CHAPTER IV

MOLECULAR STUDY

4.1 Introduction

Molecular data have had a profound impact on the field in plant systematics. Previously, the majority of data used in plant molecular phylogenetic studies mostly derives from chloroplast DNA and nuclear rDNA (Small et al., 2004). Phylogenetic reconstruction in the grass family, Poaceae have ongoing early in this century with proposed evolutionary hypotheses based on assessment of existing knowledge of grasses (e.g. Bew, 1929; Hubbard, 1948; Prat, 1960; Stebbins, 1956, 1982; Clayton, 1981; Tsvelev, 1983). Recently, molecular information has provided the basis for phylogenetic hypotheses in grasses at the subfamily and tribe levels. DNA sequence data are now commonly used to resolve problems which taxonomists have different opinion on relationships when using traditional characters. These molecular studies were based on information from chloroplast DNA (cpDNA) restriction sites and DNA sequencing of the *trn*L-F and the ITS region of the subtribes Ischaeminae and Rottboelliinae that have never been conducted earlier. Therefore, the analyses will provide the DNA sequence data of both subtribes in Thailand, which can be serve as a basis for further work in grass systematic.

4.1.1 Chloroplast genome

Noncoding sequences of the chloroplast genome are a primary source of data for molecular systematic, phylogeographic, and population genetic studies in plants. The chloroplast genomes from higher plants typically range in size from 120 to 170 kilobase pairs (kb), and there is a relatively high degree of conservation in size, structure, gene content, and linear order of the genes in land plants. Chloroplast (cp-) DNA genome is a circular molecule and subdivided into two single-copy regions, the large single copy (LSC) region and small single copy (SSC) region, which are separated by inverted repeats. The LSC region is slightly less conserved in sequence than the rest of the chloroplast genome (Clegg, Learn & Golenberg, 1991) (Fig. 4.1). Previous studies have suggested that the inverted repeat regions accumulate point mutations slower than the single-copy regions (Curtis & Clegg, 1984; Wolfe et al.,

1987; Wolfe, 1991; Gaut, 1998). Perry and Wolfe (2002) showed that the nucleotide substitution rate is 2.3 times higher in the single-copy regions relatively to the inverted repeats. Because of the inverted repeats evolve at relatively slower rate and the first adopted chloroplast regions (e.g., *rbc*L, *atp*B, *trn*L-F) are located in the LSC, most plant researchers using molecular tools have focused on the single-copy regions. The use of non coding chloroplast DNA sequences to generate plant phylogenies began in the early 1990s with the seminal publications of Taberlet et al. (1991), Clegg et al. (1994), Morton and Clegg (1993), and Gielly and Tabelet (1994).

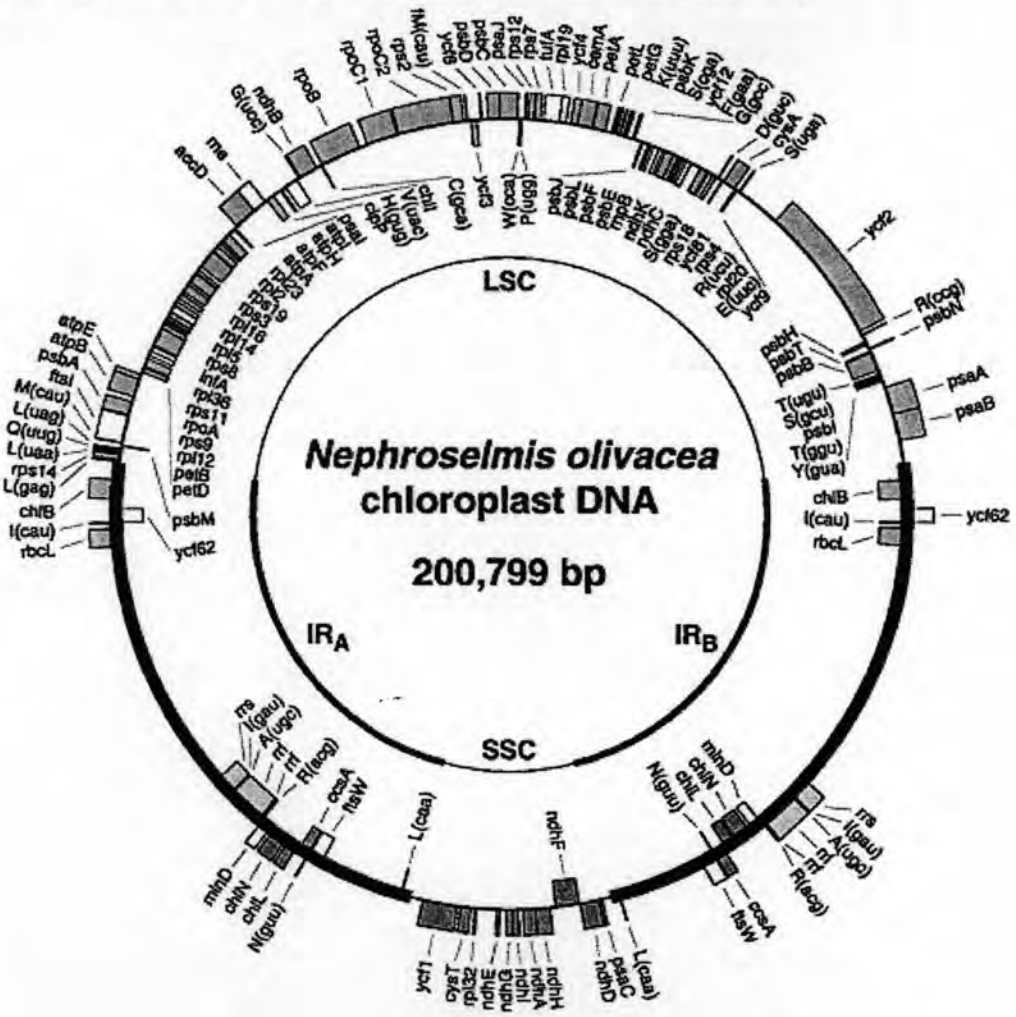

Figure 4.1 Physical and gene map of the green algae *Nephroselmis* chloroplast genome, showing the typical structural arrangement found in land plants. Genes located on the inside of the map are transcribed counter clockwise, and genes on the outside are transcribed clockwise. The inner circle shows where the Small Single-Copy region (SSC), Large Single Copy region (LSC) and Inverted Repeats (IR) are located (adapted from Turmel et al., 1999).

The cpDNA *trn*T-L-F region consists of three non-coding regions (Fig. 4.2): firstly is an intergenic spacer between *trn*T (UGU) and the *trn*L (UAA) 5' exon, second is the *trn*L (UAA) intron and the last region is another intergenic spacer between the *trn*L (UAA) 3' exon and *trn*F (GAA) (Taberlet et al., 1991). In particular the *trn*L intron and *trn*L-F spacer (collective, the *trn*L-F region) has become one of the most widely used chloroplast markers for phylogenetic analysis in plants (Quandt et al., 2004) due to their easily to amplified and their small size, with the *trn*L intron ranging from 350 to 600 bp and the *trn*L-F spacer ranging from approximately 129 to 350 bp (Soltis & Soltis, 1998). However, the different portions of the cpDNA evolve at different rates, which result in a wide range of possibilities for resolving relationships from species and genus level to family and even higher taxonomic levels but limit its applicability among closely related species and populations (Soltis & Soltis, 1998).

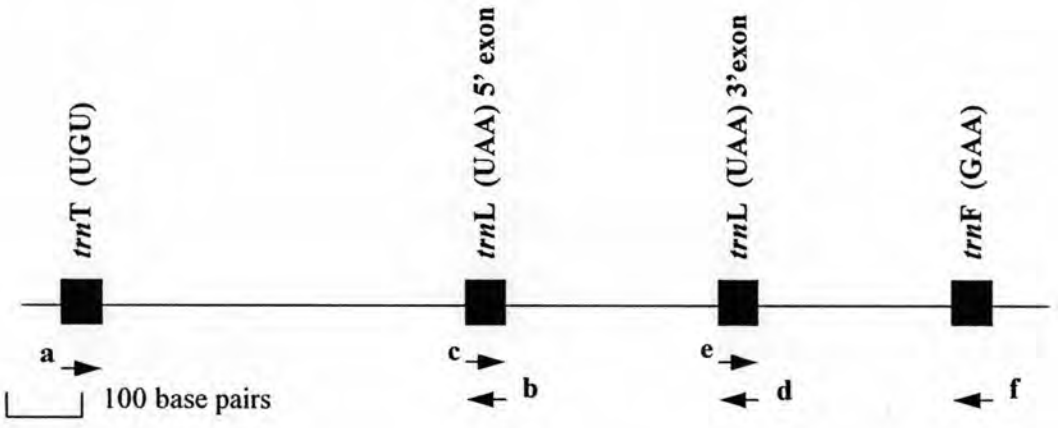

Figure 4.2 Organisation of three non-coding regions of cpDNA. Tips of arrows indicate 3' ends of the primers (adapted from Taberlet et al., 1991).

4.1.2 Nuclear sequences

Nuclear ribosomal DNA (nrDNA) is organized as individual chromosomal units that are repeated thousands of times in most of the higher plant genomes. Each repeat contains a transcribed region, which is separated from the adjacent repeat by a long non-transcribed intergenic spacer (IGS) (Hamby & Zimmer, 1992). Within the transcribed region are three conserved ribosomal RNA regions and the two non-coding spacer regions. The conserved regions, comprising the ribosomal RNA gene

(Fig. 4.3), are arranged in 5'-18S-5.8S-26S-3' order with the ITS spacers (designated as ITS1 and ITS2) flanking the 5.8S region, an evolutionary highly conserved sequence. In addition, an external transcribed spacer region (ETS) is situated at the beginning of the 5' end of the transcribed unit.

The occurrence of high nucleotide variability in combination with low length variation found in both of the internal transcribed spacer regions (ITS1 and ITS2), suggested that this part of the rDNA gene could be used in reconstructing phylogenies within plant genera. Therefore, sequences could be readily aligned across related species and yet would contain sufficient variations for resolution of phylogeny (Baldwin et al., 1995).

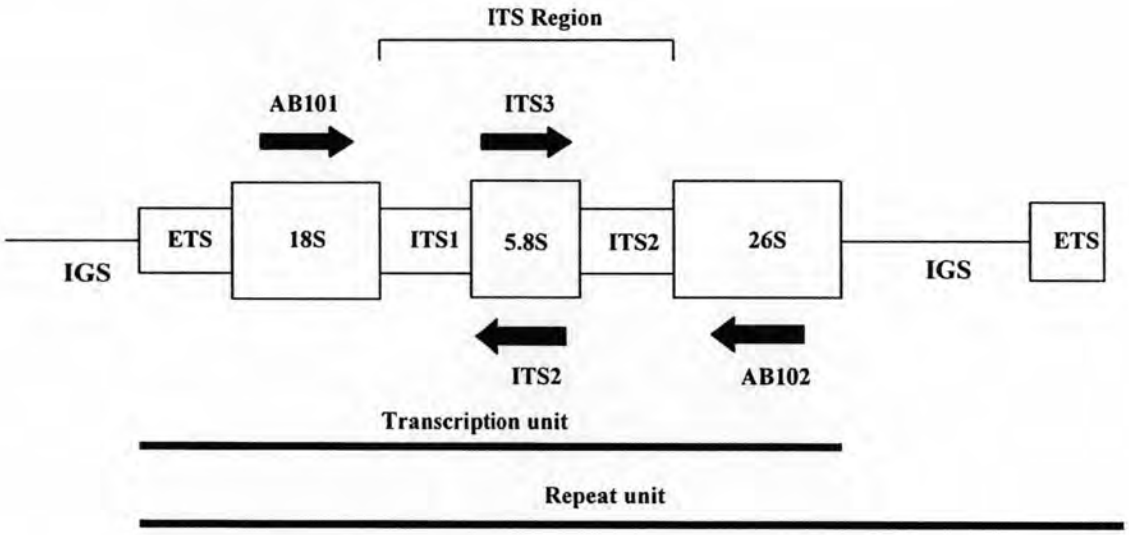

Figure 4.3 Organisation of a repeat unit of nrDNA showing the ITS, ETS, 18S, 5.8S and 26S regions. Arrows indicate the orientation and approximate position of primer sites. Primer names and sequences are from Sun et al. (1994) (adapted from Baldwin et al., 1995).

It has been relatively easy to develop primers for the amplification of the ITS region because it is flanked by the highly conserved regions 18S and 26S and the set of primers designed by White et al. (1990) has been used successfully for amplification across a wide range of plant families. Together with the rest of the nrDNA gene, ITS undergoes rapid concerted evolution, such that all members of the gene family exhibit the same sequence within an individual (Amheim, 1983) via

unequal crossing over and gene conservation. Its small size (less than 800 bp) makes it particularly appropriate for direct sequencing amplified DNA obtained from PCR. The ITS sequences have proven to be a valuable source of characters to address phylogenetic relationships among closely related species in different plant families (Francisco-Ortega et al., 2001) and also have a powerful in revealing hybridization and reticulate evolution (Sang, Crawford & Stuessy, 1997).

4.2 Literature reviews

The molecular phylogenetic information available for the family has been summarized by Linder and Kellogg (1995). They concluded that the evidence gathered till now points to a monophyletic Panicoideae sister to the Centothecoideae. The monophyly of the Andropogoneae is also very well supported, but no conclusions can be reached as to the status of the rest of the tribes.

Clayton (1972, 1973) divided Andropogoneae into "awned" and "awnless" taxa. Later, Clayton & Renvoize (1986) divided the tribe into 11 subtribes, based largely on characters of the inflorescence.

Kellogg & Watson (1993) undertook a phylogenetic analysis of morphological characters for all genera of Andropogoneae. They found that the distinction between awned and awnless taxa was largely supported but most of the subtribes defined by Clayton & Renvoize (1986) were polyphyletic.

Spangler et al. (1999) using cpDNA from gene *ndhF* sequences to estimate the phylogeny of the tribe Andropogoneae. The investigations showed that subtribal designations are not informative due to the lack of resolution between clades in the strict consensus tree. However, there are well-supported clades in the tribe, including three *Sorghum* lineages and a "core" Andropogoneae clade, even though relationships among clades are poorly supported. A preliminary assessment of the generic limits of the genus *Sorghum* indicated a paraphyletic group, including the genera *Cleistachne*, *Miscanthus*, and a species of *Microstegium*. They have proposed subtribe Sorginae which included *Sorghum, Sorghastrum, Cleistachne, Bothriochloa, Capillipedium, Dichanthium* and *Chrysopogon* and does not form a monophyletic group.

Kellogg (2000) suggested that both molecular and morphological data supported the monophyly of the Andropogoneae. The characters that correlated with the origin of the traditional Andropogoneae are presence of a disarticulating rachis

and differentiation of the spikelets of a pair. Within the tribe, there is variation in presence of distinctive cells in the mesophyll, formation of branch complexes on the upper part of the culm (the anthotagma), timing of inflorescence branching, sex expression of spikelets, induration of glumes, and formation of awns. However, the different states of these characters do not correlate with the molecular phylogeny. Hence, the subtribes of Andropogoneae are polyphyletic.

Mathews et al. (2002) added newly collected data from phytochrome B to the data from the other two genes (GBSSI and *ndh*F) in previous studies (Mathews et al., 1995). The phylogeny shows a single origin of a disarticulating rachis, which is a synapomorphy for the tribe. The combined data reject the monophyly of subtribes Andropogonenae and Anthistiriinae and provided evidence that subtribes Sorghinae, Saccharinae, and Rottboelliinae are paraphyletic or polyphyletic. A *Chionachne* and *Phcelurus* are shown to diverge early in the history of the tribe.

In this study, nucleotide sequences data from taxa of subtribes Ischaeminae and Rottboelliinae are cladistically analysed using outgroup comparison to produce a rooted phylogenetic hypothesis for the two subtribes. The sequence data (both substitution and insertion and deletion characters) are from regions of the taxa's chloroplast (cpDNA) and nuclear ribosomal DNA (nrDNA). In the cpDNA, the *trn*L-*trn*F region (composed of the *trnL* (UAA) intron, the *trn*L (UAA) 3' exon, and the intergenic spacer between the *trn*L (UAA) 3' exon and *trn*F (GAA)) was chosen because it has proved to be phylogenetically informative at the infra-and inter-generic levels of the other taxa (Gielly & Taberlet, 1994; Gielly et al., 1996). Similarly, the Internal Transcribed Spacer (ITS) region of nrDNA (composed of the spacers between the 18S and 5.8S genes (ITS1) and the 5.8S and 26S genes (ITS2) was chosen for its informativeness elsewhere at those levels (Baldwin et al., 1995).

The ITS regions has been investigated extensively and has been successfully used to resolve phylogenetic relationships at various levels. This includes studies of various genera such as *Brachisria* and *Urochloa* (González & Morton, 2005); *Bouteloua* (Columbus et al., 1998, 2000); *Chloris* (Alice et al., 2000); *Deschampsia* (Chiapella, 2007); *Deschampsia Antarctica* (Fernández Souto et al., 2006); *Eleusine* (Neves et al., 2005); *Elymus* (Liu et al., 2006); *Hordeum* (Blattner, 2004); *Paspalum* (Souza-Chies & Essi, 2007); *Schizostachyum* and its allies (Yang, Peng & Li, 2007); *Spartina* (Baumel et al., 2001); *Sporobolus* (Ortiz-Diaz & Culham, 2000) and *Sporobolus indicus* complex (Shrestha et al., 2003). Studies have also been conducted

on a related group of grasses known as the Festucoid grass (Catalán et al., 2004); Stipoid grasses (Jacobs et al., 2000); Chloridoid grasses (Roodt-Wilding & Spies, 2006); the Chloridoid tribe Triodieae (Mant et al., 2000); the Panicoid tribe Androponeae (Hodkinson et al., 2002b); the Pooi tribes Aveneae (Grebenstein et al., 1998) and Triticeae (Hsiao et al., 1995a). The ITS regions has also been used to a lesser extent to investigate subfamily relationships in the Arundinoideae (Hsiao et al., 1998); tribe Danthonieae by Barker et al., 2000 in a combined analysis of morphology and *rpo*C2, *rbc*L and ITS sequence data and the Pooideae (Hsiao et al., 1994, 1995b). The phylogeny of the entire grass family (Hsiao et al., 1999) was also investigated using ITS.

In Poaceae the *trn*L-F regions has been also used in many levels such as *Poa jemtlandica* (Brysting et al., 2000); *Spartina anglica* (Baumel et al., 2001); *Axonopus* (Gomez-Martinez & Culham, 1997); *Bouteloua* (Columbus et al., 2000); Chloridoid grasses (Roodt-Wilding & Spies, 2006); *Deschampsia* (Chiapella, 2007); *Elymus* (Liu et al., 2006); Festucoid grass of subtribe Loliinae (Catalán et al., 2004); *Miscanthus* (Hodkinson et al., 2002a); *Paspalum* (Souza-Chies & Essi, 2007); *Spartina* (Ferris et al., 1997); the Panicoid tribes Andropogoneae (Hodkinson et al., 2002b); Paniceae (Gomez-Martinez & Culham, 2000; Doust & Kellogg, 2002); the Pooi tribe Triticeae (Mason-Gamer et al., 2002) and *trn*L-F has also demonstrated its suitability for clarifying the systematic problems in the Gramineae (Neves et al., 2005).

Separate analyses were conducted and the data sets were then combined for combined analyses. It was decided to combine the two regions in many studies in recent years have indicated that combined molecular data sets that utilize different levels of variation (as provide by ITS and *trn*L-F) provide resolution at different levels of the cladogram and, therefore, phylogenetic resolution and bootstrap values are improved by combining separate data sets (Chase & Cox, 1998; Soltis & Soltis, 1998; Whitten et al., 2000).

4.3 Material and methods

4.3.1 Taxon sampling and ingroup selection

Forty-two taxa representing 12 genera of subtribes Ischaeminae and Rottboelliinae were sampled. The species were chosen to represent variation within each genus. At least 10% of the species in each genus were included to avoid any excess heterogeneity in rates of molecular evolution that may be found. The

remaining one genus in subtribe Rottboelliinae, as circumscribed by Nanakorn & Norsaengsri (2001), i.e. *Vossia* Wall. & Griff. is a monotypic genus could not be included owing to lack of material.

Field collected leaves and vouchers were collected between June 2004 and August 2007. All materials were preserved in silica gel.

4.3.2 DNA extractions

Total genomic DNA was extracted from 0.3 g of leaves from individual plants preserved in silica gel or 0.2 g of leaf material from herbarium specimens using the modified CTAB (hexadecyltrimethylammonium bromide) method of Doyle & Doyle (1987) and precipitated in isopropanol. The crude total genomic DNA was purified using the QIAquick PCR Purification kit, namely Spin column Technique and DNA levels were checked on the 1% agarose gel, and stored at -20°C or -80°C (Appendix E).

4.3.3 Amplification of target DNA regions

The nrDNA internal transcribed spacer region (ITS) and further sequencing was amplified with Polymerase Chain Reaction (PCR) using both the external primers AB101F (5'-ACG AAT TCA TGG TCC GGT GAA GTG TTC TTC G-3'), AB102R (5'-TAG AAT TCC CCG GTT CGC TCG CCG TTA C-3') and internal pair primers ITS 2 (5'-GCT GCG TTC TTC ATC GAT GC-3'), ITS 3 (5'-GCA TCG ATG AAG AAC GCA GC-3') (Sun et al., 1994). The PCR reactions were carried out in a total volume of 25 µl. Each reaction consisted of 1 µl cellular DNA, 0.5 µl dNTP (10 mM), 1.5 µl $MgCl_2$ (50 mM), 0.5 µl of each primer (100ng/µl), 2.5 µl 10X Buffer, 2.5 µl BSA, 0.5 µl DMSO, and 0.5 µl *Taq* polymerase. Strong denaturing reaction conditions were needed because of the high GC (guanine-cytosine) content of ITS region and the presence of secondary structure (Bucker & Holtsford, 1996). This was achieved by the addition of DMSO, which reduces strand reannealing (Winship, 1989; Varadaraj & Skinner, 1994). Buckler et al. (1997) also suggested that the addition of DMSO would prevent the preferential amplification of pseudogenes. These might be preferentially amplified due to low secondary structure stability. The reactions were carried out using a GeneAmp PCR System 9700 thermocycler. The thermal cycling were run on following the program comprised of 30 cycles, 2 min

premelt at 94°C, 1 min denaturation at 94°C, 1 min annealing at 48°C, 1.30 min extension at 72°C and 4 min final extension at 4°C.

Table 4.1 List of primers used for amplification and sequencing.

Marker	Name	Sequence	Reference
ITS			
forward	AB101	5'-ACGAATTCATGGTCCGGTGAAGTGTTCTTCG-3'	Sun *et al.*, 1994
reverse	ITS2	5'-GCTGCGTTCTTCATCGATGC-3'	Sun *et al.*, 1994
forward	ITS3	5'-GCATCGATGAAGAACGCAGC-3'	Sun *et al.*, 1994
reverse	AB102	5'-TAGAATTCCCCGGTTCGCTCGCCGTTAC-3'	Sun *et al.*, 1994
***trn*L-F**			
forward	c	5'-CGAAATCGGTAGACGCTACG-3'	Taberlet *et al.*, 1991
reverse	d	5'-GGGGATAGAGGGACTTGAAC-3'	Taberlet *et al.*, 1991
forward	e	5'-GGTTCAAGTCCCTCTATCCC-3'	Taberlet *et al.*, 1991
reverse	f	5'-AATTGAACTGGTGACACGAG-3'	Taberlet *et al.*, 1991

The non-coding regions of cpDNA were PCR amplified using primers 'c' (5'-CGA AAT CGG TAG ACG CTA CG-3') and 'd' (5'-GGG GAT AGA GGG ACT TGA AC-3') for the *trn*L intron, and primers 'e' (5'-GGT TCA AGT CCC TCT ATC CC-3') and 'f' (5'-AAT TGA ACT GGT GAC ACG AG-3') for the *trn*L-F intergenic spacer (Taberlet et al., 1991). The PCR reactions were carried out in a total volume of 25 µl. Each reaction consisted of 1 µl DNA, 0.5 µl dNTP (10 mM), 2.5 µl $MgCl_2$ (50 mM), 0.5 µl of each primer (100ng/µl), 2.5 µl Buffer, 2.5 µl BSA and 0.5 µl *Taq* polymerase. The thermal cycling were run on following the program comprised 30 cycles, 2 min premelt at 94°C, 1 min denaturation at 94°C, 1 min annealing at 50°C, 1.30 min extension at 72°C and 4 min final extension at 4°C.

Three to five µl of each PCR product for both the nuclear and chloroplast regions were run on a 1% agarose gel placed in 1X TAE (Tris-Acetic acid-EDTA) buffer. Amplified products were purified using Magic mini-columns (Promega, Southampton, Hampshire, UK) or QIAquick columns (Qiagen, Crawley, West Sussex, UK), protocols were provided by manufacturers of the minicolumns (Appendix E). Cleaned PCR products were stored at -20°C after checking the DNA level on the 1% agarose gel.

4.3.4 Cycle Sequencing

The sequencing primers for both regions were the same primers used for the amplification in PCR. The cycle sequencing profile was as follows: rapid thermal ramp to 96°C followed by 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min, performed for 26 cycles and held at 4°C. Sequencing of the PCR products was carried out using cycle sequencing with ABI Big Dye terminators run on an ABI 377A automatic sequencer (according to the manufacturer's protocols; Perkin-Elmer Applied Biosystems, Inc., Warrington, Cheshire, UK) at Jodrell Laboratory, the Royal Botanic Gardens, Kew, Uk.

4.3.5 Sequence assembly and alignment

Automated sequence output files were edited and assembled using Sequence Navigator and AutoAssembler (Applied Biosystems Inc.). Sequences of both regions were unambiguously aligned by visual inspection. The final aligned matrix has 818 characters of ITS region and 1,211 characters of *trn*L-F region. The alignments were directly submitted to parsimony analysis.

4.3.6 Outgroup selection

Outgroup taxa represented different degrees of relatedness to subtribe Ischaeminae and Rottboelliinae: *Eragrostis biflora* and *Centotheca lappacea* from the broader subfamily Chloridoideae and Centothecoideae, respectively.

4.3.7 Phylogenetic analyses

Heuristic parsimony analyses of the aligned sequence matrix was performed using PAUP* 4.0b10 software (Swofford, 1998) on a Macintosh G4. All characters were unordered and equally weighted (Fitch parsimony; Fitch, 1971) and individual gap positions were treated as missing data. Most-parsimonious trees were obtained using 1000 replicates of random taxon addition with equal weights and tree-bisection-reconnection (TBR) branch-swapping, with MULTREEs in effect, but holding only one tree per step and saving no more than 100 trees per replicate to minimize time swapping on suboptimal trees. Internal support was assessed using 1000 bootstrap replicates with TBR swapping and simple addition of taxa, retaining groups with frequencies above 50% in the bootstrap consensus tree. The following categories were used to describe levels of bootstrap support: weak = 50–74%; moderate = 75–84%;

and strong 85–100% (Simpson et al., 2003.). Also computed using PAUP* were the strict consensus trees, consistency index (CI) retention index (RI) and rescaled consistency index (RC). All resulting trees were rooted using both species of *Eragrostis biflora* (subfamily Chloridoideae) and *Centotheca lappacea* (subfamily Centothecoideae) as the outgroups.

4.4 Results

Sampling was designed to include 12 genera of the two subtribes in Thailand and one sequences of outgroup taxa (*Eragrostis biflora*) was taken from GenBank.

4.4.1 ITS region

The length of ITS1 ranged from 215 (*Ischaemum indicum* 1) to 403 bp (*Apluda mutica*) and ITS2 from 283 (*Hemarthria pratensis*) to 465 (*Ischaemum barbatum* 1) bp. Aligned sequences of the entire region of 44 specimens (42 ingroup and 2 outgroup specimens) consist of 818 characters. For the entire region, 113 characters (16.10%) were potentially parsimony-informative and 93 were variable. The heuristic search with ITS data produced 199 trees with the tree length of 475 steps, a consistency index (CI) = 0.6253, a retention index (RI) = 0.7367 and a rescaled consistency index (RC) = 0.4606. One of these trees with bootstrap values is shown in Fig. 4.4.

The subtribes Ischaeminae and Rottboelliinae are resolved as a polyphyletic and paraphyletic, respectively with poor bootstrap support. Within both subtribes, there are five major clades as well supported:

Clade A: Formed by all the *Ischaemum* species plus *Kerriochloa siamensis*, except *Ischaemum tenuifolium*.

Clade B: Consisted of seven genera with poor bootstrap: *Mnesithea, Hackelochloa, Hemarthria, Ophiuros, Thelepogon, Rottboellia* and *Phacelurus*.

Clade C: *Mnesithea laevis* as a sister to taxa in genus *Eremochloa* but with poor bootstrap support (<50%). Resolution is strong and fully supported (100%) monophyletic group for internal node of the genus *Eremochloa*.

Three basal lineages, the single taxon *Apluda mutica* in clade E, with poor bootstap support, a clade D (<50% bootstrap support), which include *Ischaemum tenuifolium* and *Sehima nervosum*.

4.4.2 *trn*L-F region

The length of the *trn*L intron from 498 (*Sehima nervosum*) to 531 bp (*Ischaemum hirtum*) and the *trn*L-F spacer from 381 (*Mnesithea mollicoma*) to 615 bp (*Ischaemum barbatum* 2). Aligned sequences of the entire *trn*L-F region of 44 specimens (42 ingroup and 2 outgroup specimens) consist of 1,211 characters. For the entire region 61 characters (7.15%) were potentially parsimony-informative and 53 were variable. Parsimony analysis excluding the coded gaps, resulted in 95 equally most parsimonious trees (tree length = 140 steps; CI = 0.8857; RI = 0.9797; RC = 0.8677). One of these trees with bootstrap values is shown in Fig. 4.5.

The analysis showed the polyphyly of both subtribes with weak to high support. Within both subtribes, there are two major clades.

Clade A: Formed by all taxa in genus *Ischaemum*, *Hemarthria*, *Rottboellia cochichinensis*, *Mnesithea laevis* and unresolved at its base as form a polytomy of several lineages with a weakly bootstrap support (57%).

Clade B: This clade consisted of the genera *Eremochloa, Ophiuros, Mnesithea, Hackelochloa* and *Apluda* with a strongly bootstrap support (100%). The analysis showed that the genus *Eremochloa* is a mophyletic as same as the analysis of ITS data.

4.4.3 Combined analysis

Because of the two regions investigated produced poorly identical results. Therefore, the sequence data of ITS and *trnL*-F were combined for additional analysis.

After alignment, the combined data matrix consisted of 2,029 characters. With indel information included the matrix consisted of 1,555 characters, of which 174 (11.19%) were parsimony-informative and 146 were variable. Analysis of this combined matrix resulted in 4 equally most parsimonious cladograms with a tree length of 632 steps, CI of 0.6661, RI of 0.8559 and RC of 0.5701. Although some polytomies are present. One of these trees with bootstrap values is shown in Fig. 4.6.

Both subtribes form a polyphyletic group with a weakly bootstrap support 53% and 52%, respectively, which composed of two main clades:

Clade A: Comprised of taxa of genus *Ischaemum, Mnesithea laevis, Sehima nervosum, Kerriochloa siamensis, Thelepogon elegans, Rottboellia cochinchinensis,*

Hemarthria and *Phcelurus zea*, which was recognized as a sister taxon to the remaining of this clade.

Clade B: Consisted of the genera in subtribe Rottboelliinae: *Mnesithea, Hackelochloa, Ophiuros* and *Eremochloa* except *Apluda mutica*, which is from the subtribe Ischaeminae. Well supported internal nodes were formed in this analysis.

Table 4.2 Values and statistics from PAUP* analyses of separate and combined data sets.

Sequence characteristic	Combined ITS & *trn*L-F	ITS	*trn*L-F
Number of characters	1,555	702	853
Invariant characters	1,235	496	739
Uninformative variable characters	146	93	53
Parsimony-informative variable characters	174	113	61
Consistency index (CI)	0.6661	0.6253	0.8857
Retention index (RI)	0.8559	0.7367	0.9797
Rescaled consistency index (RC)	0.5701	0.4606	0.8677
Maximum parsimony tree length	632	475	140
Number of trees	4	199	95

4.5 Discussion and conclusion

This study presents the first molecular phylogenetic hypothesis of the relationships among species and genera of the subtribes Ischaeminae and Rottboelliinae. The data sets of the DNA sequences of chloroplast DNA (*trn*L-F region) and the nuclear ribosomal internal transcribed spacer regions (ITS) were analyzed separately before being combined into a single data set. The analysis shows that the combined tree exhibits the best phylogenetic reconstruction of both subtribes, but the analysis of *trn*L-F resulted in phylogenies with highest retention index (RI). Two main lineages can be identified as identical to the analysis of *trn*L-F data (clades A and B; Fig. 4.6), while the analysis of ITS provides evidence for phylogenetic examination of closely related taxa. Recent studies have indicated that the use of non-

coding chloroplast regions may only be useful at intergeneric but not intrageneric levels due to a lack of informative sequence variation (Roodt-Wilding & Spies, 2006). This was seen in this study with the *trn*L-F region providing resolution among, but not within, genera in both subtribes.

The results show that the subtribes Ischaeminae and Rottboelliinae are polyphyletic. Previous studies using phytochrome B, GBSSI and *ndh*F also concluded that both subtribes are polyphyletic (Mathews et al., 2002). Kellogg & Watson (1993), using morphological data, placed *Ischaemum* and its sister *Digastrium* a sister to many genera of Rottboelliinae, whereas they considered *Apluda* a sister to *Apocopis*. The *ndh*F phylogeny (Spangler et al., 1999; Spangler, 2000) also places *Ischaemum* as an early diverging branch, whereas *Apluda* is part of the "core" Andropogoneae. Hence, these two groups of subtribes do not support the subtribal designations of Clayton & Renvoize (1986).

Implications for Classification

***Ischaemum* L.**

Resolution within the *Ischaemum* clade is moderate to high (most subclades having of bootstrap support (≥78%) showing that *Ischaemum* is paraphyletic. The *Ischaemum* clade also contains three morphologically very similar genera, *Mnesithea, Sehima* and *Kerriochloa*. The species *Ischaemum tenuifolium* forms a subclade separate from the other species in the genus, but with weak support. The division of *Ischaemum* into sections, proposed by Clayton & Renvoize (1986), is supported by the present molecular study. The ITS, *trn*L-F and combined trees divide the genus into three major subclades that correspond to the morphology of the lower glume of the sessile spikelet. In the combined tree, subclade I, with a strong bootstrap support of 98%, comprises all taxa which share the character of an entirely sessile lower glume. Subclade IV is formed by the group of taxa which have nodules on the sessile lower glume, with a robust bootstrap support of 99%.

Subclades II and III are derived from species of four different genera, *Ischaemum tenuifolium*, *Mnesithea laevis*, *Sehima nervosum* and *Kerriochloa siamensis,* all of which are now placed in the genus *Ischaemum*. *K. siamensis* is a sister to the remainder of this subclade, with poor bootstrap support, indicating that this group is closely related to the genus *Ischaemum*. This result is in agreement with

Clayton & Renvoize (1986), who suggested that *Kerriochloa* is related to *Ischaemum*, particularly *I. decumnens*.

All the three analyses supported the segregation of *Ischaemum tenuifolium* from the genus *Ischaemum*. This finding is also supported by the taxonomic treatments (chapter V) that moved *I. tenuifolium* to the genus *Andropogon* and Clayton & Renvoize (1986) who noted that the genus *Ischaemum* is often quite difficult to distinguish from *Andropogon*. Unfortunately, the genus *Andropogon* is not sampled here. The two genera are circumscribed on the basis of their difference in inflorescences and lower glume of the sessile spikelet characters, with the *Ischaemum* defined by 2 rarely digitate racemes that carry a spikelet with convex or flattened lower glume, whereas the *Andropogon* species are characterized by having 3–7 racemes with a concave lower glume (Clayton & Renvoize, 1986). The leaf-blades anatomical study in chapter II concluded that papillae are present in all members of *Ischaemum*, except in *I. tenuifolium,* which shows that this species is distinct from the other species in the genus. These results confirm the conclusion of Watson & Dallwitz (1992), who stated that many of the core Andropogoneae have a single oblique papilla on each of the intercostal cells, overlying the stomata. Of the genera included in their study, *Andropogon* and *Schizachyrium* lack such papillae. Both genera include a large number of species, and *Andropogon* is likely to be polyphyletic (Mathews et al., 2002); future studies will thus have to investigate more species of each genus to test the connection demonstrated here.

However, the relationships within some of the taxa in the genus *Ischaemum,* namely, *I.* sp.1, *I.* sp.2, *I. barbatum* complex and *I. indicum* complex are unresolved; there is still a low resolution between the subclades of *Ischaemum* as they were arranged in a polytomy.

***Hemarthria* R. Br. & *Phazelurus* Griseb.**

Most of the generic groupings including more than one species are well supported. This includes, for example, the genus *Hemarthria* which forms well supported lineages (≥98%) in all analyses, which shows that the genus *Hemarthria* is monophylic. The generic status of *Hemarthria* is clearly justified and its close relationship to *Phacelurus zea* (combined tree) is also evident with a weak bootstrap support of 53%. This is in agreement with Clayton & Renvoize (1986), who stated

that the fused pedicel and basic chromosome number of 9 are evidence that *Hemarthria* is more closely related to *Phacelurus* than to *Heteropholis*.

Hemarthria compressa is a sister to *H. pratensis* with a 100% bootstrap support in combined and ITS trees and a 98% bootstrap support in *trn*L-F tree. The morphology of the two taxa is also clearly different; they are distinguished by leaf-blades and joints. In *H. compressa*, the leaf-blade is not deciduous and the joints are not long cuneate, in contrast to the deciduous leaf-blades and long cuneate joints in *H. pratensis* (Van Den Heuvel & Veldkamp, 2000).

Placement of *Phacelurus zea* (bootstrap support 53%) as sister to clade A (group of *Ischaemum*) is consistent with Clayton & Renvoize (1986), who pointed out that the two genera are clearly related. In practice, there is not much difficulty in distinguishing it from *Ischaemum*, but it is sometime quite hard to find unequivocal diagnostic characters.

***Mnesithea* Kunth**

The results presented here show that the taxa *Ischaemum tenuifolium, Mnesithea laevis, Sehima nervosum* and *Kerriochloa siamensis* are nested within *Ischaemum*, although this subclade is not strongly bootstrap supported (<50%). *Kerriochloa siamensis* appears to be a sister to the other members of both subclades. Clayton & Renvoize (1986) proposed that the genus *Mnesithea* has only one species: *M. laevis*, and retained *Coelorachis* as distinct genus. The genus *Coelorachis* differs from the other genera by a free pedicel and a pedicelled spikelet. A re-assessment of the generic limits of *Mnesithea* sensu lato by Veldkamp et al. (1986) acknowledged the addition of *Coelorachis* in to *Mnesithea,* because several species of *Coelorachis* have the lower part of pedicel completely fused to the rachis node. The pedicelled spikelet varies from well developed to minute vestige, and its absence in traditional *Mnesithea* is therefore not enough to conserve the genus (Veldkamp et al., 1986). The results of this molecular study seem to conflict with the results of the morphological study in chapter II and the investigation by Veldkamp et al. (1986), who included all members of *Coelorachis* into *Mnesithea*. A majority of the *Mnesithea* species (*Coelorachis* sensu Clayton & Renvoize, 1986) form a group in a subclade VIII which is distinctly related to *Hackelochloa* and *Ophiuros*, however with a weak bootstrap support of 61%. The arguments mentioned above clearly show that this genus is problematic and should be studied more in details.

The species *Mnesithea glandulosa* is segregated from most other species in the genus and appears to be basal-most to all others of this group with a weak bootstrap support of 61%; correlated with a distinct morphological character; a lateral hook on the lower glume of the sessile spikelet, which is a characteristic of *Eremochloa*. In 1986, Clayton & Renvoize depicted the genus *Eremochloa* as derived from *Coelorachis* (=*Mnesithea*).

Within the remaining *Mnesithea* are three subclades with a week bootstrap support less than 50%. The first topology, with a 63% bootstrap support, suggested that *M. heferi* is sister to *M. mollicoma*, sharing the pubescence leaf surface. The second topology is one of the strongest (100% bootstrap support) supported in both ITS and combined analyses: *M. striata* var. *striata* and *M. striata* var. *pubescens*. Therefore, *M. striata* var. *pubescens* might be the same species as *M. striata* var. *striata*. However, the two subspecies can be distinguished by the culm and leaf-blade, which are covered in hairs in var. *pubescens*, whereas glabrous in var. *striata* (Hackel, 1889). The last topology confirmed that *M. cancellata* is sister to *M.* sp.1 with a weak bootstrap support of 52%. In this study most materials of *M.* sp.1 were previously identified as *M. striata*. *M.* sp1 is different from *M. striata* by having an oblong lower glume of the sessile spikelet, indurate, 5–5.5 by 1.2–1.5 mm; an apex with unequal wings; 7–8 longitudinal rows of small pits between the nerves of the back; and margins at base puberulous. *M. striata* has an ovate lower glume of the sessile spikelet, indurate, 4.5–5 by *ca.* 1.5 mm; an apex with 2 apical wings; and a continuous ridge along the length and interrupted by tubercles or tubercle-based hairs on the back, or glabrous. Hence, the molecular as well as morphological evidences indicated that *M.* sp.1 should be proposed as a new species in the genus *Mnesithea*.

***Hackelochloa* Kuntze**

A subclade VIII grouping (Fig. 4.6) comprises of *Mnesithea*, *Hackelochloa*, *Ophiuros* with *Mnesithea glandulosa* being a sister of this group. The small genus *Hackelochloa*, represented in this study by a couple of species, is determined as monophyletic group but with a weak bootstrap support. Further sampling of these two taxa is required in order to confirm their monophyly. From the combined analysis it revealed that *Hackelochloa* is closer to *Ophiuros* than to *Mnesithea*; however with a poor bootstrap support less than 50%. The molecular results from this study are

contrary to Veldkamp, et al. (1986) who treated *Hackelochloa* as a synonym of *Mnesithea*.

Hackelochloa granularis and *H. porifera* are morphologically similar and have been grouped together previously (Veldkamp, et al. 1986). In contrast, the morphological study separates the two taxa by the difference in size and sculpture on the lower glume. While the *trn*L-F tree (Fig. 4.5) shows that *H. granularis* and *H. porifera* are in the same clade, but with 79% bootstrap support, and both having a globose lower glume. Therefore, the two species are not identical. This is in agreement with Bor (1960), who stated that *H. porifera* differed from *H. granularis* by its larger spikelets and a more pronounced sculpture on the lower glume of the sessile spikelet.

***Eremochloa* Buse**

The results of all phylogenetic analyses presented here provide broadly consistent estimates of the phylogeny of *Eremochloa*. The genus *Eremochloa* comprises of a monophyletic group, with *Apluda* (represented by *A. mutica*) as sister group to the genus, with a strongest bootstrap support (100%). According to Buitenhuis & Veldkamp (2001), this genus is recognized by a single spike-like raceme and a submarginally pectinate lower glume. Clayton (1973), in a numerical analysis of awnless genera of the Andropogoneae, found *Eremochloa* in the Coelorachidastrae close to *Rhytachne* (not sampled here).

The genus *Eremochloa* can be divided into two main groups with a strongest bootstrap support (100%), however no obvious morphological character supports this division. The first group is composed of *E. bimaculata*, *E. maxwellii*, *E. ciliatifolia* and *E. attenuata* as a sister of this group with a 99% bootstrap support. The second group is formed by *E. ciliaris*, *E. eriopoda*, *E. lanceolata* and *E.* sp.1 as a sister of this group with 64% bootstrap support.

The morphological study shows that *Eremochloa* sp.1 is very closely related to *E. maxwellii* but differs in having an obovate and long cordate pedicel, in contrast to the obliquely-lanceolate pedicel in *E. maxwellii*. The species is found in north-eastern Thailand. The present molecular study confirmed that *E.* sp.1 should be a new species, but not a variety of *E. maxwellii*.

***Apluda* L.**

The strong support for the clade containing *Eremochloa* and *Apluda* is one of a surprised result produced by this molecular study, since there are no obvious morphological characters that unite the two genera. *Apluda* is separated from *Eremochloa*, so the closed relationship of the two genera might have been predicted.

Although not all genera and species in the subtribes Ischaeminae and Rottboelliinae were included in this analysis, the results observed here indicate the phylogenetic trends in both subtribes in Thailand. Additional taxonomic sampling in the future will result in better resolution of the relationships. Moreover, more gene regions and more samples from various geographic regions should be included in further studies in order to define the relationships among taxa in both subtribes.

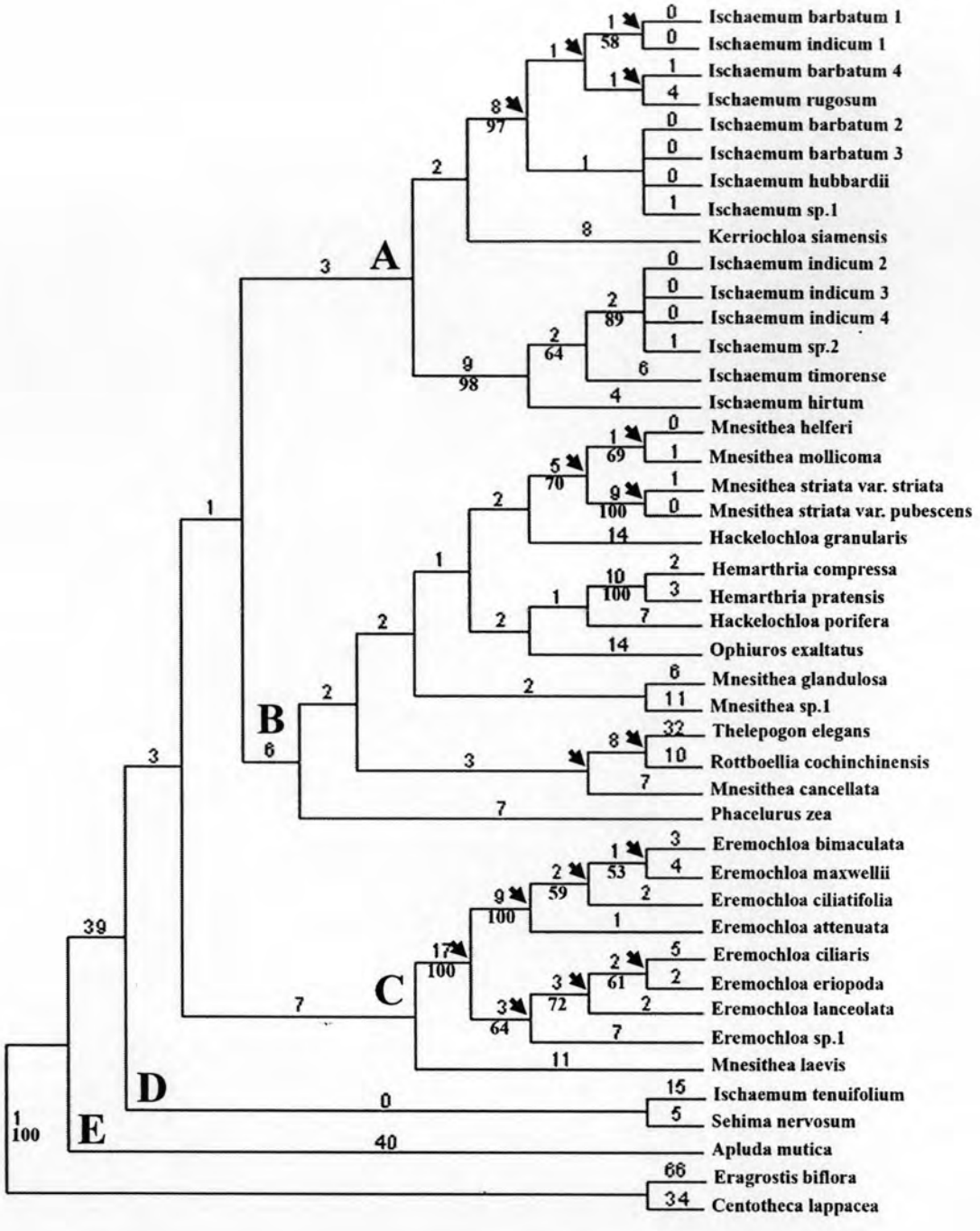

Figure 4.4 One of 199 equally most parsimonious trees of ITS region of the subtribes Ischaeminae and Rottboelliinae and outgroup taxa with a length of 475, CI of 0.63 and RI of 0.74. Branch lengths obtained using sequences are indicated above branches, bootstrap percentage values >50% are indicated below branches. Arrows show clades recovered in the strict consensus tree.

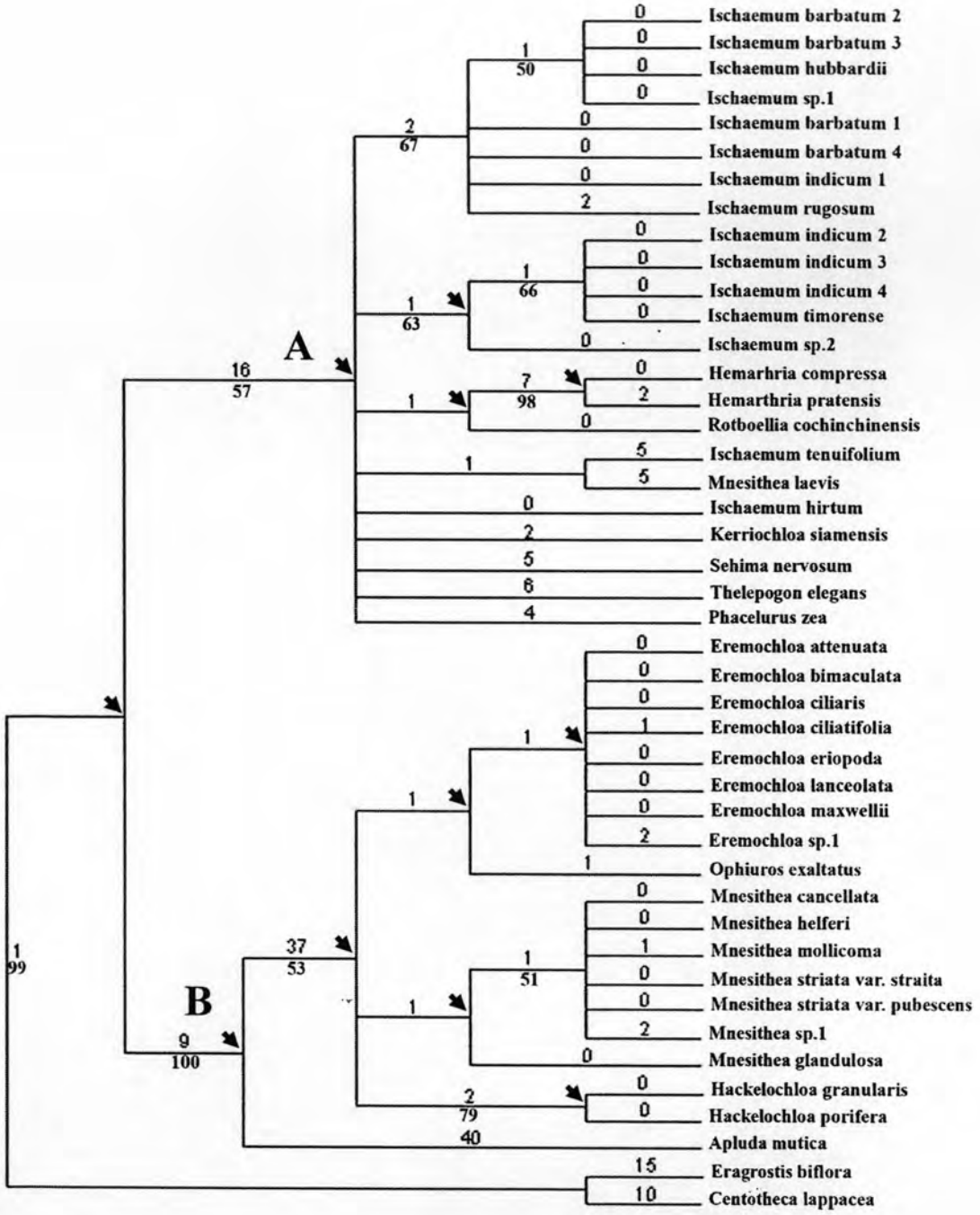

Figure 4.5 One of 95 equally most parsimonious trees of *trn*L intron and the *trn*L-F spacer of the subtribes Ischaeminae and Rottboelliinae and outgroup taxa, with a length of 140, CI of 0.89 and RI of 0.98. Branch lengths obtained using sequences are indicated above branches, bootstrap percentage values >50% are indicated below branches. Arrows show clades recovered in the strict consensus tree.

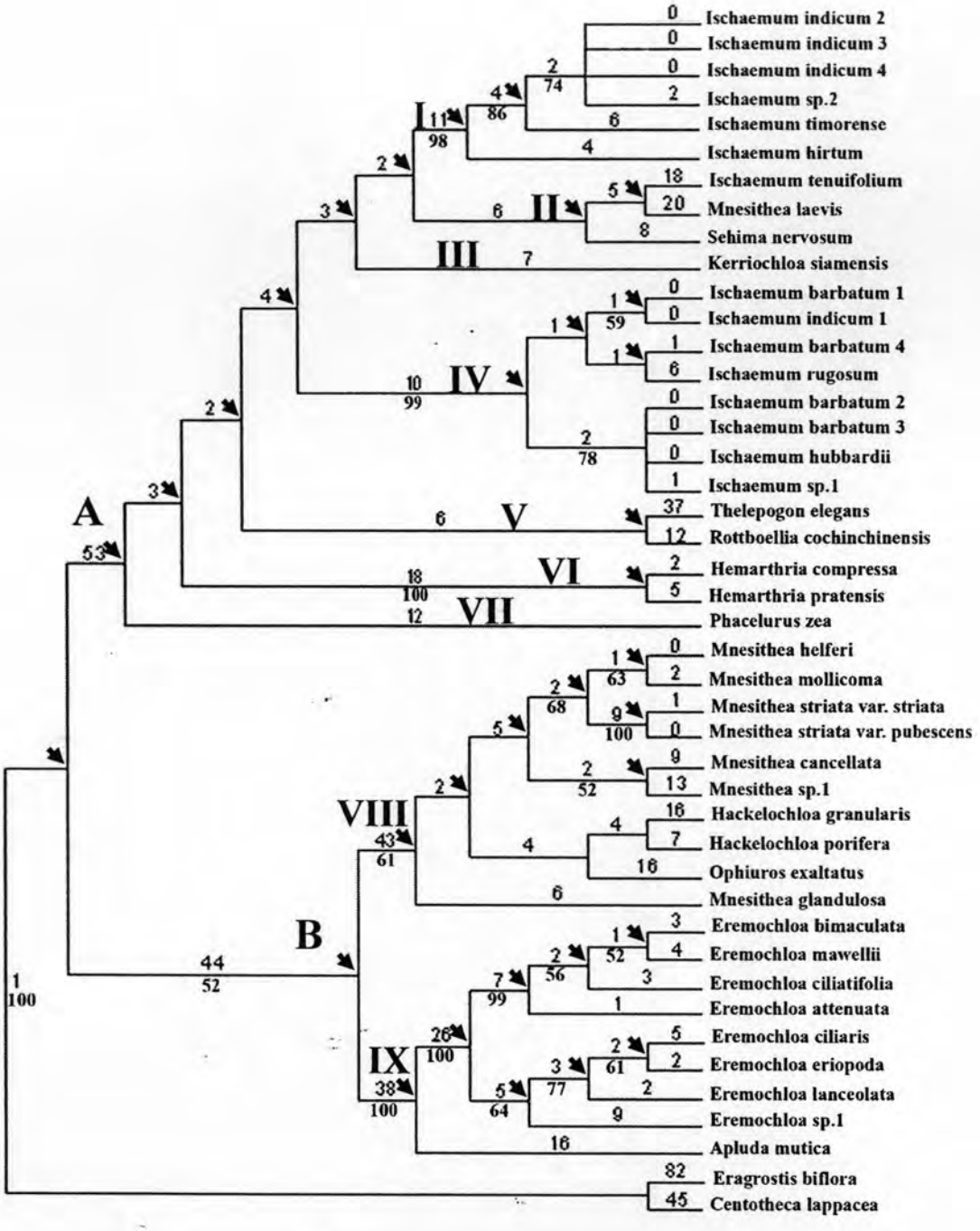

Figure 4.6 One of 4 equally most parsimonious trees of ITS region and *trn*L intron and the *trn*L-F spacer of the subtribes Ischaeminae and Rottboelliinae and outgroup taxa, with a length of 632, CI of 0.67 and RI of 0.86. Branch lengths obtained using sequences are indicated above branches, bootstrap percentage values >50% are indicated below branches. Arrows show clades recovered in the strict consensus tree.