Chapter 2

MOLECULAR PHYLOGENY OF COELOGYNE (EPIDENDROIDEAE, ORCHIDACEAE) BASED ON PLASTID RFLPS, *MATK* AND NUCLEAR RIBOSOMAL ITS SEQUENCES: EVIDENCE FOR POLYPHYLY

BARBARA GRAVENDEEL[1], MARK W. CHASE[2], ED F. DE VOGEL[1], MARCO C. ROOS[1], TED H.M. MES[3] & KONRAD BACHMANN[4]

SUMMARY

Subtribe Coelogyninae (Epidendroideae, Orchidaceae) presently comprises 16 genera. To evaluate the monophyly of one of these genera, *Coelogyne* Lindl., and reveal sectional relationships and relations to allied genera, we collected PCR RFLPs from 11 plastid regions for 42 taxa in Coelogyninae (28 *Coelogyne* species and 14 representatives of other genera) and three outgroups from Bletiinae and Thuniinae. In addition, we sequenced a large portion of the plastid *trnK* intron (mostly *matK*) and the nuclear ribosomal DNA internal transcribed spacers ITS1 and ITS2 (including the 5.8S gene). Separate phylogenetic analyses on each dataset using maximum parsimony produced mainly congruent (except for the position of *Panisea*) but weakly supported clades. Parsimony analysis of the combined data clearly identified three main clades in Coelogyninae: I) *Bracisepalum*, *Chelonistele*, *Dendrochilum*, *Entomophobia*, *Geesinkorchis* and *Nabaluia* nested within *Coelogyne*; II) *Neogyna* and *Pholidota* nested within the remainder of species of *Coelogyne* sampled; III) *Pleione*. Whereas Coelogyninae are monophyletic, *Coelogyne* is polyphyletic, with species falling into at least two well supported clades. The utility of some morphological characters used in traditional classifications was explored by reconstructing character state evolution on the combined molecular consensus tree. Lip base and petal shape appeared to be homoplasious, whereas ovary indumentum and flower number were highly congruent with well supported groups. The implications of our results for the classification of *Coelogyne* are discussed and a reorganisation of the genus by including *Neogyna* and *Pholidota* and removing several species is proposed.

Key words: Orchidaceae, Coelogyninae, *Coelogyne*, molecular phylogeny, plastid DNA RFLPs, *matK*, nuclear rDNA ITS.

INTRODUCTION

The orchid genus *Coelogyne* Lindl. comprises over 200 species distributed throughout southeast Asia with main centers of diversity in Borneo, Sumatra and the Himalayas (Butzin, 1992a). Most species are epiphytes, occuring in tropical lowland and montane rainforests. In open, humid environments, some species may also grow as lithophytes

1) Nationaal Herbarium Nederland, Universiteit Leiden branch, P.O. Box 9514, 2300 RA Leiden, The Netherlands.
2) Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3DS, United Kingdom.
3) Institute for Systematics and Ecology, Experimental Plant Systematics, University of Amsterdam, The Netherlands.
4) Institute for Plant Genetics and Crop Plant Research, D06466, Gatersleben, Germany.

or even as terrestrial plants (Comber, 1990). Most species are characterised by medium-sized to large flowers with a sweet scent and are pollinated by bees (Van der Pijl & Dodson, 1966), beetles (O'Byrne, 1994) or wasps (Carr, 1928; Dressler, 1981). The number of recent artifical hybrids published indicates the growing commercial interest in this group (Erfkamp & Gruβ, 1996).

Although revisions of several sections of *Coelogyne* were published in the last decade, a comprehensive treatment of all species is still lacking. This is partly caused by the problematic delimitation of groups within the genus. Pfitzer & Kraenzlin (1907d) grouped the species of *Coelogyne* into 14 sections. In contrast, Holttum (1964) proposed only 4 and De Vogel (1994) and Clayton (in press) 23 subdivisions. These large differences in opinion are due not only to the rather large number of species in the genus, but also the relative lack of morphological characters available to define groups of species. For example, the presence of hairs on the ovary has been used to define sect. *Tomentosae* (De Vogel, 1992). However, this character is likely to have evolved convergently in section/subgenus *Coelogyne*, *Cyathogyne*, *Rigidiformes*, *Veitchiae* and *Verrucosae*. The naturalness and relationships of the sections and subgenera of *Coelogyne* were not previously examined in a phylogenetic context.

Coelogyne is one of the 16 genera in subtribe Coelogyninae (tribe Coelogyneae, subfamily Epidendroideae) with a total of approximately 550 species (Pedersen et al., 1997). Synapomorphies of the subtribe are sympodial growth, pseudobulbs of one internode, terminal inflorescences, a winged column and massive caudicles (Dressler, 1981; De Vogel, 1986; Butzin, 1992b). *Coelogyne* Lindl. is defined by a free, never-saccate lip with high lateral lobes over the entire length of the hypochile and papillose, toothed or warty keels (Seidenfaden & Wood, 1992). The genus is defined merely by the absence of characters, such as a saccate lip base (present in all other genera of the subtribe) or a lip adnate to the column (present in *Neogyna* Rchb.f. and *Gynoglottis* J.J. Sm.; Butzin, 1992b). In addition, many characters intergrade among the genera of the subtribe. For example, a lip with small, inconspicuous lateral lobes characterises both *Chelonistele* Pfitzer and *Panisea* (Lindl.) Steud. (De Vogel, 1986; Lund, 1987).

A phylogenetic survey of *Coelogyne* and related genera of Coelogyninae using molecular characters can provide a preliminary phylogenetic classification and serve as a historical framework for evaluating hypotheses of morphological character evolution. The aims of this study are to use phylogenetic analyses of molecular data to:

1) address the generic circumscription and sectional and subgeneric relationships within *Coelogyne*;
2) investigate the relationships of *Coelogyne* with its allies in subtribe Coelogyninae;
3) determine whether some previously used morphological key characters are phylogenetically informative.

To accomplish these goals, parsimony analyses were conducted on PCR RFLP data of 11 regions of the plastid genome and sequence data from both the *trnK* intron (mostly *matK*) and the nuclear rDNA ITS regions.

PCR RFLPs were expected to be useful in reconstructing phylogenetic relationships within the genus *Coelogyne* based on previous RFLP studies in Orchidaceae (Chase & Palmer, 1992; Yukawa et al., 1993; Freudenstein & Doyle, 1994). PCR RFLPs provide a rapid way of sampling many parts of the genome, which have evolved at different rates and under different constraints (Gielly & Taberlet, 1994). They provide informa-

tion on multiple DNA regions, which in our view is better than having only two gene sequence data sets.

The *trnK* intron has been used for phylogeny reconstruction at a variety of taxonomic levels in angiosperms (Soltis & Soltis, 1998). In Orchidaceae, it has been used at generic (Whitten et al., in press) and species level (Ryan et al., 2000). The nuclear rDNA ITS regions have been used extensively to infer phylogenetic relationships in Orchidaceae at both tribal (Douzery et al., 1999), generic (Pridgeon et al., 1997) and species level (Cox et al., 1997).

MATERIALS AND METHODS

Plant samples

To determine the position of *Coelogyne* in subtribe Coelogyninae and relationships within *Coelogyne*, 45 taxa were analyzed. The sampling includes 18 of the 23 sections/subgenera currently recognised within *Coelogyne* and 11 of the 16 genera of Coelogyninae. Morphologically uniform sections/(sub)genera are represented by a single taxon only, whereas more variable groups are represented by several species. Not included were five small sections of *Coelogyne* (sect. *Ancipites* Pfitzer, *Fuscescentes* Pfitzer & Kraenzl., *Micranthae* Pradhan, *Ocellatae* Pfitzer and *Proliferae* Lindl.) and five mostly monotypic genera (*Bulleya* Schltr., *Dickasonia* L.O. Williams, *Gynoglottis* J.J. Sm., *Ischnogyne* Schltr. and *Otochilus* Lindl.). Outgroups were sampled from tribe Arethuseae, subtribes Bletiinae and Thuniinae, based on the placement of representatives of these subtribes as sister taxa to *Coelogyne* using morphological data (Burns-Balogh & Funk, 1986), *ndhF* (Neyland & Urbatsch, 1996), *rbcL* (Cameron et al., 1999), *nad1* b–c (Freudenstein et al., 2000) and *matK* evidence (Chase et al., unpubl.). Voucher specimens are listed in Table 2.1 and deposited at K or L.

DNA extractions

Total genomic DNA was extracted from 50 mg of leaf tissue following the 2x CTAB method of Doyle & Doyle (1987). Some samples were purified through a cesium chloride/ethidium bromide gradient (1.55 g ml^{-1}). Leaf material was taken from one individual per species.

PCR RFLPs

RFLPs were detected by digesting three coding (16S, *psbA*, *psbD*) and eight non-coding regions (*trnT-trnL*, *trnL*, *trnL-trnF*, *trnC-trnD*, *trnS-psaA*, *atpB-rbcL*, *psbA-trnH*, *petA-psbE*) of the plastid genome using 19 restriction enzymes: *Bam*HI, *Bcl*I, *Bgl*II, *Bsm*I, *Cla*I, *Dra*I, *Eco*RI, *Eco*RV, *Hind*III, *Nde*I, *Nsi*I, *Pst*I, *Pvu*II, *Sac*I, *Sca*I, *Ssp*I, *Xba*I (six base cutters), *Dde*I (five base cutter), and *Hinf*I (four base cutter). Primers used were from Demesure et al. (1995), Fofana et al. (1997), Sang et al. (1997), Savolainen et al. (1995), Tsumura et al. (1995) and Taberlet et al. (1991). The thermal cycling protocol comprised 3 min. denaturation at 94 ºC, followed by 35 cycles, each with 45 sec. denaturation at 94 ºC, 45 sec. annealing at 50–57 ºC and an extension of 2 min. at 72 ºC, concluding with an extension of 10 min. at 72 ºC. Digested PCR products were separated on 1.5–2% agarose gels and stained with ethidium bromide to detect polymorphisms. The sizes of the fragments were determined with reference to two markers, a *Hind*III-*Eco*RI digested lambda bacteriophage DNA marker and a 100-bp marker.

Table 2.1. List of species analysed. Arranged by (sub)tribe, section and (sub)genus according to Dressler (1990), Butzin (1992), De Vogel (1994) and Clayton (in press).

Tribe	Subtribe	Genus and species	Section/ subgenus	Geographic origin	Voucher
Arethuseae	Bletiinae	*Arundina graminifolia* (D. Don) Hochr.		unknown	*Chase 395* (K)
Arethuseae	Bletiinae	*Bletia purpurea* (Lam.) DC.		Mexico	*Chase 581* (K)
Arethuseae	Thuniinae	*Thunia alba* (Lindl.) Rchb.f.		Nepal	*Chase 589* (K)
Coelogyneae	Coelogyninae	*Bracisepalum selebicum* J.J. Sm.		Sulawesi	*Leiden cult. 20446* (L)
		Chelonistele amplissima Ames & C. Schweinf.		Brunei	*Leiden cult. 26834* (L)
		Chelonistele sulphurea (Blume) Pfitzer		unknown	*Leiden cult. 21528* (L)
		Dendrochilum glumaceum Lindl.		unknown	*Leiden cult. 950648* (L)
		Dendrochilum longifolium Rchb.f.		PNG	*Leiden cult. 32110* (L)
		Entomophobia kinabaluensis (Ames) de Vogel		Sarawak	*Leiden cult. 970404* (L)
		Geesinkorchis phaiostele (Ridl.) de Vogel		Borneo	*Leiden cult. 30700* (L)
		Nabaluia angustifolia de Vogel		Sabah	*Leiden cult. 26217* (L)
		Neogyna gardneriana (Lindl.) Rchb.f.		unknown	*Leiden cult. 970729* (L)
		Panisea tricallosa Rolfe		China	*Leiden cult. 970828* (L)
		Pholidota carnea (Blume) Lindl.		Sumatra	*Leiden cult. 25469* (L)
		Pholidota imbricata Hook.		unknown	*Leiden cult. 21540* (L)
		Pleione bulbocodioides (Franch.) Rolfe		unknown	*Leiden cult. 990010* (L)
		Pleione formosana Hayata		unknown	*Leiden cult. 91051* (L)
		Coelogyne bicamerata J.J. Sm.	*Bicellae*	Sulawesi	*Leiden cult. 931067* (L)
		Coelogyne virescens Rolfe	*Brachypterae*	unkown	*Clayton cult. s.n.* (L)
		Coelogyne cristata Lindl.	*Coelogyne*	unknown	*Leiden cult. 2214* (L)
		Coelogyne foerstermannii Rchb.f.	*Coelogyne*	Sarawak	*Leiden cult. 970591* (L)
		Coelogyne sanderiana Rchb.f.	*Coelogyne*	unknown	*Leiden cult. 30765* (L)
		Coelogyne multiflora Schltr.	*Cyathogyne*	Sulawesi	*Leiden cult. 21747* (L)
		Coelogyne barbata Lindl. ex Griff.	*Elatae*	India	*Leiden cult. 990040* (L)
		Coelogyne stricta (D. Don) Schltr.	*Elatae*	unknown	*Leiden cult. 30695* (L)
		Coelogyne flaccida Lindl.	*Flaccidae*	unknown	*Leiden cult. 940707* (L)
		Coelogyne trinervis Lindl.	*Flaccidae*	unknown	*Leiden cult. 26940* (L)
		Coelogyne fimbriata Lindl.	*Fuliginosae*	unknown	*Leiden cult. 30759* (L)
		Coelogyne miniata (Blume) Lindl.	*Hologyne*	Java	*Leiden cult. 990287* (L)
		Coelogyne eberhardtii Gagnep.	*Lawrenceanae*	Vietnam	*Leiden cult. 970803* (L)
		Coelogyne chloroptera Rchb.f.	*Lentiginosae*	Philippines	*Leiden cult. 23511* (L)
		Coelogyne bilamellata Lindl.	*Longifoliae*	Philippines	*Leiden cult. 25164* (L)
		Coelogyne cuprea H. Wendl. & Kraenzl.	*Longifoliae*	Brunei	*Leiden cult. 914768* (L)
		Coelogyne harana J.J. Sm.	*Moniliformes*	Kalimantan	*Leiden cult. 970290* (L)
		Coelogyne kelamensis J.J. Sm.	*Moniliformes*	Kalimantan	*Leiden cult. 930568* (L)
		Coelogyne flexuosa Rolfe	*Ptychogyne*	unknown	*Leiden cult. 19937* (L)
		Coelogyne plicatissima Ames & C. Schweinf.	*Rigidiformes*	Sarawak	*Leiden cult. 980409* (L)
		Coelogyne beccarii Rchb.f.	*Speciosae*	PNG	*Leiden cult. 32230* (L)
		Coelogyne macdonaldii F. Muell. & Kraenzl.	*Speciosae*	Vanuatu	*Leiden cult. 25836* (L)
		Coelogyne dayana Rchb.f.	*Tomentosae*	unknown	*Leiden cult. 20247* (L)
		Coelogyne rhabdobulbon Schltr.	*Tomentosae*	Sabah	*Leiden cult. 26597* (L)
		Coelogyne rochussenii de Vriese	*Tomentosae*	unknown	*Leiden cult. 27060* (L)
		Coelogyne velutina de Vogel	*Tomentosae*	Peninsular Malaysia	*Leiden cult. 25835* (L)
		Coelogyne veitchii Rolfe	*Veitchiae*	PNG	*Leiden cult. 22277* (L)
		Coelogyne asperata Lindl.	*Verrucosae*	PNG	*Leiden cult. 22279* (L)
		Coelogyne pandurata Lindl.	*Verrucosae*	unknown	*Leiden cult. 21532* (L)

matK and ITS amplifications

The *trnK* intron (mostly *matK*) was amplified with the following four primers: -19F (5'-CGTTCTGACCATATTGCACTATG-3') and 881R (5'- TMTTCATCAGAA-TAAGAGT-3'); 731F (5'-TCTGGAGTCTTTCTTGAGCGA-3') and 2R (5'-AACTA-GTCGGATGGAGTAG-3'). All primers were designed at the Royal Botanic Gardens, Kew, except for 2R (Johnson & Soltis, 1994). The thermal cycling protocol comprised 28 cycles, each with 1 min. denaturation at 94 ºC, 30 sec. annealing at 48 ºC, an extension of 1 min. at 72 ºC, concluding with an extension of 7 min. at 72 ºC. All PCR products were sequenced directly after purification with QIAquick purification columns (QIAGEN, Amsterdam, The Netherlands). Four sequencing reactions were performed for each completed sequence, one with each of the four PCR primers, and these generated nearly complete overlapping single strand sequences for the *trnK* intron fragments.

ITS1 and ITS2 spacers along with the 5.8S gene were amplified with the primers 17 SE (5'-ACGAATTCATGGTCCGGTGAAGTGTTCG-3') and 26SE (5'-TAGAAT-TCCCCGGTTCGCTCGCCGTTAC-3') from Sun et al. (1994). The thermal cycling protocol comprised 26 cycles, each with 10 sec. denaturation at 96 ºC, 5 sec. annealing at 50 ºC and extension of 4 min. at 60 ºC. All PCR products were cloned following the protocol of Promega's pGEM-T Easy Vector System and then reamplified from transformed bacterial colonies by touching them with a sterile pipet tip and using that as template. Two sequencing reactions were performed for each completed sequence, one with each of the two PCR primers, and these generated nearly complete overlapping single strand sequences for the entire ITS fragments.

All amplified, double-stranded DNA fragments were purified using Wizard PCR minicolumns (Promega, Madison, Wisconsin, USA) and sequenced on an ABI 377 automated sequencer (PE Applied Biosystems, Inc.), using standard dye-terminator chemistry following the manufacturer's protocols.

Phylogenetic analyses

Variable restriction sites were coded as present or absent. Length variations were not included as characters in the analyses. Sequences were aligned by using MegAlign version 4.03 (DNASTAR, Inc. 1999) with subsequent adjustment by hand. Characters at position 143–170 bp were excluded from the ITS sequence data due to ambiguous alignment. Sequences are deposited in GenBank (AF302692 untill AF302761) and TREEBASE (SN570). The *matK* and ITS alignments and the PCR RFLPs data set are available from the first two authors upon request: e-mail gravendeel@nhn.leidenuniv.nl or m.chase@rbgkew.org.uk.

Maximum parsimony (MP) analysis was performed on the RFLP and sequence data with PAUP* version 4.0b64 (Swofford, 1999) using heuristic search, random addition with ten replicates and TBR swapping. *Arundina graminifolia*, *Bletia purpurea* and *Thunia alba* were specified as outgroups in all analyses. All molecular characters were assessed as independent, unordered and equally weighted using Fitch parsimony (Fitch, 1971). Indels were coded as missing data only. Number of transversions and their CIs and RIs were calculated on one of the MPTs of the combined analysis by using a stepmatrix with zero weights for transitions and the TREE SCORE command (ACCTRAN optimisation). From these data the number of transitions and their CIs

and RIs were calculated. To evaluate monophyly, trees were constrained using the enforce topological constraints option in PAUP*. The relative robustness for clades found in each parsimony analysis was assessed by performing 1000 replicates of bootstrapping (Felsenstein, 1995), using simple stepwise additions, SPR swapping, MULTREES on, and holding only 10 trees per replicate. The decay index (Bremer, 1994) was also calculated using the branch and bound option to examine trees up to six steps longer than the shortest tree found for each data set. Congruence of the separate data sets was assessed by visual inspection of the individual bootstrap consensus trees. Bootstrap trees were considered incongruent only if they displayed 'hard' (i.e. bootstrap percentages >80) incongruencies (Weins, 1998).

To explore the phylogenetic utility of some traditionally used morphological characters in classifications of the Coelogyninae, character state evolution of the shape of the lip base and petals, presence of hairs on the ovary and flower number per inflorescence was reconstructed using the assumptions of maximum parsimony with the Trace Character facility in MacClade version 3.04 (Maddison & Maddison, 1992). A complete phylogenetic analysis with morphological characters in *Coelogyne* and allied genera will be addressed in a separate publication.

RESULTS

PCR RFLP analysis

Four of the amplified regions were uninformative (16S, *psbA*, *psbD*, *trnL-trnF*). A total of 38 restriction sites was observed in the remaining seven regions. Of these, 15 were invariant, three were autapomorphies and the remaining 20 were potential synapomorphies (Table 2.2). MP analysis yielded >10,000 most parsimonious trees (length = 61, CI = 0.56, RI = 0.77; Table 2.3).

The RFLPs bootstrap consensus tree shows little resolution. Five weakly supported (<50%) clades are present: *Chelonistele*, *Coelogyne foerstermannii* plus *C. sanderiana* (sect. *Coelogyne*), sect. *Verrucosae*, *C. fimbriata* (sect. *Fuliginosae*) plus *C. stricta* (sect. *Elatae*), and *Pleione*.

matK sequence analysis

Length ranges of the *matK* gene and its flanking *trnK* sequences for Coelogyninae were 1536–1544 bp and 221–245 bp respectively. Boundaries of the *matK* gene were taken from Johnson & Soltis (1994). The final alignment has a total length of 1939 sites (1554 and 385 sites, resp.), of which 272 are variable and 119 potentially phylogenetically informative; there is one autapomorphic indel of 8 bp in the *matK* gene and five synapomorphic indels in the flanking *trnK* sequences, ranging in size from 4–19 bp. The transition/transversion ratio is 0.83, higher than the ratios found in Orchidaceae so far (Whitten et al., in press), but lower than the ratios found in dicots (Soltis & Soltis, 1998). Third-codon positions contributed the most steps (163 on the combined tree), slightly more than first or second positions, but all three sites displayed equal CI and RI values (Table 2.4 & 2.5). The average number of changes per variable site is 1.4 (Table 2.3). The MP analysis yielded >10000 most parsimonious trees (length = 394, CI = 0.77, RI = 0.79; Table 2.3). The *matK* bootstrap consensus tree is congruent with the results of the RFLP data, but shows more resolution at the (sub)generic level.

Table 2.2. Restriction site data used in the phylogenetic analysis.

Region	Informative restriction enzymes	Length and variation (bp)
16S	–	1400
psbA	–	1000
psbD	–	1100
trnL-trnF	–	500 ± 50
trnC-trnD	*Bgl*I, *Cla*I, *Dde*I, *Eco*RI	4500 ± 100
trnS-psaA	*Bam*HI, *Cla*I, *Eco*RI	4200 ± 50
petA-psbE	*Bam*HI, *Cla*I, *Dde*I, *Dra*I, *Ssp*I	2000 ± 100
atpB-rbcL	*Dra*I	1400 ± 50
trnL	*Dra*I, *Eco*RI, *Eco*RV	750 ± 50
trnH-psbA	*Hinf*I	700 ± 50
trnT-trnL	*Bcl*I, *Bgl*II, *Eco*RI	670 ± 50

Therefore, all plastid data were combined in a single analysis. The bootstrap consensus topology and the corresponding bootstrap percentages and decay values of this analysis are indicated in Fig. 2.1.

According to the combined plastid data the Coelogyninae excluding *Pleione* are monophyletic, but bootstrap support for placing *Pleione* outside Coelogyninae is low (<50%). Two sister clades within the subtribe are moderately supported. The first clade consists of species of *Bracisepalum*, *Chelonistele*, *Dendrochilum*, *Entomophobia*, *Geesinkorchis*, *Nabaluia*, *Coelogyne* sect. *Coelogyne*, *Cyathogyne*, *Tomentosae*, *Veitchiae* and *Verrucosae* (60%). Four smaller sets of taxa in this first major clade are recovered in all bootstrap replicates: *Bracisepalum selebicum* together with *Dendrochilum*, *Chelonistele*, *Coelogyne dayana* plus *C. rhabdobulbon* (sect. *Tomentosae*),

Table 2.3. Values and statistics from parsimony analyses of separate and combined data matrices.

	RFLPs	*matK*	all plastid data	ITS1-5.8S-ITS	combined
number of included positions in matrix	23	1939	–	729	–
number of variable sites	23	272 (14%)	–	436 (66%)	–
number of phylogenetically informative sites	20	119	–	224	–
number of MPTs	10,000+	10,000+	174	32	4
tree length	61	394	474	1355	1729
CI	0.56	0.77	0.71	0.57	0.60
RI	0.77	0.79	0.75	0.53	0.57
average number of changes per variable site	3.4	1.4	–	2.5	–
length on combined tree	73	377	–	1092	–
number of clades in bootstrap consensus with >80% support	0	7	7	8	11

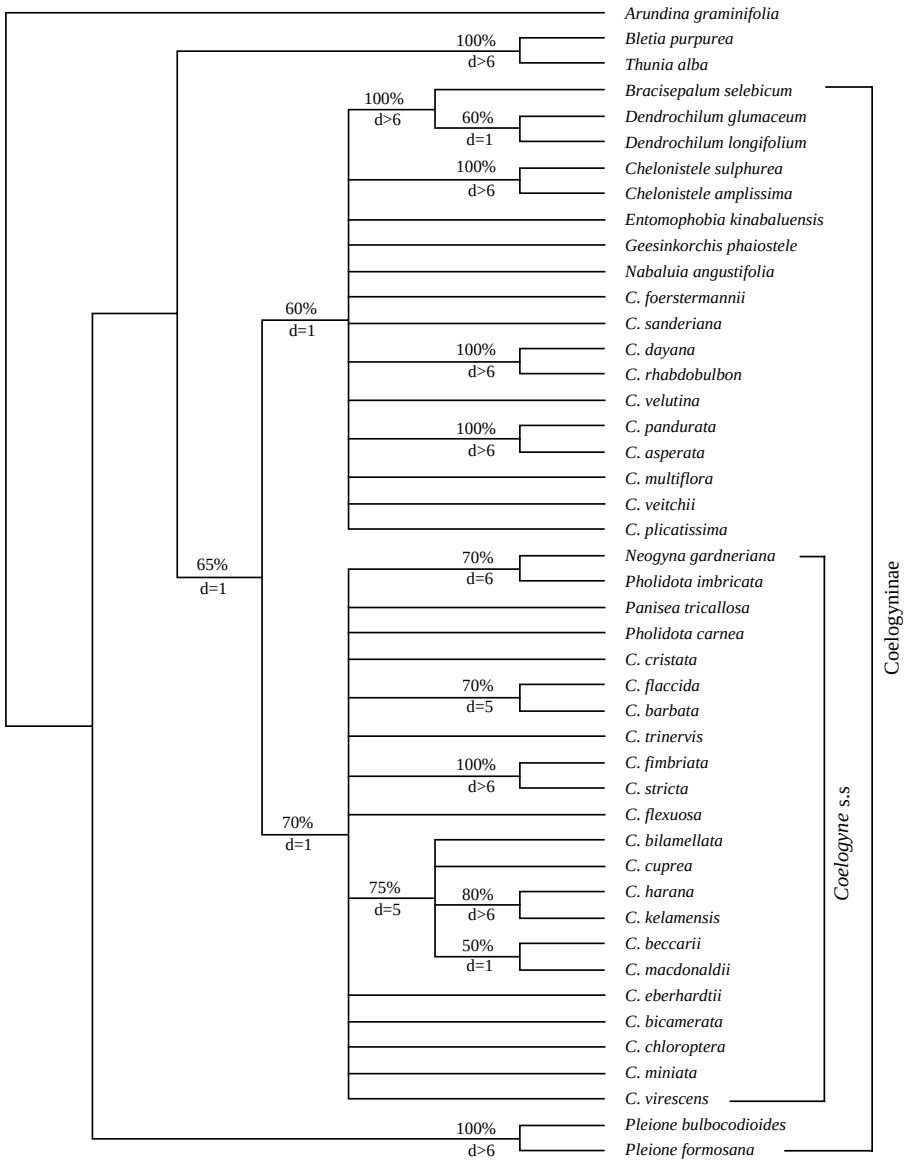

Fig. 2.1. Bootstrap consensus of 174 trees from parsimony analysis of all plastid data with bootstrap percentages and decay values (only percentages > 50% are given).

and sect. *Verrucosae*.

The second major subclade consists of species of *Neogyna*, *Panisea*, *Pholidota*, *Coelogyne* sect. *Bicellae*, *Brachypterae*, *Coelogyne*, *Elatae*, *Flaccidae*, *Fuliginosae*, *Hologyne*, *Lawrenceanae*, *Lentiginosae*, *Longifoliae*, *Moniliformes*, *Ptychogyne* and *Speciosae* (70%). Two strongly supported smaller sets of taxa are present in this second major clade: *Coelogyne fimbriata* (sect. *Fuliginosae*) plus *Coelogyne stricta* (sect. *Elatae*) (100%), and sect. *Moniliformes* (80%).

ITS sequence analysis

Length ranges of nuclear rDNA ITS sequences for Coelogyninae were 204–253 bp,159–163 bp, and 242–271 bp respectively. Boundaries from the 5.8S gene are taken from Hershkovitz & Lewis (1996). One region of 30 bp in ITS1 was considered unalignable and was excluded. The final alignment has a total length of 759 sites (286, 165, and 308 sites for ITS1, 5.8S and ITS2, resp.). Of the included positions, 436 are variable and 224 potentially phylogenetically informative, which is in accordance with variation levels in most angiosperms (Baldwin et al., 1995). The transition / transversion ratio is 1.68 in the ITS1 spacer and 1.56 in the ITS2 spacer region (Table 2.4), which is in accordance with ratios found in Orchidaceae so far (Cox et al., 1997; Pridgeon et al., 1997; Whitten et al., in press). The average number of changes per variable site is 2.5 (Table 2.3). The MP analysis yielded 32 most parsimonious trees (length = 1355, CI = 0.57, RI = 0.53; Table 2.3). The bootstrap consensus topology and the corresponding bootstrap percentages and decay values are indicated in Fig. 2.2.

The ITS bootstrap consensus tree shows less resolution than the combined plastid data. The monophyly of Coelogyninae including *Pleione* is strongly supported (90%). Six strongly supported clades are present: *Bracisepalum selebicum* together with *Dendrochilum* (100%), *Chelonistele sulphurea* plus *Entomophobia kinabaluensis* (100%), sect. *Verrucosae* (100%), sect. *Moniliformes* (100%), *C. eberhardtii* (sect. *Lawrenceanae*) plus *C. miniata* (subgenus *Hologyne*) (90%), and *Pleione* (100%). Many other clades are weakly supported (< 80%).

Combined analysis

Bootstrap consensus trees of the three individual data sets revealed no hard incongruences except for the placement of *Panisea*. To improve resolution, a combined analysis of all three data sets (excluding *Panisea*) was performed (Huelsenbeck et al., 1996). The combined matrix yielded four most parsimonious trees (length = 1729; CI = 0.60; RI = 0.57; Table 2.3). Bootstrap analyses of the combined data set (excluding *Panisea*) provided more resolution and higher internal support for relationships (> 80%) than did any of the individual data sets (Table 2.3). One of the four most parsimonious trees with the corresponding bootstrap percentages and decay values are shown in Fig. 2.3.

In the combined analysis of all molecular data, the Coelogyninae including *Pleione* receive strong support (90%). Three main clades in Coelogyninae are present. In the first major clade, species of *Bracisepalum*, *Chelonistele*, *Dendrochilum*, *Entomophobia*, *Geesinkorchis* and *Nabaluia* are nested within species of *Coelogyne* sect. *Coelogyne*, *Cyathogyne*, *Tomentosae*, *Veitchiae* and *Verrucosae* (80%). Four smaller subsets of taxa are recovered in all bootstrap replicates: *Bracisepalum selebicum* with *Dendrochilum*, *Chelonistele sulphurea* with *Entomophobia kinabaluensis*, *Coelogyne dayana* with *C. rhabdobulbon* (sect. *Tomentosae*), and sect. *Verrucosae*. In the second major clade, species of *Neogyna* and *Pholidota* are nested within species of *Coelogyne* sect. *Bicellae*, *Brachypterae*, *Coelogyne*, *Elatae*, *Flaccidae*, *Fuliginosae*, *Hologyne*, *Lawren-*

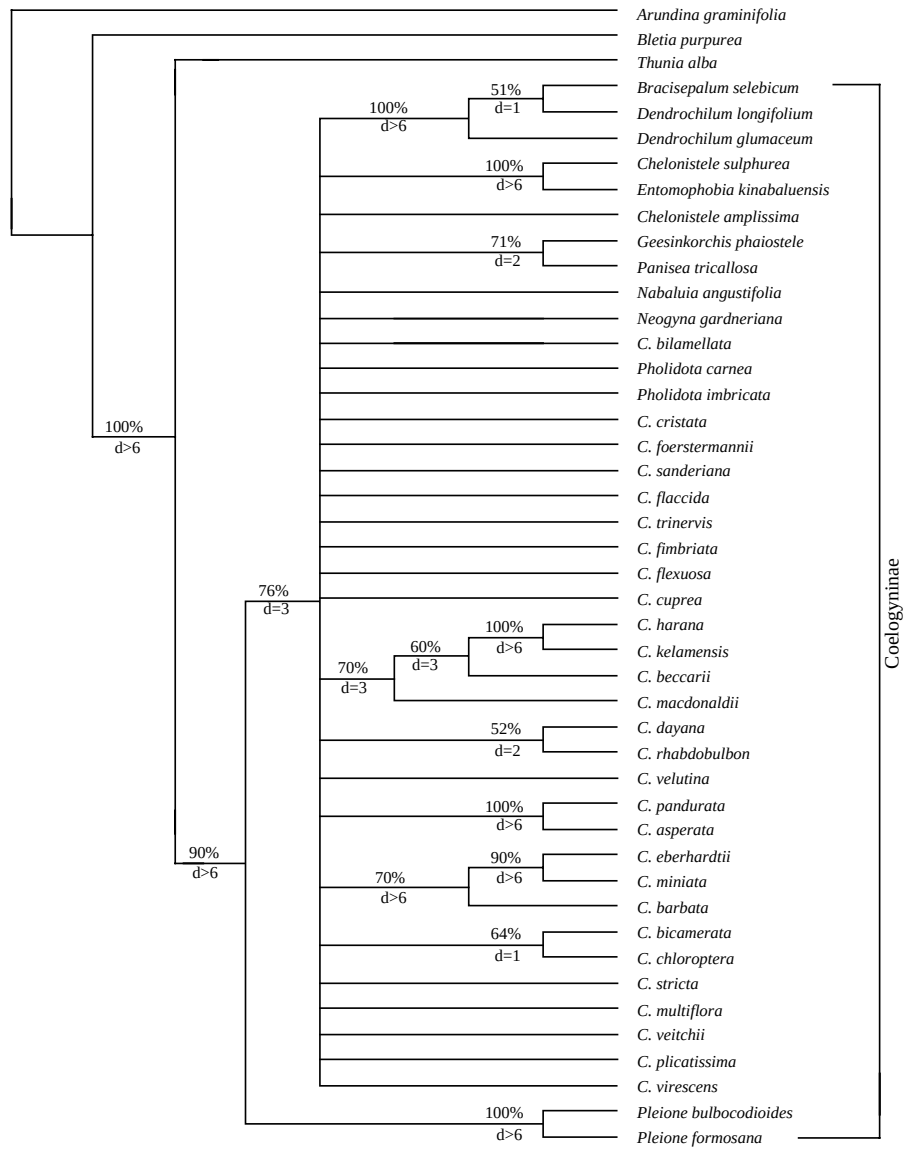

Fig. 2.2. Bootstrap consensus of 32 trees from parsimony analysis of ITS1-5.8S-ITS2 sequence data with bootstrap percentages and decay values (only percentages > 50% are given).

ceanae, *Lentiginosae*, *Longifoliae*, *Moniliformes*, *Ptychogyne* and *Speciosae* (100%). Four subsets of taxa receive strong support within this second major clade: *Neogyna gardneriana* with *Pholidota imbricata* (80%), *C. fimbriata* (sect. *Fuliginosae*) with *C. stricta* (sect. *Elatae*) (90%), sect. *Moniliformes* (100%), sect. *Moniliformes* with sect. *Speciosae* (95%), and *C. eberhardtii* (sect. *Lawrenceanae*) with *C. miniata*

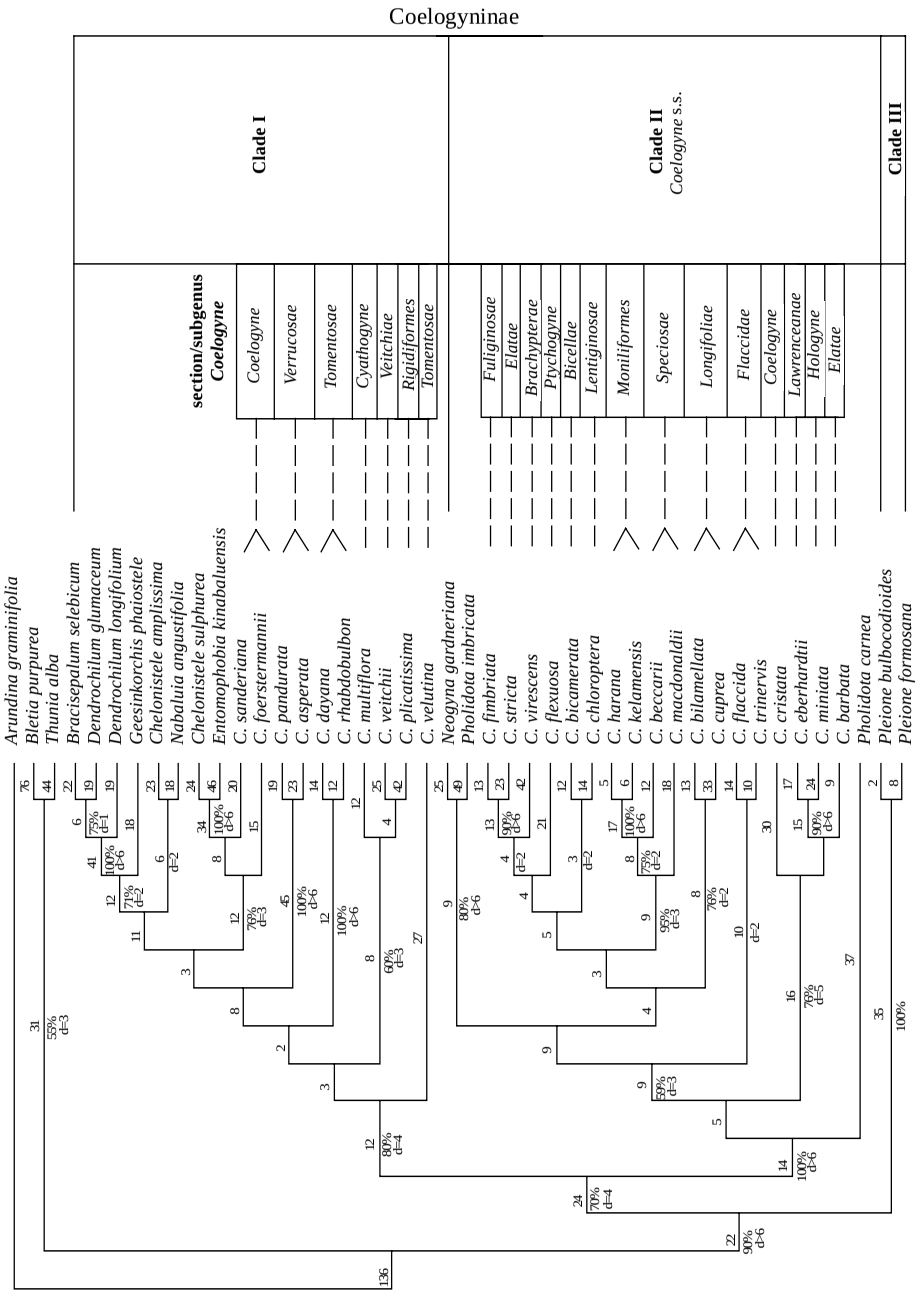

Fig. 2.3. One of the four MTPs from parsimony analysis of the combined molecular data. Branch length (number of nucleotide substitutions) is indicated above the branches and bootstrap percentages and decay values below them (only percentages > 50% are given).

Table 2.4. Number of steps, CI and RI for each codon position in *matK* based upon one of the four MPTs from the combined analysis.

Codon position	number of steps	CI	RI
1	84 (25%)	0.67	0.75
2	86 (26%)	0.74	0.75
3	163 (49%)	0.67	0.71

(subgenus *Hologyne*) (90%). The third major clade consists of species of *Pleione* (100%).

DISCUSSION

Comparative phylogenetic utility of separate data sets

The number of phylogenetically informative characters produced by the RFLPs was less succesful than predicted (20 only; Table 2.3), compared to studies of restriction enzyme digestions of PCR products in other plant groups (Tsumura et al., 1995; Mes et al., 1997; Asmussen & Liston, 1998). Apparently, the rate of plastid DNA evolution is relatively low in Coelogyninae.

The excess of transversions (Table 2.5), the excess of substitutions at third-codon positions (Table 2.4), and the multiple stop-codons present in the alignment, all indicate a loss of function for *matK* in Coelogyninae. For dicots so far, *matK* data demonstrate ts/tv ratios greater than 1.0 and a relatively even distribution of substitutions across codon positions (Xiang et al., 1998). A loss of function is also suspected for Maxillarieae (Whitten et al., in press).

The number of phylogenetic informative characters of the nuclear rDNA ITS sequences is much higher than those of the RFLPs and *matK* data (224 vs. 20 and 119, resp.; Table 2.3). However, many clades in the ITS bootstrap consensus are only weakly supported. The much lower CI and RI of the nuclear data (0.57 and 0.53) compared with the *matK* sequences (0.77 and 0.79) indicate a higher internal conflict, which might be caused by saturation at this taxonomic level of comparison. This is also consistent with the higher rate at which the variable sites in this region evolve compared with the *matK* sequences (2.5 steps/site vs. 1.4 steps/site, resp.; Table 2.3).

New generic circumscription of Coelogyne

Many recent studies have indicated that combined molecular data sets utilizing

Table 2.5. Number of steps, CI and RI for transitions and transversions for each region based upon one of the four MPTs from the combined analysis.

	matK		ITS1		ITS2	
	ts	tv	ts	tv	ts	tv
number of steps	189	228	433	258	318	203
CI	0.74	0.62	0.41	0.55	0.42	0.62
RI	0.82	0.64	0.27	0.35	0.53	0.51
ts/tv	0.83		1.68		1.56	

regions with different levels of variation provide resolution at different levels of the cladogram, and that phylogenetic resolution and bootstrap percentages are improved by directly combining independent molecular data sets (Chase & Cox, 1998; Soltis & Soltis, 1998; Whitten et al., in press). The high level of congruence among the three data sets, representing both coding and non-coding as well as plastid and nuclear regions, and the high bootstrap percentages in the combined analysis support the combined tree as a good hypothesis of phylogenetic relationships of *Coelogyne* and the Coelogyninae.

Coelogyne as currently circumscribed is polyphyletic, with species falling into at least two well-supported clades. Constrained analyses of the combined data set showed that the number of additional steps needed to achieve monophyly of *Coelogyne* is relatively great (at least 154 steps longer).

There are two possible taxonomic solutions for a new phylogenetic classification of *Coelogyne* in which only monophyletic groups are recognised. The first would be to include all sampled species of Coelogyninae (excluding *Pleione*) within a single genus. According to the rules of priority, this genus should be called *Coelogyne*. However, given the large differences in floral morphology, the creation of many synonyms in widespread horticultural use, and the high number of species *Coelogyne* would then encompass (approximately 530), this option is not satisfactory. A second possibility would be to reduce *Coelogyne* to one of the three main clades found. The type species of *Coelogyne* (*C. cristata*) belongs to clade II (*Coelogyne* s.s.). The best option for reorganising *Coelogyne* seems therefore to take the following actions:

1) restriction of *Coelogyne* to the *Coelogyne* s.s. clade, including *Neogyna* and *Pholidota*. These two genera were already considered to be just sections of *Coelogyne* by Lindley, Griffith and Reichenbach f. (Lindley, 1830; Griffith, 1851; De Vogel, 1988). All species sampled of *Neogyna* and *Pholidota* have glabrous ovaries and a hypochile with broad, erect lateral lobes. These characters are also present in the other species of the *Coelogyne* s.s. clade;
2) removal of the species of *Coelogyne* sect. *Coelogyne* (in part), *Cyathogyne*, *Rigidiformes*, *Tomentosae*, *Veitchiae* and *Verrucosae*. The main morphological characters distinguishing these species from *Coelogyne* s.s. are the relatively high number of simultaneously opening flowers with persistent floral bracts, hairy ovaries and ovate-oblong petals. These characters are also present in *Bracisepalum*, *Chelonistele*, *Dendrochilum*, *Entomophobia*, *Geesinkorchis* and *Nabaluia*.

Sectional and subgeneric relationships within Coelogyne

From the 18 different sections of *Coelogyne* considered here, just two (with only two sampled species each) form strongly supported monophyletic groups in the combined analysis: sect. *Moniliformes* and *Verrucosae* (both 100% supported). This is consistent with the clear morphological synapomorphies that characterise both sections. All species of sect. *Moniliformes* have elongated, unifoliate pseudobulbs, a rhachis with distinctly swollen, short internodes and many flowers that open in succession (Carr, 1935). All species of sect. *Verrucosae* have rounded to strongly flattened bifoliate pseudobulbs, inflorescences with many, simultaneously opening flowers and a rhachis with a few sterile bracts at the base and scattered minute scale-like hairs, which are also present on the pedicel, ovary and the abaxial side of the sepals and petals (Sierra

et al., 2000).

Coelogyne sect. *Longifoliae* is also monophyletic, although support for this clade is weak (76%). The species of sect. *Longifoliae* all have bifoliate pseudobulbs, long and stiff inflorescences, a rhachis with long internodes, and intermediate-sized flowers that open in succession (Clayton, in press). *Coelogyne* sect. *Flaccidae* is monophyletic in all shortest trees, but bootstrap support for this clade is low (< 50%). This is in accordance with the few and not unique synapomorphies that define this section. *Coelogyne* sect. *Flaccidae* is characterised by a low number of simultaneously opening flowers with deciduous floral bracts and undulating keels on the lip, a combination of characters that also defines sect. *Ocellatae* (Clayton, in press).

Coelogyne sect. *Tomentosae* is not monophyletic, but none of the branches separating its two parts receives even low internal support.

Coelogyne sect. *Coelogyne* and sect. *Elatae* are clearly paraphyletic. This is in accordance with the high variety in pseudobulb shape, inflorescence type, flower size and morphology of the keels on the lip in both sections. The only character that is present in all species currently assigned to sect. *Cristatae* is the colour of the flowers: white, with yellow/brown spots. The species currently assigned to sect. *Elatae* only share the sterile bracts at the base of the rhachis and the simultaneously opening flowers, a combination of characters present in many other Coelogyninae species.

Coelogyne sect. *Lawrenceanae* and sect. *Speciosae* are well separated, which is not in accordance with Seidenfaden (1975), who suggested they should be combined. Molecular data support our view that they should be considered different sections because of their clear morphological differences. All species of sect. *Lawrenceanae* have shiny green, smooth pseudobulbs, hysteranthous inflorescences and flowers with deeply incised, glabrous keels on the lip. All species of sect. *Speciosae* are characterised by angular, dull green pseudobulbs, synanthous or proteranthous inflorescences and flowers with hairy or warty keels on the lip (Gravendeel & de Vogel, 1999).

A well-supported subset of species is formed by *C. multiflora* (subgenus *Cyathogyne*), *C. plicatissima* (sect. *Rigidiformes*) and *C. veitchii* (sect. *Veitchiae*). Morphological synapomorphies for this clade are the hairy ovaries, persistent floral bracts and small flowers. Another clade with high support consists of *C. fimbriata* (sect. *Fuliginosae*) and *C. stricta* (sect. *Elatae*). Both species have sterile bracts on the scape and intermediate-sized flowers. A third group of taxa supported by high bootstrap percentages is made up of *C. eberhardtii* (sect. *Lawrenceanae*) and *C. miniata* (subgenus *Hologyne*). Both species have bifoliate pseudobulbs and deciduous floral bracts. However, in other characters, such as plant size, leaf texture, inflorescence type and keel morphology, they show considerable differences. To investigate whether these three clades warrant the status of new sections, a much larger sampling within *Coelogyne* is needed.

Naturalness and content of subtribe Coelogyninae

Analysis of the combined RFLP, *matK* and ITS data set indicates that Coelogyninae are monophyletic and diverged early into three major clades. Clade I comprises species of *Coelogyne* sect. *Coelogyne*, *Cyathogyne*, *Rigidiformes*, *Tomentosae*, *Veitchiae* and *Verrucosae*, from which *Bracisepalum*, *Chelonistele*, *Dendrochilum*, *Entomophobia*, *Geesinkorchis* and *Nabaluia* were split by various authors. Synapomorphies for this

group of species are the simultaneously opening flowers (with the exception of *Geesinkorchis*) and inflorescences with relatively many flowers. Many other characters, such as small flower size, persistent floral bracts, hairy ovaries, ovate-oblong petals and a hypochile with inconspicuous lateral lobes, although present in most taxa of clade I, are not perfectly coincident, probably due to considerable convergent evolution in this group of species. The presence of many different generic names in clade I can be explained by the high number of autapomorphies present, such as the presence of stipes in *Geesinkorchis* and a transverse callus on the lip of *Entomophobia*. *Bracisepalum selebicum* and both *Dendrochilum* species sampled form a well-supported subset of taxa in clade I. These three species have unifoliate pseudobulbs, pendulous inflorescences with sterile bracts on the base of the rhachis and small flowers with a hypochile with inconspicuous lateral lobes. Another well-supported subclade in clade I comprises *Chelonistele sulphurea* and *Entomophobia kinabaluensis*. Both species have erect inflorescences with sterile bracts on the base of the rhachis and small flowers with a relatively short column. Generic boundaries within clade I are not clear yet, as most internal nodes of this clade are only poorly supported. More data should be collected to improve resolution. Additional taxon sampling is also needed to find the limits of new monophyletic groups.

Clade II (*Coelogyne* s.s.) subsequently diverged into species of *Neogyna* and *Pholidota* nested within species of *Coelogyne* sect. *Bicellae*, *Brachypterae*, *Elatae*, *Flaccidae*, *Fuliginosae*, *Hologyne*, *Lawrenceanae*, *Lentiginosae*, *Longifoliae*, *Moniliformes*, *Ptychogyne* and *Speciosae*. Synapomorphies for this group of species are the glabrous ovaries, linear petals (with the exception of *Pholidota*) and broad, erect lateral lobes of the hypochile. Many other characters, such as a small flower number, deciduous floral bracts and large flower size are not present in all taxa of clade II. *Neogyna gardneriana* and *Pholidota imbricata* form a strongly supported subset of taxa in clade II. Both species have an epichile with semi-orbicular, widely retuse lateral lobes.

Clade III consists of species of *Pleione*. The relatively isolated position of *Pleione* is consistent with the purplish pink colour of the flowers, short-lived pseudobulbs and annually deciduous leaves of many species in this genus, which do not occur in any of the other Coelogyninae (Cribb et al., 1983).

The position of *Panisea* differs in the plastid and the ITS trees. In all of the plastid cladograms, *Panisea* is placed in the *Coelogyne* s.s. clade (clade II), in some of them as sister species to *C. fimbriata* and *C. stricta*. In contrast, *Panisea* appears as sister species to *Geesinkorchis* in the majority of the ITS trees (clade I). A combined analysis including *Panisea* (not shown) results in a nearly complete loss of internal support for clades I and II, an indication that its position is incongruent in the trees from each genome. Therefore, we removed *Panisea* so that clear patterns could be discerned. Hard incongruencies between nuclear and organellar phylogenetic trees are often attributed to introgression of a cytoplasmic genome from one species into the nuclear background of another species (Wendel & Doyle, 1998). *Panisea tricallosa*, *C. fimbriata* and *C. stricta* show an overlap in distribution area in northern India and Nepal. We suggest that *Panisea tricallosa* shares a similar *matK* sequence with these species as a result of introgression. However, introgression is not the only process that could produce such incongruence. A second cause might be coalescence of alleles antedating

species divergence (lineage sorting). There are relatively few examples of plastid DNA polymorphisms that transcend species boundaries, probably because of the generally slow rate of plastid DNA evolution (Wendel & Doyle, 1998). Therefore, introgression due to hybridisation appears to be the most probable explanation for the incongruence caused by *Panisea*. However, it is difficult to distinguish between introgression and lineage sorting, because they both may produce similar phylogenetic patterns (Hardig et al., 2000).

Phylogenetic utility of traditionally used key characters

The shape of the lip base and petals, presence of hairs on the ovary and number of flowers per inflorescence have been used for diagnosing genera within Coelogyninae and sections / subgenera within *Coelogyne* (Pfitzer & Kraenzlin, 1907b, d; De Vogel, 1992; Pedersen et al., 1997; Clayton, in press). To evaluate their phylogenetic significance, we reconstructed their distribution on the strict consensus of the four cladograms from the combined analysis.

Lip base shape

A saccate lip base is present in all Coelogyninae sampled except for the species of *Coelogyne*. Fig. 2.4A shows the most parsimonious derivation of a saccate lip base. A saccate lip base is gained at least four times and appears not to be phylogenetically useful at the generic level. The evolutionary lability of this character might be caused by a close association with pollination systems, which can be homoplasious in Orchidaceae (Dressler, 1981; Chase & Palmer, 1992; Hapeman & Inoue, 1997). Moreover, lip bases in Coelogyninae might not be derived from the same structure for all taxa studied. For instance, in *Bracisepalum* the base of the lip has two sac-like extensions, which might not be homologous with the saccate lip base of the other Coelogyninae. Studies of floral development may give additional insight as to whether different lip base types are derived by common descent or as a result of parallelism.

Petal shape

Petals are ovate-oblong in *Bracisepalum*, *Dendrochilum*, *Entomophobia*, *Pholidota*, *Pleione* and species of *Coelogyne* sect. *Coelogyne*, *Cyathogyne*, *Rigidiformes*, *Tomentosae*, *Veitchiae* and *Verrucosae*. All other taxa sampled in Coelogyninae have linear petals. The ancestral habit for the Coelogyninae seems to be ovate-oblong petals, and linear petals are gained in *Coelogyne* s.s., *Chelonistele*, *Geesinkorchis* and *Nabaluia* (Fig. 2.4B). Linear petals are derived at least three times and appear not to be phylogenetically useful at the generic level within the principal lineages of Coelogyninae.

Ovary indumentum

Hairy ovaries are present in all species of clade I except for *Geesinkorchis*, whereas all other Coelogyninae have glabrous ovaries (see also Fig. 2.4C). The presence of hairs on the ovary is a uniquely derived character state supporting the view of clade I as a separate group in Coelogyninae.

Flower number

A clear gap is present in the number of flowers per inflorescence of the taxa analyzed.

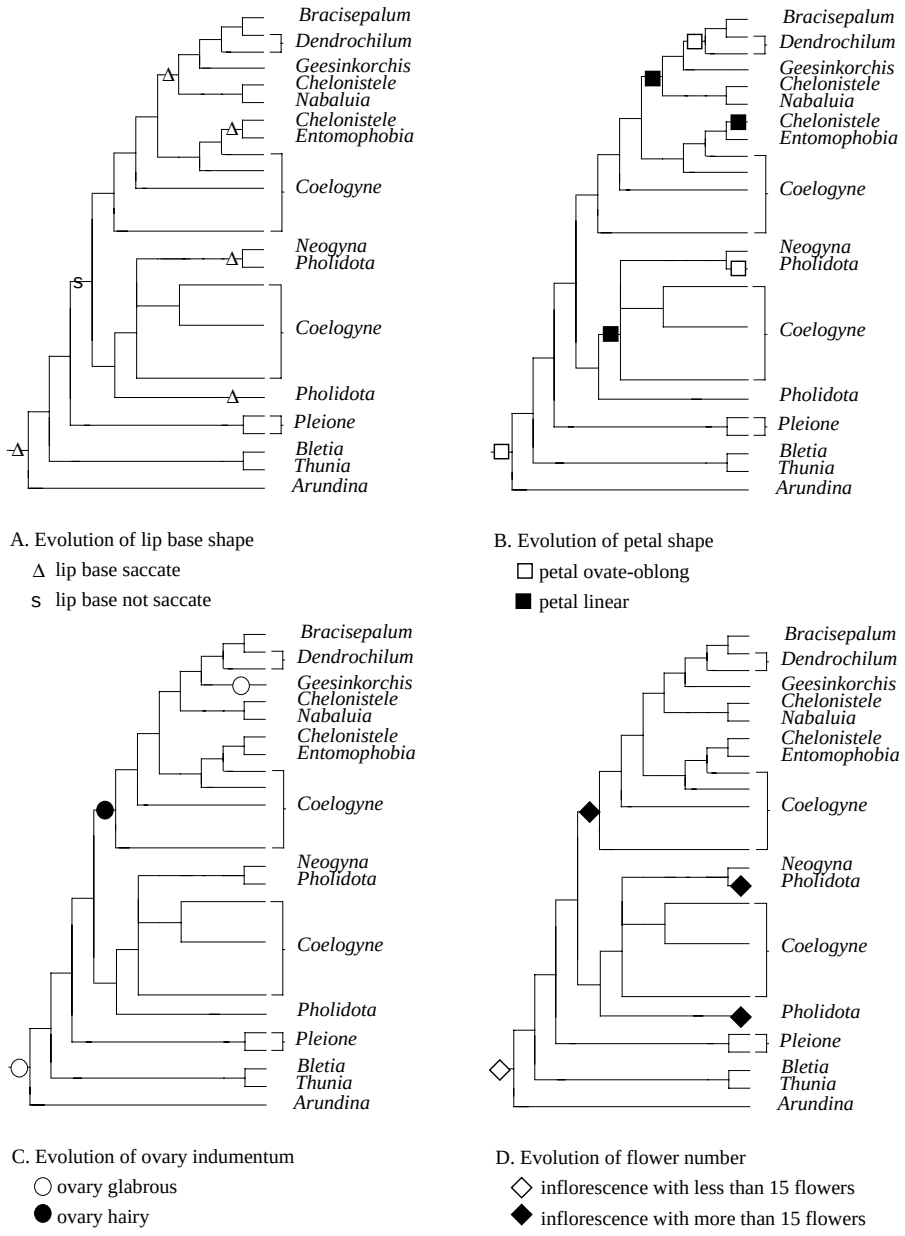

Fig. 2.4. A most parsimonious reconstruction of character state evolution for three characters used in traditional classi•cations of *Coelogyne* and the Coelogyninae. — A. Lipbase shape; B. petal shape; C. ovary indumentum; D. flower number.

Inflorescences with 15 or more flowers are present in all species of clade I, whereas all other Coelogyninae (with the exception of *Pholidota*) have much lower flower numbers (see also Fig. 2.4D). A high flower number represents a good synapomorphy for clade I.

Reorganisation of Coelogyne

The traditional generic circumscription of *Coelogyne* is mainly based on the absence of a saccate lip base, which is present in all other genera of Coelogyninae. Absence of characters often indicates symplesiomorphy and generally makes the group defined by such a character paraphyletic. Sectional delimitations were previously based on such characters, which also intergrade considerably among closely related species. The poly- and paraphyletic nature of *Coelogyne* and several of its sections according to molecular data clearly shows how convergent floral morphology has confounded traditional taxonomy. Traditionally used classifications of *Coelogyne* and Coelogyninae are not supported by the molecular data presented here and should be abandoned.

When plotted on the molecular cladograms, some of the traditionally used key characters for generic delimitation in Coelogyninae, such as lip base and petal shape, seem unacceptably homoplasious. In contrast, ovary indumentum and flower number are good diagnostic characters. We propose to redefine the genus *Coelogyne* by the following two actions:

1) inclusion of *Neogyna* and *Pholidota*. These two genera fit perfectly in *Coelogyne* when a new definition of the genus consists of glabrous ovaries only, a lip with a saccate or flat base, and a hypochile with broad, erect lateral lobes;
2) removal of the species of *Coelogyne* sect. *Coelogyne* (in part), *Cyathogyne*, *Tomentosae*, *Rigidiformes*, *Veitchiae* and *Verrucosae*. These species fit better in clade I, because they share several synapomorphies with other genera in this clade, such as a relatively high number of simultaneously opening flowers with persistent floral bracts and hairy ovaries. Our phylogenetic analyses indicate that approximately 160 species would be left in *Coelogyne*.

In contrast with the *Coelogyne* s.s. clade, a good morphological delimitation of clade I is still difficult. Many characters, although present in most taxa of clade I, do not map perfectly on the molecular cladograms due to a substantial amount of convergent evolution in this group. In addition, generic boundaries within clade I are not yet clear, as most internal nodes have only low support. Additional sampling in clade I is needed to find the limits of new monophyletic groups, and to justify any formal proposals for nomenclatural changes.

ACKNOWLEDGEMENTS

The authors thank Dudley Clayton (UK), Mark Clements (CSIRO, Australia), Phillip Cribb (Royal Botanic Gardens, Kew), Anton Sieder (Botanical Gardens, Vienna), Art Vogel (Hortus Botanicus, Leiden), Sofie Mursidawati and Dwi Murti Puspitaningtyas (Kebun Raya Bogor, Indonesia), Julaihi Lai (Semengoh Botanical Gardens, Sarawak) and Peter O'Byrne (Singapore) for plant specimens. Bertie Joan van Heuven, Jeffrey Joseph, Martyn Powell and Peter Kuperus assisted with sequencing. Fieldwork for this study was supported by grants from the Stichting Fonds Christine Buisman, Alberta M.W. Mennega Stichting, Dutch Scientific Fund for Research in the Tropics (Treub Maatschappij) and the Netherlands Foundation for the Advancement of Tropical Research (WOTRO).