J.(eds.), PCR protocols: a guide to methods and applications. 315-322. Academic Press:New York.

Whittall, J., Liston, A., Gisler, S. and Meinke, R. J. 2000. Detecting nucleotide additivity from direct sequences is a SNAP: an example from *Sidalcea* (Malvaceae). Plant Biol. 2: 1-7.

Widmer, A. and Baltisberger, M. 1999. Molecular evidence for allopolyploid speciation and a single origin of the narrow endemic *Draba ladina* (Brassicaceae). Am. J. Bot. 86: 1282–1289.

Winship, P. R. 1989. An improved method for direct sequencing PCR amplified material using dimethyl sulphoxide. Nucleic Acids Res. 17: 1266.

Wolfe, K. H., Li, W. H. and Sharp P. M. 1987. Rate of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNA. Proc. Natl. Acad. Sci. U S A. 84: 9054-9058.

Wolfe, K.H. and Sharp, P.M. 1988. Identification of functional open reading frames in chloroplast genomes. Gene 66: 215-222.

Wolters, J. and Erdmann, V. A. 1986. Cladistic analysis of SS rRNA and 16S rRNA secondary and primary structure. J. Mol. Evol. 24: 152-166.

Wurdack, K. J., Hoffmann, P. and Chase, M.W. 2005. Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae sensu stricto) using plastid *rbc*L and *trn*L-F DNA sequences. Am. J. Bot. 92(8): 1397-1420.

Xu, D. H., Abe, J., Gai, J. Y. and Shimamoto, Y. 2002. Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multi cultivated soybean. Theor. App. Genet. 105: 645-653.

Xue, H. G., Zhoe, S. D., He, X. J. and Yu, Y. 2007. Molecular authentication of the tradition Dai Medicinal plant *Croton caudatus*. Planta Med. 73: 251-251

Yang, D. Y., Fushimi, H., Cai, S. Q. and Komatsu, K. 2004. Molecular analysis of *Rheum* species used as rhei rhizoma based on the chloroplast *mat*K gene sequence and its application for identification. Biol. Pharm. Bull. 27(3): 375-383.

Yang, T. W., Yang, Y. A. and Xiong, Z. G. 2000. Paternal inheritance of chloroplast DNA in the genus *Larrea* (Zygophyllaceae). Am. J. Bot. 87: 1452-1458.

Yang, Z., Kumar, S. and Nei, M. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics 141: 1641-1650.

Yamada, K., Shinozaki, K. and Sugiura, M. 1986. DNA sequences of tobacco chloroplast genes for $tRNA^{Ser}$ (GGA), $tRNA^{Thr}$ (UGU), $tRNA^{Leu}$ (UAA), $tRNA^{Phe}$ (GAA): the $tRNA^{Leu}$ gene contains a 503 bp intron. Plant Mol. Bio. 6: 193-199.

Yokota, Y., Kawata, T., Iida, Y., Kato, A. and Tanifuji, S. 1989. Nucleotide sequences of the 5.8S rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. J. Mol. Evol. 29: 294-301.

Zhang, D. X. and Hewitt, G. M. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol. Eco. 12: 563-584.

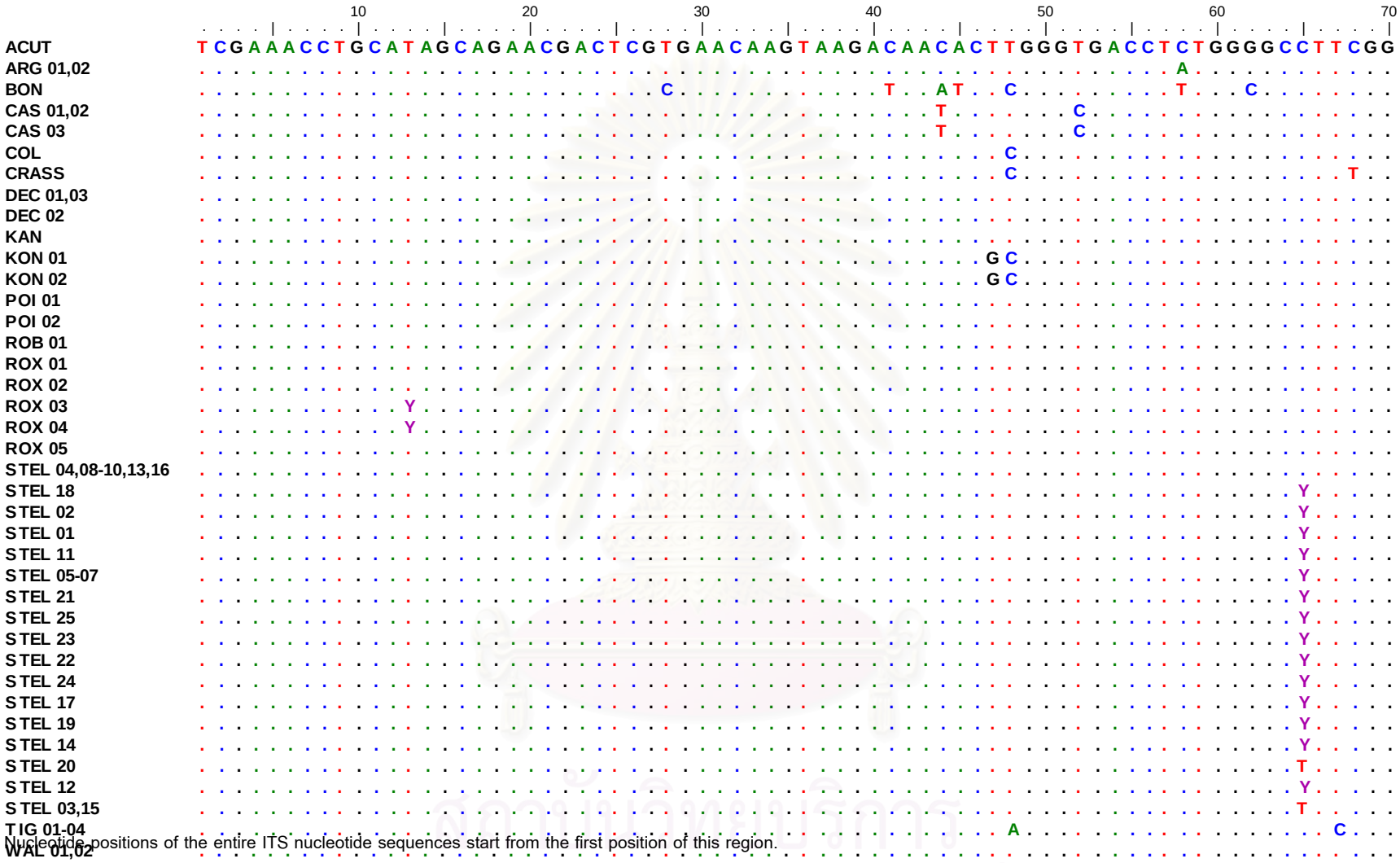
Zhang, J. L., Zhang, C. Q., Gao, L. M., Yang, J. B. and Li, H. T. 2007. Natural hybridization origin of *Rhododendron agastum* (Ericaceae) in Yunnan, China: inferred from morphological and molecular evidence. J. Plant Res. 120: 457-463.

APPENDICES

APPENDIC A

ITS nucleotide alignment of *Croton* Species

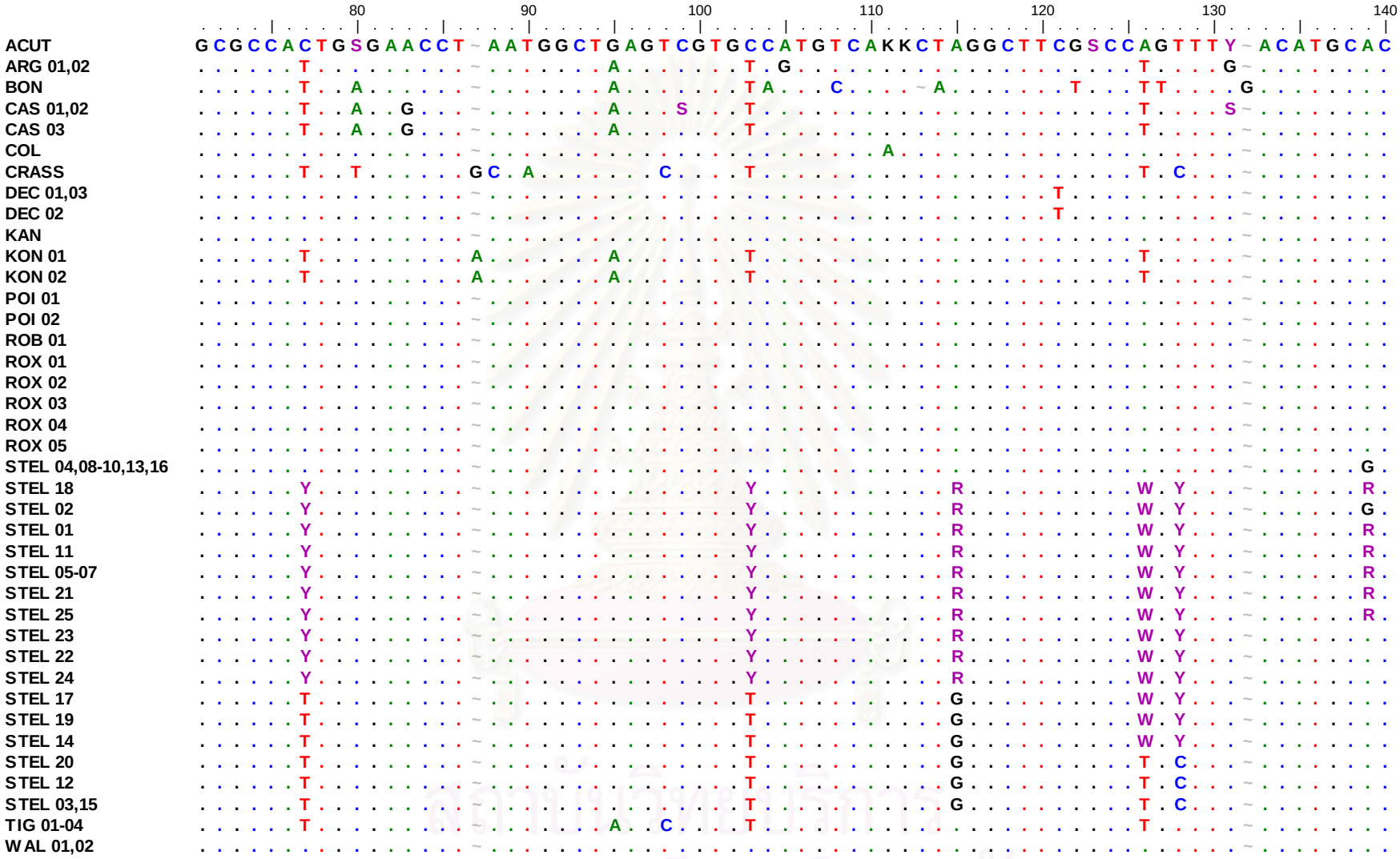
ITS1 region : position 1-276
5.8S region : position 277-436
ITS2 region : position 437-629



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

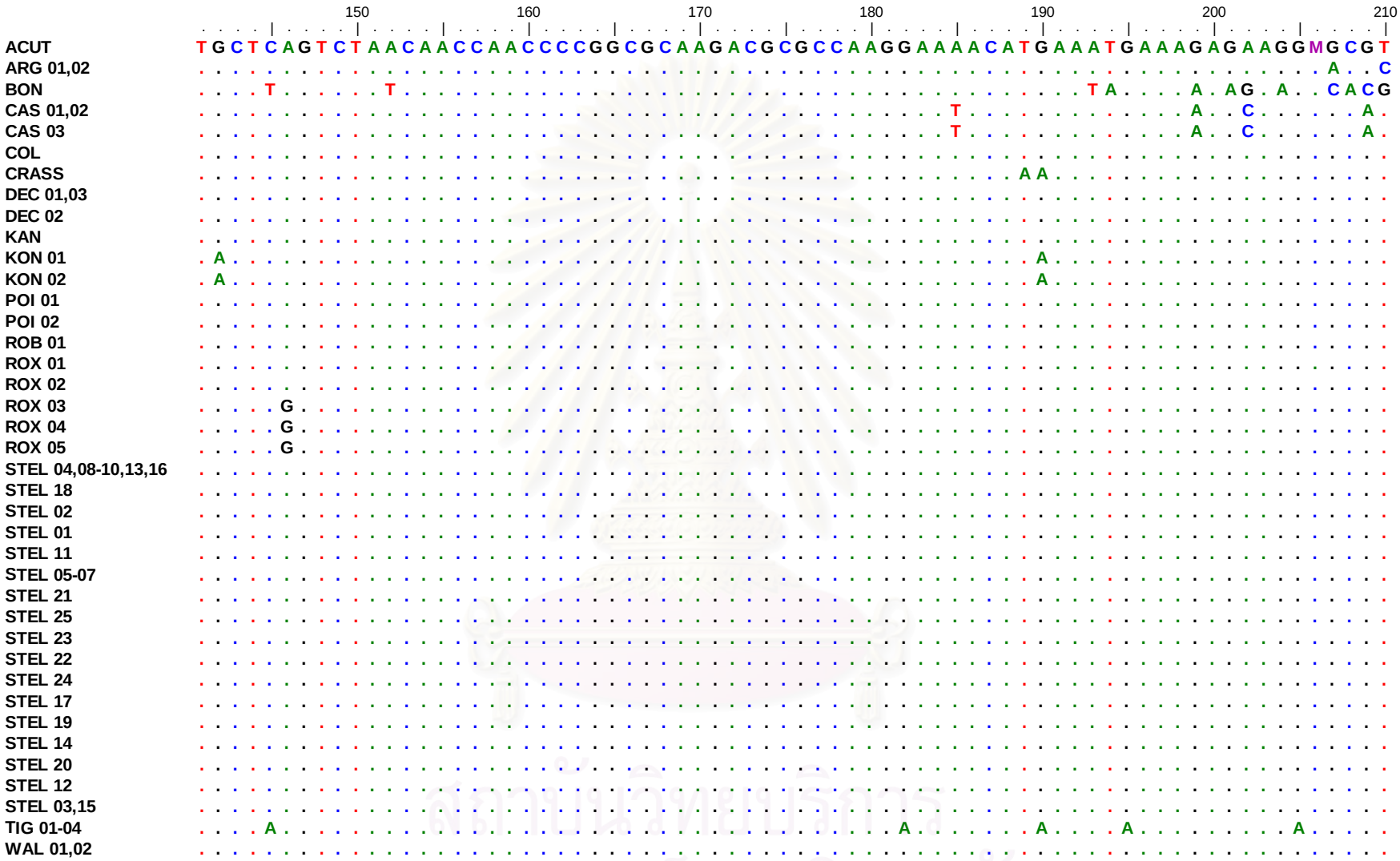
The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

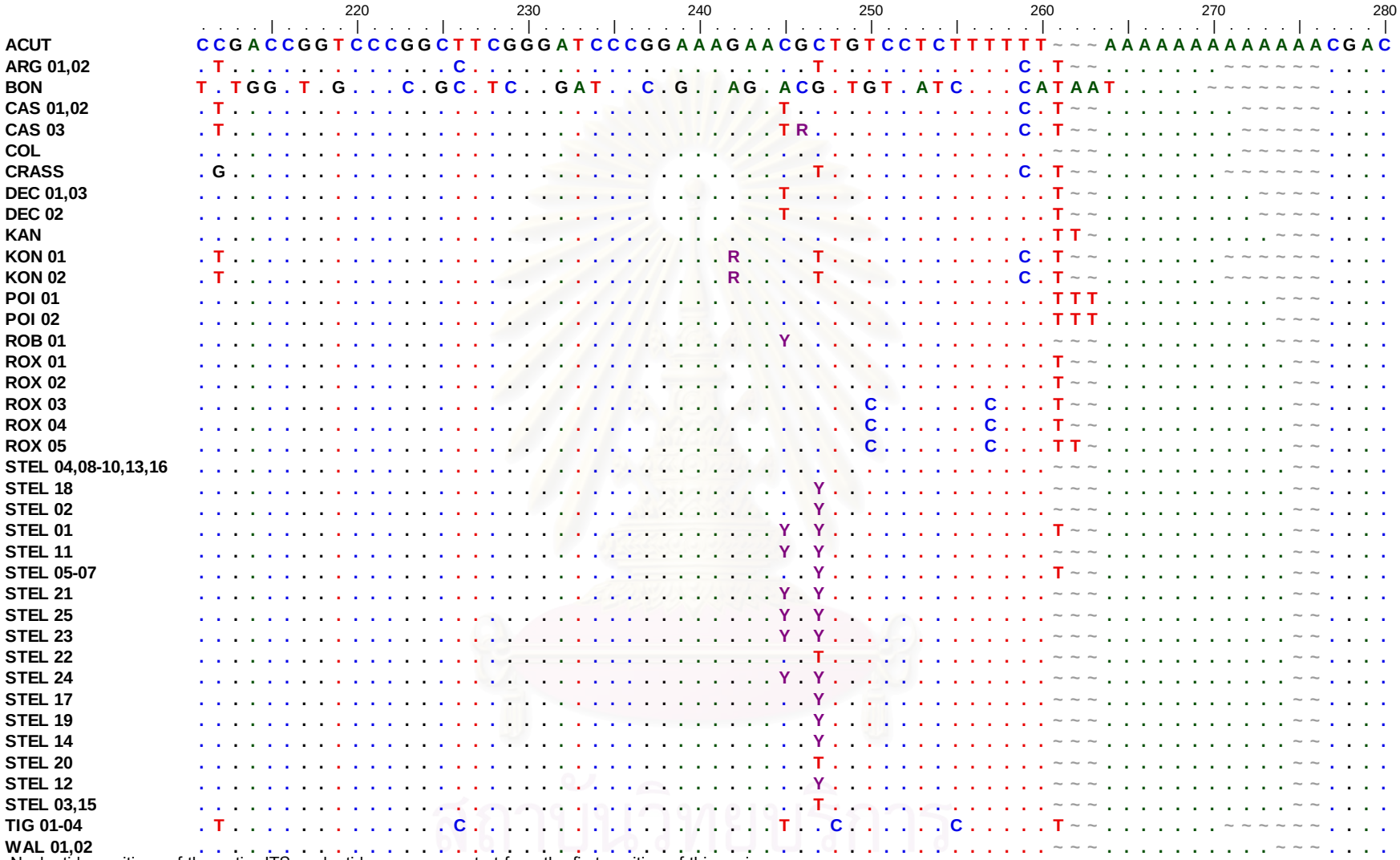
The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

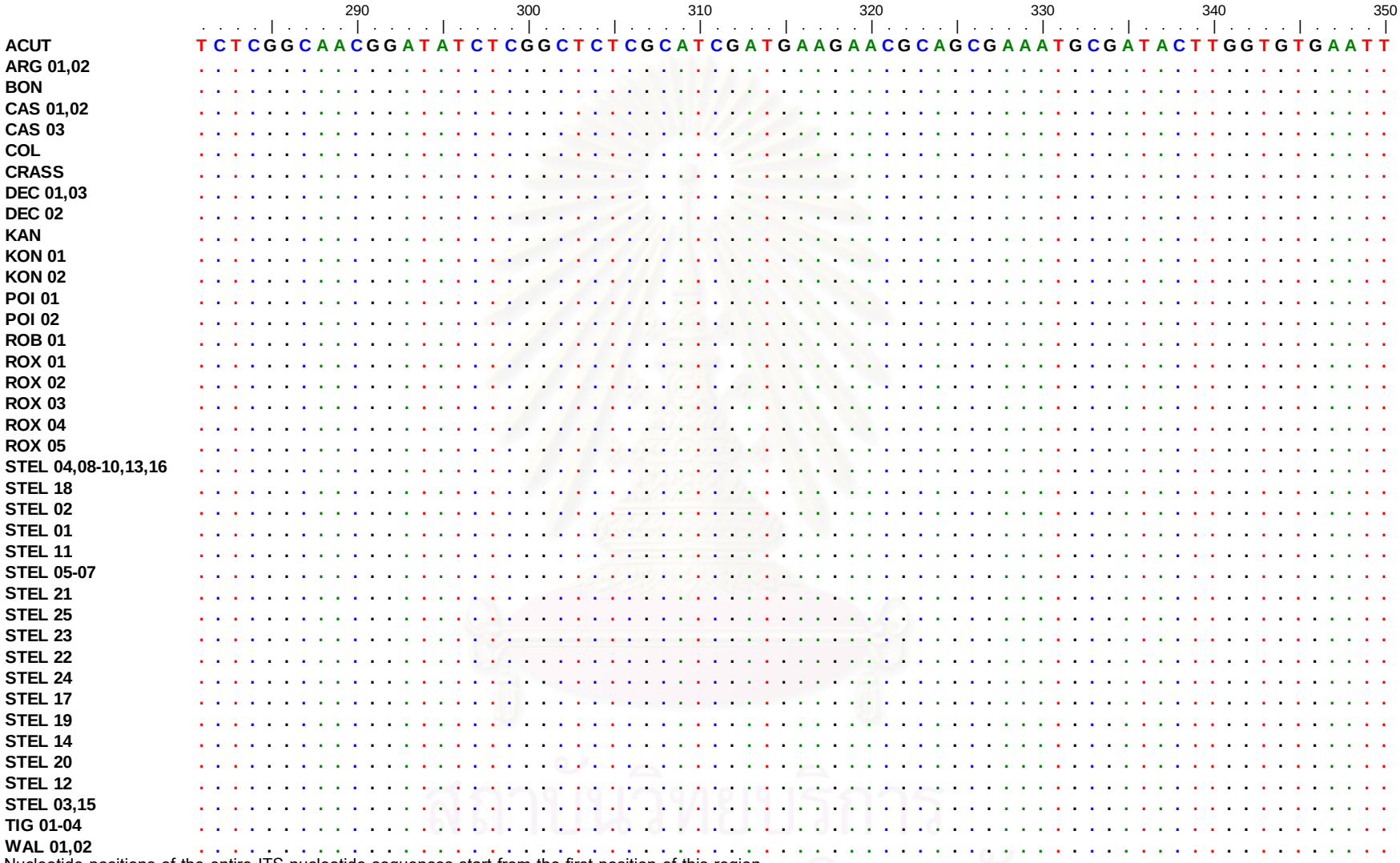
The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

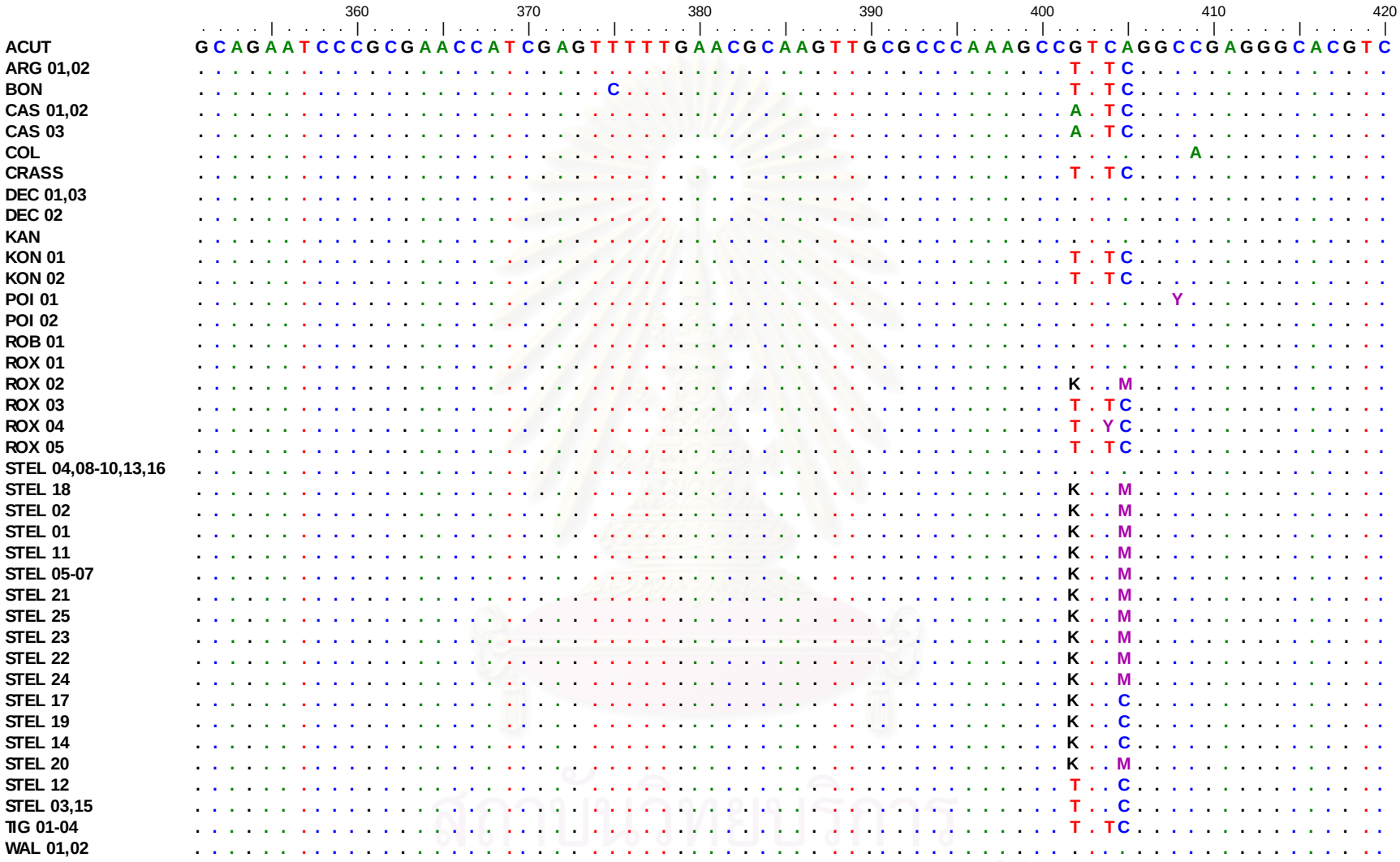
The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

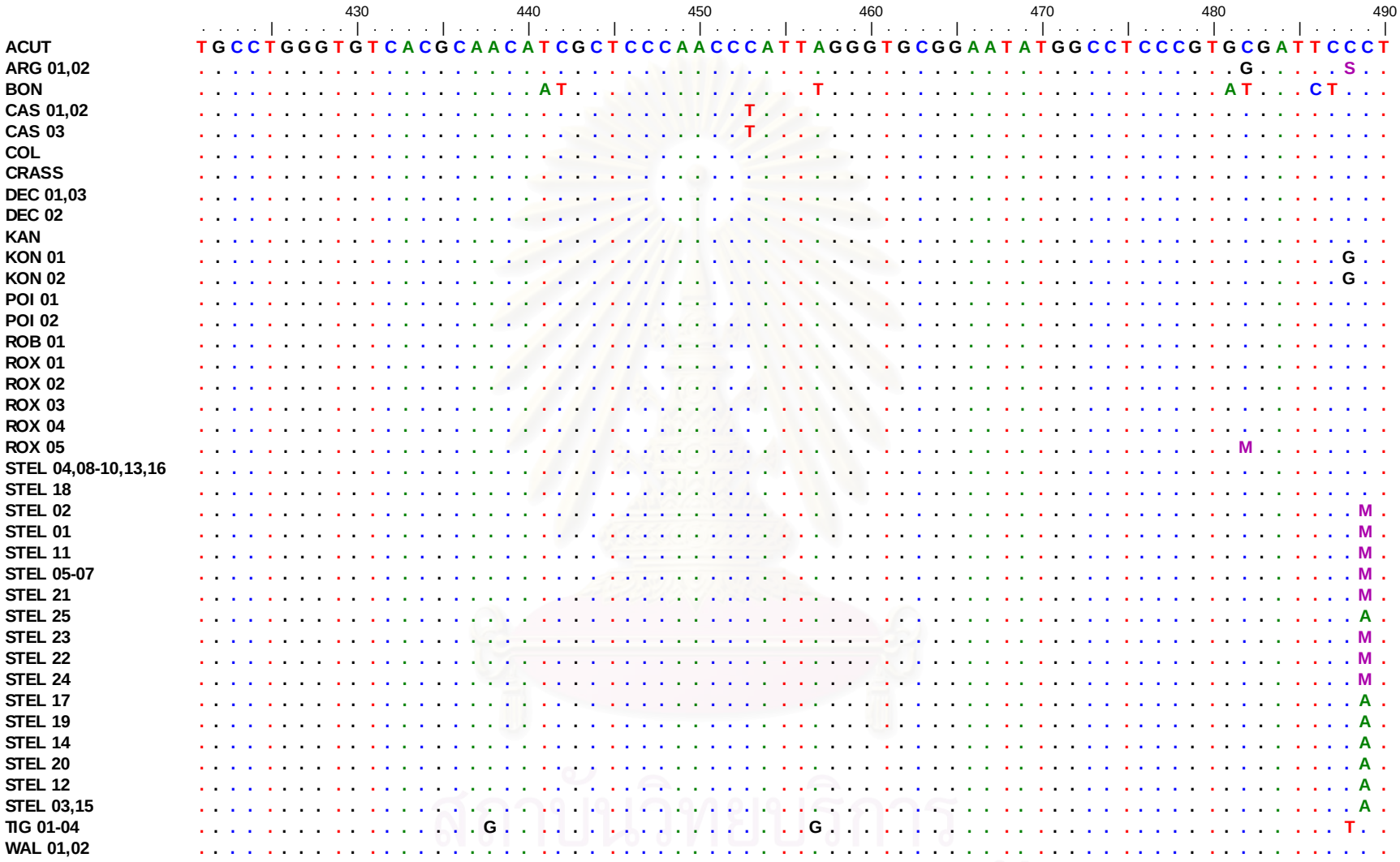


Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

140

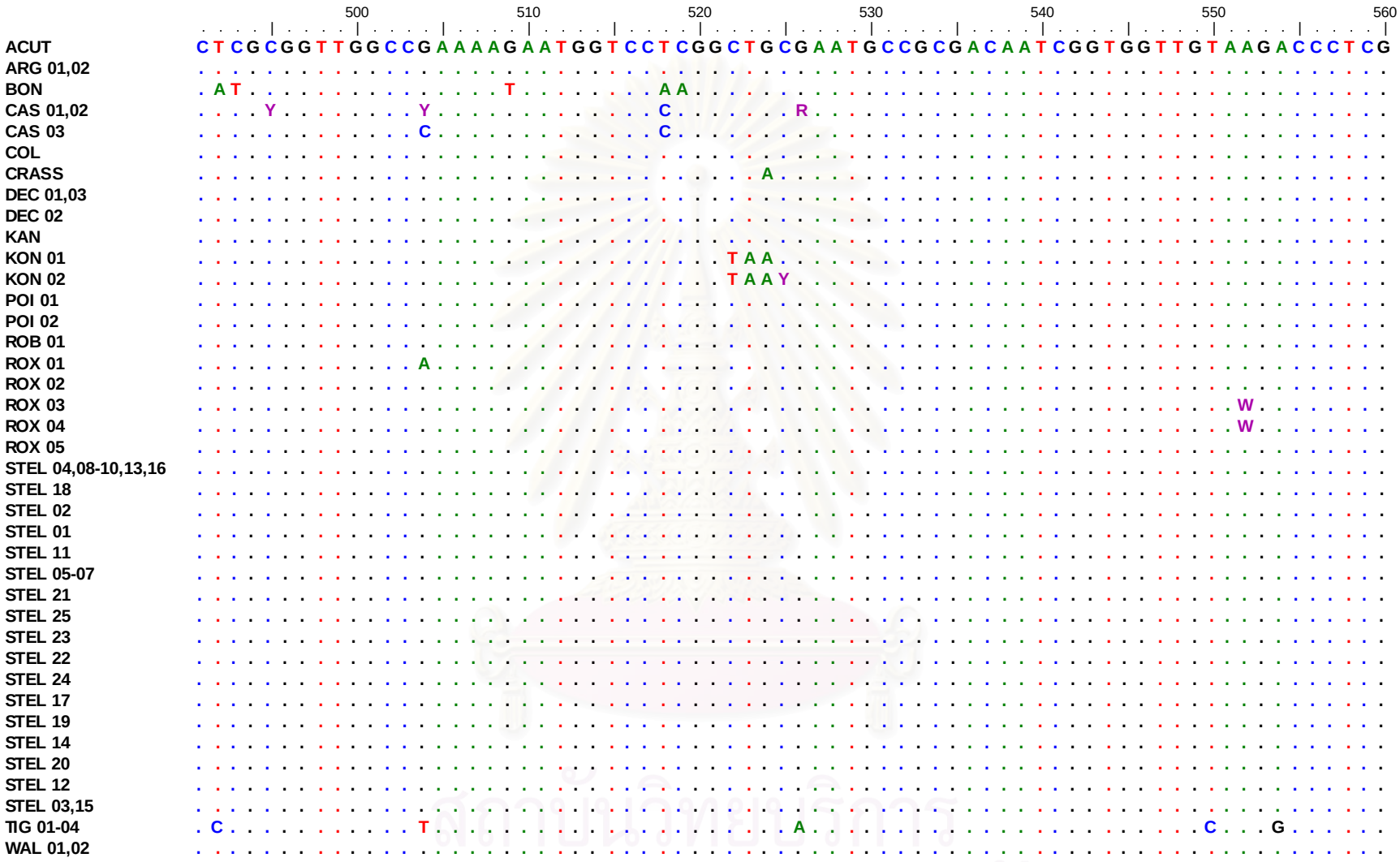


Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

141

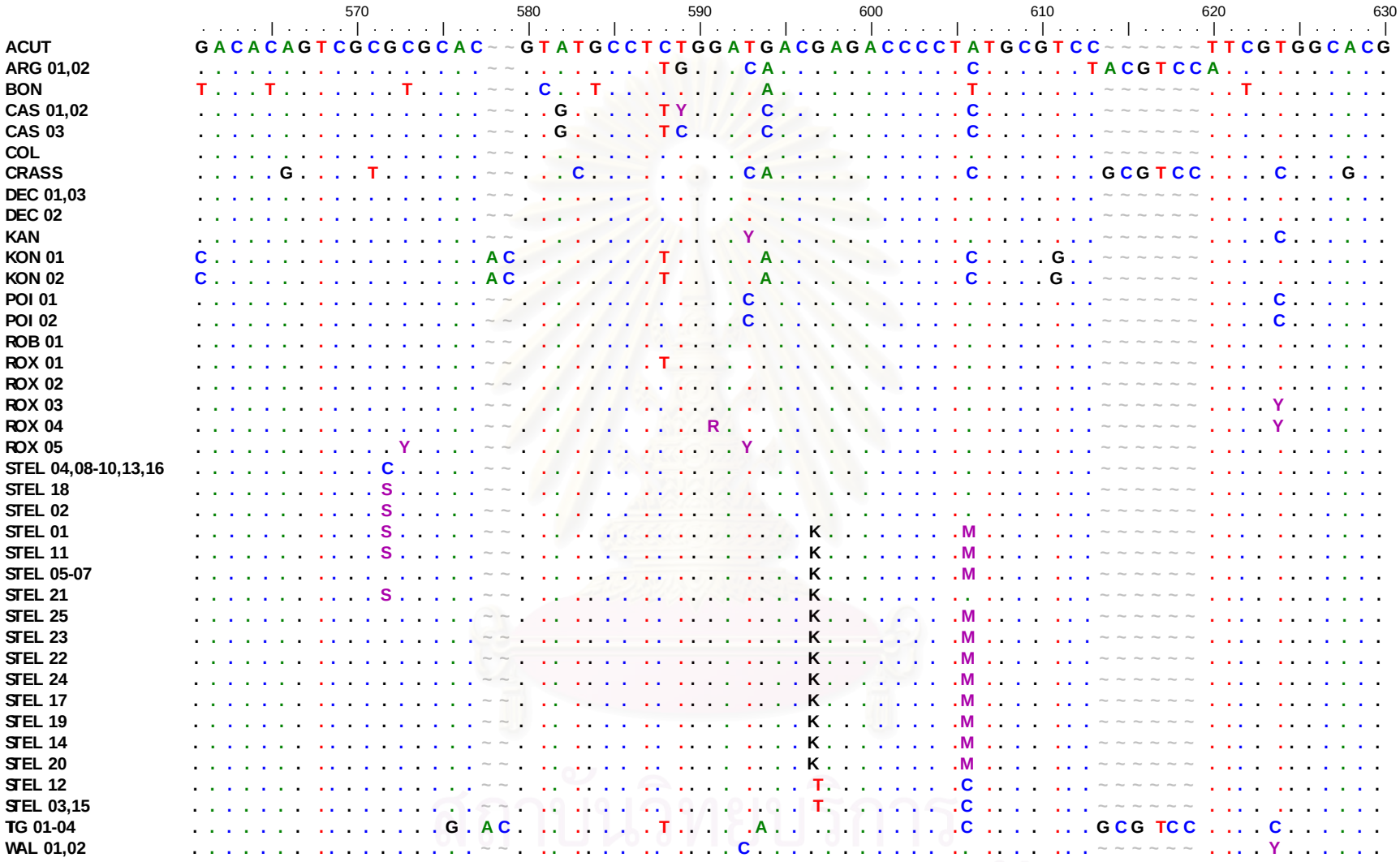


Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

142



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

143

```
                    . . . . | . . .
ACUT                C T C A C A T C
ARG 01,02           . . . . . . . .
BON                 . . . . . T . .
CAS 01,02           . . . . . . . .
CAS 03              . . . . . . . .
COL                 . . . . . . . .
CRASS               . . . . . . . .
DEC 01,03           . . . . A . . .
DEC 02              . . . . M . . .
KAN                 . . . . . . . .
KON 01              . . . . . . . .
KON 02              . . . . . . . .
POI 01              . . . . . . . .
POI 02              . . . . . . . .
ROB 01              . . . . A . . .
ROX 01              . . . . . . . .
ROX 02              . . . . . . . .
ROX 03              . . . . . . . .
ROX 04              . . . . . . . .
ROX 05              . . . . . . . .
STEL 04,08-10,13,16 . . . . A . . .
STEL 18             . . . . A . . .
STEL 02             . . . . M . . .
STEL 01             . . . . M . . .
STEL 11             . . . . M . . .
STEL 05-07          . . . . M . . .
STEL 21             . . . . M . . .
STEL 25             . . . . . . . .
STEL 23             . . . . . . . .
STEL 22             . . . . . . . .
STEL 24             . . . . . . . .
STEL 17             . . . . . . . .
STEL 19             . . . . . . . .
STEL 14             . . . . . . . .
STEL 20             . . . . . . . .
STEL 12             . . . . . . . .
STEL 03,15          . . . . . . . .
TIG 01-04           . . . . . . . .
WAL 01,02           . . . . . . . .
```

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

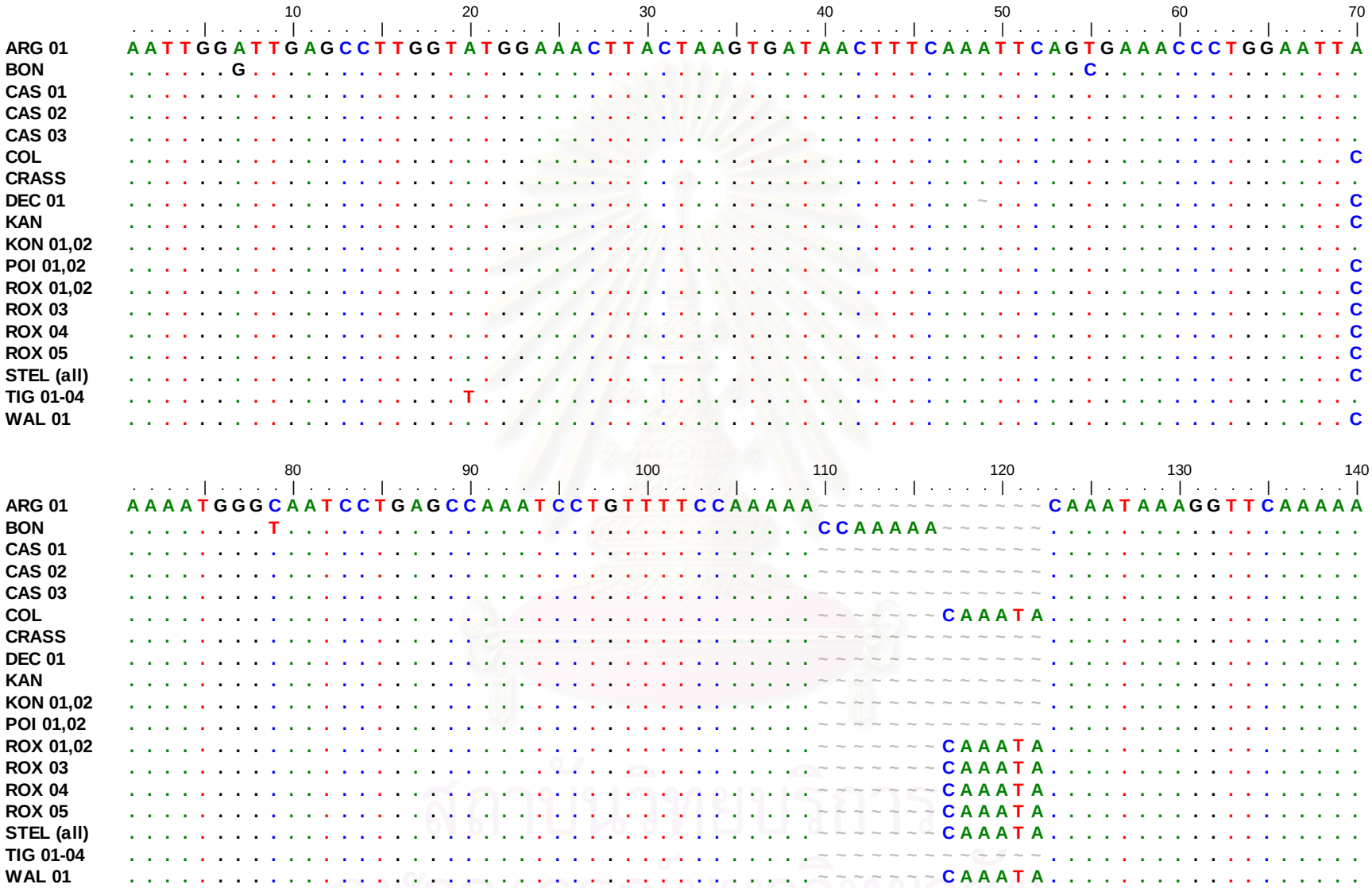
The IUPAC Codes: Y = C/T, R = A/G, W = A/T, K = G/T, M = A/C, S = C/G

144

APPENDIC B

*trn*L-F nucleotide alignment of *Croton* Species

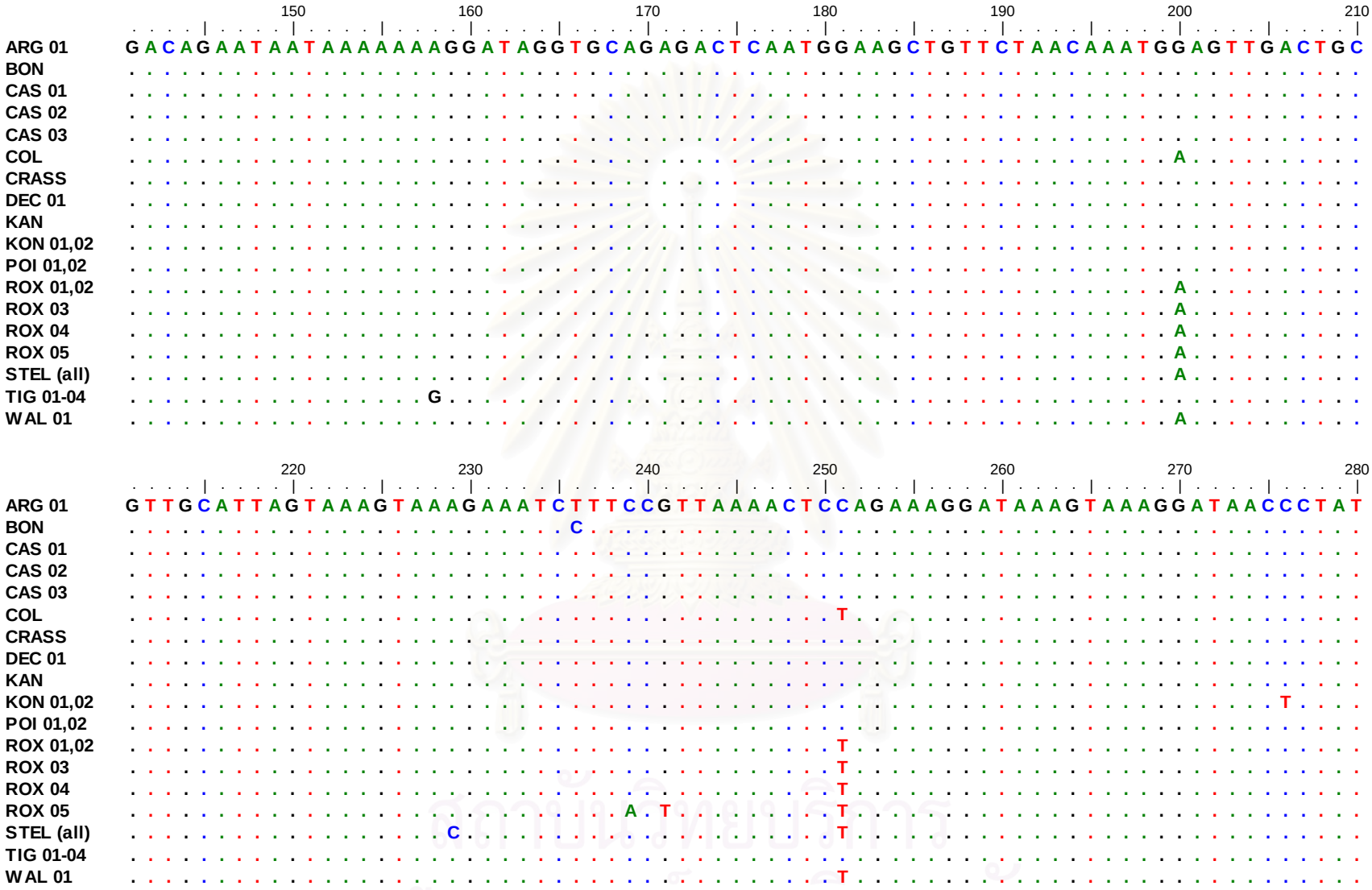
*trn*L-F intron : position 1-573
*trn*L-F intergenicspacer : position 574-1079



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

146



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

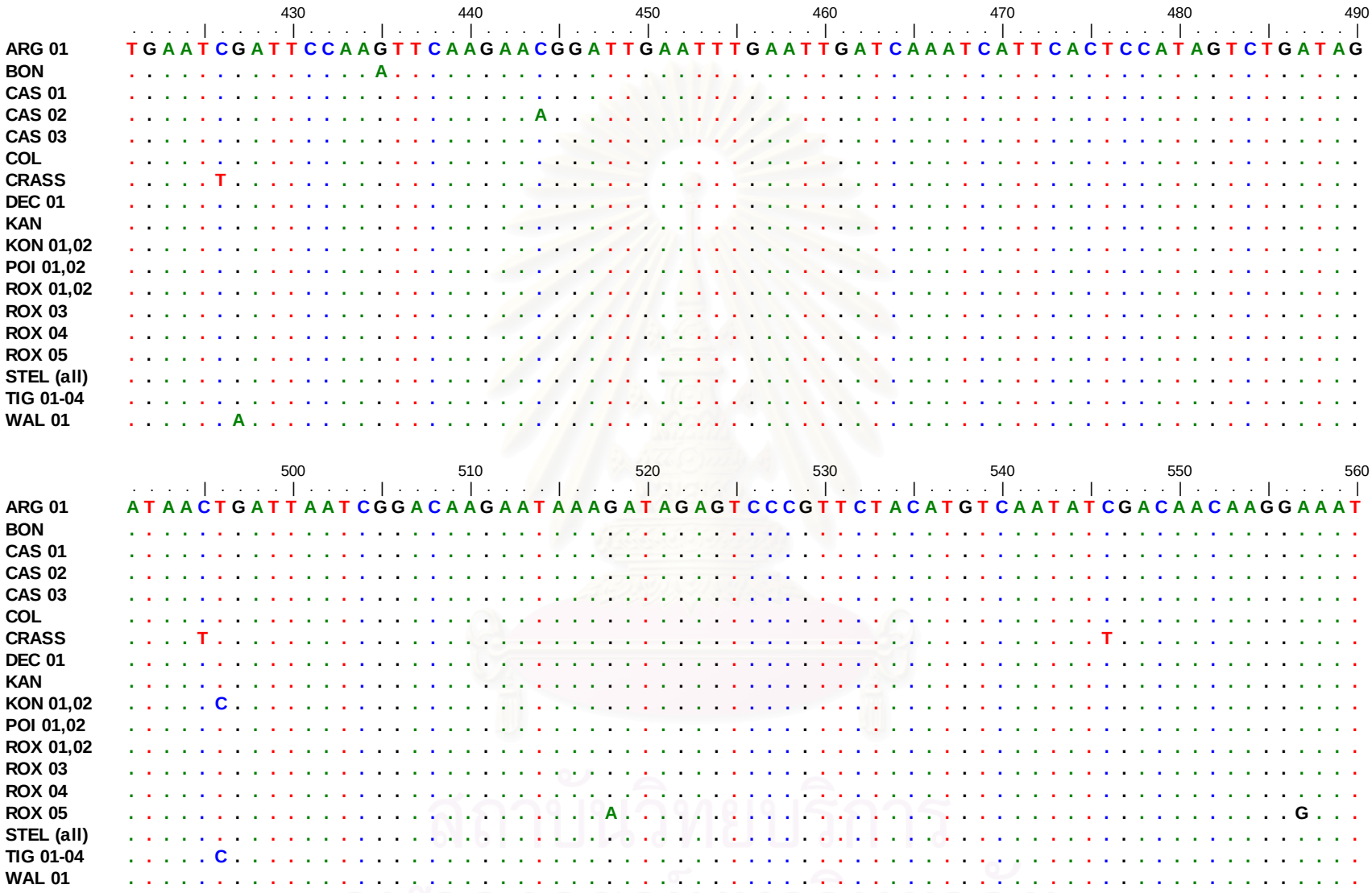
Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

```
              290       300       310       320       330       340       350
ARG 01     ATT-CATACATATTCGTATTGAAATACTATCTCAAATGATTAATTAA~~~~~~~~~~~~~TAGGGCCCGA
BON        ...T.....G.....................................~~~~~~~~~~~~~..........
CAS 01     ...~.....G.....................................~~~~~~~~~~~~~..........
CAS 02     ...~.....G...........................C.........~~~~~~~~~~~~~..........
CAS 03     ...~.....G.....................................~~~~~~~~~~~~~..........
COL        ...~...........................................~~~~~~~~~~~~~..........
CRASS      ...~.....G.....................................~~~~~~~~~~~~~..........
DEC 01     ...~...........................................TTAATTAA~~~~~..........
KAN        ...~..........................................~TTAATTAAATTAA..........
KON 01,02  ...~.....G.....................................~~~~~~~~~~~~~..........
POI 01,02  ...~..........................................~TTAATTAAATTAA..........
ROX 01,02  ...~...........................................~~~~~~~~~~~~~..........
ROX 03     ...~...........................................~~~~~~~~~~~~~..........
ROX 04     ...~...........................................~~~~~~~~~~~~~..........
ROX 05     ...~...........................................~~~~~~~~~~~~~..........
STEL (all) ...~...........................................~~~~~~~~~~~~~..........
TIG 01-04  ...~.....G.....................................~~~~~~~~~~~~~..........
WAL 01     ...-......................T....................~~~~~~~~~~~~~..........

              360       370       380       390       400       410       420
ARG 01     ATCAATATTTTTTTATCTTT--ATTTTTTTATCTTTCTATGAAAAATGAAATAAAAAAAAAAAATTGTTT
BON        ...C................CT..................T...................~~~~......
CAS 01     ......~~~~~~~~~~~~~~~~......................................~~~~......
CAS 02     ......~~~~~~~~~~~~~~~~......................................~~~~......
CAS 03     ......~~~~~~~~~~~~~~~~......................................~~~~......
COL        ......~~~~~~~~~~~~~~~~......................................~~~~......
CRASS      ......~~~~~~~~~~~~~~~~..................A...................~~~~.G....
DEC 01     ......~~~~~~~~~~~~~~~~.............................A........~~~~.G....
KAN        ......~~~~~~~~~~~~~~~~......................................~~~~......
KON 01,02  .....G~~~~~~~~~~~~~~~~......................................~~~~......
POI 01,02  ......~~~~~~~~~~~~~~~~......................................~~~~......
ROX 01,02  ......~~~~~~~~~~~~~~~~......................................~~~~......
ROX 03     ......~~~~~~~~~~~~~~~~......................................~~~~......
ROX 04     ......~~~~~~~~~~~~~~~~......................................~~~~......
ROX 05     ......~~~~~~~~~~~~~~~~..........T...........................~~~~......
STEL (all) ......~~~~~~~~~~~~~~~~......................................~~~~......
TIG 01-04  ......~~~~~~~~~~~~~~~~......................................~~~~......
WAL 01     ......~~~~~~~~~~~~~~~~......................................~~~~......
```

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

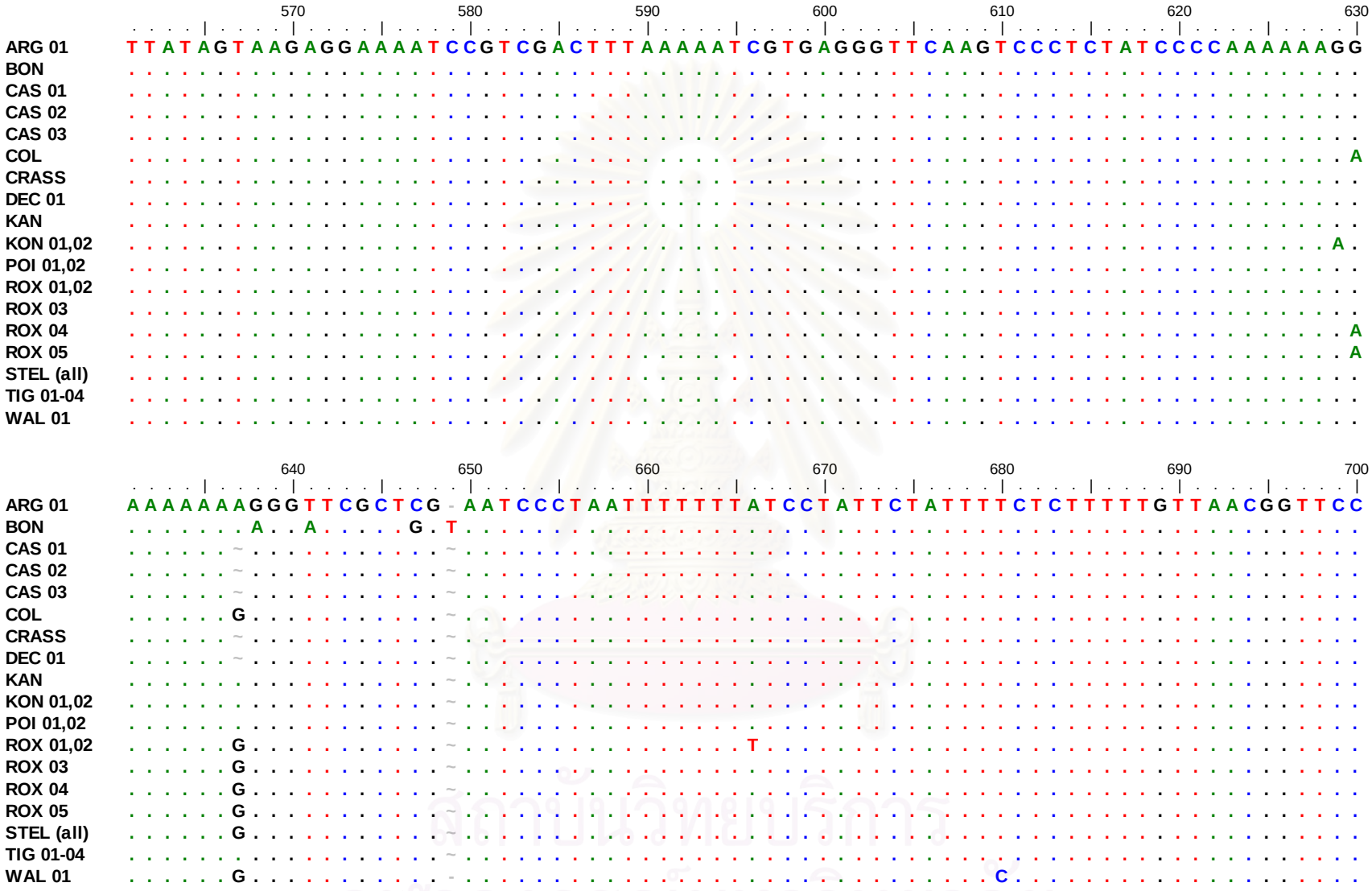
Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

148



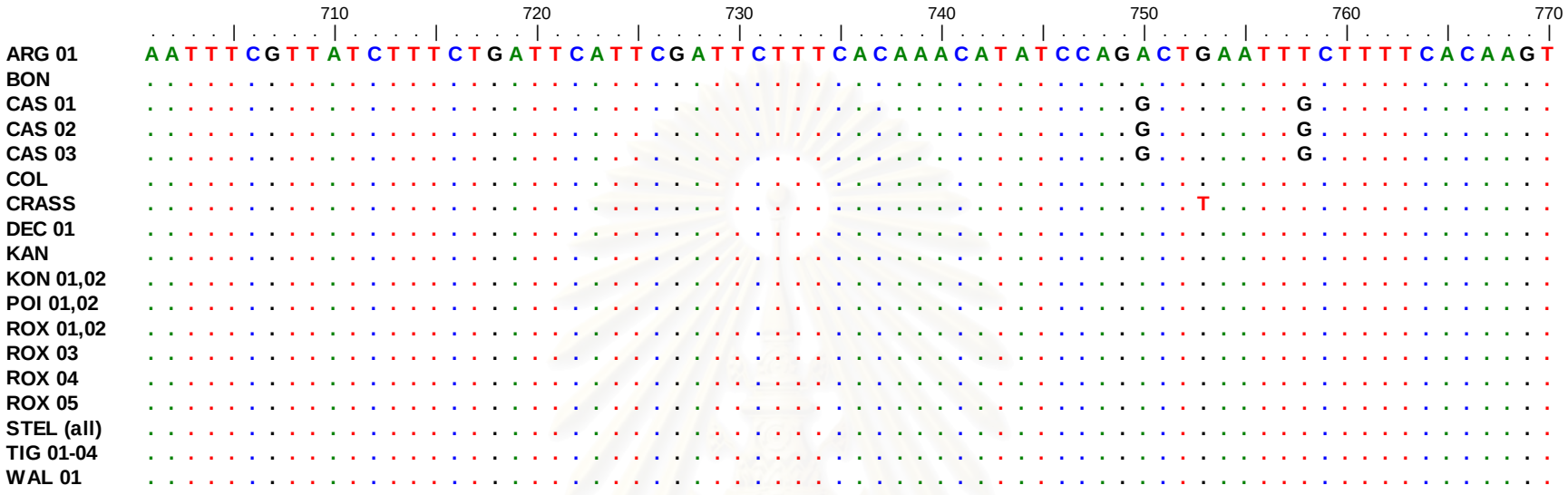
Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

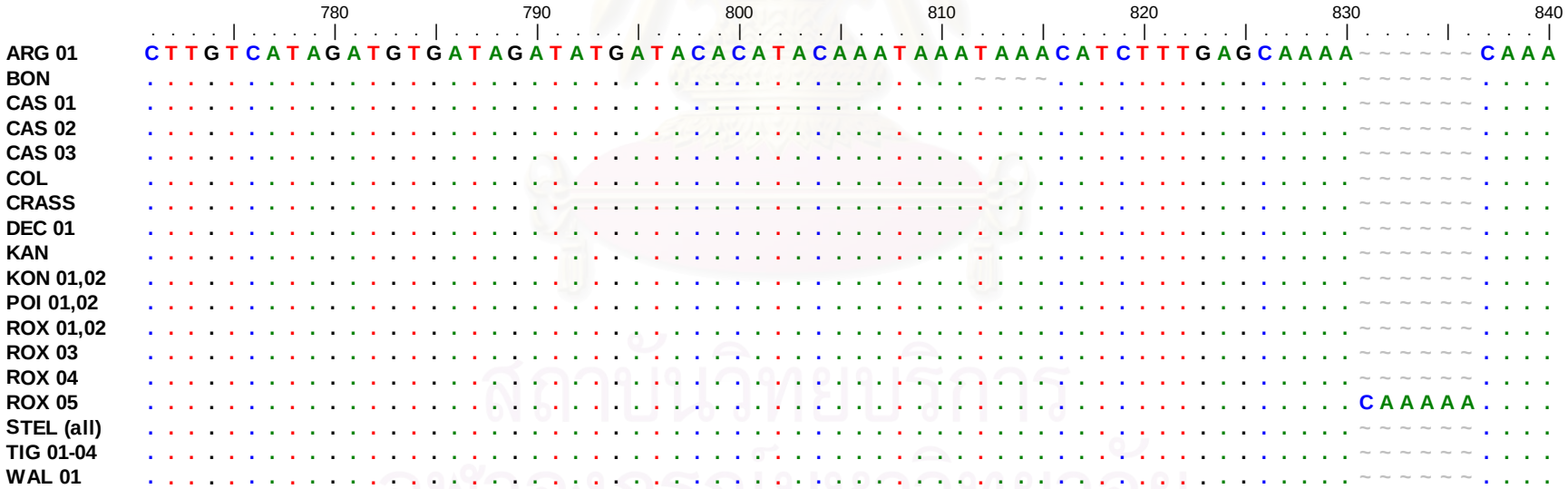
Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

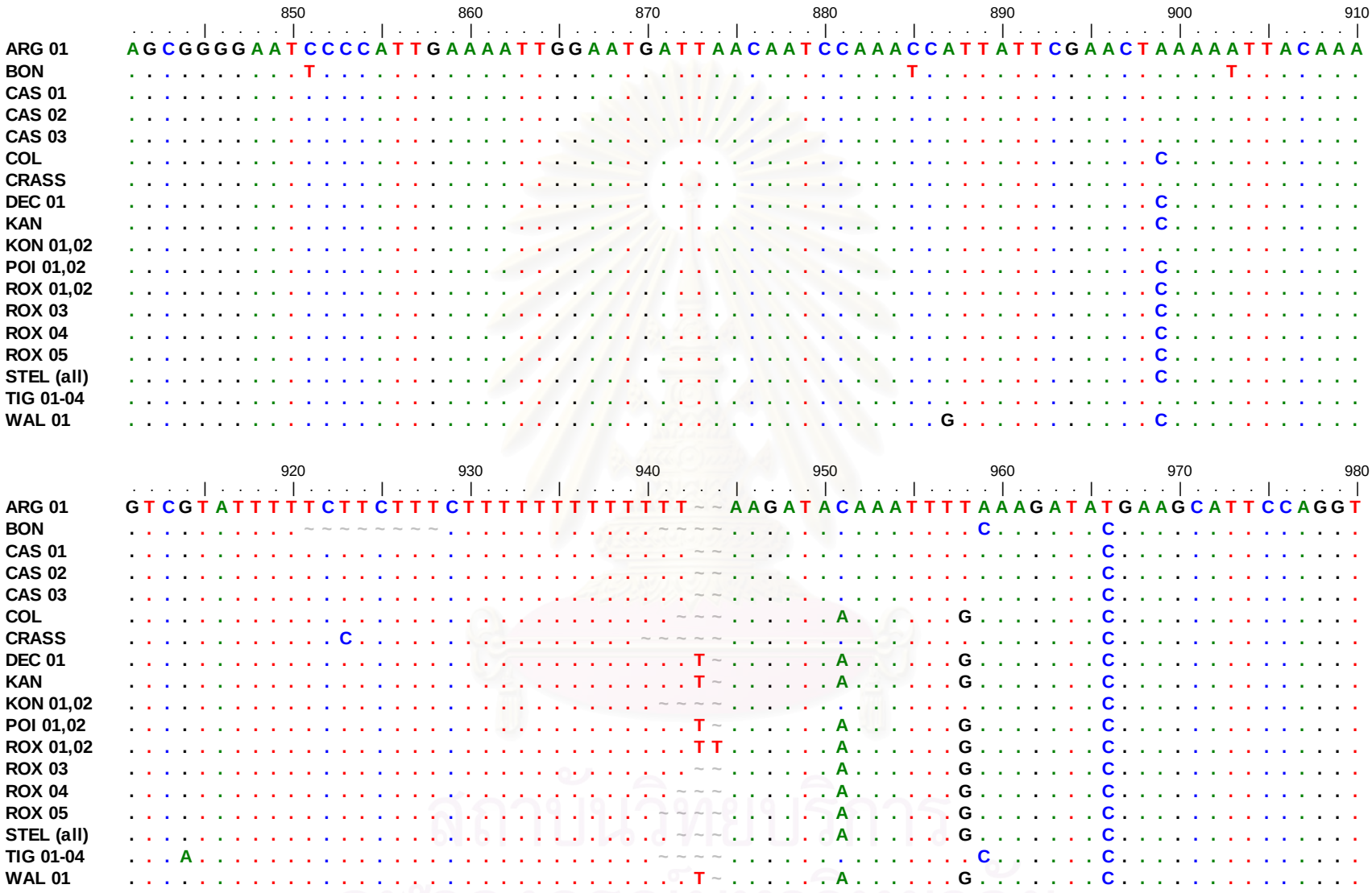
Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.





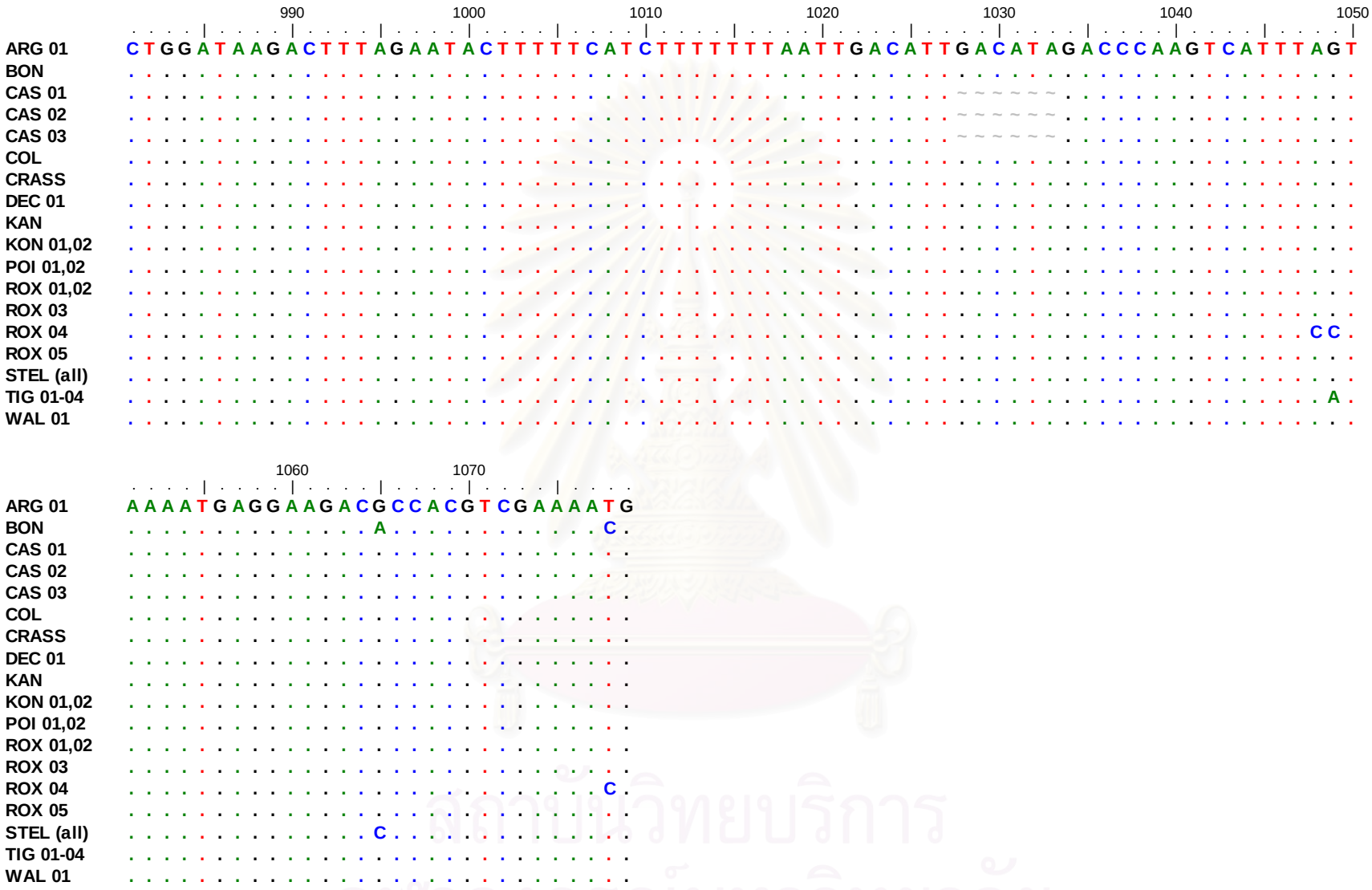
Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



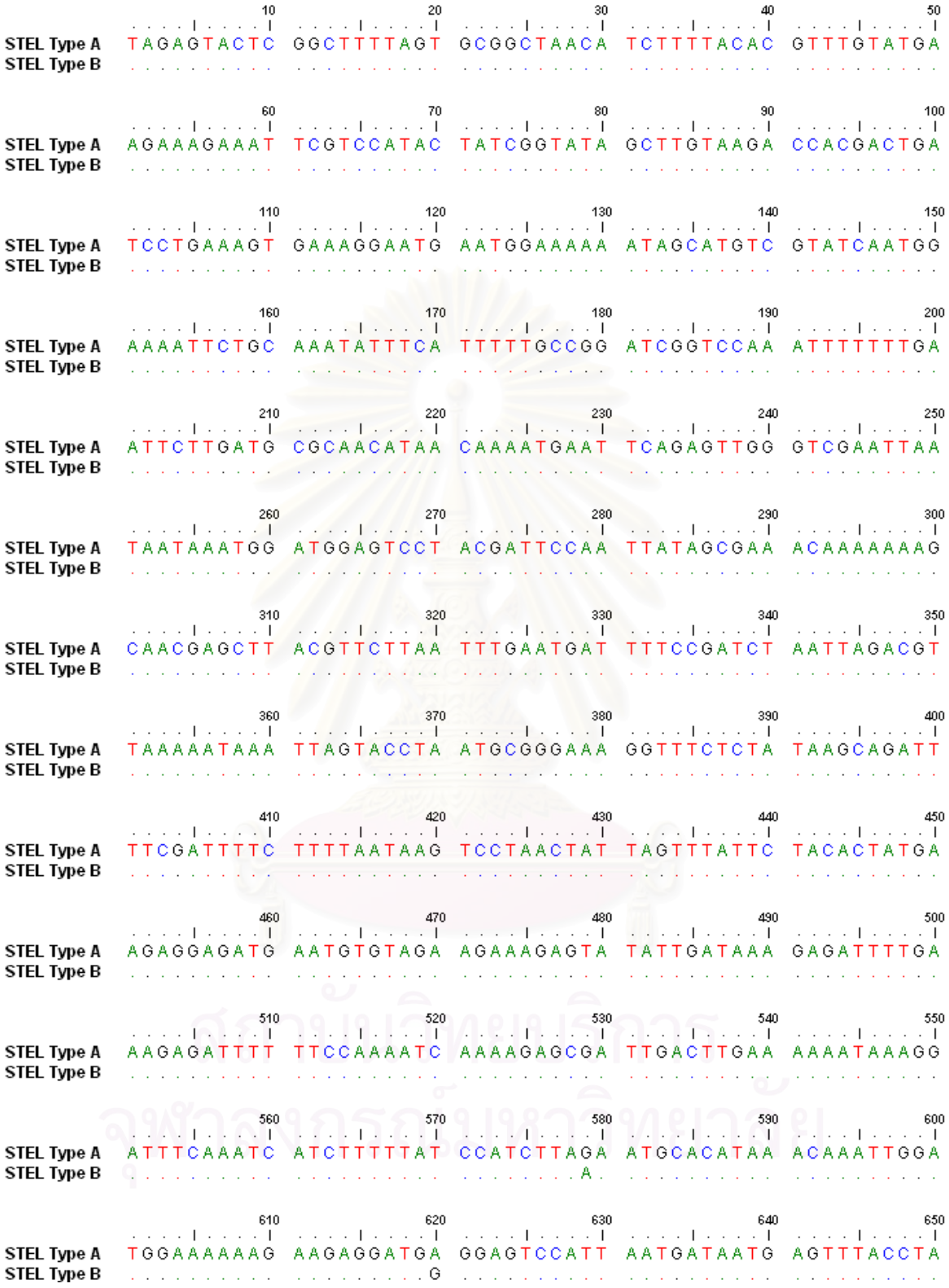
Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

APPENDIC C

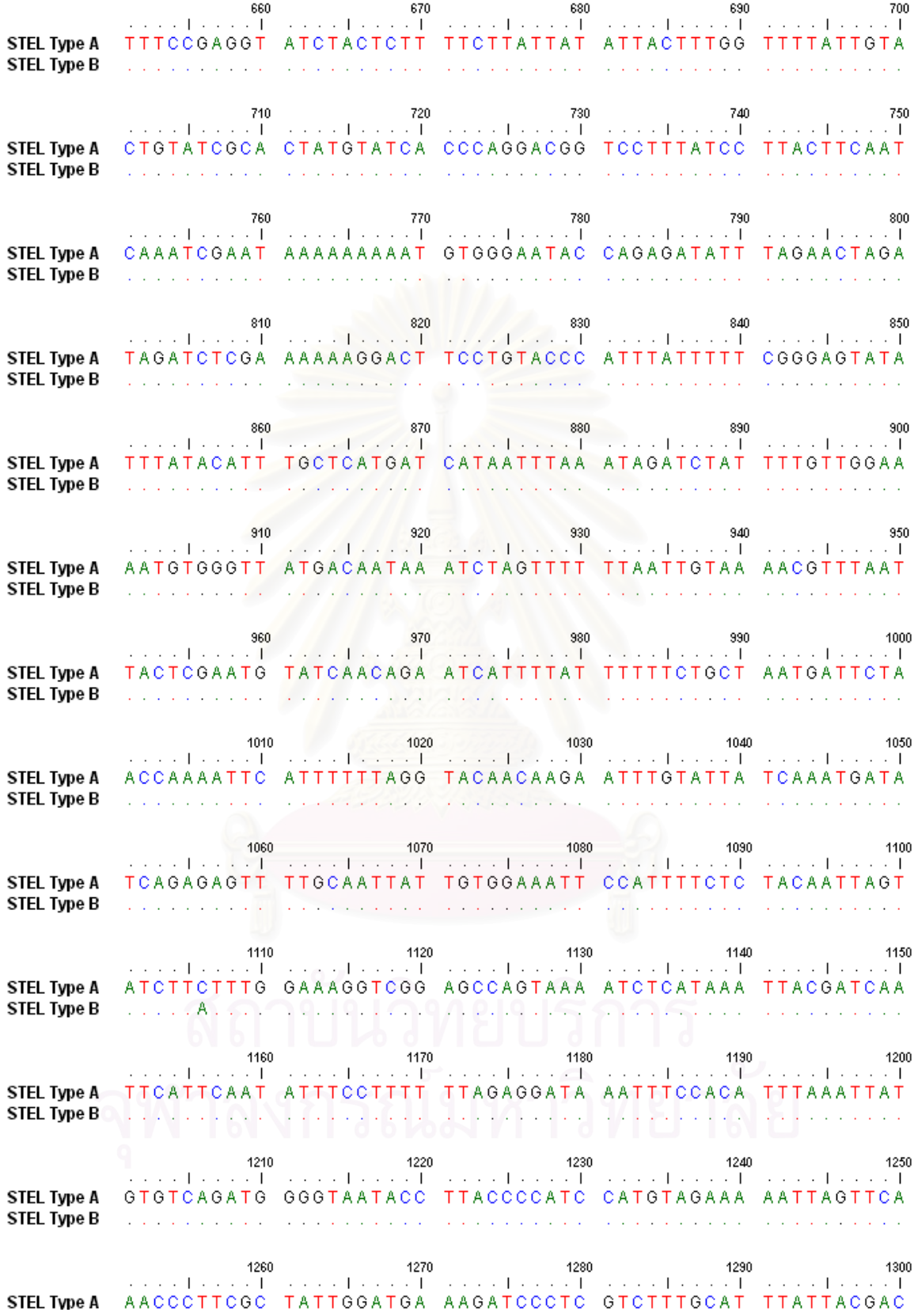
The Complete *trn*K Intron Nucleotide Alignment
of *C. stellatopilosus* Type A and Type B

*trn*K 5' intron : position 1-767
*mat*K gene : position 768-2287
*trn*K 3' intron : position 2288-2527



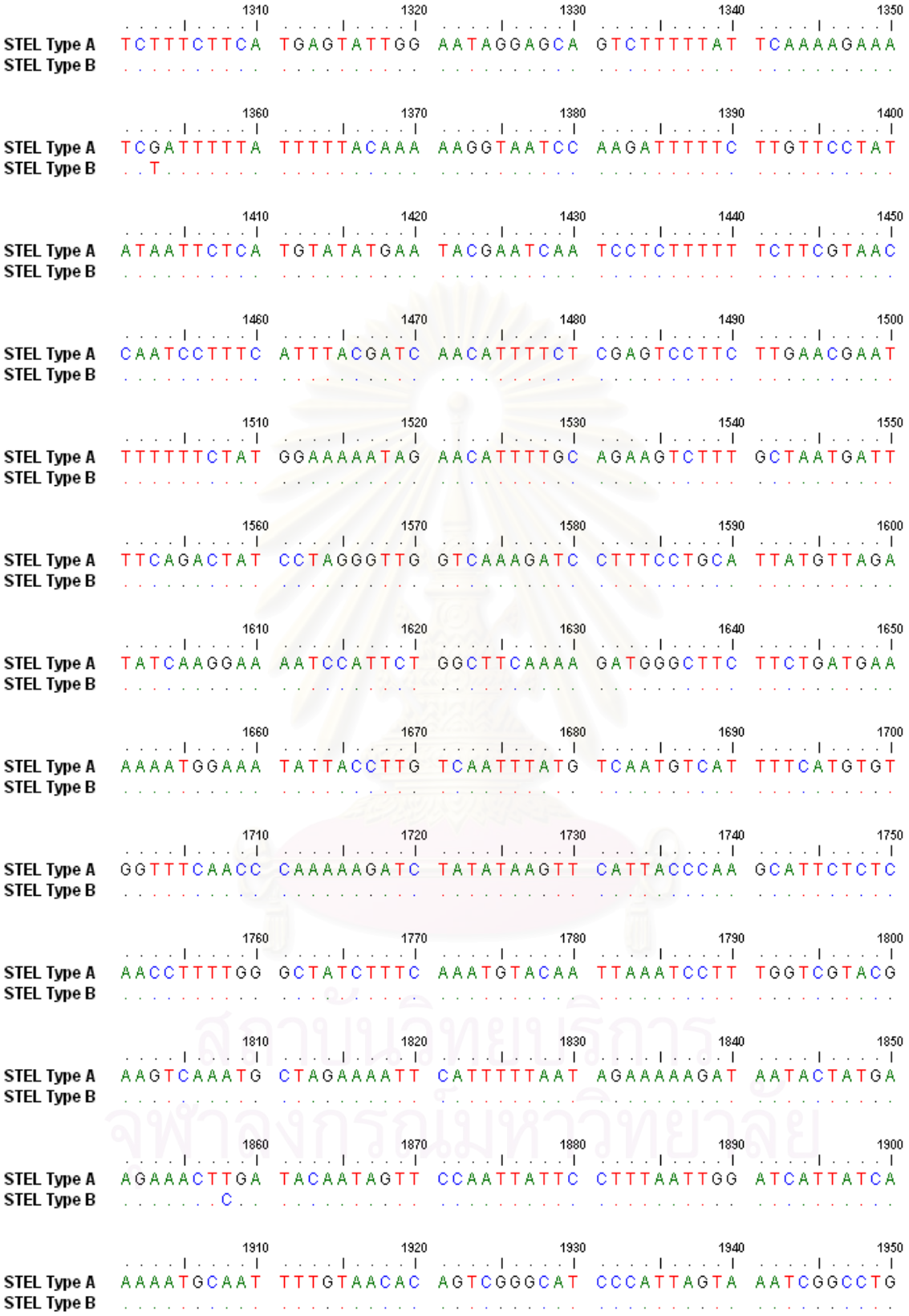
Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



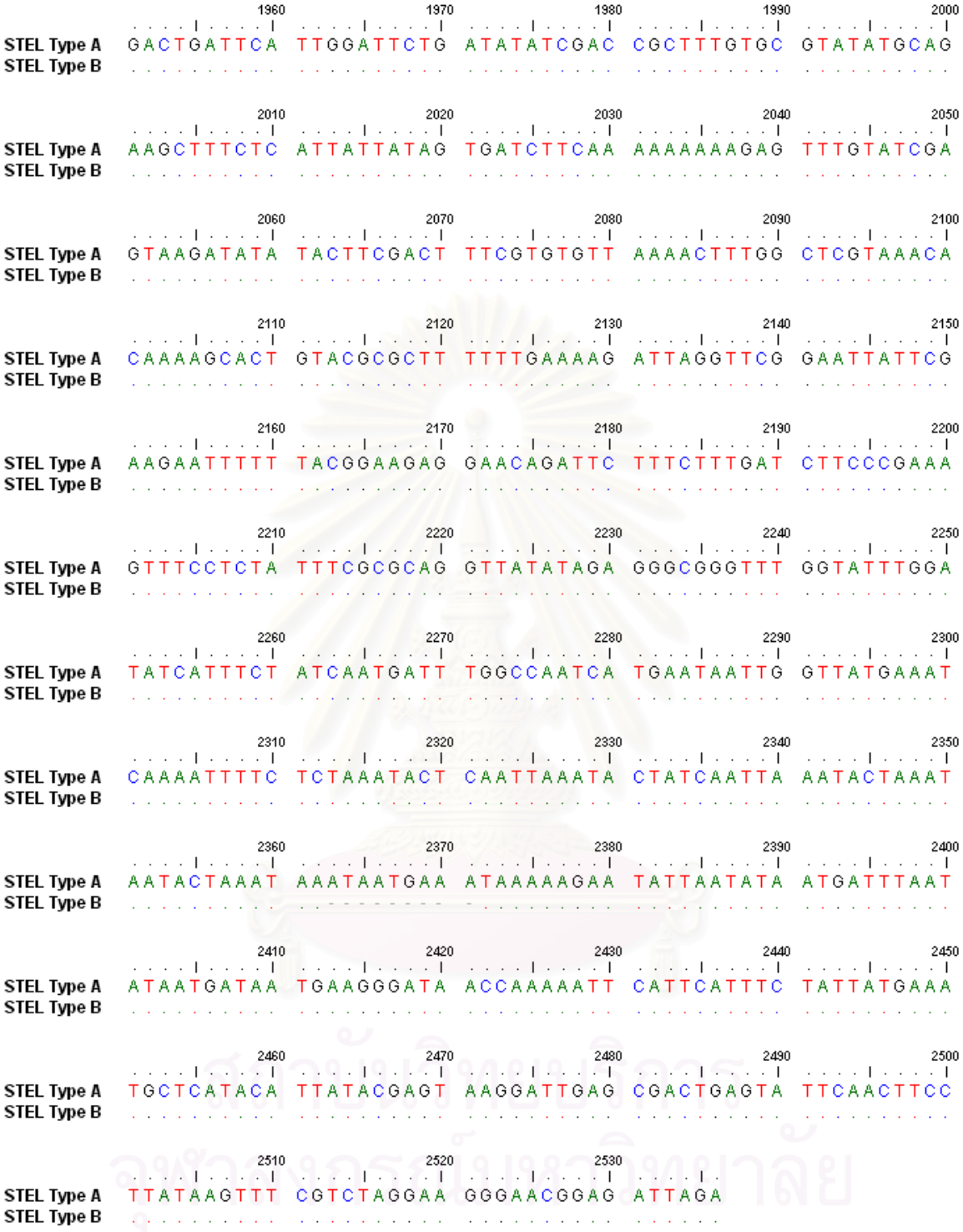
Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

APPENDIC D

Details of TLC Densitometric Analysis
of *C. stellatopilosus* Leave Ethanolic Extracts

Sample	AUC1	AUC2	AUC3	Average	AUC SD	Plaunotol %w/w	SD %(w/w)
STEL 01	7838.8	7928.7	7740.3	7835.93	94.233	0.416	0.0050
STEL 02	12478.1	12455	12425.4	12452.83	26.417	0.662	0.0014
STEL 03	9322.7	9392.7	9348.2	9354.53	35.427	0.497	0.0019
STEL 04	7828.5	7842.5	7898.9	7856.63	37.267	0.418	0.0020
STEL 05	7863.7	7843.4	7927.9	7878.33	44.110	0.419	0.0023
STEL 06	9441	9463.1	9494.9	9466.33	27.095	0.503	0.0014
STEL 07	7296.6	7367.2	7400.9	7354.90	53.227	0.319	0.0028
STEL 08	5443.8	5422.8	5468.4	5445.00	22.824	0.289	0.0012
STEL 09	4827.4	4834.2	4816.9	4826.17	8.716	0.256	0.0005
STEL 10	9305.4	9226.3	9265.4	9265.70	39.551	0.492	0.0021
STEL 11	13371.8	14440.7	13331.4	13714.63	629.117	0.729	0.0334
STEL 12	9395.6	9442.7	9437.5	9425.27	25.823	0.501	0.0014
STEL 13	8402.6	8435.4	8425.6	8421.20	16.837	0.448	0.0009
STEL 14	11336.2	11342.5	11321.6	11333.43	10.721	0.602	0.0006
STEL 15	8301.4	8347	8326.2	8324.87	22.829	0.004	0.0013
STEL 16	3914.8	3964.3	3926.1	3935.07	25.940	0.209	0.0014
STEL 17	13361.1	13371.3	13401.9	13378.10	21.233	0.711	0.0011
STEL 18	7420.8	7395.6	7436.7	7417.70	20.725	0.394	0.0011
STEL 19	15954.7	15952.1	15966	15957.60	7.390	0.848	0.0004
STEL 20	9324.8	9352.1	9347.5	9341.47	14.616	0.496	0.0008
STEL 21	8135	8174.5	8167.3	8158.93	21.037	0.434	0.0011
STEL 22	13458	13368.2	13293.4	13373.20	82.414	0.711	0.0044
STEL 23	13252.1	12942.7	13212.8	13135.87	168.437	0.698	0.0090
STEL 24	15472.1	15308.6	15297.3	15359.33	97.822	0.816	0.0052
STEL 25	16943.4	16835.3	16804.2	16860.97	73.063	0.896	0.0039
STEL 26	16645.3	16528.7	16677.1	16617.03	78.134	0.883	0.0042
STEL 27	13787.2	13647.7	13597.2	13677.37	98.413	0.727	0.0052
STEL 28	18797.5	18847.5	18822.5	18822.50	25.000	1.000	0.0013
STEL 29	13074	12923.7	12104.4	12700.70	521.850	0.675	0.0277
STEL 30	10068.3	9713.9	9843.4	9875.20	179.327	0.525	0.0095

APPENDIC E

The Partial *trn*K Intron Nucleotide Alignment
of *C. roxburghii* (Codes ROX 01-05)

*mat*K gene : position 1-1527
Partial *trn*K 3' intron : position 1528-1773

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

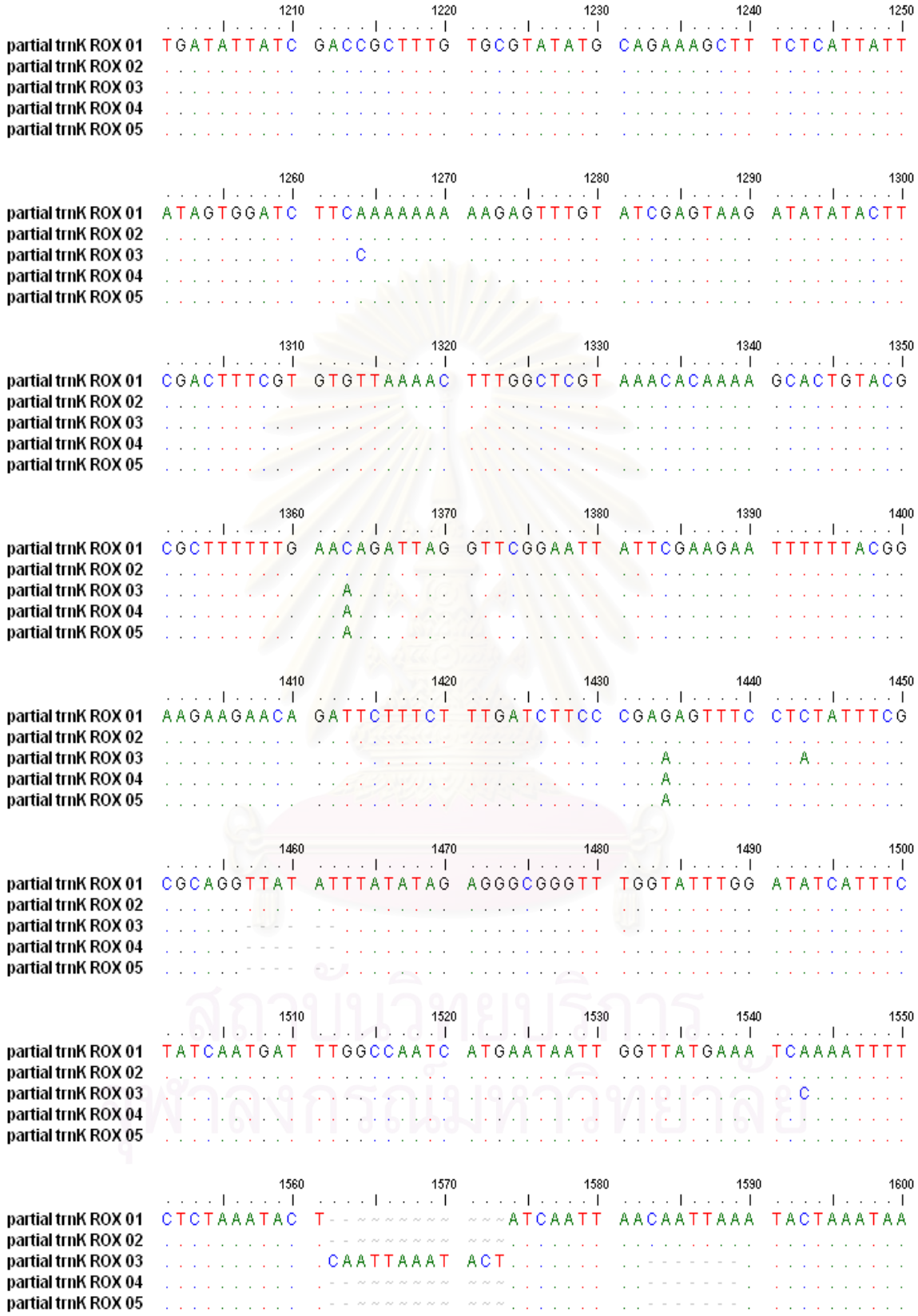
Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

```
                          810        820        830        840        850
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 TGGTCAAAGA TCCTTTTCTG CATTATGTTA AATATCAAGG AAAATCCATT
partial trnK ROX 02 .......... .......... .......... .......... ..........
partial trnK ROX 03 .......... ......C... .......... G......... ..........
partial trnK ROX 04 .......... ......C... .......... G......... ..........
partial trnK ROX 05 .......... ......C... .......... G......... ..........

                          860        870        880        890        900
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 CTGGCTTCAA AAGATGGGCT TCTTCTGATG AAAAAATGGA AATATTACCT
partial trnK ROX 02 .......... .......... .......... .......... ..........
partial trnK ROX 03 ...T...... .......... .......... .......... ..........
partial trnK ROX 04 .......... .......... .......... .......... ..........
partial trnK ROX 05 .......... .......... .......... .......... ..........

                          910        920        930        940        950
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 TGTCAATTTA TGTCAATGTC ATTTTCATGT GTGGTTTCAA CCCCAAAAGA
partial trnK ROX 02 .......... .......... .......... .......... ...A......
partial trnK ROX 03 .......... .......... .....T.... ....G..... ...A......
partial trnK ROX 04 .......... .......... .......... .......... ...A......
partial trnK ROX 05 .......... .......... .......... .......... ...A......

                          960        970        980        990       1000
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 TCTATATAAG TTCATTACCC AAGCATTCTC TCAACCTTTT GGGCTATCTT
partial trnK ROX 02 .......... .......... .......... .......... ..........
partial trnK ROX 03 .......... .......... .......... .......... ..........
partial trnK ROX 04 .......... .......... .......... .......... ..........
partial trnK ROX 05 .......... .......... .......... .......... ..........

                         1010       1020       1030       1040       1050
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 TCAAATGTAC AATTAAATCC TTTAGTCGTA CGAAGTCAAA TGCTAGAAAA
partial trnK ROX 02 .......... .......... .......... .......... ..........
partial trnK ROX 03 .......... .......... ...G...... .......... ..........
partial trnK ROX 04 .......... .......... ...G...... .......... ..........
partial trnK ROX 05 .......... .......... ...G...... .......... ..........

                         1060       1070       1080       1090       1100
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 TTCATTTTTA ATAGAAAAAG ATAATACTAT GAAGAAACTC GATACAATAG
partial trnK ROX 02 .......... .......... .......... .......... ..........
partial trnK ROX 03 .......... .......... .......... .......... ..........
partial trnK ROX 04 .......... .......... .......... .......... ..........
partial trnK ROX 05 .......... .......... .......... .......... ..........

                         1110       1120       1130       1140       1150
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 TTCCAATTAT TCCTTTAATT GGATCATTAT CAGAAATGCA ATTTTGTAAC
partial trnK ROX 02 .......... .......... .......... ..A....... ..........
partial trnK ROX 03 .......... .......... .......... ..A....... ..........
partial trnK ROX 04 .......... .......... .......... ..A....... ..........
partial trnK ROX 05 .......... .......... .......... ..A....... ..........

                         1160       1170       1180       1190       1200
                   ....|....| ....|....| ....|....| ....|....| ....|....|
partial trnK ROX 01 ACAGTAGGGC ATCCCATTAG TAAATCGGCC TGGACTGATT CATTGGATTC
partial trnK ROX 02 .......... .......... .......... .......... ..........
partial trnK ROX 03 .......... .......... .......... .......... ..........
partial trnK ROX 04 .......... .C........ .......... .......... ..........
partial trnK ROX 05 .......... .C........ .......... .......... ..........
```

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.



Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

Nucleotide positions of the entire ITS nucleotide sequences start from the first position of this region.

Dots indicate the nucleotide sequences that are the same as the sequence in the first line and hyphens indicate indels.

VITA

Miss Prasoborn Rinthong was born on November 9, 1978 in Buriram, Thailand. She received her Bachelor's Degree of Science in Pharmacy, Second Class Honors in 2001 from the Faculty of Pharmaceutical Sciences, Khon Khaen University, Thailand. During April 2006 to March 2007, she was an exchange student in Department of Pharmacognosy, Institute of Natural Medicine, University of Toyama, Toyama, Japan. Throughout her Ph.D. program, she was full grant supported from The Commission of Higher Education, Ministry of Education, Thailand. Since, April 2001, Prasoborn has been worked as lecturer at Faculty of Pharmaceutical Sciences, Maha Sarakham University, Maha Sarakham, Thailand.