Phylogeny of *Curcuma* (Zingiberaceae) based on plastid and nuclear sequences: Proposal of the new subgenus *Ecomata*

Eliška Záveská,[1] Tomáš Fér,[1] Otakar Šída,[2] Karol Krak,[3] Karol Marhold[1,4] & Jana Leong-Škorničková[5]

1 *Department of Botany, Faculty of Science, Charles University, Benátská 2, 128 01 Praha 2, Czech Republic*
2 *Department of Botany, National Museum in Prague, Cirkusová 1840, 193 00 Praha 9, Czech Republic*
3 *Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, 25243 Průhonice, Czech Republic*
4 *Institute of Botany, Slovak Academy of Sciences, Dúbravská cesta 9, 845 23 Bratislava, Slovak Republic*
5 *The Herbarium, Singapore Botanic Gardens, 1 Cluny Road, 259569, Singapore*
Author for correspondence: *Jana Leong-Škorničková, jana_skornickova@seznam.cz*

Abstract *Curcuma* comprises 120 species that occur throughout tropical and subtropical Asia. The taxonomy of the genus is haunted by polyploid speciation and homoploid hybridization, making it the most challenging genus in Zingibereae (Zingiberaceae). *Curcuma* is best known for turmeric (*C. longa*), but numerous species are extensively used as medicinal plants, ornamentals, and sources of starch, among many other uses. The delimitation of the genus has been a matter of dispute since its establishment by Linnaeus (1753), and further conflict has arisen from recent molecular and morphological studies suggesting either paraphyly of *Curcuma* or the necessity to broaden the genus to include four small genera (*Laosanthus*, *Paracautleya*, *Stahlianthus*, *Smithatris*) as well as several species currently placed in *Kaempferia* and *Hitchenia*. All previous infrageneric classifications were based on limited material that did not include species from the Indochinese floristic region, and these classifications are unable to unequivocally accommodate all currently known members of the genus. To test the monophyly and delimitation of *Curcuma* and to gain more insight into infrageneric relationships, three plastid regions (*trnL-trnF*, *psbA-trnH*, *matK*) and the internal transcribed spacer (ITS) of nuclear ribosomal DNA were sequenced. Fifty *Curcuma* species covering the morphological and geographic variation of the genus and 12 *Curcuma*-like species currently or previously treated as members of other genera were included in this study. In addition, four Zingibereae and three other Zingiberaceae species were used as outgroups. The results of maximum parsimony and Bayesian analyses clearly support a broad generic boundary for *Curcuma*, with inclusion of *Laosanthus*, *Paracautleya*, *Stahlianthus*, *Smithatris* and some species of *Kaempferia* and *Hitchenia* (*K. scaposa*, *K. candida*, *H. caulina*, *H. glauca*). Four main groups in *Curcuma* s.l. were detected, and their importance for classification at the subgenus level is discussed. A new infrageneric classification is proposed here with a formal description of a new subgenus. Cloning uncovered a broad range of variation of ITS sequences within individuals, particularly in the terminal 'Curcuma' group containing representatives of the nominal subgenus *Curcuma*. This 'intra-individual ITS polymorphism' increases with ploidy level and is coupled with preferred vegetative reproduction. Additional studies are needed to further uncover highly complex relationships in this subgenus.

Keywords cpDNA; generic delimitation; internal transcribed spacers; ITS; phylogeny; polyploidy; Zingiberoideae

Supplementary Material The alignment is available in the Supplementary Data section of the online version of this article (http://www.ingentaconnect.com/content/iapt/tax).

■ INTRODUCTION

Curcuma L. (Zingiberaceae), with at least 120 species, is the third-largest genus in the primarily tropical Zingiberaceae (ca. 50 genera, >1500 spp.). It is a diverse polyploid complex containing many taxa of economic, medicinal, ornamental and cultural importance, the type species of the genus, *C. longa* L. (turmeric) being the best-known example. *Curcuma* is distributed throughout South and Southeast Asia, with a few species extending to China, Australia and the South Pacific. The highest diversity is concentrated in monsoonal Asia from India to Indochina (Škorničková, 2007). Despite the considerable economic potential of this genus, its phylogeny and taxonomy remain poorly understood, mainly due to extensive polyploidization (2*x*–15*x*) and hybridization (Leong-Škorničková & al., 2007; Záveská & al., 2011) resulting in different levels of genetic and morphological variation among species and in blurred species boundaries (Škorničková, 2007). The nomenclatural complexity of the group, together with the cultivation and spread by humans of numerous species across the tropics, has further contributed to the misapplication of names (Leong-Škorničková & al., 2008a, b, 2010). Recent explorations of the

This paper was a contribution to the XI Symposium of the International Organization of Plant Biosystematists, held in Aurangabad, India from 2 to 4 September 2010 (see Taxon 60: 298–299. 2011).

Indian and Indochinese floristic regions have resulted in many new taxa (e.g., Sirirugsa & Newman, 2000; Mood & Larsen, 2001; Škorničková, 2007; Leong-Škorničková & Lý, 2010; Leong-Škorničková & al., 2010), with many more awaiting formal description (Leong-Škorničková & al., in prep.).

The delimitation of *Curcuma* has been a matter of dispute since its establishment in 1753 by Linnaeus, who based the genus on *Curcuma rotunda* L. (now in *Boesenbergia* Kuntze) and *Curcuma longa* L. (now the type species of the genus). Diagnostic characters used by various authors to define *Curcuma* (e.g., pouched inflorescences, presence of anther spurs, presence of coma bracts) are neither unique nor universal to all members of the genus as perceived today. Moreover, there are several small or monotypic genera closely related to *Curcuma*, e.g., *Hitchenia* Wall., *Smithatris* W.J. Kress & K. Larsen and *Stahlianthus* Kuntze, that were separated from *Curcuma* based on morphology. The first phylogenetic study of Zingiberaceae based on analyses of nrDNA sequences and the plastid *matK* region (Kress & al., 2002) confirmed the long-suspected complexity of generic concepts in Zingiberaceae. Kress & al. (2002) concluded that many genera accepted now are paraphyletic (e.g., *Alpinia* Roxb. and *Amomum* Roxb.), and suggested that *Curcuma* was paraphyletic with respect to *Hitchenia*, *Stahlianthus* and *Smithatris*, which also share cone-like inflorescences with few-flowered, congested bracts. These authors favoured splitting *Curcuma* into several smaller genera, rather than accepting a wide concept of *Curcuma* to include these associated genera. The other study which considered the phylogeny of tribe Zingibereae (Ngamriabsakul & al., 2004) used the same nuclear gene and a different chloroplast region, *trnL-trnF*. This study also confirmed that the *Curcuma*-like genera *Hitchenia*, *Paracautleya* R.M. Sm., *Smithatris*, and *Stahlianthus* are well nested in the *Curcuma* complex, and its authors argued that the characters supporting separation of these genera were likely to be autapomorphic. They suggested that these genera should be regarded as a part of a single genus (Fig. 1). Shortly afterwards, *Paracautleya* (Škorničková & Sabu, 2005a), *Kaempferia scaposa* (Leong-Škorničková & al., 2007), and *K. candida* (Techaprasan & Leong-Škorničková, 2011) were transferred to *Curcuma*. Škorničková (2007) also suggested that the monotypic genus *Laosanthus* K. Larsen & Jenjitt. (Larsen & Jenjittikul, 2001), which was not included in any previous analysis, was likely to be yet another highly specialized member of *Curcuma*. As the above-mentioned molecular studies analysed only a limited number of *Curcuma* and *Curcuma*-like representatives, the generic concepts of *Curcuma*, *Hitchenia*, *Stahlianthus*, *Smithatris* and *Laosanthus* remain to be re-examined based on molecular analyses of a broader range of samples.

There have been several attempts to define infrageneric groups in *Curcuma* (Roxburgh, 1810; Horaninow, 1862; Baker, 1890; Valeton, 1918; Velayudhan & al., 1996, 1999), with two subgenera, subg. *Curcuma* and subg. *Hitcheniopsis* (Baker) K. Schum. (Schumann, 1904) currently recognized. However, based on the current state of knowledge with broad geographical sampling of *Curcuma* species, none of the previous classifications seems to be suitable. The majority of proposed classifications suffer from being based on material that came only from the western part of the distribution range of the genus, while omitting several morphological types occurring in Indochina. Thus, numerous species, especially those from Thailand and Indochina, are unclassifiable in either of the two subgenera. Moreover, the morphological characters employed for the classifications (position of the inflorescence and presence/absence of anther spurs) are ambiguous in many species (Škorničková, 2007). In a cytological study of Indian representatives of *Curcuma* subg. *Curcuma* (Leong-Škorničková & al., 2007), extensive polyploidy was detected. This polyploidy accounts for the long-standing taxonomic problems in this group (Škorničková, 2007 and references there). The impact of polyploidy on the complexity of species relationships in other groups in *Curcuma* is not yet known, but molecular markers could provide some insights. To date, few molecular studies have focused on relationships within species groups of *Curcuma* (e.g., Cao & al., 2001; Cao & Komatsu, 2003; Syamkumar & Sasikumar, 2007; Záveská & al., 2011), and the only study covering some of the above-mentioned 'unclassifiable' species and representatives of the *Curcuma*-like genera (Sirirugsa & al., 2007) suffered from incomplete sampling. No comprehensive phylogenetic study of the widely conceived genus *Curcuma*, focusing on the delimitation of the genus and evaluating previously proposed infrageneric classifications, is available.

We chose to use the ITS marker for this study, because it had already been successfully employed for resolving phylogenetic relationships in Zingiberaceae (e.g., Pedersen, 2004; Williams & al., 2004; Kress & al., 2005). In addition, we analysed three regions of cpDNA (*trnL-trnF*, *psbA-trnH*, *matK*). As several *Curcuma* species are expected to be of polyploid/hybrid origin (e.g., Leong-Škorničková & al., 2007), the combination of a biparentally inherited rDNA region with maternally inherited cpDNA was selected for its potential to detect past cases of reticulation and hybrid speciation, as well as parentage of polyploids (Alvarez & Wendel, 2003).

The goals of the present study were (1) to examine the paraphyletic genus *Curcuma* and assess the taxonomic position of previously questioned *Curcuma*-like species often placed in *Hitchenia*, *Paracautleya*, *Stahlianthus*, *Smithatris*, *Kaempferia*, and *Laosanthus*; (2) to resolve phylogenetic relationships in *Curcuma* in its broad sense; and (3) to compare the importance of polyploidy and/or homoploid hybridization among different groups in *Curcuma*.

Fig. 1. Representatives of *Curcuma* **(A–L)** and *Curcuma*-like species currently classified as members of various genera **(M–P). A,** *C. angustifolia*; **B,** *C. longa*; **C,** *C. roscoeana*; **D,** *C. rubrobracteata*; **E,** *C. alismatifolia*; **F,** *C. gracillima*; **G,** *C. harmandii*; **H,** *C. sparganifolia*; **I,** *C. ecomata*; **J,** *C. bicolor*; **K,** *C. pierreana*; **L,** *C. singularis*; **M,** *Laosanthus graminifolius*; **N,** *Kaempferia candida*; **O,** *Stahlianthus campanulatus*; **P,** *Smithatris supraneanae*. — Photos by Otakar Šída (I), Martin van den Bult (J) and Jana Leong-Škorničková (A–H, K–P). ▶


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P

TAXON 61 (4) • August 2012: 747–763

■ MATERIALS AND METHODS

Plant material. — Sixty-two species of *Curcuma* s.l., including 12 *Curcuma*-like species previously placed in other genera (Kress & al., 2002; Ngamriabsakul & al., 2004; Leong-Škorničková, unpub. data), were used. Samples of these species were often collected at or near the type localities and covered a broad spectrum of morphological variation (Fig. 1) as well as geographical distribution (Fig. 2). Seven outgroup species, including *Pyrgophyllum yunnanense* and *Camptandra parvula*, placed basally to the *Curcuma* clade by Kress & al. (2002), and the more distantly related *Zingiber capitatum*, *Globba* sp., *Alpinia conchigera*, *Alpinia galanga* and *Larsenianthus careyanus* (Kress & al., 2010), were also included in the analyses (Appendix). All taxa were identified by Jana Leong-Škorničková, who is currently revising *Curcuma* (Škorničková & al., 2003a, b, 2004, 2007; Škorničková & Sabu, 2005a, b, c). The names of the taxa follow Škorničková (2007) and Leong-Škorničková & al. (2010).

Each collection number indicates one population of a taxon at the given locality. One individual per population was analysed, and one to three populations per species were selected for analyses to detect variation within species.

DNA extraction, PCR and sequencing. — Total genomic DNA was extracted from 0.5 g dried leaf tissue using the DNeasy Plant Mini Kit (Qiagen, Valencia, California, U.S.A.).

The internal transcribed spacers (ITS1, ITS2) and 5.8S of the rDNA repeat array were amplified using forward and reverse primers ITS 5 and ITS 4 (White & al., 1990). PCR amplifications were performed in a total volume of 20 µl. Each reaction contained 1× concentrated PCR buffer (Sigma Aldrich, St. Louis, Missouri, U.S.A.), 0.5 U REDTaq JumpStart DNA Polymerase (Sigma Aldrich), 0.2 mM dNTP, 6.25 pmol each of forward and reverse primer (Sigma Aldrich), and 5 ng of genomic DNA.

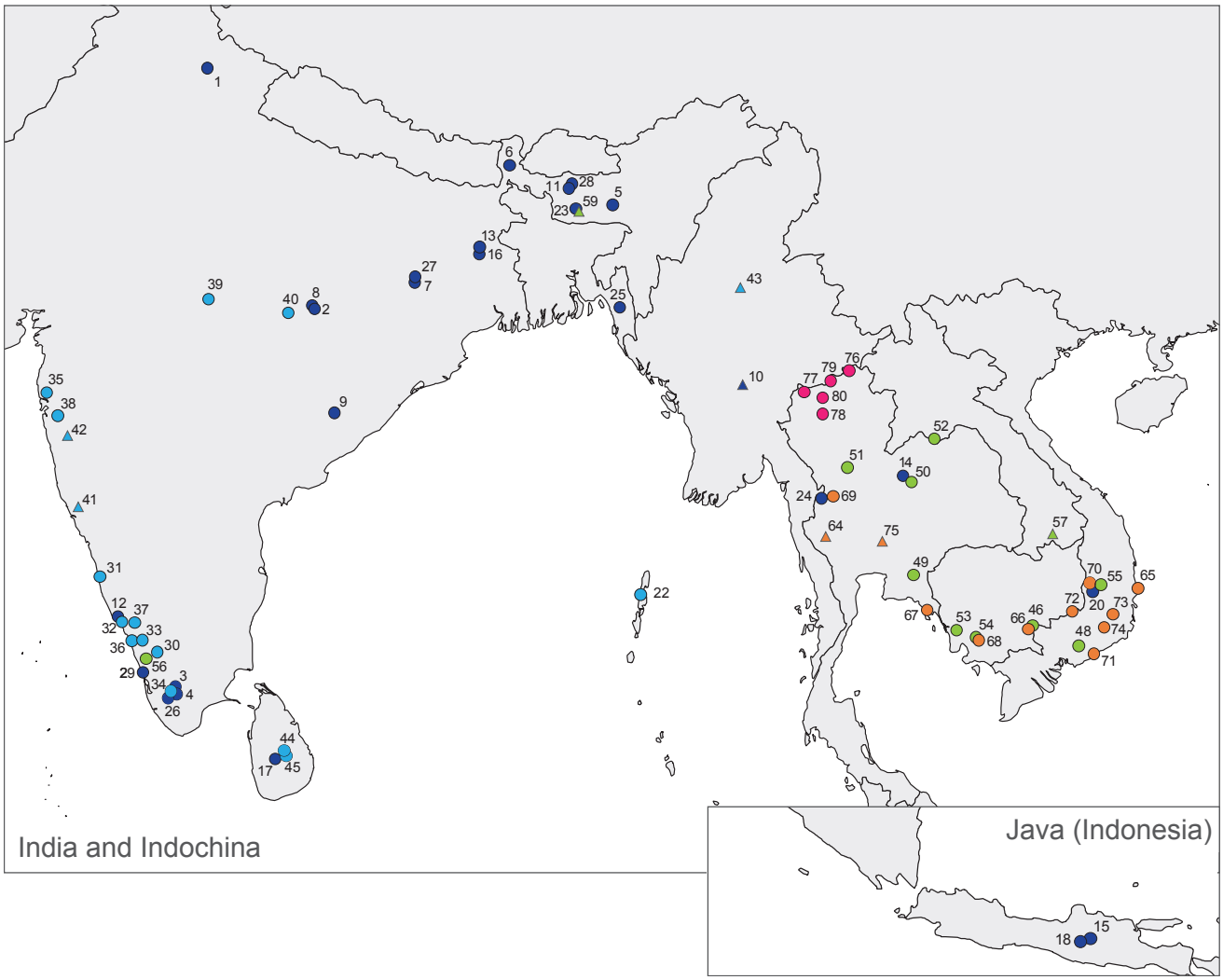


Fig. 2. Geographical distribution of the populations studied. The numbering of the dots (*Curcuma* species) and triangles (*Curcuma*-like species currently classified as members of various genera) corresponds to the population numbers in the Appendix. The colours of the dots and triangles indicate the affiliation of individuals to the main evolutionary lineages detected by sequence analyses of nrDNA and cpDNA (magenta, 'Eco-mata' group; orange, 'Pierreana' group; green, 'Hitcheniopsis' group; dark and light blue, 'Curcuma-I' and 'Curcuma-II' groups, respectively). Origin of following specimens is known only at country level and these were therefore not mapped: 19, 21, 47, 58, 60, 61, 62 and 63.

The chloroplast *trnL-trnF* region was amplified using the trnL-c and trnF-f primers (Taberlet & al., 1991); the *psbA-trnH* region was amplified using the psbA and trnH primers (Sang & al., 1997; Tate & Simpson, 2003); and the *matK* region was amplified using the trnK1F (Manos & Steele, 1997) and trnK2R (Steele & Vilgalys, 1994) primers. PCR amplifications were set up in the same way as for the ITS region with the exception of the reaction mix for the *matK* region, in which an additional 2.5 mM of Mg^{2+} was included. All reactions were run on an Eppendorf Mastercycler Gradient with initial denaturation at 94°C for 1 min followed by 35 cycles of 94°C for 60 s, with annealing temperatures of 50, 55, 55 and 63°C for ITS, *trnL-trnF*, *psbA-trnH* and *matK*, respectively, for 60 s, and 72°C for 60 s. The cycling ended with incubation at 72°C for 10 min. The sequencing primers included the amplification primers and the Zingiberaceae-specific internal *matK* primers mIF and mIR (Kress & al., 2002).

PCR products were purified with the Jetquick kit (GENOMED, Löhne, Germany) and directly sequenced at Macrogen Inc. (http://www.macrogen.com) or Seqlab of the Biological Section of Charles University with the original PCR primer sets in both directions.

Cloning. — Due to the occurrence of within-individual polymorphisms in most of the directly sequenced PCR products of the ITS region, all PCR products for ITS were cloned using the pGEM-T Easy Vector System (Promega, Madison, Wisconsin, U.S.A.) following the manufacturer's instructions but downscaled to half-volume reactions. After overnight culture at 37°C on LB ampicillin/IPTG/X-gal selective plates (prepared according to the manufacturer's instructions), colonies carrying the insert were identified by colour. Five to twenty-three colonies from each individual were used as templates for PCR using universal primers for the pGEM-T Easy vector (pUC/M13Forward and Reverse) and the same conditions as the initial PCR. Purified PCR products were sequenced in one direction using the universal primer pUC/M13Forward.

Data analyses. — The sequences were aligned using MAFFT v.6 (http://mafft.cbrc.jp/alignment/server/), and then the alignment was improved manually in BioEdit v.7.0.0 (Hall, 2004).

For the primary ITS dataset, a large number of sequences, including chimeric ones, were accumulated during the cloning procedure. In vitro recombination of DNA sequences may cause problems when cloning PCR products in which multiple alleles or paralogous gene copies have been amplified (Cronn & al., 2002; Anthony & al., 2007; Kelly & al., 2010). Therefore, several steps were performed to analyse the sequences in detail. First, an alignment of all cloned sequences that originated from the same individual ("intra-individual alignments") was constructed and inspected for the presence of duplicated sequences (identical sequences present in more than one copy in the alignment). Duplicates were subsequently deleted from the dataset. Autapomorphies found in only one clone in the entire intra-individual alignment were considered as polymerase errors and corrected (Popp & al., 2005). The numbers of polymorphic sites within these alignments were counted to describe the degree of intra-individual ITS polymorphism. Second, a NeighbourNet (Bryant & Moulton, 2004) network of the clones from each intra-individual alignment was generated to visualize reticulate signals in the data and to suggest potential recombinant sequences (Huber & Moulton, 2005) and plausible parental ones. The networks were constructed based on pairwise (uncorrected) distance matrices in SplitsTree4 v.4.9.1 (Huson & Bryant, 2006). The NeighbourNet network was used as a guide to screen the submatrix of cloned sequences by eye in BioEdit for evidence of recombination. Recombinant sequences were identified as those that combined characteristic mutations (single nucleotide polymorphisms or indels) from each of the relevant parental sequences (Russell & al., 2010; Salmon & al., 2010). We tried to apply recombination detection programs included in the RDP3 v.3.44 package (Martin & al., 2005). However, these failed, as they were usually unable to detect chimeric sequences that were obvious to the eye, similarly to the outcomes reported by Anthony & al. (2007) and Russell & al. (2010). To be effective, such formalized recombination detection approaches usually require higher levels of sequence divergence (Posada & Crandall, 2001) and longer sequence length than we had in our dataset (5–40 sequences within the submatrix with a length of about 700 bp from each cloned sample). Detected recombinants were finally discarded from the dataset because they behave erratically in phylogenetic analyses (Popp & Oxelman, 2001; Popp & al., 2005; Grimm & Denk, 2008). To reduce the size of the final matrix, a maximum parsimony (MP) analysis (see below) based on all clones from the previously inspected alignments was performed and majority-rule consensus sequences were constructed for clones that were from the same accession and formed monophyletic groups. The number of different clones and/or consensus sequences originating from a single individual were counted and then used for subsequent phylogenetic analyses and represented as multiple terminals in the final multi-labeled phylogenetic tree (Popp & al., 2005).

Phylogenetic analyses. — Two frequently used methods for sequence data analysis—maximum parsimony and Bayesian analyses—were used for phylogeny inference of the ITS and cpDNA datasets. For MP analysis, indels in a particular alignment (ITS, cpDNA) were coded using the simple indel coding method (Simmons & Ochoterena, 2000) with GapCoder (Young & Healy, 2003). In this analysis, a heuristic search was done, with 100 random sequence additions and no more than 100 trees of length greater than or equal to 1 saved per replicate, tree bisection-reconnection (TBR) branch swapping, and the MulTrees option in effect, using PAUP* v.4.0b10 (Swofford, 2002). The consistency index (CI) (Kluge & Farris, 1969) and retention index (RI) (Farris, 1989) were calculated to estimate levels of homoplasy. Support for the nodes was evaluated with bootstrap analyses (Felsenstein, 1985). A heuristic search was done, with the MulTrees option off, on 1000 bootstrap replicates with random addition of sequences (100 replicates for each bootstrap replicate) followed by TBR branch swapping. Bootstrap support was categorized according to the criteria of Kress & al. (2002): strong (>85%), moderate (70%–85%), weak (50%–69%), or poor (<50%).

Bayesian analysis was accomplished with MrBayes v.3.1.2. (Huelsenbeck & Ronquist, 2001) performed on Bioportal

(Kumar & al., 2009). Optimal analysis settings were chosen based on initial hierarchical likelihood ratio tests (hLRT) and the approximate Akaike information criterion (AIC) calculated by MrModeltest v.2.3 (Nylander, 2004). Two parallel runs with four chains each were used, sampling every 1000th tree for 10 and 20 million generations for the cpDNA and ITS datasets, respectively. For the ITS data, the first 25% of samples (= 5000 trees) were discarded as burn-in, and the remaining 15,000 trees per run were summarized. For the cpDNA data, the first 25% of samples (= 2500 trees) were discarded as burn-in, and the remaining 7500 trees per run were summarized. Nodes with posterior probability (PP) values of 0.95 and above were regarded as significant and those with PP values below 0.95 regarded as non-significant.

For presentation of species relationships, an arbitrarily chosen phylogram of one of the most parsimonious trees was used. MP bootstrap support (BS) values higher than 50% and PP values above 0.9 were mapped onto the corresponding branches of the most parsimonious tree.

■ RESULTS

Phylogenetic analyses. — In the MP analysis, in general, the level of homoplasy was lower for the cpDNA dataset (CI excluding uninformative characters = 0.530) than for the ITS dataset (CI = 0.364) (Table 1). Both the ITS and cpDNA datasets were analysed with the Bayesian approach, using settings for the suggested best-fitting model GTR+I+G (posterior probabilities presented in Fig. 3A, B). In general, Bayesian majority-rule consensus trees differed only slightly in topology from the strict consensus trees yielded by the MP analyses. However, all well-supported clades in the MP analyses (of the ITS as well as cpDNA datasets) comprised the same representatives as the groups that were well supported in the Bayesian analysis.

ITS phylogeny. — Both the MP strict consensus tree and the Bayesian majority-rule tree resolved similar structures of basic relationships differing mainly in the BS and PP values for particular groups (Fig. 3A). An ingroup that was strongly supported in relation to the outgroup comprised all *Curcuma* and *Curcuma*-like species (100% BS, 1.00 PP). All species currently classified as members of the genera *Hitchenia*, *Smithatris*, *Stahlianthus* and *Laosanthus* as well as a few species from other genera recently transferred to *Curcuma* (former *Paracautleya bhatii*, *Kaempferia candida* and *K. scaposa*) were resolved close to particular ingroup species of *Curcuma*.

Based on both analyses, the ingroup taxa were further subdivided into four strongly or moderately supported main groups—'Ecomata', 'Pierreana', 'Hitcheniopsis' and 'Curcuma' (Fig. 3A; the 'Curcuma' group is further subdivided into the subgroups 'Curcuma-I' and 'Curcuma-II')—plus *C. candida* and *C. vamana* (indicated by a broken line in Fig. 3A). These two species were only poorly supported as stem lineages of the 'Pierreana' and 'Hitcheniopsis' groups, respectively. The 'Ecomata' and 'Pierreana' groups were resolved as sister lineages in some of the most parsimonious trees, whereas they appeared in a basal polytomy in the MP strict consensus tree and the 50% majority rule consensus tree from the Bayesian analysis, together with *C. candida*. The 'Hitcheniopsis' and 'Curcuma' groups were weakly supported as sister lineages in MP (56% BS), but had significant support in the Bayesian analyses (0.99 PP).

The most basally placed group, 'Ecomata' (98% BS, 1.00 PP), was represented by four species—*C. ecomata*, *C. bicolor*, *C. glans* and *C. flaviflora*—most of which are native to southern China and northern Thailand. The second group at the base of the tree, 'Pierreana' (81% BS, 0.89 PP), is formed by *C. singularis*, *C. pambrosima*, *C. vitellina*, *C. rhomba*, *C. pierreana*, and four new species that are not yet formally described (*C.* sp. nov. – 73334; *C.* sp. nov. 'newmanii' – JLS 365, *C.* sp. nov. 'xanthella' – Ly 348 and *C.* aff. *cochinchinensis* – JLS213; Leong-Škorničková & al., in prep.). These species are morphologically similar to species of the 'Ecomata' group and are distributed mainly in southern Indochina, from Thailand to Vietnam. One of the two species of *Smithatris* (*S. supraneanae*) was also placed in this group. Relationships among particular

Table 1. Summary of the ITS and cpDNA data matrices and the maximum parsimony/Bayesian analyses of the two datasets.

	ITS	cpDNA (*matK*+*trnL-trnF*+*psbA-trnH*)
No. of terminals/no. of individuals	146/86	69/69
Aligned positions (nucleotides + *gaps*)	828 (738+*90*)	4669 (2653+968+933+*113*)
Maximum parsimony analyses		
No. of parsimony-informative characters (including gaps)	231	114
Number/lengths of MP trees	100/1021	1700/625
CI (excluding uninformative characters)/RI	0.364/0.815	0.530/0.836
Bayesian analyses		
Best-fitting model (AIC/hLRT)	(GTR+I+G)/(GTR+I+G)	(GTR+I+G)/(GTR+I+G)
No. of generations	20,000,000	10,000,000
No. of Bayesian trees after burn-in (used for computation of consensus tree)	15,000	7500

species within this group were only poorly resolved, with the exception of a strongly supported (94% BS, 1.00 PP) branch with three species from south and central Vietnam consisting of *C.* sp. nov. 'newmanii', *C. pierreana* and *C. rhomba*.

Curcuma candida, previously placed in *Kaempferia* (Techaprsan & Leong-Škorničková, 2011), was resolved as a stem lineage of the 'Pierreana' group in some of the most parsimonious trees (Fig. 3A). As support for this relationship was obtained neither in the MP bootstrap nor the Bayesian analysis, this species should be regarded as a lineage of the basal polytomy.

The relationship of the 'Hitcheniopsis' group with its stem lineage, *C. vamana*, was poorly resolved (<50% BS, 0.79 PP). The 'Hitcheniopsis' group (100% BS, 1.00 PP) includes seven species originally regarded as members of *Curcuma* (*C. thorelii*, *C. gracillima*, *C. alismatifolia*, *C. parviflora* agg., *C. harmandii*, *C. rhabdota*, *C. sparganifolia*), along with the monotypic genus *Laosanthus* (*L. graminifolius*), and five accessions of *Stahlianthus*. Four accessions of *Stahlianthus* formed a strongly supported branch (99% BS, 1.00 PP) in the terminal part of the 'Hitcheniopsis' group, while the fifth (*S. involucratus*) was resolved as sister lineage to other *Stahlianthus* accessions together with *C. thorelii*. A different hierarchical structure of species relationships within this group was indicated only in the MP strict consensus tree, but particular branches were not well supported by BS or PP.

The 'Curcuma' group, represented by *C. longa*, *C. aeruginosa*, and *C. zedoaria* (see Appendix), was resolved in a sister position to the 'Hitcheniopsis' group (89% BS, 1.00 PP). The stem position in this group is held by the second species of *Smithatris* (*S. myanmarensis*). Besides *S. myanmarensis*, three species traditionally classified as members of the *Curcuma*-like genera *Hitchenia* (*H. glauca*, *H. caulina*) and *Kaempferia* (*K. scaposa*) were nested in this group. In the 'Curcuma' group, the terminal branch (marked as the 'Curcuma-I' group in Fig. 3A) was not significantly supported in the Bayesian tree (0.82 PP) and was present but not supported in the MP strict consensus tree. The remaining species of this group were resolved in the basal polytomy (marked as 'Curcuma-II' group in Fig. 3A) in both the MP and Bayesian consensus trees.

cpDNA phylogeny. — The MP tree based on concatenated *trnL-trnF*, *psbA-trnH* and *matK* sequence data provided strong support for the monophyly of a broadly conceived *Curcuma*, encompassing *Curcuma* and *Curcuma*-like genera and species (100% BS). The support of the same branch was not significant in Bayesian analysis (0.90 PP). Four weakly to strongly supported main groups corresponding to those resolved by the ITS phylogeny were detected in the cpDNA phylogeny ('Ecomata': 98% BS, 1.00 PP; 'Pierreana': 74% BS, 1.00 PP; 'Hitcheniopsis': 88% BS, 1.00 PP; 'Curcuma': 61% BS, 0.82 PP in Fig. 3B). In general, these groups contained the same species in the cpDNA and ITS phylogenies and ITS. The only exceptions to this are *C. candida* and *C. vamana* that were placed outside the main groups in the ITS phylogeny but were strongly supported as representatives of the 'Curcuma' group by the cpDNA data.

In contrast to the ITS phylogeny, the 'Ecomata', 'Pierreana' and 'Hitcheniopsis' groups were resolved as a clade in the MP and Bayesian analyses, but without support for this relationship. As in the ITS data, no clear structure of species relationships within particular groups was detected.

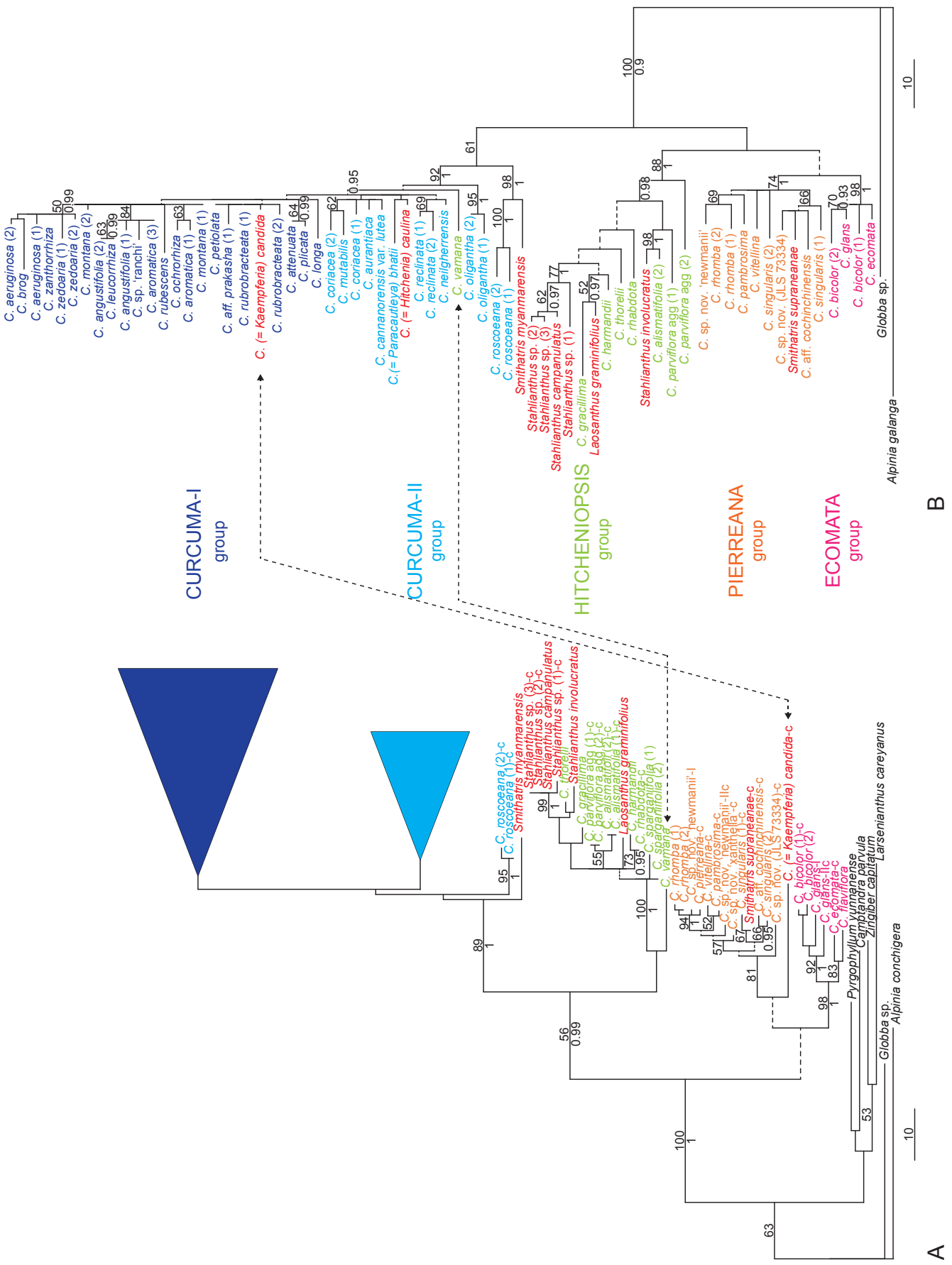
The 'Curcuma' group split into the stem lineage represented by *C. roscoeana* and *Smithatris myanmarensis* (98% BS, 1.00 PP) and the rest of the group (92% BS, 1.00 PP), at the base of which *Curcuma oligantha* and *C. vamana* were separated. Relationships within the 'Curcuma' group were not satisfactorily resolved by either the Bayesian or the MP bootstrap analyses. In some of the most parsimonious trees from the cpDNA analysis, the terminal group of species corresponded to the 'Curcuma-I' group of the ITS phylogeny (marked as 'Curcuma-I' group in Fig. 3B), with the species in the basal polytomy corresponding to the 'Curcuma-II' group of the ITS phylogeny (marked as 'Curcuma-I' group in Fig. 3B), but this arrangement was not supported.

Major incongruencies between cpDNA and nrDNA phylogenies. — The first incongruence between the ITS and cpDNA phylogenies lies in the topology of the four main groups of the genus. The nrDNA data placed the 'Ecomata' and 'Pierreana' groups in the basal polytomy together with *C. candida* (in the MP strict consensus and Bayesian trees), while the 'Hitcheniopsis' and 'Curcuma' groups, sister to each other, formed the terminal part of the phylogeny. However, based on the cpDNA data, 'Ecomata', 'Pierreana' and 'Hitcheniopsis' form a monophyletic group, with the 'Curcuma' group resolved as sister to this assembly.

The second apparent contradiction between the chloroplast and nuclear markers involves the positions of *C. candida* and *C. vamana*. According to the ITS data (and in concert with overall morphology), *C. candida* was placed close to species from the 'Pierreana' group, and *C. vamana* was placed in a stem position relative to the 'Hitcheniopsis' group. In contrast, based on the cpDNA data, both species were strongly supported as representatives of the 'Curcuma' group.

The third incongruence is in species relationships within the main groups. In general, relationships within the main groups were not well supported in either the ITS or cpDNA phylogenies, resulting in large polytomies. Moreover, even when relationships of particular clades of species were supported by the ITS data, the same relationships were neither supported by nor generally observable in the cpDNA dataset.

ITS intra-individual polymorphism. — Because direct sequencing of the ITS region led to unreadable sequences for most individuals, cloning was employed to uncover ITS polymorphism in these individuals. As a result, a broad range of intra-individual ITS variation was detected in the ITS dataset, ranging from 0 to 87 polymorphic sites within intra-individual alignments. To describe this variation across the genus, we defined four classes of polymorphism according to the number of polymorphic sites in intra-individual alignments. The species with low levels of intra-individual polymorphism were categorized as class I or II (0–5 and 6–15 polymorphic sites, respectively), while the species with high levels of intra-individual polymorphism were designated as class III or IV (16–25 and ≥26 polymorphic sites, respectively; Fig. 4; Appendix).



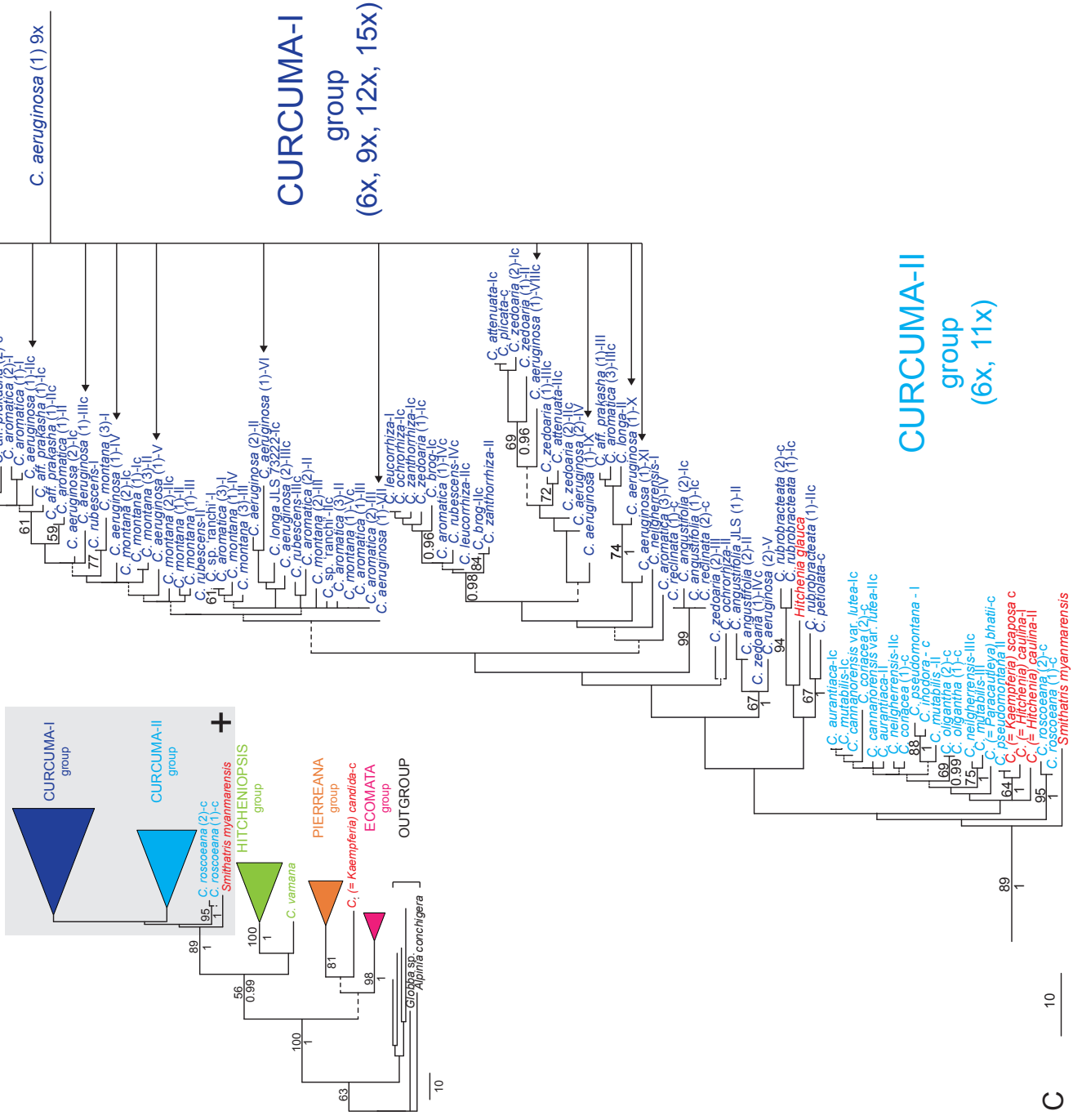


Fig. 3. MP trees based on ITS and cpDNA sequences. **A,** an arbitrarily chosen phylogram of one of the 100 most parsimonious trees based on ITS sequence data. Multiple terminals from a single individual are indicated by Roman numerals, 'c' indicates a majority-rule consensus sequence (see the Methods section for details); **B,** an arbitrarily chosen phylogram of one of the 1700 most parsimonious trees based on cpDNA sequence data (the concatenated regions *trnL-trnF*, *psbA-trnH* and *matK*); **C,** detail of the ITS phylogram (A) showing the terminal part of the phylogeny. Ploidy levels in the 'Curcuma-I' and 'Curcuma-II' groups are indicated. The exemplar species *C. aeruginosa* (9*x*) with the highest observed within-individual ITS polymorphism is indicated by arrows. The four groups are indicated by colours (magenta, 'Ecomata'; orange, 'Pierreana'; green, 'Hitcheniopsis'; dark and light blue, 'Curcuma-I' and 'Curcuma-II', respectively). Curcuma-like species are highlighted in red. The differing positions of two species in the ITS and cpDNA phylograms are indicated by broken lines. Bootstrap values (BS) above 50% and Bayesian posterior probabilities (PP) values above 0.9 are indicated above and below branches, respectively. Dotted branches collapse in the MP strict consensus tree.

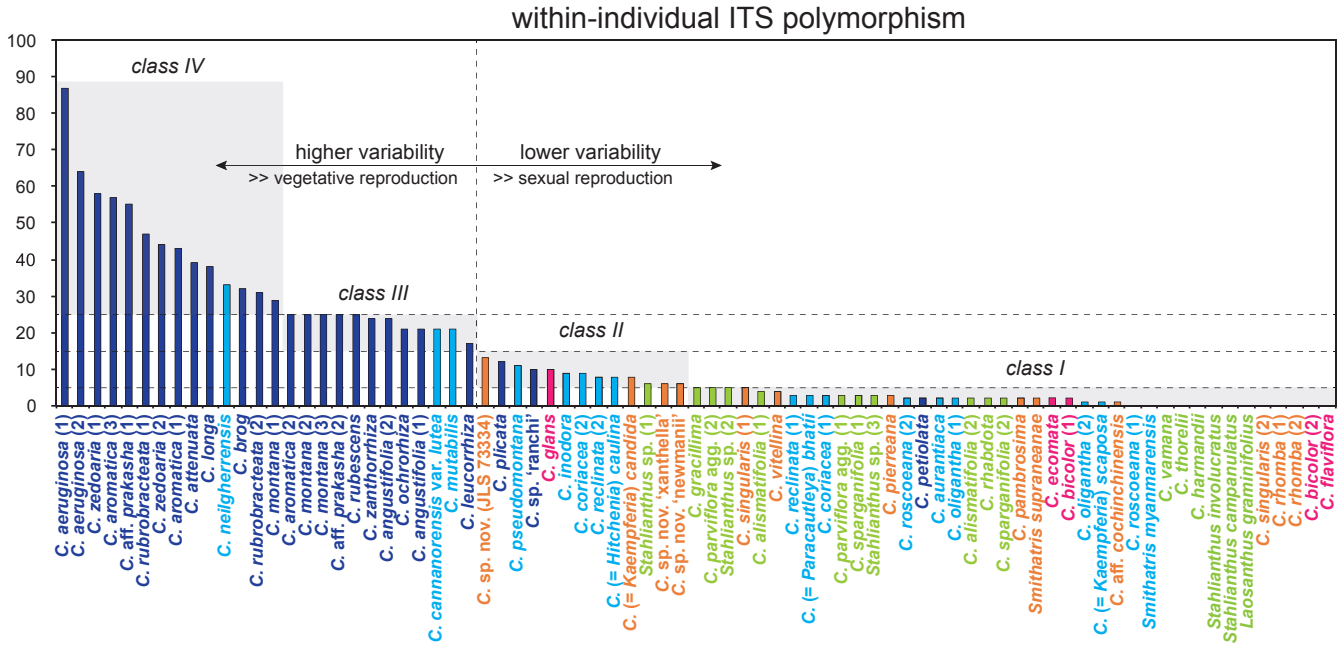


Fig. 4. Description of within-individual ITS sequence variation across species and individuals of *Curcuma* s.l. A single bar represents a single individual in the ITS sequence analysis (horizontal axis). The height of the bar corresponds to the number of polymorphic sites in intra-individual alignments (vertical axis). Colours of bars correspond to colours in the phylograms (Fig. 3) and indicate the affiliation of individuals to the main evolutionary lineages detected by sequence analyses. Four classes of polymorphism are depicted (see Results section for details). The link between degree of ITS sequence variation and mode of reproduction of species is highlighted (see Discussion for details).

Classes I and II include all species from the 'Ecomata', 'Pierreana' and 'Hitcheniopsis' (including *C. candida* and *C. vamana*) groups, together with several species belonging to the 'Curcuma' group (mainly species from the 'Curcuma-II' group in Fig. 3). The majority of these species are represented by single terminals in the final phylogenetic tree, implying that all different ITS sequences from a single individual were more closely related to each other than to any sequence from different individuals, and that they formed monophyletic groups in preliminary phylogenetic analyses. The only exceptions were two species belonging to the 'Curcuma' group (*C.* sp. 'ranchi', 12*x*; *C. pseudomontana*, 6*x*), one species belonging to the 'Pierreana' group (*C.* sp. nov. 'newmanii'), and one species belonging to the 'Ecomata' group (*C. glans*), each of which had two terminals in the final phylogenetic tree (Fig. 3A; Appendix).

Classes III and IV include only species from the 'Curcuma' group, where ploidy levels range from hexaploidy to pentadecaploidy (Leong-Škorničková & al., 2007). Class III comprises mostly hexaploids that predominantly reproduce sexually (with some of them being facultatively asexual) as well as nonaploids, which reproduce vegetatively. None of the species in this class has more than four terminals in the final phylogenetic tree. Class IV comprises mostly nonaploids and pentadecaploids, which are obligate asexual species, as well as all remaining hexaploids, which are facultatively sexual. Representatives of this class have the highest numbers of terminals in the final phylogenetic tree—up to eleven terminals in the nonaploid *C. aeruginosa* (Fig 3A, Appendix 1).

■ DISCUSSION

Incongruence between nrDNA and cpDNA phylogenies. — Concatenated sequences from three cpDNA regions resulted in a phylogeny with poorly resolved relationships among the major groups. In case of the ITS data, the main groups are well supported, but relationships among the basal groups ('Ecomata', 'Pierreana'), relationships of *C. vamana* and *C. candida* to the 'Hitcheniopsis' and 'Pierreana' groups, respectively, and relationships within the main groups are only poorly supported and particularly incongruent with the topology resulting from the cpDNA data. In recent studies, conflicting phylogenetic signals between datasets are often explained either by lineage sorting or by ancient hybridization (e.g., Rieseberg & Wendel, 1993; Rieseberg & al., 1996; Okuyama & al., 2005). To identify the precise cause of such incongruencies, information from more nuclear gene regions and the approximate times of divergence among particular groups are needed (e.g., Sang & Zhong, 2000).

Based on the cpDNA data, the main groups 'Ecomata', 'Pierreana' and 'Hitcheniopsis' form a monophyletic lineage, and the 'Curcuma' group another (in both MP and Bayesian consensus trees). In contrast to this, the 'Hitcheniopsis' and 'Curcuma' groups are sister in the ITS phylogeny, with unresolved relationships of both of these to 'Ecomata' and 'Pierreana'. Support for backbone relationships was very weak in both the cpDNA and ITS phylogenies. Thus, with this state of knowledge, we regard relationships among the main lineages as

unresolved, and the incongruencies among nuclear and plastid data as non-significant.

The observed topology of the cpDNA tree, despite its low support, reflects the geographical distribution of the main groups (Fig. 2). Representatives of the 'Ecomata' and 'Pierreana' groups are vicariant, the former being restricted to the northern part of the Indochinese floristic region (sensu Takhtajan, 1986), from N Thailand to S China, whereas the latter occurs in the southern and central parts of this region. The 'Hitcheniopsis' group occurs in a slightly larger geographic area extending from the Indochinese region to Myanmar and NE India, with one isolated occurrence in S India. In contrast, whereas the seed-setting representatives of the 'Curcuma' group are most diverse in the area from India to western Thailand, the vegetatively reproducing members (polyploids) of this group are distributed (but not necessarily native) across the whole range of the genus, from India to Australia and New Guinea.

The different positions of two species, *Curcuma vamana* and *C. candida*, in the ITS and cpDNA phylogenies are well supported, and therefore the lack of sequence variability can be eliminated as a plausible reason for their incongruent relationships (Wolfe & Randle, 2004). Hybridogenic origins of these species from distantly related parental species (generally from different main groups, 'Hitcheniopsis' and 'Curcuma' in case of *C. vamana* and 'Pierreana' and 'Curcuma' in case of *C. candida*) is a plausible explanation for the conflicts. The geographical distributions and, to varying extents, the morphological characters of the species might also indicate their hybrid origin.

Curcuma vamana, which is stenoendemic to the southernmost part of the Indian subcontinent, is isolated from the main 'Hitcheniopsis' distribution area by more than 2000 km. This phenomenon could have been caused by long-distance dispersal or by the reduction of an originally wider distribution range of the 'Hitcheniopsis' group. Subsequent hybridization or introgression with representatives from the 'Curcuma' group could have caused the capture of the plastid genome (e.g., Rieseberg & al., 1996; Okuyama & al., 2005) by *C. vamana*, and homogenization of ITS sequences by concerted evolution might have led to its basal position in the ITS phylogeny. Surprisingly, this species has typical morphological characters of the 'Hitcheniopsis' group (lack of anther spurs and epigynous glands) and there are almost no morphological characters indicating a relationship to the 'Curcuma' group (the only exceptions being creeping branches of the rhizome and bright yellow flowers).

Curcuma candida is geographically restricted to the area along the border of Thailand and Myanmar (Jenjittikul & Larsen, 2000), where hybridization events between a species from the 'Pierreana' group, as pollen donor, and a species from the 'Curcuma' group, as maternal genome donor, could have occurred in the past. The inflorescence of *C. candida* resembles that of other members of the 'Pierreana' group, whereas its leafy shoot is like that of typical members of the 'Curcuma' group. The flowers of *C. candida* are, however, the largest in the genus and have a fairly atypical anther shape, likely adaptations to a highly specialized pollinator, which could obscure possible phylogenetic relationships.

Hybrids between distantly related taxa can seriously distort an otherwise hierarchical tree structure (e.g., by McDade, 1992). If hybridization is responsible for the conflicting positions of these species, concerted evolution probably has had sufficient time to homogenize the ITS sequences to the paternal type (assuming maternal inheritance of plastids), as neither species exhibits any extensive intra-individual ITS polymorphism. Lineage sorting, i.e., random extinction of ancestral alleles or paralogues, is an alternative explanation, but the possible influences of these two factors cannot be distinguished from each other on the basis of topology alone (e.g., Sang & Zhong, 2000). Further information about diversification times, chromosome numbers, genome sizes and ploidy levels for representatives of the 'Hitcheniopsis' and 'Pierreana' groups is needed to test hypotheses about the hybrid origin of these species or the impact of lineage sorting.

Finally, a number of incongruencies between the ITS and cpDNA phylogenies can be observed in relationships among species in major groups. The most notable example is the terminal 'Curcuma' group, which includes a number of highly polyploid species (such as *C. aeruginosa*) with single individuals that yield multiple terminals in the ITS tree but single terminals in the cpDNA tree. The combination of hybridization and the multiple-copy character of the ITS region means that several ITS repeat types can be maintained in the genome and/or further undergo complete or incomplete homogenization (e.g., Soltis & Soltis, 1991; Soltis & al., 1995; Alvarez & Wendel, 2003). These processes separately and/or collectively result in a network of paralogous sequence relationships potentially confounding accurate phylogenetic reconstruction based on ITS data. The high degree of homoplasy in the ITS dataset, exemplified by the low CI values, might be evidence for reticulate evolution caused by hybridization and/or polyploidization. Moreover, lineage sorting might operate on an abundance of ITS alleles, resulting in a pattern that differs from the maternally inherited cpDNA haplotypes. Due to the extreme complexity of reticulate evolution of *Curcuma* (Leong-Škorničková & al., 2007; Záveská & al., 2011), this study focused mainly on resolving a broad pattern that would make possible an infrageneric classification of *Curcuma*. Relationships within the main groups should be further refined by more extensive sampling and the addition of more variable single-copy nuclear markers (Záveská & al., in prep.).

Delimitation of *Curcuma*. — Our study presents a phylogeny of the 'Curcuma clade' in the sense of Kress & al. (2002), based on extensive and geographically broad sampling of species of *Curcuma* and *Curcuma*-like genera that were previously suggested by Kress & al. (2002) to be paraphyletic and by Ngamriabsakul & al. (2004) to be well nested in the *Curcuma* complex. Four gene regions were used for phylogeny reconstruction: the chloroplast *trnL-trnF*, *psbA-trnH*, and *matK* regions as well as the internal transcribed spacer (ITS) region of the nuclear ribosomal DNA. Our results confirm that all *Curcuma*-like genera are nested in the *Curcuma* ingroup. None of the *Curcuma*-like genera, as originally defined, forms a monophyletic group. All of the *Curcuma*-like species have highly derived autapomorphic flower morphologies that misled

previous authors into classifying them in separate genera. These distinctive morphological deviations are probably the result of specific ecological preferences and specific pollination systems. For instance, *Hitchenia caulina*, *H. glauca*, *Kaempferia scaposa* and *Smithatris supraneanae* are night-flowering species that are pollinated by moths, in contrast to the majority of *Curcuma* species, which are bee-pollinated (Valeton, 1918; Leong-Škorničková, pers. obs.). Data on the pollination biology of *Laosanthus graminifolius*, *Smithatris myanmarensis*, and the species of *Stahlianthus* are not available. It is expected that these data will shed more light on their unusual floral morphologies. Morphological arguments for the formal inclusion of *Kaempferia scaposa*, *K. candida*, and *Paracautleya bhatii* in *Curcuma* have been discussed in previous studies (Škorničková & Sabu, 2005a; Leong-Škorničková & al., 2007; Techaprasan & Leong-Škorničková, 2011).

Our molecular results confirm the suggestion by Ngamriabsakul & al. (2004) that representatives of the *Curcuma*-like genera are highly derived *Curcuma* species, and formal transfer of these species into *Curcuma*, which should be recognized in a broad sense (*Curcuma* s.l.), will be proposed (Leong-Škorničková & al., in prep.). The broadly defined genus *Curcuma* s.l. (corresponding in general to the 'Curcuma clade' in the sense of Kress & al., 2002) is well defined in molecular terms. In tribe Zingibereae, the most closely related genus, *Camptandra* Ridl., is placed as sister to *Curcuma* s.l. with poor BS support (Fig. 3A, but see also Kress & al., 2002; Ngamriabsakul & al., 2004). Morphological delimitation of *Curcuma* s.l. in tribe Zingibereae is far more challenging because this large genus contains numerous highly specialized and morphologically derived species. No character has been found that is both exclusive to *Curcuma* and present in all its species. However, there is a set of characters that is common to the vast majority of species of *Curcuma* s.l., and includes basally connate reflexed bracts, well-developed lateral staminodes, versatile anthers, and a cincinnus of two or more flowers.

Implications for the infrageneric classification of *Curcuma* s.l. — Previous infrageneric classifications of *Curcuma* were based exclusively on morphological characters that are ambiguous in many species, e.g., position of the inflorescence (Škorničková, 2007). Moreover, all previous studies suffered from limited geographical and morphological sampling that overlooked the very species-rich Indochinese floristic region (Roxburgh, 1810; Baker, 1890; Valeton, 1918; Velayudhan, 1996, 1999). The only molecular study focused on phylogenetic relationships in *Curcuma* and involving wider species sampling was based on ITS data alone (Sirirugsa & al., 2007). That study recognized four well-separated groups based on the molecular data—'*C. ecomata*', '*C. cochinchinensis*', '*C. alismatifolia*' and '*C. longa*', which in general correspond to the 'Ecomata', 'Pierreana', 'Hitcheniopsis' and 'Curcuma' groups, respectively, detected in our study. Although Sirirugsa & al. (2007) found four groups based on molecular data, they distinguished up to five groups of species based on morphology. That study hinted at hidden genetic as well as morphological variation of *Curcuma* species, but included only a limited number of species, and focused mainly on species found in Thailand. Additionally, Sirirugsa & al. (2007) did not suggest an official infrageneric classification of *Curcuma*.

The present study sampled more than 50% of the species of *Curcuma*, across its geographical range, and included representatives from the subgenera previously recognized by Schumann (1904) as well as species previously unknown and/or neglected in classifications. The results presented here support continued recognition of two currently recognized subgenera, *Curcuma* subg. *Curcuma* and *C.* subg. *Hitcheniopsis* (Schumann, 1904), but also suggest recognition of a new subgenus—*Curcuma* subg. *Ecomata* (formally described below), comprising previously unknown or neglected species of similar morphology and geographical distribution. In fact, our molecular data support four groups in *Curcuma*, but because only three groups are unambiguously distinguishable morphologically, we suggest recognition of only three subgenera. Our proposal of a new subgenus is in agreement with the recommendation of Sirirugsa & al. (2007), to establish at least one more subgenus to accommodate a distinct group of species native to the Indochinese floristic region. Below, we present the rationale for recognizing three subgenera in *Curcuma*. Detailed lists of species belonging to these subgenera, as well as their circumscription, however, are currently subject of further research.

Curcuma subg. ***Ecomata*** Škorničk. & Šída f., **subg. nov.** – Type: *Curcuma ecomata* Craib

Curcuma species with epigynous glands, anther spurs, mostly acute fertile bracts connate only at the base and lacking a conspicuous coma of sterile bracts, leaves with well-developed ligules. $2n$ = mostly 42. Indochinese floristic region from E Myanmar and S China to Cambodia and S Vietnam.

***Curcuma* subg. *Ecomata*: accommodating the 'Ecomata' and 'Pierreana' groups. —** The 'Ecomata' and 'Pierreana' groups comprise species from the two most basally placed lineages in both the ITS and cpDNA phylogenies. Their representatives have never been placed in published infrageneric classifications. Morphological synapomorphies uniting the two groups include the lack of a coma, predominantly more-or-less open flowers, the presence of epigynous glands, and the presence of variably shaped anther spurs. Species belonging to these groups are rarely found beyond the borders of Thailand and Indochina, although a few have been detected in southern China. Several new species belonging to these groups have recently been described from these areas (Leong-Škorničková & Ly, 2010; Leong-Škorničková & al., 2010), while descriptions of others are in progress (Leong-Škorničková & al., in prep.). Additional exploration of neighbouring areas will further refine the geographical distribution of these groups and their species diversity. Two species previously classified as representatives of other genera are included here as well: *Smithatris supraneanae* and *Curcuma* (= *Kaempferia) candida*. The morphological synapomorphies shared by 'Ecomata' and 'Pierreana', coupled with the lack of morphological characters distinguishing them, justify the recognition of a single subgenus comprising both groups.

***Curcuma* subg. *Hitcheniopsis*: loss of epigynous glands.** — Most of the representatives of the 'Hitcheniopsis' group resolved here correspond to the species previously classified by Schumann (1904) in *Curcuma* subg. *Hitcheniopsis*. The morphological delimitation of this subgenus was rather vague and based on the central position of the inflorescence, clearly connate bracts with recurved apices, and lack of anther spurs. Although Schumann also included *C. petiolata* and *C. roscoeana* in *C.* subg. *Hitcheniopsis*, neither of these species was supported as a representative of the 'Hitcheniopsis' group in this study; rather, they were unambiguously resolved as members of the 'Curcuma' group. Species currently classified as members of *Stahlianthus* were resolved as most derived in the 'Hitcheniopsis' group, based on ITS as well as cpDNA data, and the monotypic *Laosanthus* was also nested in this group. Based on our observations, reliable synapomorphies for the 'Hitcheniopsis' group shared by all analysed members are the lack of epigynous glands and the lack of anther spurs. The geographical distribution of the 'Hitcheniopsis' group is similar to that of the 'Ecomata' and 'Pierreana' groups, with a few species reaching further westwards to Myanmar (e.g., *C. parviflora*), northeast India, and also southern China (*Stahlianthus* spp.). *Curcuma vamana* was resolved in a stem position to the 'Hitcheniopsis' group based on nuclear markers, but it appeared in the 'Curcuma' group based on chloroplast markers. Morphologically, *C. vamana* matches the 'Hitcheniopsis' group (absence of epigynous glands and anther spurs), but its considerably disjunct distribution in southern India, sympatric occurrence with numerous seed-setting species from the 'Curcuma' group, and ambiguous position in the phylogenetic trees suggest an ancient hybridogenous origin—a hypothesis discussed above but that remains to be tested.

***Curcuma* subg. *Curcuma*: frequent polyploidization and diversification.** — The molecularly well-defined 'Curcuma' group includes species traditionally classified in subgenus *Curcuma* ('*Eucurcuma*' sensu Schumann, 1904) as well as five species from *Curcuma*-like genera (*Hitchenia caulina*, *H. glauca*, *C.* (= *Kaempferia*) *scaposa*, *C.* (= *Paracautleya*) *bhatii* and *Smithatris myanmarensis*). It is obvious that the 'Curcuma' group is genetically and morphologically more complex than *C.* subg. *Curcuma* as defined by Schumann (1904). This study also does not support morphologically defined groups by other authors (e.g., the 'Petiolata' group, composed of *C. aurantiaca*, *C. petiolata*, *C. roscoeana* and *C. rubrobracteata* and the 'Longa' group, composed of the rest of the traditionally circumscribed *C.* subg. *Curcuma*; Sirirugsa & al., 2007). Species of the *Curcuma*-like genera listed above fell in the 'Curcuma' group. They have previously been classified outside of *Curcuma* due to their highly specialized floral morphologies reflecting their unusual pollination syndromes (e.g., nocturnal anthesis in at least three of these species) and adaptations to unusual ecological habitats (Škorničková & Sabu, 2005a; Leong-Škorničková & al., 2007) and despite their vegetative morphology being well within the range of variation of *Curcuma*. The unavoidable inclusion of these highly specialized taxa makes the definition of the 'Curcuma' group by morphological synapomorphies problematic. However, for 'typical' representatives of the 'Curcuma' group, morphological synapomorphies include the presence of epigynous glands; inflorescences usually with a coma; closed, bullet-type flowers; and the presence of two forward-facing spurs. Numerous members of the 'Curcuma' group are widely distributed and cultivated in South and Southeast Asia and elsewhere in the tropics, but the centre of diversity for sexually reproducing (seed-setting) species is in India, Bangladesh and Myanmar, and that for vegetatively reproducing species likely in India.

Relationships of species within this group are very complex because polyploidization and hybridization were important for speciation (Leong-Škorničková & al., 2007; Záveská & al., 2011). The hierarchical structure within the group seems to be related to variation in genome size (Leong-Škorničková & al., 2007), which is also mostly congruent with AFLP data (Záveská & al., 2011). The basal polytomy in the 'Curcuma' group ('Curcuma-II' group in ITS and cpDNA trees) corresponds to species with greater monoploid genome size ('Genome Group II and III' sensu Leong-Škorničková & al., 2007), while the most terminal branch of the tree ('Curcuma-I' group in ITS and cpDNA trees) encompassed the species with smaller 1Cx-values ('Genome Group I' sensu Leong-Škorničková & al., 2007). More detailed analyses of species relationships within the 'Curcuma' group will be the subject of further studies.

Intra-individual ITS polymorphism: variation across the main groups. — Various levels of intra-individual ITS polymorphism were observed in the *Curcuma* species investigated. The gradient from lower to higher variation mirrored the phylogenetic relationships in the genus from the basal groups to the terminal ones. Other factors, such as ploidy level variation (polyploid origin), homoploid hybridization, and mode of reproduction, seem to have had a significant influence on this variation.

Although there is still limited information about basic chromosome numbers and chromosome numbers of species from the basal part of the phylogeny (representatives of the 'Ecomata', 'Pierreana' and 'Hitcheniopsis' groups), their within-individual ITS variation is rather low, resulting in single terminal sequences in the final phylogeny, suggesting that a single ITS paralogue prevails in their genomes. As these species mainly reproduce sexually and seem to be functional diploids (disomic inheritance observed in allozyme analyses, unpub. data), we assume that the majority of them arose by processes other than hybrid/polyploid speciation (e.g., by allopatric speciation) and that their ITS sequence divergence from the most recent common ancestor is the result of progressive accumulation of mutations in the single ITS sequence type.

In contrast, representatives of the terminal 'Curcuma' group show high intra-individual ITS polymorphism that may be attributed to extensive variation in ploidy levels, from 6*x* to 15*x* (2*n* = 42 to 105; Leong-Škorničková & al., 2007). The number of ITS paralogues in the genome is not directly proportional to the ploidy level, and other factors such as mode of reproduction or age of hybrid/allopolyploid origin might cause the presence of an uneven number of ITS paralogues in species of the same ploidy level (e.g., Kovařík & al., 2005; Dadejová & al., 2007).

The highest within-individual ITS polymorphism was detected in taxa from the terminal lineage of the 'Curcuma' group (the 'Curcuma-I' group in Fig. 3A that corresponds to the 'Genome Group I' sensu Leong-Škorničková & al., 2007) that is composed of hexaploid as well as higher polyploid taxa. A maximum of 87 polymorphic sites in a single individual alignment and up to 11 ITS paralogues in a single genome were observed in this group. This variation may be the consequence of genetic processes that can impinge on the evolution of the multiple-copy ITS region after allopolyploid formation. For instance, divergent (parental) ITS types can be maintained in an allopolyploid genome (Soltis & Soltis, 1991; Soltis & al., 1995), and maintenance of two or more ITS repeat types can be followed by processes of recombination leading to excessive chimeric ITS sequences (Alvarez & Wendel, 2003). As suggested by Baldwin & al. (1995), vegetative reproduction is likely to be a reason for maintenance of parental ITS sequences in hybrids. Indeed, Sang & al. (1995) noted that in polyploid *Paeonia* frequent reproduction via rhizomes may significantly prolong generation time, which in turn might cause slow rates of concerted evolution. As most *Curcuma* hexaploids are facultatively vegetatively reproducing species and all the higher polyploids reproduce exclusively vegetatively by rhizomes (Škorničková, 2007), it is very likely that the mode of reproduction in *Curcuma* is linked to processes allowing maintenance of high numbers of ITS paralogues in a single genome.

■ ACKNOWLEDGEMENTS

We extend sincere thanks to V. Machalová for helping with the cpDNA sequencing analyses and S. Voldrich for DNA extraction of some of the samples. We are very indebted to the following individuals and institutions that helped us to source some plant tissue material: Dr. Charun Maknoi, Dr. Piyakaset Suksathan, Dr. Pramote Triboun, Dr. Thaya Jenjittikul, Karen E. Shigematsu, Lý Ngoc-Sam, Martin van der Bult, Dr. Axel D. Poulsen and Dr. Mark F. Newman; Institute of Biological Sciences – Aarhus University, Lyon Arboretum, Prague Botanic Garden, Queen Sirikit Botanic Gardens, Royal Botanic Garden Edinburgh and Singapore Botanic Gardens. We thank the authorities of the Ministry of Environment and Forests, Government of India and the Botanical Survey of India for granting collection permits and permits to access Indian herbaria. We thank the Indian Council of Cultural Relationships and the Ministry of Education, Youth and Sport of the Czech Republic for granting the fellowship to J.L.-Š. at Calicut University, Kerala, India and to Dr. M. Sabu for his various help and support throughout her fellowship. We also gratefully acknowledge valuable suggestions of two anonymous reviewers and language corrections provided by Dr. Mark F. Newman and Dr. Jonathan Rosenthal. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, the Department of Science and Technology, Government of India (grant no SP/SO/A-20/99 dt. 9.11.2001), the National Parks Board, Singapore, the Ministry of Culture of the Czech Republic (DKRVO National Museum, Prague), the Czech Science Foundation, GAČR (grant no. P506/10/0623), and the Czech Academy of Sciences (long-term research and development project no. RVO 67985939).

■ LITERATURE CITED

Alvarez, I. & Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molec. Phylogen. Evol.* 29: 417–434.

Anthony, N.M., Clifford, S.L., Bawe-Johnson, M., Abernethy, K.A., Bruford, M.W. & Wickings, E.J. 2007. Distinguishing gorilla mitochondrial sequences from nuclear integrations and PCR recombinants: Guidelines for their diagnosis in complex sequence databases. *Molec. Phylogen. Evol.* 43: 553–566.

Baker, J.G. 1890. *Curcuma*. Pp. 209–216 in: Hooker, J.D. (ed.), *The flora of British India*, vol. 6. London: L. Reeve.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S. & Donoghue, M.J. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence of angiosperm phylogeny. *Ann. Missouri Bot. Gard.* 82: 247–277.

Bryant, D. & Moulton, V. 2004. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molec. Biol. Evol.* 21: 255–265.

Cao, H. & Komatsu, K. 2003. Molecular identification of six medicinal *Curcuma* plants produced in Sichuan: Evidence from plastid *trnK* gene sequences. *Acta Pharm. Sin.* 38: 871–875.

Cao, H., Sasaki, Y., Fushimi, H. & Komatsu, K. 2001. Molecular analysis of medicinally used Chinese and Japanese *Curcuma* based on 18S rRNA gene and *trnK* gene sequences. *Biol. Pharm. Bull.* 24: 1389–1394.

Cronn, R., Cedroni, M., Haselkorn, T., Grover, C. & Wendel, J.F. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theor. Appl. Genet.* 104: 482–489.

Dadejová, M., Lim, K.Y., Součková-Skalická, K., Matyášek, R., Grandbastien, M.A., Leitch, A. & Kovařík, A. 2007. Transcription activity of rRNA genes correlates with a tendency towards intergenomic homogenization in *Nicotiana* allotetraploids. *New Phytol.* 174: 658–668.

Farris, J.S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.

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

Grimm, G.W. & Denk, T. 2008. ITS evolution in *Platanus*: Homoeologues, pseudogenes, and ancient hybridization. *Ann. Bot. (Oxford)* 101: 403–419.

Hall, T. 2004. BioEdit: Biological sequence alignment editor for Win95/98/NT/2K/XP, version 7.0.4.1. http://www.mbio.ncsu.edu/BioEdit/bioedit.html.

Horaninow, P. 1862. *Prodromus monographiae scitaminearum*. St. Petersburg: Typis Academiae Caesareae Scientiarum.

Huber, K.T. & Moulton, V. 2005. Phylogenetic networks. Pp. 178–200 in: Gascuel, O. (ed.), *Mathematics of evolution and phylogeny*. Oxford: Oxford University Press.

Huelsenbeck, J.P. & Ronquist, F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.

Huson, D.H. & Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Molec. Biol. Evol.* 23: 254–267.

Jenjittikul, T. & Larsen, K. 2000. *Kaempferia candida* Wall. (Zingiberaceae), a new record for Thailand. *Thai Forest Bull., Bot.* 28: 45–49.

Kelly, L.J., Leitch, A.R., Clarkson, J.J., Hunter, R.B., Knapp, S. & Chase, M.W. 2010. Intragenic recombination events and evidence for hybrid speciation in Nicotiana (Solanaceae). *Molec. Biol. Evol.* 27: 781–799.

Kluge, A.G. & Farris, J.S. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 42: 1–32.

Kovařík, A., Pires, J.C., Leitch, A.R., Lim, K.Y., Sherwood, A.M., Matyasek, R., Rocca, J., Soltis, D.E. & Soltis, P.S. 2005. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* 169: 931–944.

Kress, W.J., Liu, A.Z., Newman, M. & Li, Q.J. 2005. The molecular phylogeny of *Alpinia* (Zingiberaceae): A complex and polyphyletic genus of gingers. *Amer. J. Bot.* 92: 167–178.

Kress, W.J., Mood, J.D, Sabu, M., Prince, L.M., Dey, S. & Sanoj, E. 2010. *Larsenianthus*, a new Asian genus of Gingers (Zingiberaceae) with four species. *PhytoKeys* 1: 15–32.

Kress, W.J., Prince, L.M. & Williams, K.J. 2002. The phylogeny and a new classification of the gingers (Zingiberaceae): Evidence from molecular data. *Amer. J. Bot.* 89: 1682–1696.

Kumar, S.S., Skjæveland, Å., Orr, R.J.S., Enger, P., Ruden, T., Mevik, B.-H., Burki, F., Botnen, A. & Shalchian-Tabrizi, K. 2009. AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *B. M. C. Bioinformatics* 10: 357, doi:10.1186/1471-2105-10-357.

Larsen, K. & Jenjittikul, T. 2001. *Laosanthus*, a new genus of Zingiberaceae from Laos. *Nordic J. Bot.* 21: 135–138.

Leong-Škorničková, J. & Lý, N.-S. 2010. *Curcuma pambrosima* sp. nov. (Zingiberaceae) from central Vietnam. *Nordic J. Bot.* 28: 652–655.

Leong-Škorničková, J., Šída, O., Jarolímová, V., Sabu, M., Fér, T., Trávníček, P. & Suda, J. 2007. Chromosome numbers and genome size variation in Indian species of *Curcuma* L. (Zingiberaceae). *Ann. Bot. (Oxford)* 100: 505–526.

Leong-Škorničková, J., Šída, O. & Marhold, K. 2010. Back to types! Towards stability of the names in Indian *Curcuma* L. (Zingiberaceae). *Taxon* 59: 269–282.

Leong-Škorničková, J., Šída, O., Sabu, M. & Marhold, K. 2008a. Taxonomic and nomenclatural puzzles in Indian *Curcuma*: The identity and nomenclatural history of *C. zedoaria* (Christm.) Roscoe and *C. zerumbet* Roxb. (Zingiberaceae). *Taxon* 57: 949–962.

Leong-Škorničková, J., Šída, O., Wijesundara, S. & Marhold, K. 2008b. On the identity of turmeric: The typification of *Curcuma longa* L. (Zingiberaceae). *Bot. J. Linn. Soc.* 157: 37–46.

Linnaeus, C. 1753. *Species plantarum*. Stockholm: impensis Laurentii Salvii.

Manos, P.S. & Steele, K.P. 1997. Phylogenetic analyses of ''higher'' Hamamelididae based on plastid sequence data. *Amer. J. Bot.* 84: 1407–1419.

Martin, D.P., Williamson, C. & Posada, D. 2005. RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* 21: 260–262.

McDade, L.A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.

Mood, J. & Larsen, K. 2001. New Curcumas from South East Asia. *New Plantsman* 8: 207–217.

Ngamriabsakul, C., Newman, M. & Cronk, Q.B.C. 2004. The phylogeny of tribe Zingibereae (Zingiberaceae) based on ITS (nr DNA) and *trnL-F* (cp DNA) sequences. *Edinburgh J. Bot.* 60: 483–507.

Nylander, J.A.A. 2004. MrModeltest, version 2.1. Evolutionary Biology Centre, Uppsala University, Uppsala. Program distributed by the author. http://www.abc.se/~nylander/.

Okuyama, Y., Fujii, N., Wakabayashi, M., Kawakita, A., Ito, M., Watanabe, M., Murakami, N. & Makoto, K. 2005. Nonuniform concerted evolution and chloroplast capture: Heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). *Molec. Biol. Evol.* 22: 285–296.

Pedersen, L.B. 2004. Phylogenetic analysis of the subfamily *Alpinioideae* (Zingiberaceae), particularly *Etlingera* Giseke, based on nuclear and plastid DNA. *Pl. Syst. Evol.* 245: 239–258.

Popp, M. & Oxelman, B. 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. *Molec. Phylogen. Evol.* 20: 474–481.

Popp, M., Erixon, P., Eggens, F. & Oxelman, B. 2005. Origin and evolution of a circumpolar polyploid species complex in *Silene* (Caryophyllaceae) inferred from low copy nuclear RNA polymerase introns, rDNA, and chloroplast DNA. *Syst. Bot.* 30: 302–313.

Posada, D. & Crandall, K.A. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc. Natl. Acad. Sci. U.S.A.* 98: 13757–13762.

Rieseberg, L.H. & Wendel, J.F. 1993. Introgression and its consequences in plants. Pp. 70–109 in: Harrison, R.G. (ed.), *Hybrid zones and the evolutionary process*. New York: Oxford University Press.

Rieseberg, L.H., Whitton, J. & Linder, C.R. 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Bot. Neerl.* 45: 243–262.

Roxburgh, W. 1810. Descriptions of several of the monandrous plants of India. *Asiat. Res.* 11: 318–362.

Russell, A., Samuel, R., Klejna, V., Barfuss, M.H.J., Rupp, B. & Chase, M.W. 2010. Reticulate evolution in diploid and tetraploid species of *Polystachya* (Orchidaceae) as shown by plastid DNA sequences and low-copy nuclear genes. *Ann. Bot. (Oxford)* 106: 37–56.

Salmon, A., Flagel, L., Ying, B., Udall, J.A. & Wendel, J.F. 2010. Homoeologous nonreciprocal recombination in polyploid cotton. *New Phytol.* 186: 123–134.

Sang, T. & Zhong, Y. 2000. Testing hybridization hypotheses based on incongruent gene trees. *Syst. Biol.* 49: 422–434.

Sang, T., Crawford, D.J. & Stuessy, T.F. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. *Proc. Natl. Acad. Sci. U.S.A.* 92: 6813–6817.

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

Schumann, K. 1904. Zingiberaceae. Pp. 1–458 in: Engler, A. (ed.), *Das Pflanzenreich*, IV, vol. 46, Leipzig: Engelmann.

Simmons, M.P. & Ochoterena, H. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Syst. Biol.* 49: 369–381

Sirirugsa, P. & Newman, M. 2000. A new species of *Curcuma* L. (Zingiberaceae) from S.E. Asia. *New Plantsman* 6: 196–197.

Sirirugsa, P., Larsen, K. & Maknoi, C. 2007. The genus *Curcuma* L. (Zingiberaceae): Distribution and classification with reference to species diversity in Thailand. *Gard. Bull. Singapore* 59: 203–219.

Škorničková, J. 2007. *Taxonomic studies in Indian Curcuma L.* Dissertation, Charles University, Prague, Czech Republic.

Škorničková, J. & Sabu, M. 2005a. The recirumscription of *Curcuma* L. to include the genus *Paracautleya* R.M. Sm. *Gard. Bull. Singapore* 57: 37–46.

Škorničková, J. & Sabu, M. 2005b. *Curcuma roscoeana* Wall (Zingiberaceae) in India. *Gard. Bull. Singapore* 57: 187–198.

Škorničková, J. & Sabu, M. 2005c. The identity and distribution of *Curcuma zanthorrhiza* Roxb. (Zingiberaceae). *Gard. Bull. Singapore* 57: 199–210.

Škorničková, J., Rehse, T. & Sabu, M. 2007. Other economically important *Curcuma* species. Pp. 451–467 in: Ravindran, P.N. (ed.), *Turmeric: The genus* Curcuma. Florida: CRC Press.

Škorničková, J., Sabu, M. & Prasanthkumar, M.G. 2003a. A new species of *Curcuma* L. (Zingiberaceae) from Mizoram. *Gard. Bull. Singapore* 55: 89–95.

Škorničková, J., Sabu, M. & Prasanthkumar, M.G. 2003b. *Curcuma codonantha* (Zingiberaceae)—A new species form the Andaman Islands, India. *Gard. Bull. Singapore* 55: 219–228.

Škorničková, J., Sabu, M. & Prasanthkumar, M.G. 2004. *Curcuma mutabilis* (Zingiberaceae): A new species from South India. *Gard. Bull. Singapore* 56: 43–54.

Soltis, P.S. & Soltis, D.E. 1991. Multiple origins of the allotetraploid *Tragopogon mirus* (Compositae): rDNA evidence. *Syst. Bot.* 16: 407–413.

Soltis, P.S., Plunkett, G.M., Novak, S.J. & Soltis, D.E. 1995. Genetic variation in *Tragopogon* species: Additional origins of the allotetraploids *T. mirus* and *T. miscellus* (Compositae). *Amer. J. Bot.* 82: 1329–1341.

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

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sunderland, Massachusetts: Sinauer.

Syamkumar, S. & Sasikumar, B. 2007. Molecular marker based genetic diversity analysis of *Curcuma* species from India. *Sci. Hort.* 112: 235–241.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Pl. Molec. Biol.* 17: 1105–1109.

Takhtajan, A. 1986. *Floristic regions of the World*, trans. T.J. Crovello & A. Cronquist). Berkeley: University of California Press.

Tate, J.A. & Simpson, B.B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Syst. Bot.* 28: 723–737.

Techaprasan, J. & Leong-Škorničková, J. 2011. Transfer of *Kaempferia candida* to *Curcuma* (Zingiberaceae) based on morphological and molecular data. *Nordic J. Bot.* 29: 773–779.

Valeton, T. 1918. New notes on the Zingiberaceae of Java and Malaya. *Bull. Jard. Bot. Buitenzorg* 27: 1–167.

Velayudhan, K.C., Amalraj, V.A. & Muralidharan, V.K. 1996. The conspectus of the genus *Curcuma* in India. *J. Econ. Taxon. Bot.* 20: 375–382.

Velayudhan, K.C., Mralidharan, V.K., Amalraj, V.A., Gautam, P.L., Mandal, S. & Kumar, D. 1999. *Curcuma genetic resources.* Trichur, India: NBPGR Regional Station.

White, T.J., Bruns, T., Lee, S. & Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315–322 in: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (eds.), *PCR protocols: A guide to methods and applications*. San Diego: Academic Press.

Williams, K., Kress, J.W. & Manos, P.S. 2004. The phylogeny, evolution and classification of the genus *Globba* and tribe Globbeae (Zingiberaceae): Appendages do matter. *Amer. J. Bot.* 91: 100–114.

Wolfe, A.D. & Randle, C.P. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: Implications for plant molecular systematics. *Syst. Bot.* 29: 1011–1020.

Young, N.D. & Healy, J. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. *B. M. C. Bioinformatics* 4: 6, doi: 10.1186/1471-2105-4-6.

Záveská, E., Fér, T., Šída, O., Leong-Škorničková, J., Sabu, M. & Marhold, K. 2011. Genetic diversity patterns in *Curcuma* reflect differences in genome size. *Bot. J. Linn. Soc.* 165: 388–401.

Appendix. Plant material analysed in the present study. Species names, collection numbers, localities, and GenBank accession numbers are listed. Chromosome numbers were taken from Leong-Škorničková & al. (2007). Serial numbers preceding species name correspond to numbers attached to their distributional records in Fig. 2. Serial numbers in brackets following the species name indicate multiple individuals per species and correspond to species labels in Fig. 3 and Fig. 4. Collectors: *CM* = Charun Maknoi; *JLS* = Jana Leong-Škorničková & al.; *JM* = John Mood; *KL* = Kai Larsen; *LNS* = Lý Ngoc Sam; *MB* = Martin van der Bult; *MN* = Mark Newman & al.; *THD* = Tran Huu Dang & al. A dash (–) indicates missing data; an asterisk (*) denotes sequences obtained from GenBank.

Serial number. Taxon, voucher specimen, locality, number of terminals in ITS phylogenetic tree / number of polymorphic sites in intra-individual alignment (= intra-individual ITS polymorphism), accession numbers for the ITS (terminals are separated by a slash, refer to Methods for details), *trnL-trnF*, *psbA-trnH* and *matK*.

CURCUMA-I group (2*n* = 42, 6*x*): 1. ***C. angustifolia*** Roxb. (1), *JLS 73453* (CALI, SING), India - Uttaranchal, 2/21, JQ409970/JQ409972, JQ409820, JQ409753, JQ409685; **2. *C. angustifolia*** Roxb. (2), *JLS 73480* (CALI, SING), India - Chhattisgarh, 2/24, JQ409971/JQ409973, JQ409819, JQ409752, JQ409684; **3. *C. aromatica*** Salisb. (1), *JLS 84183* (CALI, SING), India - Kerala, 4/43, JQ409910/JQ409938/JQ409958/JQ409962, JQ409834, JQ409767, JQ409699; **4. *C. aromatica*** Salisb. (2), *JLS 84170* (CALI, SING), India - Kerala, 3/25, JQ409929/JQ409937/JQ409961, –, –, –; **5. *C.* aff. *prakasha*** S. Tripathi (1), *JLS 71443* (CALI, SING), India - Meghalaya, 3/55, JQ409923/JQ409957/JQ409963, JQ409837, JQ409770, JQ409702; **6. *C.* aff. *prakasha*** S. Tripathi (2), *JLS 71476* (CALI, SING), India - W. Bengal, 1/25, JQ409982, –, –, –; **7. *C. montana*** Roxb. (1), *JLS 71484* (CALI, SING), India - Jharkhand, 5/29, JQ409932/JQ409933/JQ409939/JQ409944/JQ409945, JQ409836, JQ409769, JQ409701; **8. *C. montana*** Roxb. (2), *JLS 73479* (CALI, SING), India - Chhattisgarh, 3/25, JQ409934/JQ409936/JQ409950, JQ409835, JQ409768, JQ409700; **9. *C. montana*** Roxb. (3), *JLS 73425* (CALI, SING), India - Orissa, 3/25, JQ409935/JQ409941/JQ409955, –, –, –; **10. *Hitchenia glauca*** Wall., *Kress #00-6743* (US), Myanmar, 1/-, AF478765*, –, –, –. **CURCUMA-I group (2*n* = 63, 9*x*): 11. *C. aeruginosa*** Roxb. (1), *JLS 71431* (CALI, SING), India - Assam, 11/87, JQ409920/JQ409926/JQ409927/JQ409931/JQ409946/JQ409948/JQ409953/JQ409959/JQ409960/JQ409965/JQ409983, JQ409824, JQ409757, JQ409689; **12. *C. aeruginosa*** Roxb. (2), *JLS 84142* (CALI, SING), India - Kerala, 5/64, JQ409921/JQ409949/JQ409951/JQ409964/JQ409975, JQ409821, JQ409754, JQ409686; **13. *C. aromatica*** Salisb. (3), *JLS 71492* (CALI, SING), India - Jharkhand, 4/57, JQ409925/JQ409940/JQ409943/JQ409952, JQ409825, JQ409758, JQ409690; **14. *C. attenuata*** Baker, *JLS 73342* (E, SING), Thailand, 2/39, JQ409978/JQ409981, JQ409826, JQ409759, JQ409691; **15. *C. brog*** Valeton, *JLS 73347* (E, SING), Indonesia, 2/32, JQ409913/JQ409919, JQ409831, JQ409764, JQ409696; **16. *C. leucorrhiza*** Roxb., *JLS 71489* (CALI, SING), India - Jharkhand, 2/17, JQ409912/JQ409915, JQ409830, JQ409763, JQ409695; **17. *C. longa*** L., *JLS 73222* (PDA, SING), Sri Lanka, 2/38, JQ409924/JQ409956, JQ409823, JQ409756, JQ409688; **18. *C. ochrorhiza*** Valeton, *JLS 73348* (E, SING), Indonesia, 2/21, JQ409916/JQ409966, JQ409828, JQ409761, JQ409693; **19. *C. petiolata*** Roxb., *JLS 22* (SING), Thailand, 1/2, JQ409985, JQ409816, JQ409749, JQ409681; **20. *C. plicata*** Baker, *THD 80* (E, SING), Vietnam, 1/12, JQ409979, JQ409827, JQ409760, JQ409692; **21. *C. roscoeana*** Wall. (1), *JLS 73351* (E, SING), Thailand, 1/0, JQ409887, JQ409804, JQ409737, JQ409670; **22. *C. roscoeana*** Wall. (2), *JLS 73310* (CALI, SING), India - Andamans, 1/2, JQ409886, JQ409803, JQ409736, JQ409669; **23. *C. rubescens*** Roxb., *JLS 71457* (CALI, SING), India - Meghalaya, 4/25, JQ409911/JQ409928/JQ409930/JQ409954, JQ409833, JQ409766, JQ409698; **24. *C. rubrobracteata*** Škorničk., M. Sabu & Prasanthk. (1), *JLS 73332* (E, SING), Thailand, 2/47, JQ409984/JQ409987, JQ409817, JQ409750, JQ409682; **25. *C. rubrobracteata*** Škorničk., M. Sabu & Prasanthk. (2), *JLS 73333* (E, SING), Bangladesh, 1/31, JQ409986, JQ409822, JQ409755, JQ409687; **26. *C. zanthorrhiza*** Roxb., *JLS 84182* (CALI, SING), India - Kerala, 2/24, JQ409914/JQ409918, JQ409832, JQ409765, JQ409697. **CURCUMA-I group (2*n* = 84, 12*x*): 27. *C.* sp. 'ranchi'**, –, *JLS 71485* (CALI, SING), India - Jharkhand, 2/10, JQ409942/JQ409947, JQ409829, JQ409762, JQ409694. **CURCUMA-I group (2n = 105, 15x): 28. *C. zedoaria*** (Christm.) Roscoe (1), *JLS 71432* (CALI, SING), India - Assam, 4/58, JQ409967/JQ409974/JQ409976/JQ409977, JQ409818, JQ409751, JQ409683; **29. *C. zedoaria*** (Christm.) Roscoe (2), *JLS 84120* (CALI, SING), India - Kerala, 3/44, JQ409917/JQ409922/JQ409980, JQ409838, JQ409771, JQ409703**. CURCUMA-II group (2*n* = 42, 6*x*): 30. *C. aurantiaca*** Zijp, *JLS 84155* (CALI, SING), India - Kerala, 2/2 , JQ409904/JQ409906, JQ409806, JQ409739, JQ409672; **31. *C. bhatii*** (R.M. Sm.) Škorničk. & M. Sabu [= ***Paracautleya bhatii*** R.M. Sm.], *JLS 73446* (CALI, SING), India - Karnataka, 1/3, JQ409897, JQ409814, JQ409747, JQ409679; **32. *C. cannanorensis*** var. ***lutea*** R. Ansari, V.J. Nair & N.C. Nair, *JLS 84143* (CALI, SING), India - Kerala, 2/21, JQ409892/JQ409903, JQ409809, JQ409742, JQ409675; **33. *C. coriacea*** Mangaly & M. Sabu (1), *JLS 86140* (CALI, SING), India - Kerala, 1/3, JQ409902, JQ409807, JQ409740, JQ409673; **34. *C. coriacea*** Mangaly & M. Sabu (2), *JLS 84171* (CALI, SING), India - Kerala, 1/9, JQ409907, JQ409808, JQ409741, JQ409674; **35. *C. inodora*** Blatt., *JLS 73403* (CALI, SING), India - Maharashtra, 1/9, JQ409899, –, –, –; **36. *C. mutabilis*** Škorničk., M. Sabu & Prasanthk., *JLS 84145* (CALI, SING), India - Kerala, 3/21, JQ409890/JQ409901/JQ409905, JQ409812, JQ409745, JQ409714; **37. *C. neilgherrensis*** Wight, *JLS 84157* (CALI, SING), India - Kerala, 3/33, JQ409891/JQ409893/JQ409909, JQ409813, JQ409746, JQ409678; **38. *C. pseudomontana*** J. Graham, *JLS 73402* (CALI, SING), India - Maharashtra, 2/11, JQ409894/JQ409895, –, –, –; **39. *C. reclinata*** Roxb. (1), *JLS 73467*, India - Madhya Pradesh, 1/3, JQ409968, JQ409810, JQ409743, JQ409676; **40. *C. reclinata*** Roxb. (2), *JLS 73477*, India - Chhattisgarh (CALI, SING), 1/8, JQ409969, JQ409811, JQ409744, JQ409677; **41. *C. scaposa*** (Nimmo) Škorničk. & M. Sabu [= ***Kaempferia scaposa*** (Nimmo) Benth.], *JLS 86407*, India - Goa, 1/1, JQ409896, –, –, –; **42. *C. caulina*** J. Graham [= *Hitchenia* ***caulina*** (J. Graham) Baker], *JLS 84178* (CALI, SING), India - Maharashtra, 2/8, JQ409898/JQ409900, JQ409815, JQ409748, JQ409680; **43. *Smithatris myanmarensis*** W. J. Kress,

Appendix. Continued.

KL 47520 (AAU), Myanmar, 1/0, JQ409908, JQ409844, JQ409777, JQ409708**. CURCUMA-II group (2*n* = 77, 11*x*): 44. *C. oligantha*** Trimen (1), *JLS 73325* (PDA, SING), Sri Lanka, 1/2, JQ409889, JQ409839, JQ409772, JQ409715; **45. *C. oligantha*** Trimen (2), *JLS 73223* (PDA, SING), Sri Lanka, 1/1, JQ409888, JQ409805, JQ409738, JQ409671**. ALISMATIFOLIA group: 46. *C. alismatifolia*** Gagnep. (1), *THD 20* (E, SING), Vietnam, 1/4, JQ409860, –, –, –; **47. *C. alismatifolia*** Gagnep. (2), *JLS 427* (SING), Thailand, 1/2, JQ409859, JQ409788, JQ409721, JQ409654; **48. *C. gracillima*** Gagnep., *THD 48* (E, SING), Vietnam, 1/5, JQ409851, JQ409787, JQ409720, JQ409653; **49. *C. harmandii*** Gagnep., *JLS 73326* (E, SING), Thailand, 1/0, JQ409853, JQ409791, JQ409724, JQ409657; **50. *C. parviflora* agg.** Wall. (1), *JLS 73328* (E, SING), Thailand, 1/3, JQ409855, JQ409789, JQ409722, JQ409655; **51. *C. parviflora* agg.** Wall. (2), *JLS 73329* (E, SING), Thailand, 1/5, JQ409856, JQ409790, JQ409723, JQ409656; **52. *C. rhabdota*** Sirirugsa & M.F. Newman, *JLS 73331* (E, SING), Laos, 1/2, JQ409854, JQ409792, JQ409725, JQ409658; **53. *C. sparganiifolia*** Gagnep. (1), *MN 2382* (E, SING), Cambodia, 1/3, JQ409858, –, –, –; **54. *C. sparganiifolia*** Gagnep. (2), *MN 2403* (E, SING), Cambodia, 1/2, JQ409857, –, –, –; **55. *C. thorelii*** Gagnep., *THD 78* (E, SING), Vietnam, 1/0, JQ409852, JQ409786, JQ409719, JQ409652; **56. *C. vamana*** M. Sabu & Mangaly, *JLS 84156* (CALI, SING), India - Kerala, 1/0, JQ409867, JQ409793, JQ409726, JQ409659; **57. *Laosanthus graminifolius*** K. Larsen & Jenjitt., *KL 47445* (AAU), Laos , 1/0, JQ409861, JQ409845, JQ409778, JQ409709; **58. *Stahlianthus campanulatus*** Kuntze, *JLS 73246* (SING), China, 1/0, JQ409864, JQ409785, JQ409718, JQ409651; **59. *Stahlianthus involucratus*** (King ex Baker) Craib ex Loes., *JLS 71449* (CALI, SING), India - Meghalaya, 1/0, JQ409862, JQ409784, JQ409717, JQ409650; **60. *Stahlianthus* sp.** (1), *JLS 616* (SING), Thailand, 1/6, JQ409866, JQ409841, JQ409774, JQ409705; **61. *Stahlianthus* sp.** (2), *JLS 618* (SING), Thailand, 1/5, JQ409865, JQ409842, JQ409775, JQ409706; **62. *Stahlianthus* sp.** (3), *JLS 620* (SING), Thailand, 1/3, JQ409863, JQ409843, JQ409776, JQ409707. **PIERREANA group: 63. *C.* aff. *cochinchinensis*** Gagnep., *JLS 213* (SING), Thailand, 1/1, JQ409871, JQ409796, JQ409729, JQ409662; **64. *C. candida*** (Wall.) Techaprasan & Škorničk. [= ***Kaempferia candida*** Wall.], *JLS 606* (SING), Thailand, 1/8, JQ409988, JQ409840, JQ409773, JQ409704; **65. *C. pambrosima*** Škorničk. & N. S. Lý, *LNS 316* (SING, VNM), Vietnam, 1/2, JQ409874, JQ409800, JQ409733, JQ409666; **66. *C. pierreana*** Gagnep., *THD 19* (E, SING), Vietnam, 1/3, JQ409879, –, –, –; **67. *C. singularis*** Gagnep. (1), *JLS 73343* (E, SING), Thailand, 1/5, JQ409869, JQ409798, JQ409731, JQ409664; **68. *C. singularis*** Gagnep. (2), *MN 2413* (E, SING), Cambodia, 1/0, JQ409872, JQ409846, JQ409779, JQ409716; **69. *C.* sp. nov.**, *JLS 73334* (E, SING), Thailand, 1/13, JQ409868, JQ409795, JQ409728, JQ409661; **70. *C.* sp. nov. 'newmanii'**, *JLS 365* (SING), Vietnam, 2/6, JQ409876/JQ409877, JQ409799, JQ409732, JQ409665; **71. *C.* sp. nov. 'xanthella'**, *LNS 348* (SING, VNM), Vietnam, 1/6, JQ409875, –, –, –; **72. *C. rhomba*** Mood & K.Larsen (1), *LNS 368* (SING, VNM), Vietnam, 1/0, JQ409878, JQ409802, JQ409735, JQ409668; **73. *C. rhomba*** Mood & K.Larsen (2), *JM 97p147* (AAU), Vietnam, 1/0, JQ409880, JQ409850, JQ409783, JQ409713; **74. *C. vitellina*** Škorničk. & H. D. Tran, *THD 70* (E, SING), Vietnam, 1/4, JQ409873, JQ409801, JQ409734, JQ409667; **75. *Smithatris supraneanae*** W.J.Kress & K.Larsen, *JLS 206* (SING), Thailand, 1/2, JQ409870, JQ409797, JQ409730, JQ409663. **ECOMATA group: 76. *C. bicolor*** Mood & K. Larsen (1), *MB 1074* (CMU), Thailand, 1/2, JQ409882, JQ409847, JQ409780, JQ409710; **77. *C. bicolor*** Mood & K. Larsen (2), *JM 97p149* (AAU), Thailand, 1/0, JQ409883, JQ409848, JQ409781, JQ409711; **78. *C. ecomata*** Craib, *JLS 73353* (E, SING), Thailand, 1/2, JQ409881, JQ409794, JQ409727, JQ409660; **79. *C. flaviflora*** S.Q. Tong, *CM 356* (QBG), Thailand, 1/-, DQ395335, –, –, –; **80. *C. glans*** K. Larsen & Mood, *JM & KL 1455B (AAU)*, Thailand, 2/10, JQ409884/JQ409885, JQ409849, JQ409782, JQ409712. **OUTGROUP: 81. *Alpinia conchigera*** Griff., *JLS 86247* (CALI, SING), India, 1/-, JQ409990, –, –, –; **82. *Alpinia galanga*** (L.) Willd., A. Rangsiruji 3 (E)/ –/Kress #94-5263 (US), ex hort. , –, AY424775*, EU552528*, AF478815*; **83. *Camptandra parvula*** (King ex Baker) Ridl., *Kress #99-6328* (US), Thailand, 1/-, AF478730*, –, –, –; **84. *Globba* sp.**, *JLS 73450* (CALI, SING), India, 1/-, JQ409989, –, –, –; **85. *Larsenianthus careyanus*** (Benth.) W.J.Kress & Mood, *JLS 92631* (CALI, SING), India, 1/-, JQ409992, –, –, –; **86. *Pyrgophyllum yunnanense*** (Gagnep.) T.L. Wu & Z.Y. Chen, *Kress #00-6596* (US), China, 1/-, AF478777*, –, –, –; **87. *Zingiber capitatum*** Roxb., *JLS 73466* (CALI, SING), India - Jharkhand, 1/-, JQ409991, –, –, –.