

CHAPTER IV

MOLECULAR STUDY

The application of DNA sequence data has become one of the most utilized approaches for inferring taxonomic and phylogenetic relationships of plants. In recent years, the phylogenetic analyses using molecular have been increasingly recognized in Zingiberaceae. In this chapter, the DNA sequence data of Thai *Hedychium* from nuclear region is provided.

Materials and methods

Taxa sampling and outgroup

All of plant materials were collected from either forest habitats or plantations in Thailand. Twenty one taxa of *Hedychium* (40 samples) were sequenced for ITS regions as shown in Table 4.1. Attempts were made several times to acquire DNA sequence data of *H. aureum*, *H. coronarium* var. *forrestii*, *H. speciosum* and *H. spicatum* var. *spicatum* from fresh or dried material but DNA samples proved difficult to amplify and fail to sequence. A representative of 4 closely related taxa were included as outgroup; *Boesenbergia plicata*, *B. pulcherrima*, *Kaempferia rotunda* and *Zingiber officinale*. In addition, 30 ITS sequence accessions imported from Genbank were included for phylogenetic analysis (Table 4.2).

DNA extraction

Fresh or silica gel dried leaves of each sample were ground under liquid nitrogen to fine powder using a mortar and pestle. Total DNA extraction was carried out using a DNeasy[®] Plant Mini Kit (QIAGEN, Germany), following the manufacturer's protocol. Total genomic DNA was determined quality and quantity on 0.8 % agarose gel electrophoresis stained by ethidium bromide and visualized under UV light. A 100bp Sharp DNA Marker (RBC Bioscience, Taiwan) was used as standard molecular size. The extracted DNA was stored at -20°C, for further use as templates for PCR amplification.

PCR amplification of target ITS region

PCR amplification of nrDNA internal transcribed spacer region (ITS1, 5.8S and ITS2) was performed as a template in 50 µl reaction, each reaction containing ddH₂O (36.5 µl), 10X PCR buffer (5 µl), 25 mM MgCl₂ (3 µl) (Roche, U.S.A), 10 mM dNTPs (1 µl), 5 U FastStart Taq DNA Polymerase (0.5 µl) (Roche, U.S.A), 0.25 mM Forward primer (1 µl), 0.25 mM Reverse primer (1 µl) and template DNA (2 µl). The amplification of the ITS region was carried out using primer ITS1 (5'-GTC CAC TGA ACC TTA TCA TTT AG-3') for forward and ITS4 (5'-TCC TCC GCT TAT TGA TAT GC-3') for reverse (Urbatsch et al., 2000 and White et al., 1990 in Thong-a-ram et al., 2005).

PCR amplification was carried out in PCR Thermo Cycler, 2720 Thermo Cycler (Applied Biosystems). Each PCR reaction was started with an initial denaturation step at 95°C for 10 minutes to ensure the complete separation of the DNA strands, followed by 35 cycles of double-strand template DNA denaturation (95°C for 30 seconds), single-strand template DNA primer annealing (56°C to 58°C for 40 seconds), and primer extension (72°C for 45 seconds). The final extension step at 72°C for 7 minutes to ensure that all amplicons are fully extended, then held at 4°C and stored at -20°C, for further sequencing. The PCR products were checked on 0.8 % agarose gel electrophoresis at 100 volts, gel was stained by ethidium bromide solution for 15 minutes and destained in deionized water for 5 minutes. DNA was visualized under UV light and photographed. A 100 bp Sharp DNA Marker (RBC Bioscience, Taiwan) was used as standard molecular size.

PCR product purification and nucleotide sequencing

PCR products were purified and nucleotide sequences were determined by Molecular Informatics Laboratory, Hong Kong. Primers for sequencing were ITS1 and ITS4 as shown in Figure 4.1.

Sequence analysis and phylogenetic analysis

About 450-600 bp sequences per sample were usually obtained and the consensus sequences were assembled and analyzed using BioEdit (Hall, 2004). Multiple sequence alignments were evaluated with the freeware program Multalin (Corpet, 1998) and aligned by eyes.

Phylogenetic trees were generated by using PAUP* (Version 4.0 beta 10a, Sinauer Assoc. Inc., U.S.A.). The parsimony analysis employed a Heuristic search with tree-bisection-reconnection (TBR) branch-swapping, MULTREES, a random addition sequence of 100 replicates. Bootstrap analyses were conducted on 1000 replications to assess relative clade support. The consistency index (CI), retention index (RI) and rescaled consistency index (RC) were also calculated using PAUP. Bootstrap values $\geq 85\%$ support were considered strong, 70-84% moderate, 50-69% weak and $< 50\%$ lacking.

Table 4.1 List of specimens examined for molecular phylogenetic analysis.

Botanical name	Locality	Collector numbers
<i>H. biflorum</i> 1	Khao Pra Woa, Tak	<i>P. Wongsuwan</i> 42
<i>H. biflorum</i> 2	Khao Pra Woa, Tak	<i>P. Wongsuwan</i> 45
<i>H. coccineum</i> 1	Pang Mapha, Mae Hong Son	<i>P. Wongsuwan</i> 92
<i>H. coccineum</i> 2	Doi Suthep-Pui, Ching Mai	<i>P. Wongsuwan</i> 74
<i>H. collinum</i>	Cultivated from plants collected at Khuean Banglang, Yala	<i>P. Wongsuwan</i> 137
<i>H. coronarium</i> var. <i>coronarium</i> 1	Cultivated in Ubonrachatani	<i>P. Wongsuwan</i> 68
<i>H. coronarium</i> var. <i>coronarium</i> 2	Cultivated in Khon Kaen	<i>P. Wongsuwan</i> 66
<i>H. coronarium</i> var. <i>chrysoleucum</i> 1	Cultivated in Ching Mai	<i>P. Wongsuwan</i> 47
<i>H. coronarium</i> var. <i>chrysoleucum</i> 2	Cultivated in Khon Kaen	<i>P. Wongsuwan</i> 32
<i>H. coronarium</i> var. <i>flavescens</i> 1	Cultivated in Ching Rai	<i>P. Wongsuwan</i> 70
<i>H. coronarium</i> var. <i>flavescens</i> 2	Cultivated in Khon Kaen	<i>P. Wongsuwan</i> 50
<i>H. ellipticum</i> 1	Phu Ruea, Loei	<i>P. Wongsuwan</i> 72
<i>H. ellipticum</i> 2	Doi Angkhang, Chiang Mai	<i>P. Wongsuwan</i> 79

Table 4.1 List of specimens examined for molecular phylogenetic analysis (Cont.).

Botanical name	Locality	Collector numbers
<i>H. gomezianum</i> 1	Namtok Praan, Trang	<i>P. Wongsuwan</i> 87
<i>H. gomezianum</i> 2	Khlong Naka, Ranong	<i>P. Wongsuwan</i> 71
<i>H. khaomaenense</i>	Khao Maen, Nakon Sri Tamarat	<i>P. Wongsuwan</i> 64
<i>H. longicornutum</i>	Cultivated from plants collected in Hala-Bala, Narathiwat	<i>P. Wongsuwan</i> 44
<i>H. malayanum</i>	Cultivated from plants collected on top of Hangus mountain	<i>P. Wongsuwan</i> 89
<i>H. muanwongyathiae</i>	Doi Angkhang, Chiang Mai	<i>Picheansoonthon & Wongsuwan</i> 52
<i>H. neocarneau</i> 1	Phu Luang, Loei	<i>P. Wongsuwan</i> 73
<i>H. neocarneau</i> 2	Phu Ruea, Loei	<i>P. Wongsuwan</i> 59
<i>H. pauciflorum</i> 1	Doi Suthep-Pui, Chiang Mai	<i>P. Wongsuwan</i> 69
<i>H. pauciflorum</i> 2	Pang Mapha, Mae Hong Son	<i>P. Wongsuwan</i> 94
<i>H. phuluangense</i>	Phu Luang, Loei	<i>P. Wongsuwan</i> 51
<i>H. roxburghii</i> 1	Cultivated from plants collected on Ko Samui, Suratthani	<i>P. Wongsuwan</i> 138
<i>H. roxburghii</i> 2	Khlong Naka, Ranong	<i>P. Wongsuwan</i> 116
<i>H. roxburghii</i> 3	Ko Samui, Suratthani	<i>P. Wongsuwan</i> 139

Table 4.1 List of specimens examined for molecular phylogenetic analysis (Cont.).

Botanical name	Locality	Collector numbers
<i>H. siamense</i>	Khlong Naka, Ranong	<i>P. Wongsuwan</i> 130
<i>H. spicatum</i> var. <i>acuminatum</i> 1	Pang Mapha, Mae Hong Son	<i>P. Wongsuwan</i> 99
<i>H. ellipticum</i> 3	Doi Chiang Dao, Chiang Mai	<i>P. Wongsuwan</i> 110
<i>H. spicatum</i> var. <i>acuminatum</i> 2	Doi Suthep-Pui, Chiang Mai	<i>P. Wongsuwan</i> 140
<i>H. stenopetalum</i> 1	Doi Intanon, Chiang Mai	<i>P. Wongsuwan</i> 48
<i>H. stenopetalum</i> 2	Doi Angkhang, Chiang Mai	<i>P. Wongsuwan</i> 80
<i>H. tomentosum</i>	Doi Chiang Dao, Chiang Mai	<i>P. Wongsuwan</i> 127
<i>H. villosum</i> 1	Phu Luang, Loei	<i>P. Wongsuwan</i> 62
<i>H. villosum</i> 2	Doi Chiang Dao, Chiang Mai	<i>P. Wongsuwan</i> 122
<i>H. aff. roxburghii</i>	Cultivated from plants collected at Ranong	<i>P. Wongsuwan</i> 141
<i>H. aff. spicatum</i> var. <i>acuminatum</i>	Doi Suthep-Pui, Chiang Mai	<i>P. Wongsuwan</i> 142
<i>H. aff. siamense</i> 1	Cultivated from plants collected at Mueang Pilok, Kanchanaburi	<i>P. Wongsuwan</i> 143
<i>H. aff. siamense</i> 2	Cultivated from plants collected at Mae Sot, Tak	<i>P. Wongsuwan</i> 144

Table 4.2 List of specimens examined for molecular phylogenetic analysis from GenBank

Taxon name	GenBank accession number
Ingroup	
<i>H. acuminatum</i> Ker Gawl.	AF202379.1
<i>H. borneense</i> R.M. Sm.	AF202380.1
<i>H. bousigonianum</i> Gagnep.	AF202381.1
<i>H. carneum</i> Y.Y. Qian	AF202382.1
<i>H. coccineum</i> Ker Gawl.	AF202375.1
<i>H. coronarium</i> König	AF202383.1
<i>H. cylindricum</i> Ridl.	AF202384.1
<i>H. densiflorum</i> Wall.	AF202402.1
<i>H. densiflorum</i> 'Stephen'	AF202385.1
<i>H. ellipticum</i> Ker Gawl.	AF202386.1
<i>H. elwesii</i> Baker	AF202378.1
<i>H. flavescens</i> Roscoe	AF202401.1
<i>H. gardnerianum</i> Roscoe	AF202391.1
<i>H. glabrum</i> S.Q. Tong	AF202387.1
<i>H. gracile</i> Roxb.	AF202399.1
<i>H. gracillimum</i> A.S. Rao & Verma	AF202400.1
<i>H. greenii</i> W.W. Sm.	AF202388.1
<i>H. hasseltii</i> Blume	AF202389.1
<i>H. horsfieldii</i> Wall.	AF202374.1
<i>H. longicornutum</i> Baker	AF202377.1
<i>H. maximum</i> Roscoe	AF202392.1
<i>H. muluense</i> R.M. Sm.	AF202393.1
<i>H. puerense</i> Y.Y. Qian	AF202394.1
<i>H. spicatum</i> Ker Gawl.	AF202395.1
<i>H. stenopetalum</i> Lodd.	AF202396.1
<i>H. tenuiflorum</i> K. Schum.	AF202376.1
<i>H. thyrsiforme</i> Ker Gawl.	AF202403.1
<i>H. urophyllum</i> Lodd.	AF202397.1
<i>H. yunnanense</i> Gagnep.	AF202398.1
<i>H. gardnerianum</i> x <i>hasseltii</i>	AF202390.1
Outgroup	
<i>B. plicata</i> (Ridl.) Holttum	AJ388279
<i>B. pulcherrima</i> (Wall.) Kuntze	AF478725
<i>K. rotunda</i> L.	AF478767
<i>Z. officinale</i> Roscoe	DQ064590

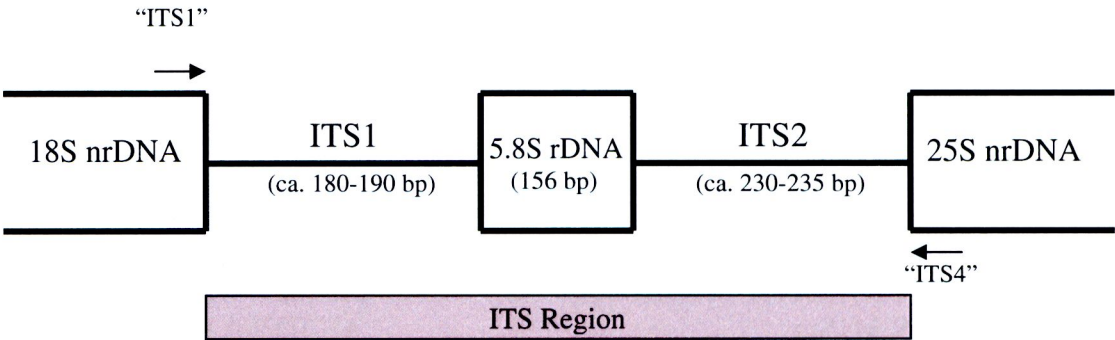


Figure 4.1 Relative positions and directions of the PCR amplification and sequencing primers in nuclear gene region ITS. Arrows (→) represent forward primer and Arrows (←) represent reward primer (modified from Takano and Okada, 2002).

Results

Analysis 1, individual ITS sequence data in Thai *Hedychium*

The aligned matrix of ITS sequences in Thai *Hedychium* analyzed resulted in data matrix contains 599 characters included which 419 of constant characters, 88 parsimony-informative variable characters and 92 parsimony-informative characters.

Parsimony analysis of ITS data set from 44 taxa resulted in 237 most parsimonious trees, with length=259, CI=0.7992, RI=0.8606 and RC=0.6878. One of most parsimonious trees is given in Figure 4.3 with bootstrap values. The tree shows monophyly of *Hedychium* is highly supported (bootstrap=100%) and four clades of *Hedychium* species were defined with moderately to strongly supported bootstrap value (bootstrap=70-100%).

Clade 1, *H. biflorum* 1, *H. biflorum* 2, *H. collinum*, *H. gomezianum* 1, *H. gomezianum* 2, *H. khaomaenense*, *H. longicornutum*, *H. malayanum*, *H. roxburghii* 3, *H. roxburghii* 2, *H. roxburghii* 1, *H. aff. roxburghii*, *H. siamense*, *H. aff. siamense* 1, *H. aff. siamense* 2, *H. villosum* 1 and *H. villosum* 2. Within clade 1, *H. biflorum* 1 and *H. biflorum* 2 are moderately supported bootstrap value (bootstrap=75%) as closely related taxa and nucleotide are identical. *H. siamense*, *H. aff. siamense* 1 and *H. aff. siamense* 2 are strongly supported bootstrap value (bootstrap=96%) and identical nucleotide. *H. gomezianum* 1 and *H. gomezianum* 2 are strongly supported bootstrap value (bootstrap=99%) and nucleotide are identical. *H. roxburghii* 3, *H. roxburghii* 2, *H. roxburghii* 1 and *H. aff. roxburghii* are moderately supported bootstrap value (bootstrap=84%) and nucleotide are identical. *H. villosum* 1 and *H. villosum* 2 are moderately supported bootstrap value (bootstrap=76%) and nucleotide are identical.

Clade 2, *H. coccineum* 1, *H. coccineum* 2, *H. coronarium* var. *coronarium* 1, *H. coronarium* var. *coronarium* 2, *H. coronarium* var. *chrysoleucum* 1, *H. coronarium* var. *chrysoleucum* 2, *H. coronarium* var. *flavescens* 1, *H. coronarium* var. *flavescens* 2, *H. muanwongyathiae*, *H. neocarneum* 1, *H. neocarneum* 2, *H. stenopetalum* 1 and *H. stenopetalum* 2. Within clade 2, *H. coronarium* 1 and *H. coronarium* 2 are weakly supported bootstrap value (bootstrap=76%) and nucleotide are identical. *H. neocarneum*

1 and *H. neocarneum* 2 are weakly supported bootstrap value (bootstrap=62%) and nucleotide are identical.

Clade 3, *H. ellipticum* 1, *H. ellipticum* 2, *H. ellipticum* 3, *H. phuluangense*, *H. pauciflorum* 1, *H. pauciflorum* 2 and *H. tomentosum*. With in clade 3, *H. ellipticum* 1, *H. ellipticum* 2, *H. ellipticum* 3 and *H. phuluangense* are moderately supported bootstrap value (bootstrap=77%). *H. pauciflorum* 1 and *H. pauciflorum* 2 are strongly bootstrap supported value (bootstrap=99%) and nucleotide are identical.

Clade 4, *H. spicatum* var. *acumimatum* 1, *H. spicatum* var. *acumimatum* 2 and *H. aff. spicatum* var. *acumimatum* are strongly supported bootstrap value (bootstrap=100%) and nucleotide are identical.

Analysis 2, combined ITS sequence data from Thailand and Genbank

The aligned matrix of ITS sequences in Thai *Hedychium* and data from Genbank analyzed resulted in data matrix contains 620 characters of which 400 of constant characters, 124 parsimony-informative variable characters and 96 parsimony-informative characters.

Parsimony analysis of ITS data set from 55 taxa resulted in 321 most parsimonious trees, with length=366, CI=0.6995, RI=0.7483 and RC=0.5234. One of most parsimonious trees is given in Figure 4.4 with bootstrap values. The tree shows monophyly of *Hedychium* is highly supported bootstrap value (bootstrap=100%) and four clades of *Hedychium* species were defined with weakly supported bootstrap value (bootstrap=50-60%).

Clade 1, *H. biflorum*, *H. collinum*, *H. siamense*, *H. cylindricum*, *H. gomezianum*, *H. longicornutum*, *H. longicornutum* (Genbank), *H. khaomaenense*, *H. malayanum*, *H. roxburghii*, *H. borneense*, *H. bousigonianum*, *H. hasseltii*, *H. horsfieldii*, *H. muluense* and *H. villosum*. Within clade 1, *H. longicornutum* and *H. longicornutum* (Genbank) are strongly supported bootstrap value (bootstrap=96%) as closely related taxa.

Clade 2, *H. yunnanense*, *H. glabrum*, *H. ellipticum*, *H. ellipticum* (Genbank), *H. gracillimum*, *H. phuluangense* and *H. gracile*. Within clade 2, *H. yunnanense* and *H. glabrum* are weak bootstrap supported (bootstrap=61%). *H. ellipticum* (Genbank) and *H. gracillimum* are weakly supported bootstrap value (bootstrap=62%).

Clade 3, *H. coccineum*, *H. coccineum* (Genbank), *H. coronarium* var. *coronarium*, *H. coronarium* (Genbank), *H. maximum*, *H. coronarium* var. *chrysoleucum*, *H. coronarium* var. *flavescens*, *H. flavescens*, *H. muanwongyathiae*, *H. neocarneum*, *H. stenopetalum*, *H. stenopetalum* (Genbank), *H. greenii*, *H. spicatum*, *H. thyrsoforme*, *H. carneum*, *H. elwesii*, *H. gardnerianum*, *H. puerense*, *H. tenuiflorum*, *H. urophyllum* and *H. gardnerianum* x *hasseltii*. Within clade 3, *H. stenopetalum* (Genbank), *H. greenii* and *H. spicatum* are weakly supported bootstrap value (bootstrap=55%). *H. thyrsoforme* is weakly supported bootstrap value (bootstrap=59%).

Clade 4, *H. pauciflorum*, *H. spicatum* var. *acuminatum*, *H. acuminatum*, *H. tomentosum*, *H. densiflorum* 1 and *H. densiflorum* 2. Within clad 4, *H. spicatum* var. *acuminatum* and *H. acuminatum* are weakly supported bootstrap value (bootstrap=54%). *H. densiflorum* 1 and *H. densiflorum* 2 are strongly supported bootstrap value (bootstrap=54%).

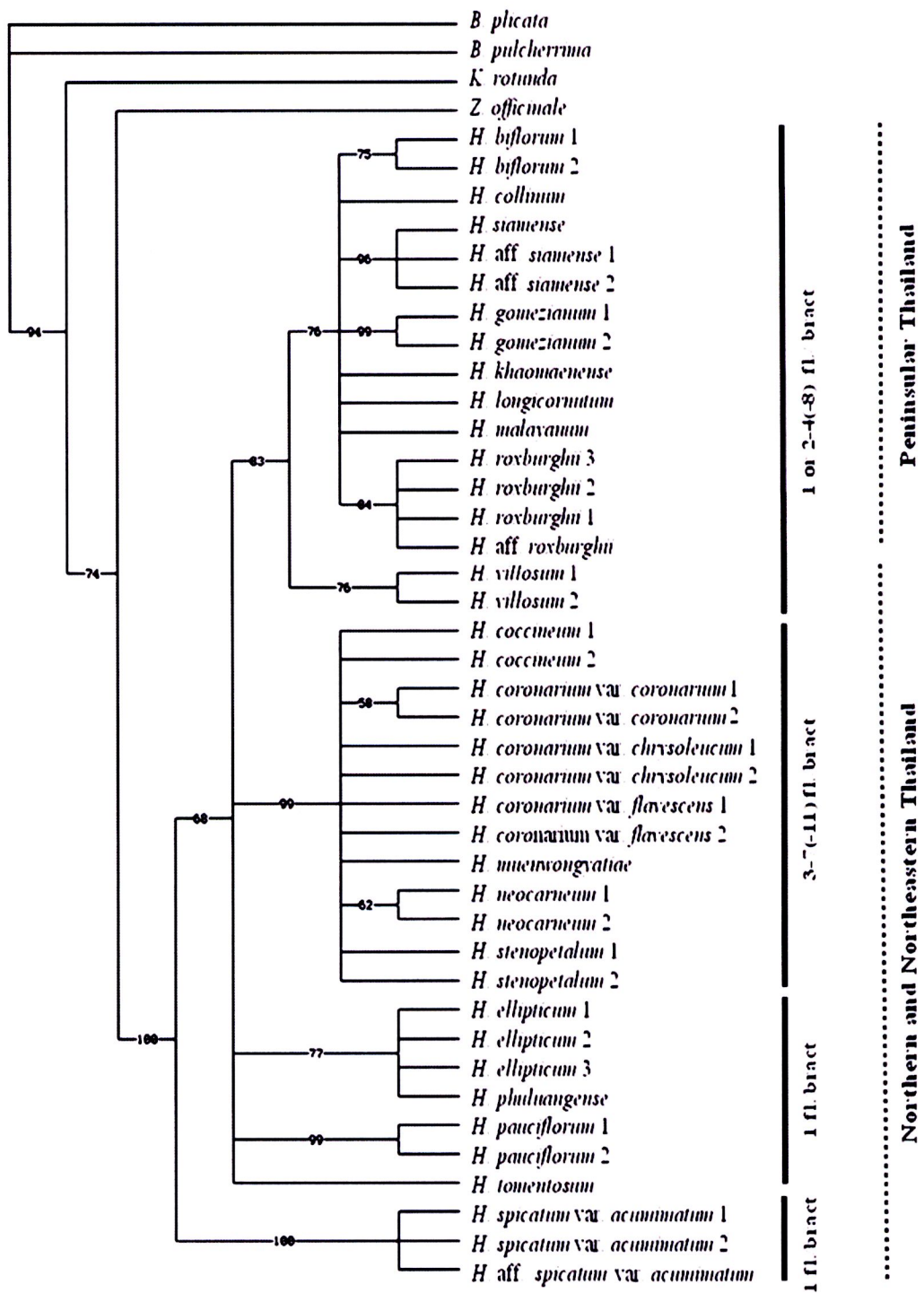


Figure 4.2 One of 237 equally most parsimonious trees of Thai *Hedychium* based on ITS sequences with length=259, CI=0.7992, RI=0.8606 and RC=0.6878. Number greater than 50% bootstrap values are given on the lines with 1,000 replicates.

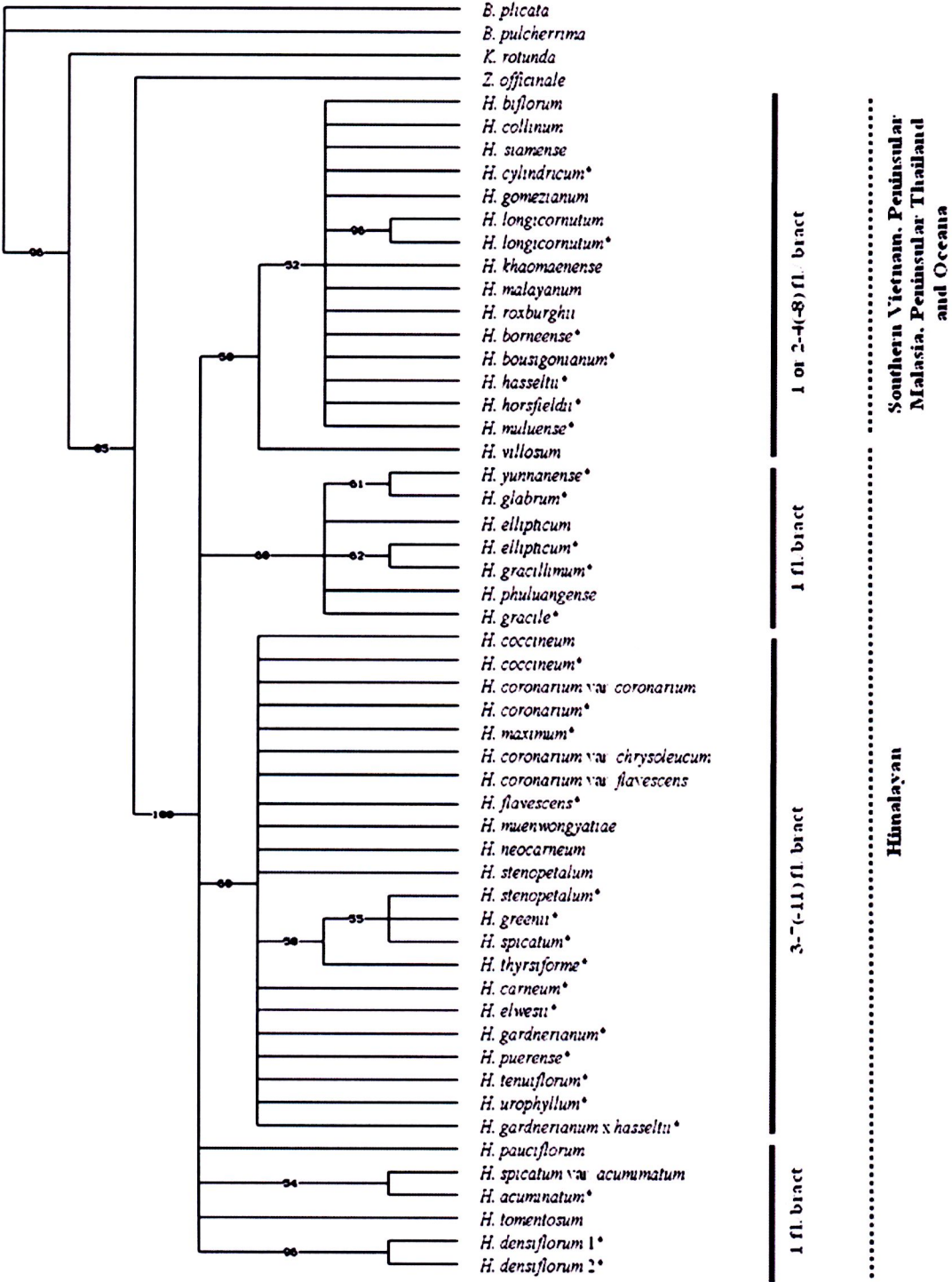


Figure 4.3 One of 321 equally most parsimonious trees of Thai *Hedychium* and data from Genbank based on ITS sequences with length=366, CI=0.6995, RI=0.7483 and RC=0.5234. Number greater than 50% bootstrap values are given on the lines with 1,000 replicates.

* sequencing data from Genbank

Discussions and Conclusions

Individual ITS sequence data in Thai *Hedychium*

Clade 1, contains imbricate bract species; *H. siamense*, *H. khaomaenense*, *H. longicornutum*, and *H. malayanum*. Most species are cylindrical imbricate bract except *H. longicornutum* and not imbricate bract species; *H. biflorum*, *H. collinum*, *H. gomezianum*, *H. roxburghii*, and *H. villosum*. Within clade 1, most species are distributed in peninsular Thailand, but *H. siamense* and *H. biflorum* are found in northern Thailand (Tak Province) and all possessing 1 or 2-4(-8) flower(s) per bract. However *H. villosum*, a Himalayan species also found in northern and northeastern Thailand, seems to be closely related to *H. collinum* in relatively small anther (only two Thai *Hedychium* possess small anther (3-5 mm longs)).

Clade 2 contains imbricate to elongated bract species; *H. coronarium* var. *coronarium*, *H. coronarium* var. *chrysoleucum*, and *H. coronarium* var. *flavescens* cultivated throughout Thailand are not imbricate bract species; *H. coccineum*, *H. muanwongyathiae*, *H. neocarneum*, and *H. stenopetalum* are found in northern and northeastern Thailand. All species possessing 3-7(-11) flowers per bract are Himalayan species.

Clade 3 contains one imbricate bract species, *H. ellipticum*; and not imbricate bract species, *H. phuluangense*, *H. pauciflorum*, and *H. tomentosum* are found in northern and northeastern Thailand. All species are Himalayan elements possessing 1 flower per bract and filament as long as or longer than labellum.

Clade 4 contains one not imbricate species; *H. spicatum* var. *acumimatum* distributed in northern Thailand. This Himalayan species possesses 1 flower per bract and filament shorter than labellum.

Combined ITS sequence data from Thailand and Genbank

Clade 1 contains *Hedychium* distributed in southern Vietnam, Peninsular Malasia, Peninsular Thailand and Oceania. This close relationship correlates to Wood *et.*

al (2000) in possessing 1 or 2-4(-8) flower(s) per bract. However, *H. villosum*, a Himalayan species, seems to be closely related to *H. collinum* in small anther.

Clade 2 contains Himalayan species and all species possess 1 flower per bract. This close relationship correlates to Wood *et. al* (2000).

Clade 3 contains Himalayan species and all species possess 3-7(-11) flowers per bract. This close relationship correlates to Wood *et. al* (2000). However, from this study *H. spicatum* var. *spicatum* possessing only 1 flower per bract is difficult to amplify and fail to sequence. The resulted inclusion of *H. spicatum* within this clade, seems to be a problem in the sampling identification from original data set for Genbank.

Clade 4 contains Himalayan species and all species possess 1 flower per bract. This close relationship correlates to Wood *et. al* (2000). However, *H. acuminatum* and *H. venustum* in this group are additions from Wood *et. al* (2000).

Conclusion of the molecular studies

1) The ITS gene sequences of Thai *Hedychium* have shown that the genus *Hedychium* is monophyletic comprising of four clades. Clade 1 contains most species with cylindrical imbricate bract, except *H. longicornutum* of which bracts are not imbricate. Within clade 1, most species are distributed in peninsular Thailand (only *H. villosum* found in northern and northeastern Thailand) and possess 1 or 2-4(-8) flower(s) per bract. Clade 2 contains imbricate to elongated bract species which are cultivated throughout Thailand and not imbricate bract species which are found in northern and northeastern Thailand. All species are Himalayan elements possessing 3-7(-11) flowers per bract. Clade 3 contains both imbricate and not imbricate bract species, found in northern and northeastern Thailand. All species are Himalayan elements possessing 1 flower per bract and filament as long as or longer than labellum. Clade 4 contains only one not imbricate species, found in northern Thailand. The species is a Himalayan element possessing 1 flower per bract and filament shorter than labellum.

2) The ITS gene sequences of combined data from Thailand and Genbank have shown that the genus *Hedychium* is monophyletic comprising of four clades. Clade

1 contains plants occurring in southern Vietnam, peninsular Malaysia, peninsular Thailand, Philippines, Borneo, Sumatra, Java Sulawesi and the Moluccas (except *H. villosum*, with circum-Himalayan distribution). All species possess 1 or 2-4 flower(s) per bract. Clade 2 and clade 4 contain all Himalayan species with 1 flower per bract. And clade 3 contains all Himalayan species with 3-7(-11) flowers per bract.

3) The ITS gene sequences data in Thai *Hedychium* of two specimens of *H. biflorum*, *H. coccineum*, *H. coronarium* var. *coronarium*, *H. coronarium* var. *chrysoleucum*, *H. coronarium* var. *flavescens*, *H. gomezianum*, *H. neocarneum*, *H. pauciflorum*, *H. stenopetalum*, *H. villosum*, three specimens of *H. siamense*, *H. spicatum* var. *acuminatum* and four specimens of *H. roxburghii* were determined and have shown to have identical nucleotides sequence despite of different collection localities.