Charles University in Prague
Faculty of Science, Department of Botany

Master Thesis

Evolution of Indian *Curcuma* L.

(Studium evoluce indických zástupců rodu *Curcuma*)

Eliška Záveská

Supervisor: Mgr. Tomáš Fér, PhD.
Advisor: RNDr. Jana Škorničková, PhD.

Prague, 2009

Diplomová práce **Elišky Záveské** byla obhájena na katedře botaniky PřF UK v Praze dne 22.9.2009 a ohodnocena klasifikačním stupněm **výborně**.

Acknowledgement

This study was financially supported by the Grant Agency of Charles University (GA UK No. 112608, Eliška Záveská, Tomáš Fér). I thank my supervisor Tomáš Fér for guidance, Jana Škorničková for introducing me into the problematics of the genus *Curcuma*, Karol Marhold for his advice on data analyses and Otakar Šída for suggestions and comments regarding cytological data included in this work. I also thank to my family and friends for their endless support.

Abstrakt

Ekonomicky významný polyploidní rod *Curcuma* L. (Zingiberaceae) je z taxonomického hlediska problematický. Druhy s rozdílným stupněm ploidie (2x, 6x, 9x, 11x, 12x, 15x) preferují různé rozmnožovací strategie, které mají za následek rozdílnou míru morfologické variability, která znesnadňuje identifikaci jednotlivých druhů. Problémy v nomenklatuře a absence celkové revize rodu navíc vede k nesprávnému užívání druhových jmen a spolu s neobjasněnými vztahy v rámci rodu neumožnují spolehlivou vnitrorodovou klasifikaci. Molekulární studie na úrovni čeledi Zingiberaceae navíc naznačuje, že rod *Curcuma* je parafyletický s několika příbuznými rody (Kress et. al. 2002). V recentní cytologické studii indických zástupců rodu *Curcuma* (Leong-Škorničková et al., 2007) bylo navrženo členění podrodu *Curcuma* podle rozdílů v homoploidní velikosti genomu. Cílem této magisterské práce je 1) ověřit monofyletický původ třech „genomových skupin" navržených studií Leong-Škorničková et al. (2007) a nastínit fylogenetické vztahy druhů jak v rámci těchto skupin, tak mezi nimi a 2) zhodnotit fylogenetické vztahy mezi rodem *Curcuma* a čtyřmi zástupci blízce příbuzných rodů (*Hitchenia caulina, Monolophus scaposus, Stahlianthus involucratus, Stahlianthus campanulatus*).

Pro řešení těchto cílů byly vybrány dva typy molekulárních markerů – AFLP fingerprinting a klonování úseku ITS ribosomální DNA. Metodou AFLP bylo analyzováno 19 indických zástupců rodu *Curcuma* (celkem 115 jedinců), zatímco pro sekvenační analýzy bylo vybráno 24 druhů rodu *Curcuma*, 4 druhy representující příbuzné rody a 4 druhy byly vybrány jako outgroup (mimoskupina). Tento výběr druhů pokryl dva současně rozeznávané podrody rodu *Curcuma* (subg. *Curcuma* a subg. *Hitcheniopsis*), tři genomové skupiny (I, II, III) definované v podrodu *Curcuma* (Leong-Škorničková et al., 2007) a šest ploidních úrovní u indických zástupců rodu.

Rekonstrukce fylogenetických vztahů na základě analýz dvou nezávislých molekulárních markerů odhalily podobnou strukturu vztahů v rámci rodu. V rámci podrodu *Curcuma* (převážně indičtí zástupci rodu) byly odhaleny dvě hlavní vývojové větve. První z nich zahrnovala všechny zástupce genomové skupiny I a jeden druh ze skupiny II. Druhá vývojová větev zahrnovala všechny ostatní zástupce skupiny II a všechny analyzované zástupce skupiny III. Fylogenetické vztahy v rámci jednotlivých vývojových větví nebyly jednoznačně vyřešeny, ale je zřejmé, že mezidruhové vztahy jsou komplikovány allopolyploidizací, hybridizací a retikulární evolucí. Složitý charakter evoluce úseku ITS v rámci genomu byl brán v úvahu při interpretaci fylogenetických vztahů mezi studovanými jedinci a krátce diskutován. *Hitchenia caulina, Monolophus scaposus, Stahlianthus involucratus* a *Stahlianthus campanulatus* byly dle analýzy ITS jednoznačně podpořeny jako zástupci rodu *Curcuma* což je v souladu s výsledky dvou nezávislých studií Ngambriabsakul et al. (2004) a Leong-Škorničková et al. (2007)

Klíčová slova: *Curcuma*, Zingiberaceae, polyploidie, fylogeneze, retikulární evoluce, velikost genomu, molekulární markery, AFLP, ITS

Abstract

Economically important genus *Curcuma* is taxonomically critical polyploid complex. Taxa of various ploidy levels (2x, 6x, 9x, 11x, 12x and 15x) prefer different modes of reproduction (vegetative, sexual or both) and consequently exhibit different levels of morphological variability. With no recent revision and extremely complicated nomenclatoric issues incorrect identifications are very common. So far there is no suitable infrageneric classification. Morover, concept of the genus *Curcuma* is vague, as indication exists, that *Curcuma* is paraphyletic with several closely related genera (Kress et al. 2002). Based on ongoing taxonomic revision of Indian *Curcuma* and recent extensive cytological study, that have suggested naturall grouping of subgenus *Curcuma* based on differences in homoploid genome sizes, the goals of present study are: 1) to test the monophyletic character of „genome size groups“ sensu Leong-Škorničková et al. (2007) and outline phylogenetic relationships within and among these groups and 2) to assess taxonomic position of four *Curcuma*-like species often placed into separate genera: *Monolophus scaposus*, *Hitchenia caulina*, *Stahlianthus campanulatus* and *Stahlianthus involucratus*.

Two molecular markers, AFLP fingerprinting and ITS cloning and sequencing, were used. Total of 19 *Curcuma* species with representation of 115 individuals, were analysed by AFLP. By ITS cloning and sequencing 24 *Curcuma* species, 4 species belonging to *Curcuma*-related genera and 4 outgroup species were analysed. This sampling covered two currently distinguished subgenera of the genus *Curcuma* (subg. *Curcuma* and subg. *Hitcheniopsis*), three genome size groups defined within subg. *Curcuma* (Leong-Škorničková et al., 2007) and six ploidy levels detected within genus *Curcuma* in India.

Phylogenetic reconstructions based on two independent molecular markers revealed in general same pattern within genus *Curcuma.* Within subg. *Curcuma* two lineages that probably evolved independently were detected. First lineage corresponding with genome group I defined in (Leong-Škorničková et al., 2007), second lineage corresponding with genome group II and III, with only few exceptions, which are discussed. Phylogenetic relationships within particular lineages were not unambiguously resolved, however allopolyploidization, hybridization and reticulate evolution in *Curcuma* species were well documented. Complex character of evolution in ITS region and its implication for resolving *Curcuma* phylogeny is discussed. The four *Curcuma*-like species often classified to related genera (*Hitchenia caulina, Monolophus scaposus, Stahlianthus involucratus, Stahlianthus campanulatus*) were well supported to be a true *Curcuma* species. Therefore, previous suggestion by Ngambriabsakul et al.(2004) and by Leong-Škorničková et al. (2007), that these species should be classified as members of genus *Curcuma*, are justified by present study.

Key words: *Curcuma*, Zingiberaceae, polyploidy, phylogeny, reticulate evolution, genome size, molecular markers, AFLP, ITS

Curcuma aeruginosa Roxb. (reproduction from *Monandrian Plants*, Roscoe 1824-1828)

CONTENT

Acknowledgement....2
Abstrakt....3
Abstract....4
CONTENT....6
1. INTRODUCTION....7
2. Literature search....10
2.1. Polyploid genus *Curcuma* L....10
2.1.1 Taxonomic classification....10
2.1.2 Distribution and species richness....10
2.1.3 Ecology and habitat....11
2.1.4 Morphology....12
2.1.5 Polyploidy in genus *Curcuma* and family Zingiberaceae....14
2.1.6 Problematic questions in genus *Curcuma*....15
2.1.7 Molecular markers in *Curcuma*....17
2.2. Molecular markers used in present study....18
2.2.1 Amplified fragment length polymorphism (AFLP)....18
2.2.2 Sequencing of Internal transcribed spacer (ITS) of nrDNA....18
3. MATERIAL AND METHODS....21
3.1. Plant material....21
3.2. AFLP fingerprinting....24
3.2.1 Species in study....24
3.2.2 DNA extraction....24
3.2.3 AFLP analysis....25
3.2.4 Data analysis....26
3.3. ITS cloning and sequencing....27
3.3.1 Species in study....27
3.3.2 DNA extraction and PCR reaction....27
3.3.3 Cloning and sequencing....28
3.3.4 Data analysis....28
4. RESULTS....31
4.1. AFLP fingerprinting....31
4.1.1 Characteristics of AFLP data....31
4.1.2 Phylogenetic relationships....33
4.2. ITS cloning and sequencing....38
4.2.1 Characteristics of ITS data....38
4.2.2 Phylogenetic relationships....41
4.3. Comparison of results from AFLP fingerprinting and ITS sequencing....50
5. DISCUSSION....53
5.1. Phylogenetic relationships within genus *Curcuma*....53
5.1.1 ITS polymorphism in *Curcuma* species....54
5.1.2 Subg. *Hitcheniopsis*....57
5.1.3 Subg. *Curcuma:* Genome size groups in the light of molecular phylogeny....58
5.1.4 Hypothesis about evolution of *Curcuma* species....63
5.2. Taxonomic position of *Curcuma*-like species....64
6. Conclusions....65
7. References....67
8. Apendix....73

1. INTRODUCTION

The genus *Curcuma* L. (Zingiberaceae) is a diverse polyploid complex containing many taxa of economic, medicinal, ornamental and cultural importance. It is distributed throughout South and South-east Asia with a few species extending to China, Australia and the South Pacific. The highest diversity is concentrated in India and Thailand, with more than 40 species in each area, followed by Burma, Bangladesh, Indonesia and Vietnam (Leong-Škorničková et al., 2007). Despite the considerable economical potential of this genus, its phylogeny and taxonomy are still poorly understood mainly due to extensive polyploidization and hybridization. Polyploidy has played a significant role in evolution and diversification of various members of Zingiberaceae (e.g. Mukherjee, 1970; Lim, 1972a,b; Poulsen, 1993; Chen and Chen, 1984; Takano, 2001; Takano and Okada, 2002). Particularly in *Curcuma* various ploidy levels (2x, 6x, 9x, 11x, 12x and 15x) developed various modes of reproduction and exhibit different levels of morphological variability, which blurs species boundaries. Moreover, Škorničková and Sabu (2005b) and Škorničková et al. (2007) observed in the field that some species may hybridize in the wild. Unclear nomenclature also contributes to taxonomic complexity of the group (Škorničková, 2007). Several authors have attempted to propose infrageneric classification of *Curcuma* (Roxburgh, 1810; Horaninow, 1862; Baker, 1890; Valeton, 1918; Velayudhan & al., 1996), which were mostly based on morphological characters. According to Škorničková (2007) none of the previously proposed infrageneric classifications is, however, suitable. For practical reasons this study followed infrageneric classification according to Schumann (1904), where two subgenera of genus *Curcuma* are recognized based on presence or absence of anther spurs.

Previous attempts to establish infrageneric classification of genus *Curcuma* or even infer phylogenetic relationships between individual species based on morphological characters only failed. Thus other markers have to be applied to reconstruct the phylogeny of the genus *Curcuma*. The application of molecular markers is one of the most promising ways to achieve this aim. In spite of several molecular studies dealing with species identification or particular evolutionary relationships within genus *Curcuma* in the past decade, (Apavatjrut et al., 1999; Chen et al., 1999; Nayak et al., 2006; Cao et al., 2001; Cao and Komatsu, 2003), no major works were published on phylogeny of *Curcuma*. Two recent phylogenetic studies on the higher taxonomic level that included *Curcuma* species were published. The phylogeny of family Zingiberaceae (Kress et al., 2002) as well as phylogeny of tribe Zingiberae (Ngamriabsakul et al., 2004) suggests that the genus *Curcuma* as accepted nowadays is paraphyletic with several *Curcuma*-like genera (e.g. *Hitchenia*, *Stahlianthus* and *Smithiatris*, for details see chapter 2. Literature search) . Škorničková (2007) suggested, that also some

representatives from genera *Paracautleya, Kaempferia* and *Monolophus* are likely to belong to the genus *Curcuma* and their taxonomic position should be studied in more details.

Genome size has been repeatedly found to be a useful taxonomic marker at various ranks and in various plant groups (e.g. Doležel et al. 2007; Zonneveld, 2001). A recent study by Leong-Škorničková et al., (2007) focused on chromosome counts and analysis of genome size in Indian *Curcuma.* Five different ploidy levels were detected among about fifty Indian species from subgenus *Curcuma* (x = 7, 6x = 42, 9x = 63, 11x = 77, 12x = 84, 15x = 105) and one diploid species from subgenus *Hitcheniopsis* (x = 11, 2x = 22). Analysis of C-values (genome sizes) and 1Cx-values (homoploid genome sizes) revealed only a few species-specific C-values, but close evolutionary relationships among some cytotypes were suggested based on the similarity in homoploid genome sizes and geographical grouping and three genome size groups in the subgenus *Curcuma* were proposed. These groups were followed in present study as the sampling partially overlapped with previous study. Genome size group I contains hexaploids and higher polyploids with 1Cx-values in range 0.30 – 0.33 pg, genome size group II contains hexaploid species only with 1Cx values in range 0.36 – 0.39 pg and genome size group III covers one hexaploid and one endecaploid species with 1Cx in range 0.41 – 0.43 pg. Species belonging to subgenus *Hitcheniopsis* generally exhibit different values of homoploid genome size in comparison with subgen. *Curcuma* species (see details in chapter 2. Literature search). However, monophyletic character of these groups has not yet been tested by any molecular marker.

To solve phylogenetic relationships of closely related species, sufficiently variable molecular markers have to be employed. In *Curcuma* species the genetic structure is only poorly known and thus markers where no prior sequence information is needed to perform the analysis, are preferred. In preliminary study two non-coding cpDNA regions (*trn*L-*trn*F, *trn*K-*mat*K), which are often used to solve phylogenies on generic and infrageneric level (Palmer, 1985), only limited set of informative variable characters were found (Fér et al., unpublished data).

AFLP marker (Vos et al., 1995) was successfully used in systematic studies where DNA sequence variation was limited (e.g. Despres et al., 2003, Guo, 2005). Therefore, this marker was chosen for resolving the phylogenetic relationships of Indian *Curcuma* species and to uncover molecular origin of genome size groups in present study.

Internal transcribed spacer (ITS) region of nuclear ribosomal DNA was chosen as an additional source of genetic information to solve phylogenetic relationships within genus *Curcuma*, to uncover relationships between *Curcuma* and four species belonging to related genera, and to shed light on evolutionary phenomenons that influenced complex evolutionary history of polyploid *Curcuma* species. As rDNA region is biparentally inherited, it is useful for revealing past cases of reticulation, hybrid speciation, and parentage of polyploids (Alvarez and Wendel, 2003). Additionally, ITS marker (usually in combination with another marker e.g. *trn*L-*trn*F or *mat*K) was successfully used in resolving phylogenetic relationships

in several groups within Zingiberaceae (e.g. within tribes *Hedychieae*, Searle and Hedderson, 2000; *Zingibereae*, Ngamriabsakul et al., 2004; *Globbeae*, Williams et al., 2004; or within large genera *Hedychium*, Wood et al., 2000; *Alpinia*, Rangsiruji et al., 2000b, Kress et al., 2005; *Amomum*, Xia et al., 2004; *Etlingera*, Pedersen, 2004). Phylogeny of entire family Zingiberaceae (Kress et al., 2002), as is accepted nowadays, is based on analysis of nrDNA ITS and cpDNA *matK* regions as well.

This study is divided into two parts. First part focuses on the phylogenetic analysis of selected members of the genus *Curcuma.* Two different molecular markers - AFLP and sequencing of nrDNA ITS region - were selected to answer if three genome size groups defined within subgenus *Curcuma* are of monophyletic origin and if so, could be homoploid genome size suitable taxonomic marker for infrageneric classification of genus *Curcuma?* And secondly, what are the phylogenetic relationships between species within and among these groups and which evolutionary phenomenons influenced evolutionary history of polyploid *Curcuma* species the most. Second part of the study employed sequencing of ITS nuclear ribosomal DNA region to answer if particular species classified as members of related genera *Hitchenia*, *Stahlianthus* and *Monolophus* are true *Curcuma* species or just *Curcuma*-pretenders?

2. Literature search

2.1. Polyploid genus *Curcuma* L.

2.1.1 Taxonomic classification

kingdom: Plantae
subkingdom: Tracheobionta
division: Magnoliophyta
class: Liliopsida
order: Zingiberales
family: Zingiberaceae
genus: *Curcuma*

Based on new classification of family Zingiberaceae (Kress et al., 2002) where four subfamilies and six tribes were recognized, *Curcuma* falls into tribe *Zingibereae*.

2.1.2 Distribution and species richness

The genus *Curcuma* L. is one of the largest genera of family Zingiberaceae. It is naturally distributed throughout S and SE Asia with a few species extending to China, Australia, and the South Pacific. The highest diversity is in India and Thailand, with about 40 species each, followed by Myanmar, Bangladesh, Indonesia and Vietnam. The true number of species in the genus is not known, although rough presumptions in past 30 years varied anywhere from c. 50 (Smith, 1981; Branney, 2005), 80 (Larsen et al., 1998) to 100 (Sirirugsa, 1996), while Škornickova et al., (2004) suggests, that the number will probably reach 120 as recent explorations in India and SE Asia proceeds (According to Škorničková 2007).

2.1.3 Ecology and habitat

In their native range, monsoonal Asia, *Curcuma* species are an important component of the understory semi-shaded areas like primary and secondary forests margins, shrubby formations, teak forests, plantations, coconut and arecanut grooves, riverbanks or rarely also in bamboo dominated forests and open grasslands. Numbers of them are the first pioneers in disturbed areas such as roadsides and ditches. The ploidy level is quite closely linked to their preffered habitat and while diploid species are usually found in primary kind of habitats, polyploids are found in secondary and disturbed areas. Several species are cultivated. *Curcuma* species are commonly found in lowlands, but can be found in their natural habitats also in higher altitudes up to 1500 m (According to Škorničková 2007).

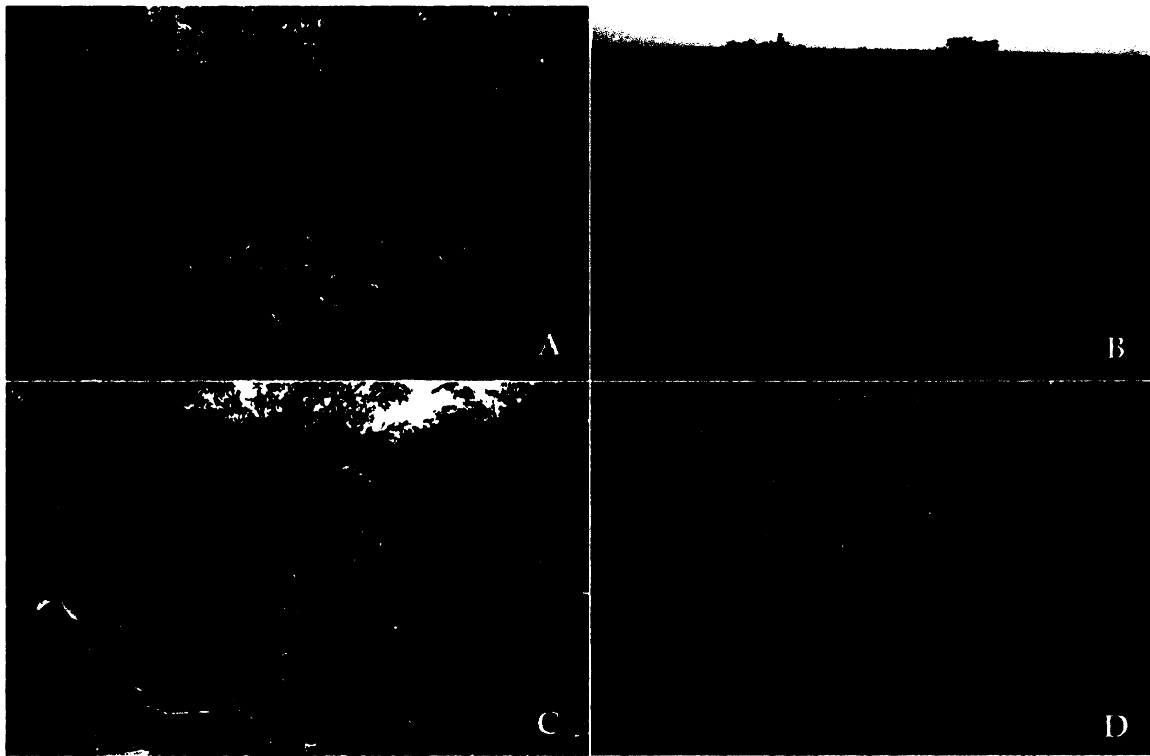


Fig. 1. Examples of *Curcuma* habitats. A. Dipterocarp forest, Chhattisgharh, Central India, *C. montana*. B. Exposed lateritic slopes, Udupi, South India, *C. bhatii*. C. Shrubby semideciduous slopes, Malwan, West India, *C. decipiens*. D. Open spaces, N. Bongaigaon, North-east India, *C. zedoaria*. (from Škorničková, 2007)

2.1.4 Morphology

The genus *Curcuma* consist of perrenial rhizomatous herbs with leafy shoots dying back during the dry period of monsoonal areas. Most are medium sized c. 0.5 to 1.5 m tall, but the smallest species are just about 10 – 20 cm, while the statliest can reach 3.5 m. Characteristic organs of *Curcuma* species are rhizomes, root tubers, pseudostem and inflorescence of highly derived flowers (Endress, 1994; Kress et al., 2002) (see Fig. 2).

The rhizome characters, as a branching (either simple ovoid or branched) or internal colour are considered as quite specific and are a good determination character for polyploid species (see Fig. 3). Root tubers are present in all *Curcuma* species and are placed at the end of the roots. They contain mostly starch and sustain the plant throughout the dry period. A pseudostem (false stem) is formed by closely embracing leaf sheaths and it is well developed in all polyploids, while in some diploids the sheets can be so loose, that the pseudostem appears to be missing. The leaves are either sessile or petiolate, having a sheath basally. Some species have a purple or violet patch on the upper side of the lamina. Its size, shape, relative placement on the lamina, color, and density is helpful in the determination of polyploid species. It is, however, variable in some diploid species. Inflorescence in *Curcuma* is always terminal (as in other Zingiberaceae, Kiew, 1977) and its relative position to the plant is frequently used for infrageneric classification. It arises either on the vegetative leafy shoot (called as inflorescence central) or on a separate non-leafy shoot (called as inflorescence lateral). Many of the seed-setting *Curcuma* species are capable of blooming twice — once in the beginning of the season (laterally) and if the monsoon rains are sufficient also centrally later in the summer. The inflorescence in *Curcuma* is conspicuous. Bracts in the lower portion of the inflorescence are usually less colored, but fertile and enclosing 2 to many flowers (rarely only one flower), the top bracts are usually much larger, longer, and more brightly colored and called coma. It is believed that they serves the purpose of attracting pollinators. The shape of the inflorescence and colour of the bracts is stable character in polyploids, but may vary tremendously in seed setting taxa. *Curcuma* flower consist of calyx, floral tube, corolla lobes, lateral staminodes and labellum (both of these are transformed anthers), stamen and pistil, epigynous glands and ovary (see flower dissection in Fig. 2). Anther and the various appendages (crest, spurs, lateral appendages) have been considered as source of reliable characters in the family as well as within the genera by many previous workers. The phylogeny of the genus *Globba* by Williams (2040) clearly showed, that appendages do matter. Also in *Curcuma*, the overall structure and shape of anther, presence and absence of anther spurs and anther crest are valuable diagnostic characters (see Fig 3). Further, presence or absence of epigynous glands (epigynous nectaries, stylodial glands) have been considered as a good character for infrageneric classification of *Curcuma*.

In *Curcuma* the number of epigynous glands is always two and they are present in majority of the species. They are absent in many of the SE Asian species Schumann (1904) placed in the subgenus *Hitcheniopsis*. They are present in all but one Indian *Curcuma* species (*C. vamana*). (According to Škorničková 2007).

Fig. 2. *Curcuma* plant (*Curcuma zedoaria* (Christm.) Roscoe): A. Plant habit; B. Flower in fertile bract (front view); C. Flower in fertile bract (side view); D. Rhizome; E. Flower dissection; Based on Škorničková.

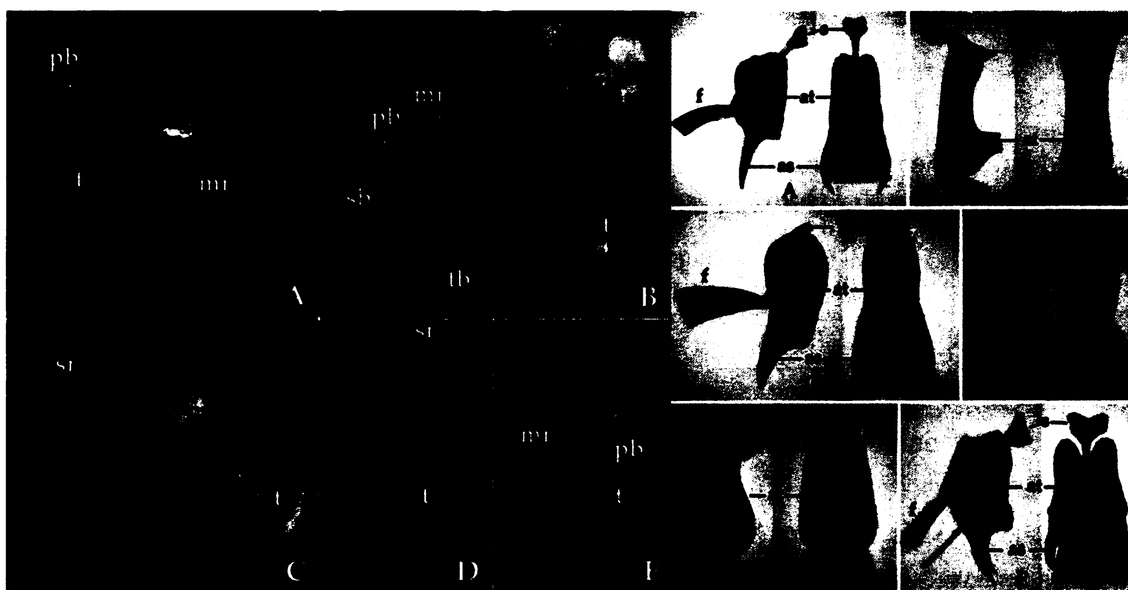


Fig. 3. *Curcuma* rhizomes and anthers – valuable characters for species determination. Rhizomes on the left: A. branched rhizome with creeping branches and sessile root tubers (*C. rubrobracteata)*; B. branched rhizome with short and stout branches and root tubers distanced from branches (*C. amarissima* aff.); C. simple ovoid rhizome, root tubers distanced from the rhizome (*C. angustifolia*); D. simple ovoid rhizome, root tubers arranged serially on one root (*Curcuma* sp.); E. branched rhizome with sessile root tubers (*C. bhatii*). Legend: mr – main rhizome, sr – simple rhizome, pb – primary branch, sb – secondary branch, tb – tertiary branch, t – tuber. Anthers of *Curcuma* spp. on the right. A. *C. codonantha*. B. *C. roscoeana*. C. *C. rubrobracteata*. D. *C. vamana*. E. *C. aurantiaca*. F. *C. reclinata*. Legend: ac – anther crest, as – anther spurs, at – anther thecae, c – connective, f – filament, s – stigma. (from Škorničková, 2007)

1.1.1 Polyploidy in genus *Curcuma* and family Zingiberaceae

Within Zingiberaceae, polyploid complexes are not rare. They have been detected mostly in the subfamily Zingiberoideae (sensu Kress et al., 2002) e.g. in the genera *Globba* (Lim, 1972a,b; Takano, 2001; Takano and Okada, 2002), *Curcuma* (e.g. Apavatjrut et al., 1996; Ardyiani, 2002; Joseph et al., 1999; Sirisawad et al., 2003), *Hedychium* (Mukherjee, 1970; Chen and Chen, 1984), *Boesenbergia* (Poulsen, 1993). Goldblatt (1980) assumes, that the families of the Zingiberales, which all have high haploid numbers, may be derived from polyploid ancestors. Producing of unreduced gamete and consequent arisal of autopolyploids has been observed in monocots (e.g. Refoufi et al., 2001; Brandizzi and Caicola, 1998) as well as in *Globba*, Zingiberaceae (Takano and Okada, 2002).

According to recent study by (Leong-Škorničková et al., 2007) in Indian *Curcuma* species, five different ploidy levels were detected among about fifty Indian species from subgenus *Curcuma* ($x = 7$, $6x = 42$, $9x = 63$, $11x = 77$, $12x = 84$, $15x = 105$) and one diploid species from subgenus *Hitcheniopsis* ($x = 11$, $2x = 22$). Based on literature records, not all

somatic chromosome numbers of *Curcuma* species seem to fit into the series of x = 7 or x = 11. Škorničková (2007) summarized that several Thai *Curcuma* species belonging to the subgenus *Hitcheniopsis* exhibit various somatic chromosome numbers (2n = 20, 24, 28, 32, 34, 36 or 56). All these species are also morphologically different from the *Curcuma* species of Indian subcontinent (e.g. lack of anther spurs and epigynous glands). It is very likely, that their chromosomal evolution was going on separately from the 'Indian' group of species and that aneuploidy plays role at least in some taxa in the subgenus.

2.1.6 Problematic questions in genus *Curcuma*

Critical problem in *Curcuma* is the identities of many species, including the names so commonly used in everyday life (e.g. *C. longa, C. aromatica*). As most species are quite variable, but many look alike, often there is one name applied to many different species or many names for one species. The types are not yet designated for most of *Curcuma* species (Škorničková, 2007). The absence of a complete revision causes great confusion among taxonomists (Velayudhan & al., 1996).

Different levels of morphological variability in various *Curcuma* species was observed in the field and seem to correlate to the ploidy level, rhizome architecture and possible mode of reproductions as elaborated in Škorničková (2007). In summary, all sterile species (higher polyploids, 2n = 63, 105) have branched rhizomes and exhibit low variability. This is not surprising as they can reproduce exclusively vegetatively. Various seed-setting (mainly diploids and hexaploids, 2n = 22, 42, 77) species however exhibit any of the three different levels of variability and either of two rhizome architectures. Seed-setting species with branched rhizomes usually exhibit medium to low level of variability, while seed-setting species with unbranched (simple ovoid) rhizomes exhibit mostly high to medium variability. This makes sense as the latter rely on sexual reproduction thus explaining their higher variability, while seed-setting species with branched rhizomes reproduces either generatively or vegetatively. It is plausible to hypothesize that vegetative reproduction evolved as a consequence of disturbed sexual reproduction of the hybridogenous taxa.

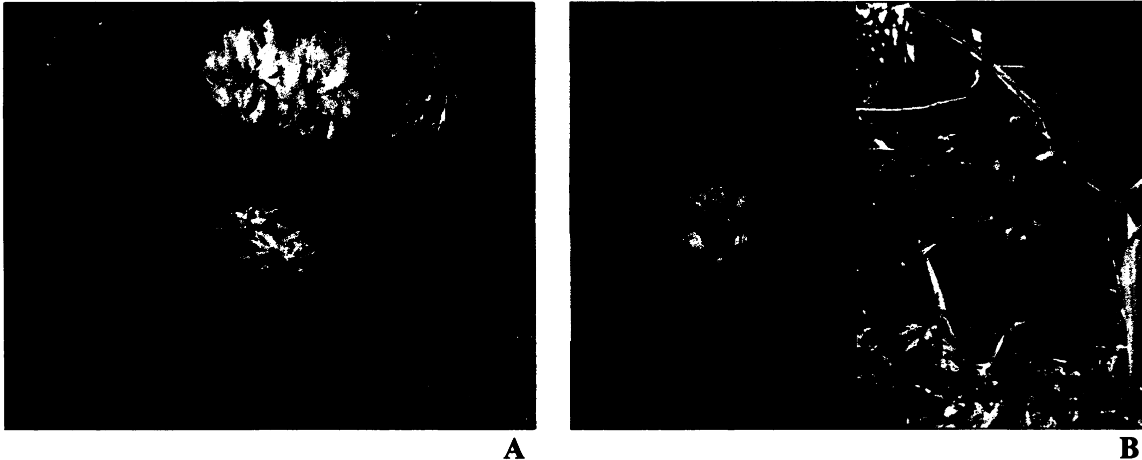


Fig. 4: Level of within-population variability of *Curcuma* species is connected to prefered mode of reproduction and ploidy level. A. High morphological variability of sexually reproducing hexaploid species. On the picture, single population of hexaploid *C.*aff. *prakasha* (Meghalaya, NE India); B. Low morhological variability of vegetatively reproducing higher polyploids. On the picture two populations of nonaploid species *C. aromatica*, which are distanced more than 2000 km (B1 – Kerala, S India; B2 – Meghalaya, NE India).

An uneven level of morphological variability in various species blurs not only specific boundaries but complicate also the matter of infrageneric classification. Most authors who ever attempted the infrageneric classification of *Curcuma*, employed position of the inflorescence (Roxburgh, 1810; Horaninow, 1862; Baker, 1890; Valeton, 1918; Velayudhan et al., 1996), which is unfortunately ambiguous character in many species. The infrageneric classification according to Schumann (1904), where two subgenera are recognized based on presence or absence of anther spurs (subgen. *Curcuma* – anther spurs present, subgen. *Hitcheniopsis* – anther spurs absent) was followed in cytological study Leong-Škorničková et al. (2007). In this study it was mentioned, that more suitable character for differentiation subgenus *Curcuma* and *Hitcheniopsis* is presence of two floral epigynous glands. This hitherto neglected character (well developed in subgenus *Curcuma* but absent in subgenus *Hitcheniopsis*) seems to be pivotal for the updated subgeneric delimitation, better reflecting the current state of knowledge than did previous taxonomic concepts (J. Leong-Škorničková et al., unpubl. res.).

A recent phylogenetic study of family Zingiberaceae based on analysis of nrDNA sequences and plastid *mat*K regions (Kress et al. 2002) suggests that the genus *Curcuma* as accepted nowadays is paraphyletic with *Hitchenia*, *Stahlianthus* and *Smithiatris,* which also share cone-like inflorescences of few flowered, congested bracts and that genus *Curcuma* itself contains three groups of species. Similarly, study of Ngamriabsakul et al. (2004) dealing with the phylogeny of tribe Zingiberae and using the same nuclear gene as Kress et al. (2002) but a different chloroplast gene, *trn*L-*trn*F revealed, that *Curcuma*-like genera (*Hitchenia*, *Paracautleya*, *Smithatris* and *Stahlianthus*) are actually well nested within *Curcuma* complex and may be regarded as a single genus. The morphological characters that

support the separation of each taxon (genus) (see Fig. 5) are according to authors, autapomorphic. Škorničková (2007) suggested, that also some species from genera *Kaempferia* and *Monolophus* are likely to be members of the genus *Curcuma* and should be study in more details.

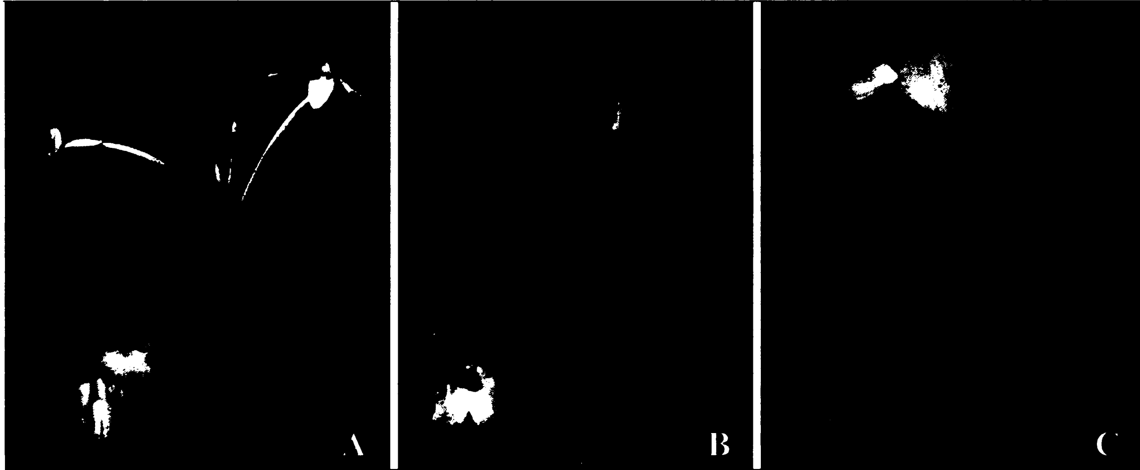


Fig. 5: Morphological differences of *Curcuma*-like species, often placed to related genera (three of four representatives analysed in this study). A. *Monolophus scaposus*; B. *Hitchenia caulina*; C. *Stahlianthus involucratus*.

1.1.2 Molecular markers in *Curcuma*

Over the past decade, several molecular techniques have been developed to provide information on diversity and genetic relationships. Apavatjrut et al. (1999) confirmed efficacy of isozymes as a tool or species identification of seven highly similar *Curcuma* taxa of the early flowering group, while Paisooksantivatana et al. (2001a,b) used isozymes as a tool to estimate the genetic diversity in *C. alismatifolia.* Chen et al. (1999) used RAPD analysis to investigate the relationship between two Chinese *Curcuma* species. Islam et al. (2004, 2005, 2007) employed RAPD analyses to investigate the level of genetic diversity within and between natural populations of *Curcuma zedoaria* in Bangladesh. Nayak et al. (2006) combined the RAPD and genome size estimation to evaluate genetic variation in 17 cultivars of turmeric (*C. longa*). Nucleotid sequencing of 18S r RNA and *trn*K was proposed by Cao et al. (2001) and Cao and Komatsu (2003) as a potential molecular identification tool to recognize six medicinally important *Curcuma* species from Szechuan, China. However, so far, there are no major published works focusing on phylogeny of Indian *Curcuma.*

2.2. Molecular markers used in present study

2.2.1 Amplified fragment length polymorphism (AFLP)

Amplified Fragmnet Length Polymorphism is a restriction fragment based method, that compare polymorphic bands aquired across whole genome (Robinson and Harris, 1999) and is a tool for effective resolving species-level phylogenies in clades that have diversified relatively recently (Beardsley, 2003; Koopman, 2005). Due to the high number of polymorphisms per assay unit, the high reproducibility, and the fact that no prior sequence information is needed to perform the AFLP technique, its popularity in systematic studies has increased rapidly over the past few years (Koopman, 2005). Although occurence of problems with nonindependence of fragments and/or the problem of identifying homologous fragments that potentially limits the phylogenetic interpretation of restriction fragment data is unquestionable (e.g. Karp et al., 1996), many studies proved that the influence of these limitations is minimized if analysed data sets are of closely related species (Koopman, 2005; O'Hanlon and Peakall, 2000).

In systematic studies using AFLPs a variety of techniques have been applied for data analysis (Koopman 2005). Neighbor joining (Hodkinson et al., 2002) as well as heuristic parsimony (Kardolus et al., 1998; Hodkinson et al., 2000; Despres et al., 2003) are commonly used. Although application of parsimony methods for AFLP data is not undisputed, Koopman (2005) have shown that Neighbor-joining analyses of AFLP data generally show a similar picture as maximum parsimony analyses at least at inter-specific level/or between closely related species. Combination of Neighbor joining and Maximum parsimony analysis for AFLP data were effectively used e.g. in Schönswetter & Tribsch 2005. For visualization of potentially conflicting signals, such as may be caused by homoplasy or hybridization, network methods using e.g. NeighbourNet algorithm (Bryant and Moulton, 2004) can be used in contrast to commonly used treebuilding methods. This method was succesfully applied to AFLP data by Weiss-Shneeweiss et al. (2008).

2.2.2 Sequencing of Internal transcribed spacer (ITS) of nrDNA

The internal transcribed spacer (ITS), region of the 18S–5.8S–26S nuclear ribosomal cistron, is one of the most popular sequences for phylogenetic inference at the generic and infrageneric levels in plants (Alvarez and Wendel, 2003, Baldwin et al., 1995) for at least three reasons. Firstly, nuclear ribosomal genes are constituents of individual 18S–5.8S–26S repeats, which are in hundreds to thousands repeats in plant genomes, therefore they are more easily isolated than most low-copy nuclear loci. Secondly, White et al. (1990) described a set of primers that are useful for amplifying ITS sequences from most plants and fungi, therefore

there is no need for primer design or prior sequence knowledge. Finally, nuclear rDNA regions are biparentally inherited, which is valuable property for revealing past cases of reticulation, hybrid speciation, and parentage of polyploids (Alvarez and Wendel, 2003). Although ITS sequence data provided many important contributions of phylogenetic understanding and knowledge of genome relationships in several plant groups, in another groups of taxa can be phylogenetic inferences misleading, especially when taxa of hybrid or allopolyploid origin are involved in the study. Number of molecular genetic processes impact ITS sequences evolution after formation of allopolyploids or hybrids, when two diverged genomes become united in a common nucleus (Alvarez and Wendel, 2003, 2005). Most genetic changes in allopolyploids involve homogenization of rRNA genes (a process called concerted evolution) whereby one particular ITS repeat type overwrites pre-existing repeat type (Dadejová et al., 2007) and subsequently only one of the parental ITS type are detectable in a allopolyploid genome (observed e.g. in *Gossypium* by Wendel et al., 1995). Second possible fate of divergent ITS types within allopolyploid genome is that divergent repeat types are maintained, evolving independently without recombination or interarray ''contact''. Thus both parental rDNA types are detectable in a genome (observed e.g. in *Tragopogon* by Soltis and Soltis, 1991 and Soltis et al., 1995). Another possible outcome of the reunion of divergent sequences following hybridization is that two or more repeat types are maintained but undergo various degrees of recombination (Alvarez and Wendel, 2003). This leads to chimeric ITS sequences that in phylogenetic analysis can behave erratically and resolve in phylogenetic position basal to either parental lineage (McDade, 1992, 1995). These phenomena separately and collectively create a network of paralogous sequence relationships potentially confounding accurate phylogenetic reconstruction. However, ITS sequence data have and may continue to provide insights into phylogenetic history, polyploid ancestry, genome relationships, historical introgression, and other evolutionary questions (Wissemann, 2002), as long as the data generated will be interpreted in light of above-mentioned molecular genetic processes.

DNA sequence data are most commonly analysed by maximum parsimony method. Alternative to parsimony, the maximum likelihood and Bayesian approach to phylogenetic inference are frequently used (Felsenstein 1988) With maximum likelihood method in practice the computational time is serious problem (Nei, 1996). On the contrary, Bayesian phylogenetic analyses are now very popular in systematics and molecular evolution because they allow the use of much more realistic models than currently possible with maximum likelihood methods (Lewis et al., 2005). In addition, distance based methods like a Neighbour-joining or Neighbour network algorithm (Bryant and Moulton, 2004) can be used for sequence data as well (Nei, 1996). Neighbour network analysis of ITS data were used e.g. for visualization of reticulate relationships within genus *Platanus* (Grimm and Denk, 2007).

Combination of AFLP fingerprinting and sequencing of ITS region were used previously by numerous authors for resolving intrageneric phylogenies. Guo et al. (2004) have infered basic phylogenetic relationship within problematic polyploid group of *Achillea* by ITS and cp DNA sequencing. Based on this analysis, they revealed, that hybrid and reticulate evolution, even on the diploid level, has been ongoing in *Achillea* since the origin of the genus. Following study by AFLP marker by Guo et al. (2005) revealed in more details mosaic genetic structure of hybrids and polyploids. Hodkinson et al.(2002) confirmed allotriploid origin of *Mischantus* species from Poaceae by using sequencing and cloning of ITS region, sequencing of cpDNA and AFLP analysis. Complex evolutionary information contained in internal transcribed spacer (ITS) sequences were well utilized also for infering evolution of genus *Platanus* (Grimm and Denk, 2008).

3. MATERIAL AND METHODS

3.1. Plant material

Total of 33 taxonomic entities were included in this study. From this 25 species belonged to genus *Curcuma* (22 assigned to species, 3 determined only tentatively according to Škorničková 2007), four species were selected as representatives of related genera (*Monolophus scaposus*, *Hitchenia caulina*, *Stahlianthus campanulatus* and *Stahlianthus involucratus*) and four species figured as an outgroup (*Hitchenia careyana, Globba* sp., *Alpinia conchigera, Zingiber capitatum*). Selection of the taxa followed in general sampling in recent extensive cytological study, as taxa included in this study are comprehensively characterized by genome sizes, chromosome numbers and ploidy levels (Leong-Škorničková et al., 2007). The samples used in this study were collected by Jana Leong-Škorničková in the wild on the Indian Peninsula (Fig. 6), Sri Lanka, China and Vietnam, often at or near the type localities of corresponding names. All taxa were determined by a specialist working on the revision of genus *Curcuma* (Škorničková et al., 2003a, b, 2004, 2005a, b, c, 2007) and taxonomic names followed Škorničková (2007). For practical reason this study used infrageneric classification based on Schumann (1904) and adjusted by Leong-Škorničková et al. (2007) (Subg. *Curcuma* – anther spurs and epigynous glands present, Subg. *Hitcheniopsis* – anther spurs and epigynous glands absent). Indian species from subgenus *Curcuma* selected for analysis well covered all three genome size groups defined in Leong-Škorničková et al.(2007) (Fig. 7).

Parts of young leaves were collected than dried in silica gel. Herbarium vouchers are deposited in CALI, with duplicates in MH and SING; incomplete sets are also kept in CAL, K and PR; vouchers of *C. oligantha* from Sri Lanka are deposited in PDA and SING. In addition, extensive photographic documentation of living material is available at SING.

Two different methods (AFLP fingerprinting and ITS sequencing) were used to resolve questions outlined in the beggining of this study. Sampling of taxa differed slightly for each method. Table 1 summarizes key characteristics for all taxa in the study and representation of individiuals in the two datasets. Each collection number indicates one population of a taxon at the locality. For AFLP study 2-10 individuals per population were analysed, while in ITS one individual per population were analysed, and subsequently 4-23 clones per individual were sequenced.

Fig. 6. Localities of Indian *Curcuma* populations under the study (blue dots) with major centres of diversity of Indian *Curcuma* species (in red). (from Škorničková, 2007)

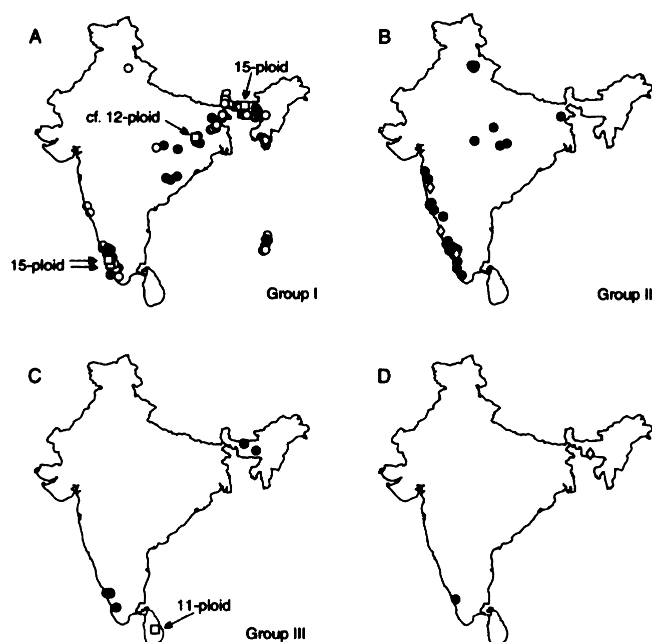


Fig. 7. Geographical origin of Indian *Curcuma* plant samples according to definitions of genome size groups. A. Species from genome group I (1Cx = 0.30–0.33 pg) and x = 7; B. species from genome group II (1Cx = 0.36–0.39 pg) and x = 7; C. species from genome group III (1Cx = 0.41–0.43 pg) and x = 7; D. species with 1Cx.0.83 pg and x = 11 (circle, *Curcuma vamana*; diamond, *Stahlianthus involucratus*). Symbol explanation (unless otherwise indicated): closed circles, hexaploids; open circles, nonaploids; open diamonds, *Curcuma*-like species often placed into separate genera; open squares, high polyploids designated by a corresponding ploidy level. (from Leong-Škorničková et al., 2007)

Table 1. Overview of *Curcuma* species in study. Collection number indicate one population of a species at the locality; basic chromosome number (x), number of somatic chromosomes (2n) and ploidy level are given for majority of populations; species of subg. *Curcuma* are classified into genome size groups: group I (1Cx = 0.30–0.33 pg), group II (1Cx = 0.36–0.39 pg), group III (1Cx = 0.41–0.43 pg) (Leong-Škorničková et al., 2007); mode of reproduction, which play important role in genus *Curcuma*, is presented for particular species (Škorničková, 2007); AFLP (N), sample size in AFLP study; ITS (N)/clones, sample size/number of clones sequenced per sample in ITS analysis; locality of origin, ex Indian localities are in small caps;

Collection number	Species	x; 2n	mode of reproduction	locality of origin	AFLP (N)	ITS (N)/clones
	Curcuma*, subgen *Curcuma	x = 7				
	Hexaploid Group I					
84183	*C. aromatica*	2n = 42	sex/veg	Kerala	3	1 / 8
84170	*C. aromatica*	2n = 42	sex/veg	Kerala	-	1 / 5
71484	*C. montana*	2n = 42	sex/veg	Jharkhand	4	1 / 8
73479	*C. montana*	2n = 42	sex/veg	Chhattisgarh	-	1 / 13
73425	*C. montana*	2n = 42	sex/veg	Orissa	-	1 / 5
73474	*C. montana*	2n = 42	sex/veg	Chhattisgarh	3	-
71443	*C.* aff. *prakasha*	2n = 42	sex/veg	Meghalaya	-	1 / 8
71476	*C.* aff. *prakasha*	2n = 42	sex/veg	W. Bengal	-	1 / 5
	Nonaploid Group I					
71431	*C. aeruginosa*	2n = 63	vegetative	Assam	9	1 / 20
84142	*C. aeruginosa*	2n = 63	vegetative	Kerala	7	1 / 7
84119	*C. aeruginosa*	2n = 63	vegetative	Kerala	10	-
71492	*C. aromatica*	2n = 63	vegetative	Jharkhand	-	1 / 9
71489	*C. leucorhiza*	2n = 63	vegetative	Jharkhand	4	-
84160	*C. longa*	2n = 63	vegetative	Kerala	4	-
73222	*C. longa*	2n = 63	vegetative	SRI LANKA	5	1 / 15
84182	*C. zanthorrhiza*	2n = 63	vegetative	Kerala	3	1 / 5
	12x Group I					
71485	C. sp. 'ranchi'	2n > 70	vegetative	Jharkhand	4	1 / 8
	15x Group I					
71432	*C. raktakanta*	2n=105	vegetative	Assam	5	1 / 23
84120	*C. raktakanta*	2n=105	vegetative	Kerala	5	-
	Hexaploid Group II					
73453	*C. angustifolia*	2n = 42	sexual	Uttaranchal	5	1 / 5
73480	*C. angustifolia*	2n = 42	sexual	Chhattisgarh	4	1 / 16
84143	*C. cannanorensis*	2n = 42	sexual	Kerala	4	1 / 15
73403	*C. inodora*	2n = 42	sexual	Maharashtra	5	1 / 15
73402	*C. pseudomontana*	2n = 42	sexual	Maharashtra	5	1 / 6
73477	*C. reclinata*	2n = 42	sexual	Chhattisgarh	2	1 / 6
73467	*C. sulcata*	2n = 42	sexual	Mad. Prad.	4	1 / 9
84155	*C. ecalcarata*	2n = 42	sexual	Kerala	4	1 / 5
73446	*C. bhatii*	2n = 42	sexual	Karnataka	5	1 / 10
84157	*C. neilgherensis*	2n = 42	sexual	Kerala	-	1 / 5
	Hexaploid Group III					
84145	*C. mutabilis*	2n = 42	sexual	Kerala	4	1 / 8
	11x Group III					
73223	*C. oligantha*	2n = 77	sex/veg	SRI LANKA	4	1 / 8

Table 1. Part 2.

Collection number	Species	x; 2n	mode of reproduction	locality of origin	AFLP (N)	ITS (N)/clones
	Curcuma*, subgen *Hitcheniopsis					
84156	*C. vamana*	2n = 22	sex/veg	Kerala	3	1 / 13
20	*C. alismatifolia*	2n = 36	sexual	VIETNAM	-	1 / 10
	***Curcuma*, subgenus** Not yet defined	**unknown**				
19	*C. pierreana*	**unknown**	sex/veg	VIETNAM	-	1 / 8
	Related genera I	**x = 7**				
86407	*Monolophus scaposus*	2n = 42	sexual	Maharashtra	-	1 / 4
84178	*Hitchenia caulina*	2n = 42	sexual	Maharashtra	-	1 / 8
	Related Genera II	**x = 11**				
C086	*Stahlianthus campanulatus*	2n = 22	sex/veg	CHINA	-	1 / 8
71449	*Stahlianthus involucratus*	2n = 22	sex/veg	Meghalaya	-	1 / 5
	Outgroup species					
C073	*Hitchenia careyana*	**unknown**		Meghalaya		1 / 8
C056	*Globba* sp.	**unknown**		Uttaranchal	-	1 / 8
C071	*Alpinia conchigera*	2n = 48		Mizoram	-	1 / 8
C058	*Zingiber capitatum*	**unknown**		Jharkhand	-	1 / 8

3.2. AFLP fingerprinting

3.2.1 Species in study

One hundred and fifteen individuals from 25 populations belonging to a total of 19 taxonomic entities were included in the study (Table 2). From the subgenus *Curcuma* 13 hexaploid (2n = 42), 7 nonaploid (2n = 63), one endecaploid (2n = 77), one dodecaploid (2n = 84) and two pentadecaploid populations (2n = 105) were sampled. From the subgenus *Hitcheniopsis* one population of diploid *C. vamana* (2n = 22) was included. Two to ten individuals per population and one to three populations per taxon were sampled.

3.2.2 DNA extraction

Total genomic DNA was extracted from 0.5 g dried leaf tissue using DNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA, USA). DNA concentration was mesured photometricaly and adjusted to 100ng/µl.

3.2.3 AFLP analysis

AFLP analysis (Vos et al., 1995) was carried out using AFLP Core Reagent Kit I (Invitrogen) and AFLP Pre-Amp Primer Mix I (Invitrogen) following instructions of the manufacturer, but downscaled to 1/5 volume. About 100 ng of genomic DNA was digested for 2.5 h at 37 °C with 0.5 U of each *Eco*RI and *Mse*I (Invitrogen), 1 µl 5× Reaction Buffer (Invitrogen) in total volume of 5 µl. Adaptors were ligated to the digested fragments by adding 4.8 µl Adaptor/Ligation Solution (Invitrogen) and 0.2 U T4 DNA Ligase (Invitrogen) to the restricted DNA (total volume 10 µl) and incubated for 8-12 h at 37 °C. Pre-amplification reactions were performed with AFLP Pre-Amp Primer Mix I (Invitrogen) in total volume of 5 µl containing 0.5 µl restricted/ligated DNA, 4.0 µl PA mix (Invitrogen), 0.5 µl 10× Buffer for RedTaq JumpStart (Sigma), and 0.1 U RedTaq JumpStart DNA Polymerase (Sigma). PCR cycling conditions were an initial denaturation of 2 min at 94 °C, 2 min at 72 °C followed by 20 cycles of 10 sec at 94 °C, 30 sec at 56 °C, 2 min at 72 °C, with a final extension of 30 min at 60 °C. Once the pre-amplification was completed, selective amplification was performed using 2.3 µl of 10× diluted pre-amplification product as a template, 1 µl 10× Buffer for RedTaq (Sigma), 0.2 mM dNTP, 0.5 pmol *Eco*RI-selective fluorescence-labelled primer, 2.5 pmol *Mse*I-selective primer, 0.2 U RedTaq JumpStart DNA Polymerase (Sigma), in total volume of 10 µl. Three primer combinations (selected after an initial screening of 10 primer combinations) were used for selective amplification: *Eco*RI-ACT (6-FAM labelled) + *Mse*I-CTT, *Eco*RI-AGC (NED labelled) + *Mse*I-CTG, *Eco*RI-ACA (HEX labelled) + *Mse*I-CAT. Cycling conditions were an initial step of 2 min at 94 °C, 2 min at 72 °C followed by 8 cycles of 1 sec at 94 °C, 30 sec at 64 °C (reduced by 1 °C per cycle), 2 min at 72 °C, followed by 23 cycles of 1 sec at 94 °C, 30 sec at 56 °C, 2 min at 72 °C with a final extension time of 30 min at 60 °C. All PCR reactions have been done in a Mastercycler ep gradient S thermal cycler (Eppendorf).

Fragments were resolved on an ABI 3100 Avant DNA sequencer (Applied Biosystems) together with GeneScan-ROX-500 size standard using GeneScan Analysis v3.7.1 (PE Applied Biosystems) and scored with Genographer.2.0 (Benham et al., 1999; http://hordeum.msu.montana.edu/genographer/). Only unambiguous fragments were scored regardless of their intensity (Tribsch et al., 2002) in the range of 100-500 bp. The resulting presence/absence matrix was used in subsequent analyses.

3.2.4 Data analysis

AFLP data characteristics

For each population total number of AFLP fragments (bands total= BT), average number of fragments per individual (band average = BA) and the percentage of polymorphic fragments (%poly) were calculated. Additionally, number of exclusive bands (EB, bands present in a given species only, but not necessarily in all its samples) and diagnostic bands (DB, bands present in all samples of a species and absent from all other species) were calculated for each species. Since the number of fragments (BT, BA) and the proportion of polymorphic fragments depended on the number of analysed individuals per population, these characteristics were calculated using special script in Scilab software (http://www.scilab.org). This procedure enable to proceed resampling of the whole dataset (special case of bootraping, see Kučera et al., 2008 for details) to achieve the same sample size in each population and prevent a loss of information, as could occur when individuals are randomly deleted from datasets with different sample size to obtain equal numbers of individuals (Schönswetter and Tribsch, 2005). In this analysis four individuals per population were selected in each step and a median value from 1,000 replicates was presented.

Phylogenetic analyses

To reveal phylogenetic relationships among samples neighbor-joining (NJ) analysis as well as maximum parsimony analysis was performed using PAUP* v. 4.0b10, Swofford, 2002).

NJ analysis was based on Nei-Li genetic distances (Nei and Li, 1979). Robustness of branches was estimated using the bootstrap approach (Felsenstein, 1985) with 5,000 replicates.

For maximum parsimony analysis, a heuristic searches was done using a starting tree(s) obtained via stepwise addition and swapped to completion using TBR branch swapping, and MULTREES on (keeping multiple, shortest trees). A strict consensus tree of most parsimonious trees was constructed. Robustness of clades was estimated with 500 bootstrap replicates with random sequence addition (10 replicates) saving no more than 500 trees per replicate.

Neighbor network

To outline the phylogenetic relationships among hexaploid taxa, NeighborNet network method (Bryant and Moulton, 2004) as implemented in SplitsTree 4 (Huson and Bryant, 2006) was used. Neighbor network was based on Nei-Li distances (Nei and Li, 1979), calculated with PAUP, of reduced dataset including 13 hexaploid and one diploid species (see Table 2). Phylogenetic networks are a generalization of phylogenetic trees that permit the representation of conflicting signal or alternative phylogenetic histories, caused e.g. by

recombination, hybridization or lateral gene transfer and are clearly useful when the underlying evolutionary history is nontreelike (Huber and Moulton, 2005). Group support was assessed via a neighbour-joining bootstrap analysis in PAUP using 5000 replicates and bootstrap values higher than 50% were maped onto network clusters, irrespective of whether these were supported by splits or not. This approach allowed direct comparison of incongruences arising from methodological differences (Weiss-Shneeweiss et al., 2008).

PCoA

Principal coordinate analysis (PCoA) using Jaccard's coefficient (Jaccard, 1908) was computed to visualize non-hierarchical structure in complete AFLP data set. PCoA was performed with SYN-TAX 2000 (Podani, 2001).

3.3. ITS cloning and sequencing

3.3.1 Species in study

Total of 37 individuals belonging to 32 species were included in ITS analysis. In addition to AFLP sampling four species (4 individuals) often classified as members of related genera (*Monolophus scaposus*, *Hitchenia caulina*, *Stahlianthus campanulatus* and *Stahlianthus involucratus*), another five *Curcuma* species (9 individuals) from different populations of India, Sri Lanka, Vietnam and China, and four outgroup species (4 individuals: *Hitchenia careyana*, *Globba* sp., *Zingiber capitatum* and *Alpinia conchigera*) have been analysed (Table 3). Conversely, representatives of 5 populations from AFLP analysis were not included in ITS study (to compare AFLP and ITS species/populations representatives see Table 1). Ploidy levels, genome size groups and other details are summarized in Table 1 and Table 3. Generally, one individual per population was analysed. For four species 2-3 individuals from different populations were analysed per species.

3.3.2 DNA extraction and PCR reaction

Leaf material dried in silica gel was used for DNA extraction in the same way as in AFLP study (see above). Plant material were obtained from same populations that were analysed by AFLP fingerprinting, but extraction of DNA was performed independently. The internal transcribed spacers (ITS1 and ITS2) and 5.8S of the rDNA repeat array were amplified and subsequently cloned for all 37 representatives. For PCR amplification forward and revers primers ITS 5 and ITS 4 (White et al., 1990) were used. PCR amplification s were performed in a total volume of 20 ul. Each reaction contained 2 µl of 10× PCR buffer (Sigma), 0.5 U Red Taq Jump Start DNA Polymerase (Sigma), 0.2 mM dNTP, 0.5 pmol of each forward and

revers primers (Sigma) and 5 ng of genomic DNA. PCR products were amplified using an initial 94°C denaturation (1 min) and 35 cycles of 94°C denaturation (45 s), 50°C annealing (50 s), and 72°C extension (50 s). Final extension was set on 10 min at 72°C. Jetquick kit (Genomed) was used to purify PCR products.

3.3.3 Cloning and sequencing

Due to the occurrence of within individual polymorphisms of directly sequenced PCR products, all of the PCR products were cloned using pGEM-T Easy Vector System (Promega) following instructions of the manufacturer. PCR amplification was used to screen which colonies contained inserts of interest and as a template for sequencing reaction using primers and conditions as for initial PCR reactions. Five to twenty-three clones from each individual were sequenced with the same primers as for initial PCR.

Purified PCR products were sequenced by Macrogen (Seoul, Korea). As in trial dataset forward and reverse sequences generated by ITS 4 and ITS 5 primers respectively did not differ in information, majority of clones were sequenced only in one direction, that means either ITS 4 or ITS 5 sequencing primer were used.

The sequences were aligned in ClustalX (Larkin et al., 2007) and then improved manually in BioEdit v.7.0.0 (Hall, 2004). Gaps were coded subsequently by GapCoder (Young and Healy, 2003) according to method of Simple indel coding (Simmons and Ochoterena, 2000).

3.3.4 Data analysis

Within-individual alignment

For each individual in analysis, multiple alignment of all cloned sequences were examined. Within this alignment duplicated sequences (identical sequences that were presened in the alignment in more then one copies), pseudogenes, recombinant sequences and unique seqence types were identified and counted (raw data in APENDIX 1). As pseudogenes and recombinant sequences behave erratically in phylogenetic analysis, these were removed from initial alignment (Popp and Oxelman, 2001, 2005, Grimm and Denk, 2008). Duplicated sequences were removed from subsequent complete alignment as well, as they do not bear any additional phylogenetic information. Based on number of unique sequence types within the genome (excluding recombinant and pseudogene sequences), within-individual ITS polymorphism were assessed.

Recombinant sequences (i.e., products of crossing of two parental sequences (Kececioglu and Gusfield, 1998), emerging *in vivo* during the meiosis, or *in vitro* during the PCR reaction (Alvarez and Wendel, 2003) were identified by the eye based on two conditions: 1) two obvious parental sequences of suspected recombinant sequence were

detected within individual alignment; 2) suspected recombinant sequence had no unique characters in comparison with the parental sequences.

Pseudogenes (non-functional nrDNA copies) were identified by the eye according to Bailey et al. (2003) by comparing conserved and relatively unconstrained regions between each two sequences within individual alignment. Sequences that may have undergone an overall increase in substitution rate, but that maintain a conserved 5.8S relative to ITS 1 and ITS 2, can be distinguished from those that have a comparable rate of change across the 5.8S, ITS 1, and ITS 2 (i.e., pseudogenes).

Dataset reduction

While first reduction step has been done by removing duplicate sequences, pseudogenes and unambiguous recombinant sequences, second reduction step has been done by removing sequence types, that likely do not bear phylogenetic information within present dataset. This reduction step was based on maximum parsimony analysis of initial dataset (= all sequence types in dataset excluding duplicates, recombinants and pseudogenes) and subsequent bootstrap analysis with 1000 replicates. All sequence types, that were in poorly supported relationships with other sequence type(s) (with bootstrap values under 50%) were removed from the initial dataset, because deficient support may be caused by absence of related sequence types in present dataset due to insufficient sampling, or by chimeric origin of the sequences. In both cases these sequences bear no additional information for remaining sequences, or can behave eraticaly in phylogeny reconstruction (Alvarez and Wendel, 2003; Popp and Oxelman, 2005) and thus "reduced dataset" was used for all of subsequent phylogenetic analysis and for presentation of relationships between species.

Phylogenetic analyses

For phylogeny inference, two frequentlly used methods for sequence data analysis were used, maximum parsimony and Bayesian analyses. Additionally, neighbour network method was used for generalization of phylogenetic trees and representation of conflicting signal or alternative phylogenetic histories, caused e.g. by recombination, hybridization or lateral gene transfer.

In maximum parsimony (MP) analysis, a heuristic search was used with random addition of sequences (100 replicates) and TBR branch swapping with the MULTREES option on, using PAUP* v. 4.0b10 (Swofford, 2002). In MP bootstrap analysis (5,000 replicates) a heuristic search was used with random addition sequences (100 replicates for each bootstrap replicate), followed by TBR branch swapping, and with MULTREES option off.

Bayesian analysis was accomplished with MrBayes 3.0(beta) (Huelsenbeck, 2000). Optimal analysis settings were chosen based on initial hierarchical likelihood ratio tests (hLRT) and calculating approximate Akaike information criterion (AIC) performed by

MrModeltest 2.3 (Nylander, 2004). According to both criterion, as a best fitting model GTR+G was chosen (nst = 6, rate = gamma). Analyses were run for 5,000,000 generations. The 'best' tree-model was determined by the highest -ln score when the Bayes analysis came to stationarity. From a total of 100,002 trees, 6,000 were set for burnin. MrBayes posterior probabilities (PP) and majority rule consensus trees were constructed based on final 94,002 trees.

Neighbor network

NeighborNet network algorithm (Bryant and Moulton, 2004) implemented in SplitsTree 4 (Huson and Bryant, 2006), was used to compute splits graphs based on uncorrected pairwise (p) distances (Grimm and Denk, 2008) based on the two ITS datasets. First dataset covers diploid and hexaploid species only (5 diploids and 14 hexaploids species from subg. *Curcuma*, subg. *Hitcheniopsis* and *Curcuma*-related genera), second dataset corresponds with final dataset used in MP an Bayesian analysis, excluding outgroup species (For species representation in NN analyses see Table 3). Support for grouping in networks was adopted from maximum parsimony bootstrap analysis (as was described above) and mapped onto network clusters, irrespective of whether these were supported by splits or not, to compare possible incongruences arising from methodological differences (Weiss-Shneeweiss et al., 2008).

4. RESULTS

4.1. AFLP fingerprinting

4.1.1 Characteristics of AFLP data

The AFLP analysis of 115 individuals representing 25 populations of 19 Indian *Curcuma* species resulted in a total of 364 scorable bands, where 350 (96%) were polymorphic. The average number of bands per individual after resampling was 72.43, with the lowest values for plants in population of diploid *C. vamana* (BA = 30.33) and the highest numbers of fragments for plants in population of nonaploid *C. aeruginosa* and pentadecaploid *C.raktakanta,* both native to Assam (BA = 108 and 103 respectively). Percentage of polymorphic loci per population ranged from 12% in vegetatively reproducing pentadecaploid *C. raktakanta* (in population native to Kerala) to 75% in sexually reproducing hexaploid *C. sulcata*. Exclusive bands (EB) varied between zero in several taxa and nine in *C. longa*. The highest number of diagnostic bands (DB) was observed in *C. bhatti* (see discussion). AFLP characteristics are summarized in Table 2.

Table 2. *Curcuma* species in AFLP study and AFLP data characteristics for 115 *Curcuma* individuals; Group I, II, III = genome size groups sensu Leong-Škorničková et al. (2007); (N), sample size; (BT), total number of bands present in a given population; bands average (BA), mean number of bands present per individual; % PPL - percentage of polymorfic bands; exclusive bands (EB), bands present in a given species only, but not necessarily in all its samples; diagnostic bands (DB), bands present in all samples of a species and absent from all other species. * Abbreviation following the name of the species correspond to locality of origin; ** Population inclluded in neighbour-network analysis.

Collection number	species / *	locality of origin	(N)	BT	BA	% PPL	EB/DB
Subg. *Curcuma*	**x = 7**						
Hexaploid Group I	**2n = 42**						
** 84183	*C. aromatica K*	Kerala	3	98	88	17	1 / 1
** 71484	*C. montana J*	Jharkhand	4	119	84.5	49	2 / 0
** 73474	*C. montana Ch*	Chhattisgarh	3	84	58.67	49	
Nonaploid Group I	**2n = 63**						
71431	*C. aeruginosa As*	Assam	9	154	108.2	34	4 / 0
84119	*C. aeruginosa K1*	Kerala	10	125	93.14	44	
84142	*C. aeruginosa K2*	Kerala	7	47	34	42	
71489	*C. leucorhiza J*	Jharkhand	4	150	85.25	54	0 / 0
84160	*C. longa K*	Kerala	4	91	76.4	74	9 / 0
73222	*C. longa SL*	SRI LANKA	5	93	73.67	24	
84182	*C. zanthorrhiza K*	Kerala	3	46	31.25	38	2 / 2
12x Group I	**2n > 70**						
71485	*C. sp. 'ranchi' J*	Jharkhand	4	93	74.8	35	1 / 0
15x Group I	**2n = 105**						
71432	*C. raktakanta As*	Assam	5	115	103	25	1 / 1
84120	*C. raktakanta K*	Kerala	5	107	81.6	12	
Hexaploid Group II	**2n = 42**						
** 73453	*C. angustifolia U*	Uttaranchal	5	108	62.8	63	3 / 0
** 73480	*C. angustifolia Ch*	Chhattisgarh	4	117	75	59	
** 84143	*C. cannanorensis K*	Kerala	4	105	69.5	50	0 / 0
** 73403	*C. inodora M*	Maharashtra	5	77	39.6	71	1 / 0
** 73402	*C. pseudomontana M*	Maharashtra	5	85	47	63	2 / 0
** 73477	*C. reclinata Ch*	Chhattisgarh	2	95	77	37	2 / 0
** 73467	*C. sulcata MP*	Mad. .Prad.	4	82	44.5	75	0 / 0
** 84155	*C. ecalcarata K*	Kerala	4	52	34	48	3 / 3
** 73446	*C. bhatii Ka*	Karnataka	5	95	60.25	33	6 / 5
Hexaploid Group III	**2n = 42**						
** 84145	*C. mutabilis K*	Kerala	4	128	98.78	63	1 / 1
11x Group III	**2n = 77**						
73223	*C. oligantha SL*	SRI LANKA	4	96	75.5	47	0 / 0
Subg. *Hitcheniopsis*	**x = 11**						
** 84156	*C. vamana K*	Kerala	3	39	30.33	32	0 / 0
TOTAL			**115**	**350**	**72.43**	**96.15**	

4.1.2 Phylogenetic relationships

NJ and MP

The neighbour-joining analysis as well as maximum parsimony analysis revealed essentialy the same structure (Fig. 8). Two groups were detected, although not supported by bootstrap values, on the unrooted trees (MP tree not presented). First group covering all taxa belonging to formerly defined genome size group I (species marked in Fig. 8 as „*Group I*“) with 2 populations of species *C. angustifolia* belonging to group II. Second group involved remaining species belonging to group II, and all species belonging to group III (marked in Fig. 8 „*Group II* “ and „*Group III* “). One species belonging to subgenus *Hitcheniopsis* was adjacent to this group.

Individuals of most of the previously recognised species clustered together with moderate to high bootstrap support (bs) 70-100%. Exceptions was species *C. longa*, of which both populations were well supported in NJ as well as in MP analysis (bs 70-100%), but did not cluster together. Some populations of previously defined species clustered together with population of another species into one well supported cluster (covering all individuals of these populations) rather than forming species specific clusters. Such situation was observed between population of hexaploid species *C. pseudomontana/C. inodora* (bs 82% in NJ, bs in MP < 50%), between *C. cannanorensis/C. mutabilis* (bs 89% in NJ and 97% in MP analysis), and between populations of nonaploid *C. leucorrhiza* and dodecaploid *C. ranchi* (bs 62% in NJ, 71% in MP).

Within „*Group I*“ species two well-supported clades were detected: 1) hexaploid and nonaploid species *C. aromatica/C. zanthorrhiza* had 87% bs in NJ and 72% bs in MP analysis; 2) three population of nonaploid *C. aeruginosa* and two population of pentadecaploid *C. raktakanta* were supported by 70% bs in NJ and 99% bs in MP analysis.

Within „*Group II* and *III*“ species also two well-supported clades were detected covering 1) populations of above mentioned *C. pseudomontana/C. inodora* and 2) populations of *C. reclinata/C. cannanorensis/C. mutabilis/C. sulcata* (bs 83% in NJ and 60% in MP). All subclades within the later clade were also well supported (see Fig. 8).

Population of species *C. bhatti, C. ecalcarata, C. oligantha* and *C. vamana* were supported as individual clusters (100% bs in NJ) with no supported relationships to other species.

Fig. 8. Neighbor-joining tree of 115 *Curcuma* individuals based on 350 informative AFLP characters; NJ and MP bootstrap value higher than 50% are presented (NJ / MP). Abbreviation following the name of the species correspond to locality of origin. In parenthesis, ploidy level of the species is given. *Group I, II, III* indicate species belonging to formerly defined genome size groups I, II and III by Leong-Škorničková et al. (2007)(details in Table 2)

Analysis of diploid and hexaploid species

Neighbor network analysis of one diploid and 11 hexaploid species (55 individuals) revealed overall the same relationships of species belonging to genome size group II and group III as was detected by MP and NJ analyses of complete dataset. Cluster of intermingling samples of *C. inodora/C. pseudomontana* and cluster common to *C. mutabilis/C. cannanorensis* were well supported by splits as well as by bootstrap analyses (70% bs and 98% bs, respectively).

In addition to analysis of complete dataset, within genome size group I (marked as „*Group I*" species in Fig. 9) poorly supported cluster of *C. montana/C. aromatica* (64% bs in NJ) was detected. Nevertheles, this relationship was not supported by any split in NN splits graph (Fig. 9). The basal relationships of hexaploid species were poorly resolved and more than two contradicting relationships in the centre of the network were detected. However, any of these basal relationships were not supported by bootstrap analysis.

Fig. 9. Neighbour-net splits graph based on AFLP data of *Curcuma* diploid (*C. vamana* only) and hexaploid species. Total of 55 individuals, 350 informative AFLP characters; bootstrap values obtained in NJ bootstrap analysis of the same dataset (5,000 replicates) are mapped on the corresponding splits. Three alternative basal relationships of the species suppoorted by splits but not by bootstrap analysis are outlined. *Group I, II, III* indicate species belonging to formerly defined genome size groups I, II and III by Leong-Škorničková et al. (2007)(details in Table 2).

Principal coordinate analysis

The non-hierarchical analysis PCoA largely confirmed the pattern detected by the neighbour-joining and MP analyses.

In a view of first two coordinates, most of the individuals of higher polyploid species (belonging to genome size group I) formed groups according to their population origin, which clearly differentiated from other species (Fig. 10A). On the other hand, hexaploid species belonging to genome size group II and III, form two subgroups, which partially corresponding with clusters in NJ, MP and NN analyses: (1) cluster of *C. cannanorensis/ C. mutabilis/ C. sulcata/ C. reclinata* and (2) cluster of *C. inodora/ C. pseudomontana/ C. bhatti/ C. vamana/ C. oligatha/ C. ecalcarata.*

Combination of the first and third coordinate (Fig. 10B) shows obvious separation of species belonging to genome size group I (black shapes) and group II (white dots). Individuals from genome size group III (grey shapes) and single representant of subgenus *Hitcheniopsis* (*C. vamana*) were integrated into group II individuals. Two populations of *C. angustifolia* (genome group II) appeared in both cases (Fig. 10A and 10B) on the border with populations of hexaploids from group I and group II.

At least combination of first three coordinates did not separate group G III or *C. vamana* from other two groups. First three coordinates explained 10.4%, 5.0% and 4.5% of the total variation respectively.

Fig. 10. Principal Coordinate Analysis based on AFLP data of 115 *Curcuma* individuals. A: Coordinate 1 versus coordinate 2; B: Coordinate 1 versus coordinate 3. First three coordinates cumulatively explaining 19.9% of the total variation. Symbol explanation: closed (black) shapes, individuals frome genome size group I („*Group I*")(●, hexaploids; ◆, nonaploids; ■, dodecaploids; ▶,pentadecaploids); open circles (○), hexaploids from genome size group II („*Group II*"); grey shapes, individuals frome genome size group III („*Group III*") (●, hexaploids; ▶, endecaploids); ☼, *C.vamana* with different Cx value („subg. *Hitcheniopsis*").

4.2. ITS cloning and sequencing

4.2.1 Characteristics of ITS data

The ITS region was sequenced for 335 clones from 37 individuals representing 32 species. Overall 179 unique sequences (further referred to „sequence types“) were identified by within individual alignment, of which 25 were evaluated as unambiguous recombinants. Remaining sequences (156) were duplicates. In entire dataset no pseudogenous sequences were detected. (see raw data in APENDIX 1, 2)

Dataset for initial maximum parsimony analysis contained 154 ITS sequence types belonging to 37 individuals. (= all 335 cloned sequences excluding 156 duplicates, 25 recombinants and 0 pseudogenes). After initial phylogenetic analysis by MP method total of 52 sequences with poorly supported relationships in bootstrap analysis were removed from the dataset (second step of reduction). Final dataset, further used for phylogenetic analyses by MP, Bayesian analysis and NN algorithm, included 102 ITS sequences representing 34 individuals from 29 species. Due to the second step of reduction all sequence types of three species were removed from final analyses (*C.* aff. *prakasha* (2), *C.* sp. „ranchi“ and *C. ecalcarata*). These species are highlighted in initial phylogenetic tree, Fig. 11. (details are summarized in Table 3).

Within-individual ITS polymorphism

The range of within-individual polymorphism in whole dataset were 1-18 different sequence types per individual. The lowest number of ITS types were detected in diploid species *Stahliantus involucratus* (1 unique ITS type), while the highest number of unique ITS types were detected within genomes of nonaploid *C. aeruginosa* and pentadecaploid *C. raktakanta* (18 unique ITS types) (see Table 3 and raw data in APENDIX 1).

Within-individual polymorphism differed in various groups of taxa. Maximum of 3 different sequence types were detected for species of subg. *Hitcheniopsis*, species belonging to *Curcuma*-related genera and for outgroup species. Different situation was in subgenus *Curcuma*. “Medium” within-individual polymorphism was detected in species of genome size group II and III where 2-7 different ITS sequence types were identified. In hexaploid taxa belonging to genome size group I was diversity also in range 2-7 sequence types per individual, while in higher polyploids from group I 3-18 sequence types per individual were detected.

Within-species ITS diversity

For four *Curcuma* species more than one individual belonging to different populations were analysed. Based on the number of unique sequences identified within the different individuals of the same species, as well as position of ITS sequence types in initial phylogenetic tree (Fig. 11) existence of unique sequences in different individuals of the same species was detected.

In two individuals of *C. angustifolia* 4 and 3 unique ITS types were detected by within-individuall alignment. After reduction of dataset, 3 ITS types remained for each individual, that occured together in two lineages on final phylogenetic tree (Fig. 12, clade I a, I b).

In three *C. montana* individuals 3, 5 and 3 ITS types respectively, were detected by within-individuall alignment. After reduction one sequence type remained in each individual and all these sequences occured in final phylogenetic tree within a single lineage (Fig. 12, clade I e).

In two individuals of *C. aromatica* 7 and 3 ITS types were identified by within-individuall alignment, where 4 and 1 type respectively, remained in the final dataset and occured in two different lieages. One lineage was formed by sequence types of both individuals (Fig. 12, clade I c), while in second lineage only sequence types of one individual have occured (Fig. 12, clade I f).

In two individuals of *C. aeruginosa* 18 and 6 ITS types before reduction and 10 and 3 types after reduction were detected. In final phylogenetic tree at lest four lineages of *C. aeruginosa* were identified (Fig. 12, clade I a, I e, I g, I i). In two of them ITS types from both individuals have occured together (clade I e and I i), while in other two lineages either ITS types from one or from another individual were present.

Table 3. *Curcuma* species in ITS study and ITS data characteristics; (N) clones, total number of clones sequenced for each individual; Uniq/dupl. clones = number of unique/duplicate sequences identified by within individual alignment; Recomb. clones = number of recombinant sequences identified by within individual alignment Within individual polymorphism = number of unique clones excluding recombinant clones (= clones included in initial dataset); Reduc. step, number of unique clones, that were removed from initial dataset after initial phylogenetic analyses (see details in methods); Final dataset, number of unique clones included in final dataset used for phylogenetic analysis; *, Accession numbers assigned to multiple accessions of the same species used in Fig. 12, 13, 14; **, Individuals included in neighbour-network analysis of diploid-hexaploid dataset

Collect. number	species (*)	(N) clones	Uniq/dupl. clones	recomb. clones	Within ind. polymorh.	Reduc. step	Final dataset
Curcuma							
Subg. *Curcuma*	**x = 7**						
Hexaploid Group I	**2n = 42**						
** 84183	*C. aromatica* (1)	8	7/ 1	0	**7**	3	**4**
** 84170	*C. aromatica* (2)	5	4 /1	1	**3**	2	**1**
** 71484	*C. montana* (1)	8	6/ 2	3	**3**	2	**1**
** 73479	*C. montana* (2)	13	8 /5	3	**5**	4	**1**
** 73425	*C. montana* (3)	5	4 /1	1	**3**	2	**1**
** 71443	*C.* aff. *prakasha* (1)	8	6 /2	0	**6**	4	**2**
71476	*C.* aff. *prakasha* (2)	5	2 /3	0	**2**	2	**0**
Nonaploid Group I	**2n = 63**						
71431	*C. aeruginosa* (1)	20	18 /2	0	**18**	8	**10**
84142	*C. aeruginosa* (2)	7	7 /0	1	**6**	3	**3**
71492	*C. aromatica*	9	5 /4	1	**4**	2	**2**
73222	*C. longa*	15	7 /8	2	**5**	1	**4**
84182	*C. zanthorrhiza*	5	5 /2	0	**3**	0	**3**
12x Group I	**2n > 70**						
71485	*C.* sp. 'ranchi'	8	8 /4	0	**4**	4	**0**
15x Group I	**2n = 105**						
71432	*C. raktakanta*	23	23 /5	0	**18**	4	**14**
Hexaploid Group II	**2n = 42**						
** 73453	*C. angustifolia* (1)	5	5 /1	0	**4**	1	**3**
** 73480	*C. angustifolia* (2)	16	7 /9	4	**3**	0	**3**
** 84143	*C. cannanorensis*	15	6 /9	2	**4**	2	**2**
** 73403	*C. inodora*	15	10 /5	3	**7**	0	**7**
** 73402	*C. pseudomontana*	6	2 /4	0	**2**	0	**2**
** 73477	*C. reclinata*	6	4 /2	1	**3**	0	**3**
** 73467	*C. sulcata*	9	4 /5	1	**3**	0	**3**
84155	*C. ecalcarata*	5	4 /1	1	**3**	3	**0**
** 73446	*C. bhatii*	10	2 /8	0	**2**	0	**2**
** 84157	*C. neilgherensis*	5	5 /0	0	**5**	2	**3**
Hexaploid Group III	**2n = 42**						
** 84145	*C. mutabilis*	8	5 /3	0	**5**	3	**2**
11x Group III	**2n = 77**						
73223	*C. oligantha*	8	3 /5	0	**3**	0	**3**

KNIHOVNA KATEDRY BOTANIKY

Table 3. Part 2.

Collect. number	species (*)	(N) clones	Uniq/dupl. clones	recomb. clones	Within ind. polymorh.	Reduc. step	Final dataset
Curcuma							
Subg. *Hitcheniopsis*	**x = 11**						
** 84156	*C. vamana*	13	2 /11	0	2	0	2
** 20	*C. alismatifolia*	10	4 /6	1	3	0	3
** 19	*C. pierreana*	8	3 /5	0	3	0	3
Related genera I	**2n = 42**						
** 86407	*Monolophus scaposus*	4	2 /2	0	2	0	2
** 84178	*Hitchenia caulina*	8	2 /6	0	2	0	2
Related Genera II	**2n = 22**						
C086	*Stahlianthus campanulatus*	8	2 /6	0	2	0	2
71449	*Stahlianthus involucratus*	5	1 /4	0	1	0	1
Outgroup species							
C073	*Hitchenia careyana*	8	2 /6	0	2	0	2
C056	*Globba* sp.	8	2 /6	0	2	0	2
C071	*Alpinia conchigera*	8	2 /6	0	2	0	2
C058	*Zingiber capitatum*	8	2 /6	0	2	0	2
Total no. of clones in datsets/reductions		**335**	179/156	25	**154**	52	**102**

4.2.2 Phylogenetic relationships

MP analysis of initial dataset

Initial dataset of nrDNA consist of 154 ITS sequences belonging to 37 individuals. Length of complete alignment was 671 characters (750 characters including gap-coding characters), where 372 characters were constant, 70 variable characters were parsimony-uninformative and 308 characters (41%) were parsimony-informative. 50% bootstrap majority-rule consensus tree is presented in Fig. 11.

Backbone of the initial tree were well resolved and were formed by basal group of outgroup taxa and ingroup taxa clade (99% bs). Ingroup clade covered three well supported clades representing from the base by: yet undefined subgenus (single representant *C. pierreana,* 100% bs), subg. *Hitcheniopsis* clade (76% bs) and subg. *Curcuma* clade (64% bs), respectively. Within subg. *Curcuma* two main clades marked as clade I and clade II were moderately supported (73% bs and 74% bs, respectively). Relationships within both main clades of subg. *Curcuma* were rather poorly resolved and supported. Base of each two clades were formed by large polytomy of sequences belonging to various individuals and species.

As this analysis resulted in a large, unwieldy tree with multiple clones of poor phylogenetic information value further analysis and descriptions were operated with „reduced dataset“ (= final dataset).

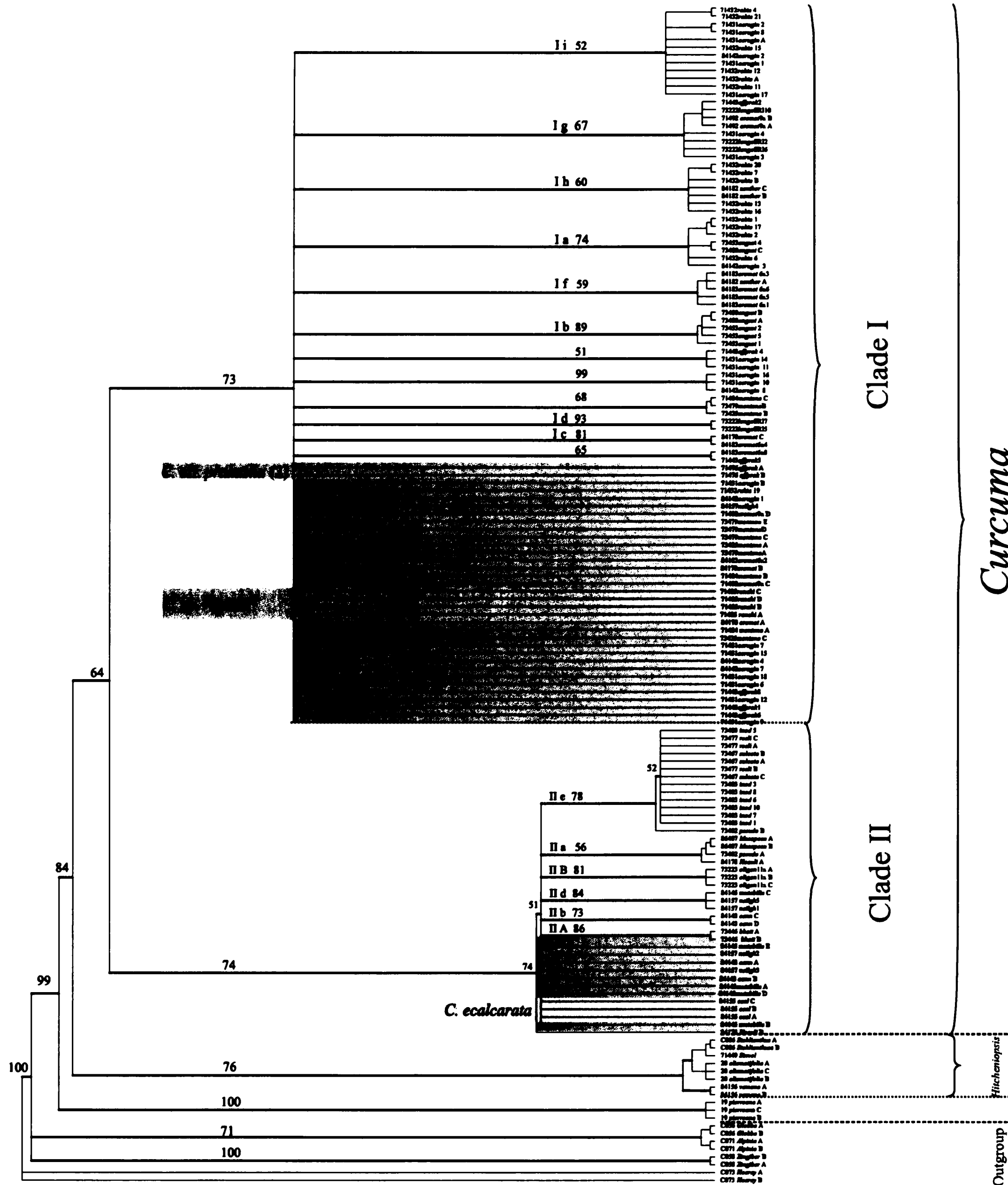


Fig. 11. Maximum parsimony 50% bootstrap majority-rule consensus tree based on initial ITS dataset (154 cloned sequences) for the 29 *Curcuma* individuals, 4 individuals belonging to related genera (*M. scaposus, H. caulina, S. involucratus, S. campanulatus*, highlighted in yellow) and 4 outgroup taxa. Collection numbers are given before name of the species. Bootstrap values over 50% are indicated. Clones that were pruned in second step of reduction are highlighted in grey. Three species that were completely removed from final analyses are highlighted in colour. Names of the accessions correspond to names in raw data (APENDIX 1, 2)

MP and Bayes analyses of final dataset

Final dataset consisted of 102 ITS sequences belonging to 34 individuals with complete length of alignment 671 characters (total of 750 characters including gap-coding characters).

For maximum parsimony analysis this matrix contained 396 constant characters, 52 parsimony-uninformative variable characters and 302 characters (40.2%) that were parsimony-informative. MP analysis resulted in 100,000 most parsimonious trees (CI=0.5929, RI=0.8701). Strict consensus tree were computed for presentation of relationships (Fig. 12). Overall topology of strict consensus tree of final dataset was highly congruent with the strict consensus tree of the initial dataset, except the absence of the large polytomies on the base of the clusters within subg. *Curcuma* clade. Bootstrap values of the individual clades were equal or higher in comparison with the support values of the same clades in initial analysis. Here were assumed bootstrap support value of 90-100% to be a well-supported node, 70-90% to be moderately supported, and 60-70% to be weakly supported node. Bootstrap value from 50 to 60% was assumed as a poorly supported, but nonrandom relationship, which will be interpreted with "special attention".

Bayesian analysis resulted in majority rule consensus tree (from 94,002 trees after burnin) based on supports of the branches by posterior probabilities (PP) higher than 50% (Bayesian tree not presented). This consensus tree differs slightly in topology in comparison with strict consensus tree from MP analyses. However, all well supported clades in MP analysis (from initial as well as final dataset) were formed by same representants as the clades that were well supported in Bayesian analysis. Here were assumed a PP values of 99% and above to be a well-supported node, 95-98% to be moderatly supported, and 90-94% to be poorly supported.

Two kinds of supported relationships between sequence types were distinguished and further described: 1) "monophyletic" clades consisting of all of the sequence types belonging to single individual/species that were detected in initial as well as final dataset. These clusters are marked in phylogenetic trees (Fig. 11, 12) in upper case leters after number of its superior clade (e.g. IIA indicate monopyletic cluster A of species *C. bhatii* within clade II of subgenus *Curcuma*). 2) "paraphyletic" clades that consist of sequence types belonging to various individuals of the same or different species and whose other sequence types are present in more than one cluster. These clusters are marked in phylogenetic trees in lower case letters after number of its superior clade (e.g. Ib indicate paraphyletic cluster b of clade I of subgenus *Curcuma*, that covered 5 sequence types of two individuals of *C. angustifolia*, however other sequence types of this species occured in another well supported clade). Topology and supported relationships of the sequences types were commented from the basal groups of the strict consensus tree (Fig. 12) to the terminate ones.

Ingroup clade of the ITS phylogeny was well supported by 100% bs (PP 99%) in the same way as in initial dataset. This clade involves taxa from both subgenera of *Curcuma* (subg. *Curcuma* and subg. *Hitcheniopsis*), as well as four species from related genera *Hitchenia* (*H. caulina*), *Monolophus* (*M. scaposus*) and *Stahlianthus* (*S. involucratus, S. campanulatus*), which were well dispersed among *Curcuma* species.

Basal group of ingroup species were represented by *Curcuma pierreana* that is not yet well classified and most probably belongs to new subgenus, species belonging to subg. *Hitcheniopsis* (*C. alismatifolia, C. vamana*) and two species belonging to related genus *Stahlianthus* (*S. involucratus and S. campanulatus*) (74% bs, PP 72%). The group of species belonging to subg. *Curcuma* was separated from subg. *Hitcheniopsis* group with bootstrap support 65% (PP 99%). Cloned ITS sequences of subg. *Curcuma* species were split into two well supported groups marked as clade II (88% bs, PP 99%), which is basal group of *Curcuma* clade and clade I (90% bs, PP 99%), which is rather terminal.

Clade II (88% bs, PP 99%) involved exclusively ITS sequence types of hexaploid taxa (and one endecaploid species), which were apriori classified to genome size groups II and III based on 1Cx values in Leong-Škorničková et al. (2007). In general this clade was represented by the same species that clustered together in AFLP analysis (although not supported by bootstrap) and that formed one of the two main groups in NJ tree (Fig. 8 and 14). In cladeII two species belonging to related genera *Hitchenia* (*H. caulina*) and *Monolophus* (*M. scaposus*) that were not sampled in AFLP analysis were involved. Due to reduction of initial dataset all of the sequence types belonging to *C. ecalcarata* were excluded from clade II (see Fig. 11, for plausible reasons see discussion).

Two monophyletic clades of species *C. bhatii* (IIA, 86% bs, PP 100%) and *C. oligantha* (IIB, 81% bs, 100%) respectively were detected within clade II. Five remaining clusters within clade II (IIa – IIe) were paraphyletic. Clade IIa (74% bs, PP 100%) covered all sequence types of *H. caulina*, *M. scaposus* and one of two sequence type of *C. pseudomontana.* Clade IIb (84% bs, PP 100%) contained 2 sequence types of *C. cannanorensis.* Remaining 2 sequence types of this species were removed by second step of reduction due to poorly supported relationships with other sequence types present in initial dataset (see Fig. 11). Clade II c (66% bs, PP 83%) contained single sequence types of *C. neilgherensis* and *C. mutabilis* respectively. Other sequence types of these species are within the clade IId (95%, PP 100%), another 2 sequence types of *C. neilgherensis* and 3 sequence types of *C. mutabilis* were removed throughout second reduction step. Clade IIe (82% bs, PP 99%) covered second sequence type of the *C. pseudomontana* which is basal to clade of all the sequence types of three species *C. inodora*, *C. sulcata* and *C.reclinata*.

Clade I (90% bs, PP 99%) involved ITS sequence types of species, which are characterized by genome sizes as members of genome size group I (Leong-Škorničková et al., 2007), with exception of taxon *C. angustifolia* that belongs to genome size group II. In comparison with AFLP results group of same species in general clustered together in NJ tree

(Fig. 8), although not supported by bootstrap analyses. As mentioned above group I is represented by taxa with high within-individual ITS polymorphism. Majority of sequence types that belongs to single individual represented more than one cluster and frequently clustered together with sequences types from different species/individuals. All of the sequence types of two species (*C.* aff. *prakasha* (2), *C.* sp. „ranchi“) and several sequence types from other species were excluded from the final analysis due to the poor support in initial phylogenetic analysis. Nine well supported clades within clade I (Ia – Ii), will be described bellow. All of these clades are "paraphyletic" in the sense mentioned above.

Clade Ia (56% bs, PP 99%) contained 4 sequence types of pentadecaploid *C. raktakanta*, one sequence type of nonaploid *C. aeruginosa* and 2 sequence types of hexaploid *C. angustifolia*. Remaining sequence types of *C. angustifolia* are involved in clade Ib (98% bs, PP 100%). Clade Ic (93% bs, PP 73%) covered two sequence types from two different individuals of hexaploid *C. aromatica*. Two of four ITS types of nonaploid *C. longa* are involved in clade Id (99% bs, PP 90%). Remaining sequence types of *C. longa* clustered together with hexaploid *C. aff. prakasha* (71443), nonaploid *C. aromatica* and nonaploid *C. aeruginosa* sequence types (cluster Ig, 86% bs, PP 100%). Clade Ie (81% bs, PP 96%) involved two well supported subclusters of *C. aeruginosa* sequence types and three seq. types from three different individuals of hexaploid *C. montana* respectively. Sequence types of hexaploid *C. aromatica* and nonaploid *C. zanthorrhiza* are strongly supported in a cluster If (99%, PP 97%). Clade Ih (60% bs, PP 98%) is formed by sequence types of *C. raktakanta* and *C. zanthorrhiza,* while another sequence types of *C. raktakanta* clustered together with *C. aeruginosa* ITS types in cluster Ii (88% bs, PP 100%). These two clades formed poorly supported terminal clade of the tree (63% bs, PP 96%).

Fig. 12. Maximum parsimony strict consensus tree of 100,000 most parsimonious trees based on „reduced" ITS dataset (26 *Curcuma* individuals, 4 individuals belonging to related genera (*M. scaposus, H. caulina, S. involucratus, S. campanulatus*, highlighted in yellow) and 4 outgroup taxa). Collection numbers are given before name of the species. Accession numbers are assigned to multiple accessions of the same species. Bootstrap values/Bayesian posterior probabilities over 50% are indicated. Subclades within *Curcuma* clades which are comented in the text are marked by roman numerals followed by upper/lowercase leters. Species where monophyletic origin of ITS sequences is predicted are highlighted in grey.

Analysis of diploid - hexaploid dataset in comparison with final dataset

Two datasets of ITS sequences have undergone neighbour network analyses in order to compare non-hierarchical relationships between sequence types of diploids and hexaploids species within dataset that do not cover higher polyploid species, and within dataset that cover all of the higher polyploid species.

NN analysis of diploid-hexaploid dataset resulted in network of 58 ITS sequence types belonging to 19 species (see Fig. 13A and Table 3)(12 hexaploid species belonging to subgenus *Curcuma*, 2 hexaploid species belonging to *Curcuma*-related genera I (*Monolophus scaposus, Hitchenia caulina*), 2 diploid species belonging to subg. *Hitcheniopsis* (*C. vamana, C. alismatifolia*), one diploid *C. pierreana* of undefined position within genus *Curcuma* and 2 diploid species classified as *Curcuma*-related genera II (*Stahlianthus campanulatus* and *Stahlianthus involucratus*).

ITS sequence types of six hexaploid species (*C. aromatica, C.* aff. *prakasha, C. angustifolia, C. pseudomontana, C. mutabilis,* and *C. neilgherensis*) occured in more than one lineage that were supported by splits and by bootstrap analysis. All sequence types of remaining 8 hexaploid taxa and 5 diploid taxa occured generally within single lineage. All sequence types of three taxa *C. sulcata/C. reclinata/C. inodora* occured in common terminal lineage, together with one sequence type of *C. pseudomontana* on the base of this lineage.

Second neighbour network dataset consisted of all sequence types involved in MP and Bayesian analysis, exluding outgroup taxa (Fig. 13B). Overall 94 sequence types from 30 species covering five ploidy levels (2x, 6x, 9x, 11x, 15x) were analysed. In general, NN revealed same structure as MP and Bayesian analysis. Therefore well supported clusters that were described above (clusters I a – I i of clade I and clusters IIa – II e of clade II) could be mapped onto splits of the network. Additionally to phylogenetic trees, lineages, from which plausibly higher polyploid taxa arised, are more clearly depicted.

Fig. 13A. NN splits graph based on diploid-hexaploid dataset (58 ITS sequences from 5 diploids and 14 *Curcuma* hexaploids species); Bootstrap support were obtained by MP bootstrap analysis (5,000 replicates) of the same dataset, bootstrap values above 50% are mapped on corresponding splits.

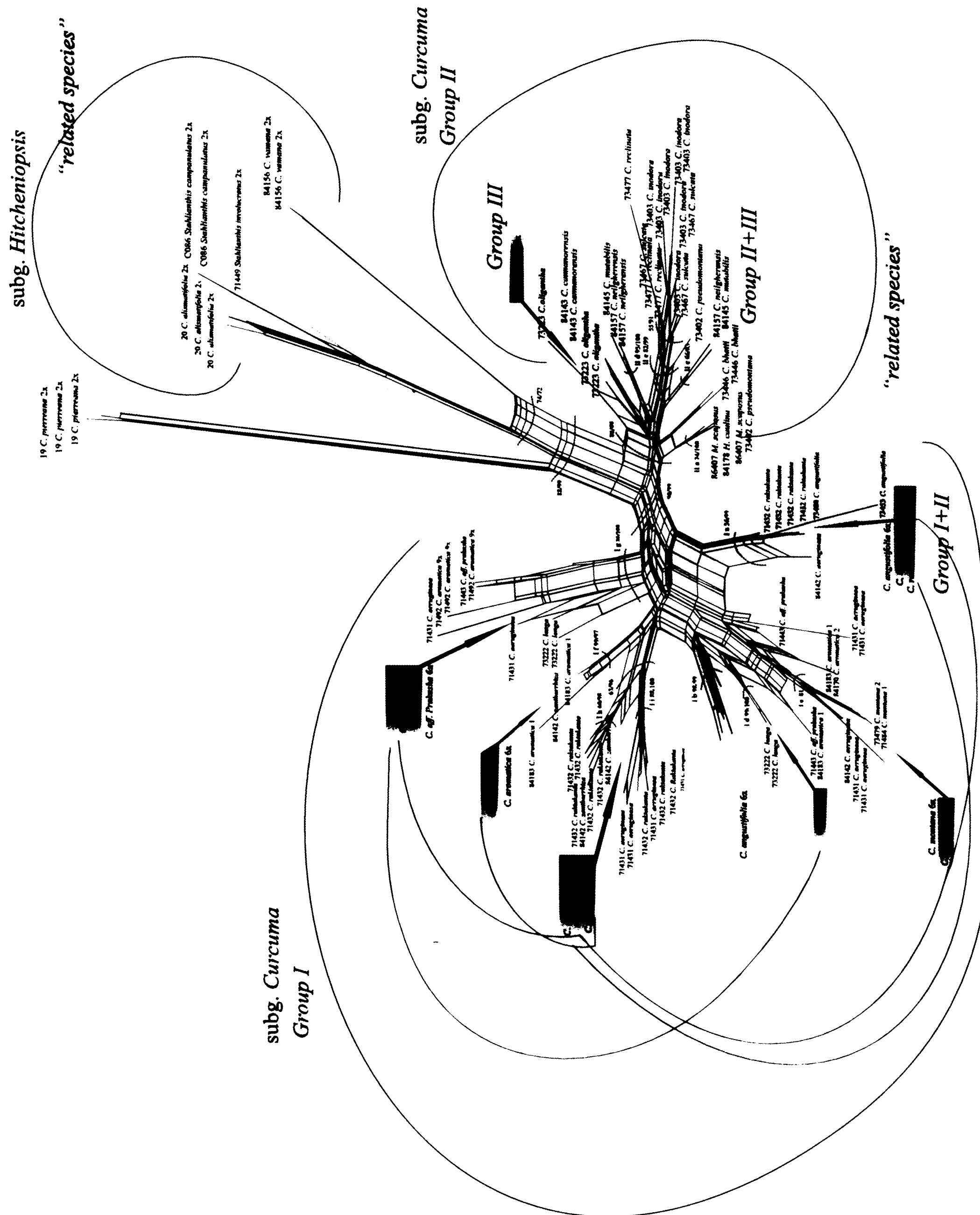


Fig. 13B. NN splits graph based on ITS „reduced dataset", excluding outgroup species (94 ITS sequences from 30 species); Bootstrap support and marking of the clades were addopted from MP strict consensus tree (Fig. 12) in order to compare reticulate pattern in the network and phylogenetic tree respectively.

4.3. Comparison of results from AFLP fingerprinting and ITS sequencing

Two different types of molecular markers were used for resolving phylogenetic relationships within genus *Curcuma* and to prove monophyletic origin of taxa belonging to genome size groups I, II, III). Results of both analyses provided in general the same insight into relationships among the genome size groups, while insights into phylogenetic relationships within these groups slightly differed in AFLP and ITS analyses respectively.

AFLP data in this study provided most reliable information at population (and within species) level, because of relationships of individuals within populations and relationships of populations within species were best resolved and supported by all AFLP data analysis methods. Relationships among species (at infrageneric-level) were well resolved, but frequently poorly supported. However, clustering of species belonging to genome size group I in one cluster and species belonging to genome size group II and III in second cluster were confirmed by NJ analysis (although not supported by bootstrap) as well as other methods of AFLP data snalysis (MP, NN, PCoA, Fig. 8, 9, 10A, B). Species *C. angustifolia* (genome size group II) exhibited border position to group I and group II species in PCoA analysis, while in phylogenetic analyses were incorporeted into genome group I species. Genome size group III were not supported as a separate group by any method of AFLP data analysis. In question how are the phylogenetic relationships between species within and among genome size groups, AFLP marker outline some well supported relationships between hexaploid species belonging to genome size group II and III and between hexaploids and higher polyploids belonging to genome size group I.

Although highly supported clusters covering species of different ploidy levels were detected (*C. raktakanta*/*C. aeruginosa, C. aromatica*/*C. zanthorrhiza*) character of origin of *Curcuma* polyploids cannot be determined unambiguously based on AFLP marker only. Similarly, two pairs of populations likely influenced by hybridization/introgression were hinted by AFLPs (*C. pseudomontana*/*C. inodora* and *C. mutabilis* /*C. cannanorensis*) but exact hybridization events cannot be proven by AFLP only. Reticulate character of evolution, were partially outined in analysis of hexaploid species represented by contradicting phylogenetic relationships in neighbour-net splits graph.

Analysis of cloned ITS sequences from several *Curcuma* individuals yielded complementary data to AFLP and provided reliable informations at within individual-level (due to analysis of several clones from single individual), within species – level (where two and more individuals of single species were cloned and analysed), infrageneric level (where several species of genus *Curcuma* were analysed) and also generic level – where taxa belonging to related genera were analysed. Relationships at all these levels were moderately to well supported by bootstrap analysis. Based on ITS results two monophyletic clades with plausible independent evolutionary history were detected. First clade covered all species

belonging to genome size group I and single species belonging to group II, *C. angustifolia*. Second clade covered remaining species of genome size group II and all species belonging to group III. This well supplement results from AFLP analysis. Relationships between species within these clades were also well resolved and were not in conflict with AFLP results.

However, contrary to AFLP data, results from cloning analysis of ITS region provided enormously complex insight into relationships between *Curcuma* species (where diploid taxa involved two or more different ITS types within the genome, while higher polyploid taxa involved up to 18 different ITS sequence types occuring in two or more different evolutionary lineages). Based on ITS results it was also possible to identified species of plausible hybrid or allopolyploid origin (majority of taxa belonging to genome group I and also hexaploids from group II), or species, that more likely evolved from single ancestor by another processes of speciation (species belonging to subg. *Hitcheniopsis*, and *C. pierreana*). Reticulate pattern of evolution in *Curcuma* species is evident from ITS results. This is particularly visible e.g. in NN analysis (Fig. 13A, B), where majotity of higher polyploids (but also some hexaploids) occured in two or more independent lineages and thus descended plausibly from more than one ancestor (Further hypothesis about evolutionary processes will be discussed).

For visualization of congruences and discongruences between AFLP and ITS results respectively, comparison of the NJ tree based on 115 *Curcuma* individuals belonging to 19 Indian species (subg. *Curcuma*) and MP strict consensus tree from 102 cloned sequences belonging to 21 *Curcuma* and 4 species of *Curcuma*-related genera is presented in Fig. 14.

Fig. 14. Comparison of AFLP fingerprinting method and ITS sequencing in genus *Curcuma*. On the left, neighbour-joining tree based on 350 informative AFLP characters covering 115 *Curcuma* individuals belonging to 19 Indian *Curcuma* species. On the right, maximum parsimony strict consensus tree of 100,000 most parsimonious trees based on 102 ITS sequences from 21 *Curcuma* species, 4 species belonging to related genera (*M. scaposus, H. caulina, S. involucratus, S. campanulatus,* highlighted in yellow) and 4 outgroup taxa (not shown). Independent evolutionary history of two groups of Indian *Curcuma* species (genome group II + III and genome group I + *C. angustifolia*, respectively) were justified by both method. Complex evolutionary history of *Curcuma* species were outlined rather by ITS cloning and sequencing analysis. Marked clades and bootstrap support values are adopted from particular analysis (Fig. 8 and 12)

5. DISCUSSION

Based on current cytological data (Leong-Škorničková et al., 2007) and with solid taxonomic background (Škorničková 2007, Škorničková et al., 2003a, b, 2004, 2005a, b, c), the aim of present study is to shed light into phylogenetic relationships of selected *Curcuma* species. Goals of this study are - to test monophyletic origin of three genome size groups of taxa defined in (Leong-Škorničková et al., 2007), which may have some value for taxonomical classification of subgenus *Curcuma* and to outline basic phylogenetic relationships within selected members of the genus *Curcuma*. Plausible evolutionary phenomenons, which influenced phylogeny of *Curcuma* species are proposed for discussion and further analysis. Second aim of the study is at least partially resolve question of generic boundaries between genus *Curcuma* and its related genera *Hitchenia, Stahlianthus* and *Monolophus.*

5.1. Phylogenetic relationships within genus *Curcuma*

Phylogenetic relationships in *Curcuma* are complex, mainly due to polyploidy, that has played a significant role in evolution of the genus, similarly as in other genera within Zingiberacea [e.g. *Globba* (Lim, 1972a,b; Takano, 2001; Takano and Okada, 2002), *Hedychium* (Mukherjee, 1970; Chen and Chen, 1984)]. As a consequence of polyploidization, different modes of reproduction (sexual and vegetative) have evolved in different species and these exhibit various levels of morphological variability. Taxonomical problems resulting from this pattern were described in chapter 2. To prove natural grouping of the genus according to Leong-Škorničková et al. (2007) AFLP fingerprinting and ITS sequencing were used. Combination of these two markers provided two slightly different insights into phylogenetic relationships in the genus *Curcuma.* AFLP method did not resolve origin of the polyploids unambiguously. On the other hand, AFLP profiles of individual *Curcuma* species corresponded well with situation that can be observed in the field in present time. For instance, based on AFLP profiles morphologicaly highly variable species exhibit significantly higher within-population genetic variability than morphologicaly nearly uniform taxa that exhibit low within-population genetic variability (data not shown, Záveská et al., in prep.). Further, individuals of species that occuring sympatrically in present time and either share recent common ancestor or are influenced by itrogression/hybrididzation, have nearly the same AFLP profiles (e.g. populations of species *C.mutabilis/C. cannanorensis, C. pseudomontana/C. inodora, C. zanthorrhiza/C. aromatica, C. aeruginosa/C. raktakanta*). Finaly, species that evolved independently in the past, e.g. due to isolation or as an endemic species with some specific ecological requirements (which can also cause considerable

morphological differences) exhibit higher number of species-specific bands in AFLP profiles (e.g. *C. bhatti* – specific habitat, that influenced overall morphology – 5 species specific AFLP bands; *C. ecalcarata* – unique shape and structure of the anther – 3 species specific bands). Nature of AFLP marker, where high number of fragments is aquired from different parts of the whole genome and subsequently is compared between individual samples (Robinson and Harris, 1999) is likely the reason of the „generallized" pattern observed in *Curcuma* species.

On the contrary, based on ITS results species of plausible hybrid/allopolyploid origin or species that more likely evolved from single ancestor was possible to identify more exactly. Reticulate pattern of evolution in *Curcuma* species is evident. However, specific evolution of nrDNA within the genome partially contributes to complexity of relationships detected by phylogenetic analyses (Gottlieb et al., 2005). Overall, character of ITS data in *Curcuma* species mirrored cascades of ancient or recent allopolyploidization or hybridization events, various degrees of homogenization of rDNA arrays (e.g. in hexaploids from group II), maintenance of some parental ITS types in allopolyploid genomes (e.g. in hexaploids and higher polyploids of group I) or parental ITS types under various degrees of recombination processes (production of chimeric ITS sequences). All these events are represented in particular species by different level of within individual ITS polymorphism. Character of nrDNA data was taken into account, when gene tree of ITS region was interpreted as an evolutionary hypothesis for *Curcuma* species.

For further studies, exact pattern of the relationship network cannot be infered only based on one rDNA region, but rather several DNA regions from different parts of genomes (cpDNA, nDNA) should be compared (similarly as in Guo et al, 2004; Hodkinson et al., 2002 and many others) and preferably more single-copy nuclear region should be employed for inference of exact relationships (Strand et al. 1997, Small et al. 2004) among *Curcuma* species.

5.1.1 ITS polymorphism in *Curcuma* species

Intra-individual variability of ITS sequences were observed in several groups of angiosperms [e.g. Grimm and Denk, 2008 (*Platanus*, Platanaceae); Denduangboripant and Cronk, 2000 (*Aeschynanthus,* Gesneriaceae); Cambell et al., 1997 (*Amelanchier*, Rosaceae); Sang et al., 1995 (*Paeonia*, Paeoniaceae)]. In the investigated *Curcuma* species, various levels of within-individual ITS polymorphism were observed. These vary according to ploidy level, hybrid origin and/or preferential mode of reproduction. The „distribution" of different within-individual ITS polymorphism partly corresponds with phylogenetic pattern of *Curcuma* species. While the species occuring in basal part of the phylogeny (subg. *Hitcheniopsis* and genome size group II and III from subg. *Curcuma*) are diploids or hexaploids reproducing mainly sexually, species in the terminal lineage (belonging to genome size group I) are either

facultative asexual species (hexaploids) or compulsory vegetatively reproducing taxa (all of the higher polyploids) (see Fig. 14).

In majority of sexually reproducing species, maximum of 3 different ITS types per individual were present in species of subg. *Hitcheniopsis* and maximum of 7 different ITS types per individual were present in hexaploids from genome group II and III of subg. *Curcuma* (clade II in phylogenetic tree, Fig. 12). Majority of these within-individual sequence types formed monophyletic clusters on the final phylogenetic tree (both species of *Stahlianthus, C. pierreana, C. alismatifolia, C. bhatii, C. oligantha,* and common cluster of *C.inodora/sulcata/reclinata,* e.g. Fig. 12). These species could be of paleo-allopolyploid origin where sequence types from ancient ancestors were almost, but not completely, homogenizied due to concerted evolution, which is facilitated by sexual reproduction (Baldwin et al., 1995). This explanation is more likely than origin of these species from single ancestor, as differences in ITS sequence types were dettected even on diploid level (see raw data in APENDIX 1). From single ancestor migt originate diploid *Stahlianthus involucratus* (subg. *Hitcheniopsis*) that possessed single unique ITS sequence type within the genome.

Within sexual species higher within-individual polymorphism and/or occurence in more than two (paraphyletic) clades observed in *C. neilgherensis* (4 sequence types) and *C. mutabilis* (5 sequence types) might be consequence of interspecific hybridization (e.g., Sites and Davis 1989; Arnold, Bennett, and Zimmer 1990; Delseny et al. 1990; Rieseberg, Carter, and Zona 1990; Crease and Lynch 1991; Rieseberg 1991; Soltis and Soltis 1991; Kim and Jansen 1994; Sang et al., 1995). In case of *C. inodora* that have 7 different ITS types but occured in monophyletic clade with *C. sulcata* and *C. reclinata,* (Fig. 12, 13A, clade II e) hybridization or introgression with closely related *C. pseudomontana* might be the reason of high within-individual polymorphism but of low diversification of these ITS sequence types.

In taxa belonging to genome size group I, which encompasses hexaploids as well as higher polyploid taxa, more extensive within-individual ITS polymorphism was detected. Maximum of 7 different haplotypes within genome were detected within hexaploid species, while maximum of 18 different haplotypes were observed within group of higher polyploids. The numbers of different ITS sequence types within group I species genomes were frequently equal or higher than ploidy level of particular species. Thus higher number of haploid sets of chromosomes is not only single source of extensive ITS polymorphism within these species, but rather specific evolutionary processes of ITS region contributed to generating extensive within-individual polymorphism of ITS sequences. Campbell et al., (1997) summarized that within-individual polymorphic nrDNA may occur in transition stages of concerted evolution; when mutation rate exceeds the rate of concerted evolution, as in length variants in the intergenic spacer; as a result of interspecific hybridization; when pseudogenes evolved or when location of nrDNA loci on nonhomologous chromosomes

potentially disrupts concerted evolution. High incidence of hybridization and/or allopolyploidization events could be possible explanation of high within-individual polymorphism in *Curcuma* species belonging to genome group I. However, many allopolyploids for which ITS sequence data are available (e.g., Baldwin 1992; Hsiao et al. 1994; Sun et al. 1994; Baldwin and Robichaux 1995; Wendel et al., 1995) do not show such extensive nucleotide site polymorphism as present in *Curcuma.* Similar pattern of polymorphic ITS sequences have been found e.g. in *Amelanchier* (Rosaceae) (Campbell et al., 1997) or *Arnica* (Asteraceae) (Ekenas et al., 2009), where in both cases the ITS polymorphism was explained by hybridization in combination with reduced concerted evolution due to agamospermy. As generalized by Baldwin et al. (1995), vegetative reproduction is likely to be a reason for maintenance of parental ITS sequences in the hybrids. Sang et al. (1995) argued, that in polyploid *Peonia* frequent reproduction via rhizomes may prolong generation time significantly, which might cause slow rates of concerted evolution. As *Curcuma* hexaploids are facultatively vegetatively reproducing species and all of the higher polyploids are reproducing exclusively vegetatively by rhizomes (see Table 1), it is very likely that mode of reproduction in *Curcuma* is linked with processes arranging maintenance of high number of ITS haplotypes within single genome. Additionally, Campbell et al., (1997) mentioned that polymorphism does not have to be clearly associated with vegetative reproduction or polyploidy, as the impact of hybridization on polymorphism may depend on the sequence divergence of the parents.

Nevertheless, the level of homogenization (concerted evolution) and other processes that impact evolution of ITS region does vary between species and probably are not strictly correlated with the age of the group. Dadejová et al. (2007) summarized that allopolyploids of *Arabidopsis* and *Brassica* apparently have no rDNA unit loss or concerted evolution, despite an evolutionary history lasting several thousands of generations (e.g. Alvarez and Wendel, 2003). By contrast, studies on some recently formed allopolyploids (Kovarík et al., 2005), synthetic interspecific hybrids and polyploids (Lin et al., 1985; Cluster et al., 1996; Weiss and Maluszynska, 2000; Skalická et al., 2003) reveal evidence of astonishingly fast genetic change influencing thousands of rDNA units perhaps in just a few generations following allopolyploidy. The reasons for interspecies differences remain enigmatic but evolutionary change probably occurs through a range of processes, including locus loss/gain, amplification/reduction in repeat copy number, gene conversion and recombination. Thus for realisation of possible options of evolutionary processes in ITS arrays within Curcuma´s genomes further extensive studies should be performed. Then more accurate interpretations of ITS gene tree could be applied for inferring evolutionary relationships of *Curcuma* species.

Intergenic recombination between different repeat types is another possible factor which increased the number of haplotypes within the genome of *Curcuma* species. Following hybridization, two or more repeat types can be maintained but undergo various degrees of

recombination (Alvarez and Wendel, 2003). In several investigated *Curcuma* species recombinant sequences unambiguously identifiable by comparison with parental sequence types were detected (see Table 3 and raw data in APENDIX 1). However, in several individuals, mainly higher polyploids (e.g. pentadecaploid *C. raktakanta* and nonaploid *C. aeruginosa*) occurence of plausible recombinants that exhibiting unique characters (mutations/indels) was also detected. These recombinants may represent transition stages in the homogenization of the ITS region (see Strachan, Webb, and Dover 1985). On the other hand, these sequence types could be related to ITS types of some ancestor that was not included in this study or has declined. For further studies enlarge sampling of hexaploid species (possible ancestors of nonaploids as well as higher polyploids) and enlarge sampling of clones from single individuals might help with identification of "true ITS types" originate from ancestral genomes and their recombinants.

5.1.2 Subg. *Hitcheniopsis*

Two species currently recognized as members of the subgenus *Hitcheniopsis* were included in this study - *C. vamana* and *C. alismatifolia.* Major characters of the subg. *Hitcheniopsis* is absence of epigynous glands and anthers without anther spurs. *C. vamana* is the only representative of subgenus *Hitcheniopsis* in India, and single diploid species in India (2x = 22). Although its geographical distribution overlaped with distribution of typical Indian species from subg. *Curcuma*, there is no evidence of any participation of *C. vamana* diploid genome in any of the genomes of hexaploids or higher polyploids from subg. *Curcuma. C. alismatifolia* (2n = 2x = 36), which is native of Thailand, Laos and Vietnam clustered perfectly with *C. vamana*. Two species of the genus *Stahlianthus* (2n = 2x = 22), *S. involucratus* and *S. campanulatus*, also clustered together with members of the subgenus *Hitcheniopsis* (see below section "Taxonomic position of *Curcuma*-like species"). The genus *Stahlianthus* is currently recognized by highly reduced bracts, of which only two form an inflorescence head. However, anther structure as well as absence of epigynous glands, which are considered as rather conservative morphological characters, support this grouping.

Novel subgenus?

Curcuma pierreana clustered at the base of all other *Curcuma*. This species belong to a small group of taxa native to Indochina. These taxa have well developed epigynous glands as species in subg. *Curcuma*. Their anther structure is also closer to subg. *Curcuma* with well develloped connective formed into two tiny hair-like anther spurs. Preliminary study of Sirirugsa et al., (2007) suggest that these taxa form separate clade and should be recognized as separate subgenus. Up to date there are no chromosome counts nor genome size information available on these taxa.

All above mentioned species are probably not influenced by such extensive polyploidization and hybridization as was detected in species belonging to subg. *Curcuma* and probably evolved in bifurcate model of evolution, that is possible to depicted by phylogenetic trees. (see species highlighted in grey in Fig.12).

5.1.3 Subg. *Curcuma:* Genome size groups in the light of molecular phylogeny

Genome size group I

Previously defined genome size group I is covered by four hexaploid species and seven higher polyploid species (five nonaploids, one dodecaploid and one pentadecaploid) in present study. All species had same interval of 1Cx value 0.30 – 0.33 pg.

All the methods of AFLP data analysis (PCoA, MP and NJ phylogenetic trees) as well as phylogenetic analysis of nuclear ribosomal DNA revealed in general the same pattern. All taxa belonging to the genome size group I form monophyletic clade, however, with inclusion of species *C. angustifolia* that belongs to genome size group II. According to 1Cx-values this species is border-line case between genome group I and group II species (O. Šída, pers. comm.). Reasons for inclusion of the species within group I remained unresolved, however, evidence about capability of *C.angustifolia* to hybridize with species from group I (Fér et al., unpubl. data) suggested that genome size of *C.angustifolia* could increase as a consequence of genome reorganization after hybridization followed by hybrid speciation (e.g. Rieseberg, 1997).

Hexaploids from genome size group I

Relationships between hexaploid species from group I were not clearly resolved. This could be caused by absence of recent common ancestor in the present dataset, and the question is whether this ancestor still exists in present time or has declined.

Relationship of *C. aromatica* and *C. montana* were supported in analyses of hexaploid datasets by both AFLP (65% bs, Fig. 9) as well as by ITS (50% bs, Fig. 13A) marker. This relationship however were not supported by any analyses of complete dataset. The reason of this poorly supported but non-random relationship might be ancient common ancestor of these species and their disjoint geographical distribution [*C. aromatica* from South-west of India (Kerala) and all individuals of *C. montana* from Central India (Jharkhand, Chhattisgarh, Orissa)]. According to ITS analyses hexaploid *C. aromatica* is probably of paleo-allopolyploid origin, as its ITS sequence types occured in at least two independent lineages on phylogenetic tree (Fig. 12, clades I c, I f) and neighbour network split graphs (Fig. 13A, B).

In addition to AFLP sampling, two hexaploid species tentatively identified as *C.* aff. *prakasha* (1) (71443) and *C.* aff. *prakasha* (2) (71476) were included in ITS study. *C.* aff. *prakasha* (2) (71476) was excluded from the final analysis because of no occurence of supported relationships with any ITS sequences types from different individuals. On the contrary sequence types of *C.* aff. *prakasha* (1) (71443) have occured in at least two lineages (Fig. 12, cluster Ig, basal to cluster Ic and Id) where one lineage was shared with sequence types of *C. aromatica* (74% bs in hexaploid dataset, Fig. 13A). Therefore, paleo-allopolyploid origin of this taxon is very likely.

Hexaploid *C. angustifolia*, previously defined as a member of genome size group II, have been according to molecular analysis included within species from genome size group I (88% bs in hexaploid dataset, Fig. 13A, 90% in reduced dataset, Fig. 12). Sequence types of this species occured in two slightly diversified lineages. However these probably originate from single recent ancestor (see hexaploid dataset analysis, Fig. 13A).

All of above mentioned species [excluding *C. aff. prakasha* (2)] have participated, in different manners and frequencies, on genomes of various higher polyploids.

Hexaploid *C. aromatica* and its "relatives"

Hexaploid *C. aromatica* and nonaploid *C. zanthorrhiza* (in AFLP 87/89% bs, Fig. 8) occured sympatrically in south-western India (Kerala, Idukki Dt.). According to ITS data well-supported relationships between some sequence types of these two taxa were detected (cluster I f, 99% bs, Fig. 12). As *C. zanthorrhiza* probably contained part of the genome, which is nearly the same as part of the genome in *C. aromatica* (the most similar sequence types of *C. aromatica* and *C. zanthorrhiza* differed by maximum of 3 point mutations, see raw data in APENDIX 2) it is likely that these two taxa have a common ancestor or that *C. aromatica* is one of direct parental species of *C. zanthorrhiza.* Participation of another genomes in *C. zanthorrhiza* is plausible, as other sequence types of *C. zanthorrhiza* occured in another lineage together with *C. raktakanta* (cluster Ih, 60% bs, Fig. 12, 13B).

Different ploidy levels (6x and 9x) were documented in *C. aromatica* (Škorničková, 2007). In addition to AFLP sampling, nonaploid species of *C. aromatica* belonging to *C. aromatica* polyploid complex were included in ITS study. Contrary to the presumption, none of ITS sequence types of nonaploid *C. aromatica* were in supported relationship with sequence types of hexaploid *C. aromatica.* It is not surprising, because taxa belonging to "*C. aromatica* complex" were established based on morphological match with rather scanty original description and nomenclatoric resolution for the name *C. aromatica* is wanting (Škorničková, pers. comm.). Therefore "*C. aromatica* complex" is artificial group of species without any recent phylogenetic connections and should be the point of further studies. Alternative parental taxa or taxa which share common ancestor with nonaploid *C. aromatica* is *C.* aff. *prakasha* (1) (cluster I g, 82% bs, Fig. 12, 13B).

Nonaploid *C. aeruginosa* and its „relatives“

Individuals from two populations of nonaploid *C. aeruginosa* were analysed by both AFLP and ITS markers. Results obtained from both markers strongly supported relationship with pentadecaploid species *C. raktakanta* (Fig. 8, 12 and 14). *C. aeruginosa* and *C. raktakanta* occur sympatrically on two very distant localities (2000 km faraway Assam in North-east of India and Kerala in South-west of India) and morphologically are highly simmilar. They probably shared a common ancestor(s). In ITS analyses, two species related to both *C. aeruginosa* and *C. raktakanta* were detected in two independent lineages [hexaploid *C. angustifolia* lineage (clade I a, Fig. 12) and nonaploid *C. zanthorhiza* lineage (63% bs and PP 96% of cluster covering I h and I i clades, Fig. 12, 13B].

In addition, sequence types of *C. aeruginosa* were detected in another two independent lineages. Firstly, in moderately supported relationships with hexaploid *C. montana* (Fig. 12, clusters I e), secondly in well supported relationship with nonaploids *C. longa* (sampled in Sri Lanka) and *C. aromatica* and hexaploid *C.* aff. *prakasha* (1) (clade I g, Fig 12, 13B). As is visible from Fig. 13A and 13B, *C.* aff. *prakasha* (1) is one of the potencial parental species of these higher polyploids or at least share with them common ancestor. Geographically, neither hexaploid *C. montana*, nor hexaploid *C.* aff. *prakasha* have been detected in sympatry with above mentioned higher polyploids. Nevertheless, data about geographical origin of the populations/species could be misleading for inferring the evolutionary history of the species, as many of nonaploids (higher polyploids) are economicaly used and human factor take part in their current distribution.

Finaly, in AFLP analysis there was observed relationship between nonaploid (*C. leucorrhiza*) and rare dodecaploid species (*C.* sp. „ranchi“) from nearby localities that are morphologicaly closely related vegetatively reproducing species (BS 61% in NJ and 69% in MP, Fig. 8). As in ITS study no individual of *C. leucorrhiza* were analysed and all haplotypes of *C.* sp. „ranchi“ were excluded due to no supported relationships with ITS sequence types in present dataset no conclusions can be drawn about origin of *C.* sp. „ranchi“ and its relationship to *C. leucorrhiza*.

Overall, all indications supports hypothesis formulated in Leong-Škorničková et al. (2007) that higher polyploids and hexaploids with same 1Cx values originate from the same lineage. Plausible ways of polyploids formation will be discussed below.

Genome size group II and III

Genome size group II previously defined by Leong-Škorničková et al., (2007) contained nine hexaploid species with interval of 1Cx value 0.36 – 0.39 pg. Species of this group together with two species belonging to genome size group III (*C. mutabilis*, *C. oligantha*) and two species belonging to *Curcuma*-related genera (*Monolophus scaposus*, *Hitchenia caulina*)

were highly supported as a monophyletic group in ITS analysis (Fig. 12, 88% bs, PP 99%). In AFLP analysis, in general same species (excluding species from *Curcuma*-related genera) clustered together (although not supported by bootstrap) in NJ and MP trees (Fig. 8) and have formed the group in PCoA analysis (Fig. 10B).

Genome size group III with homoploid genome size in range 0.41 – 0.43 pg comprised of only few taxa in India and was thus represented in the study by one hexaploid (*C. mutabilis*) and one endecaploid (*C. oligantha*) species. Both AFLP and ITS data did not support group III as independent monophyletic group. On the contrary, some relationships between species belonging to group III and group II were strongly supported.

Species originally belonging to related genera, *Monolophus scaposus* and *Hitchenia caulina,* (analysed only in ITS study) were strongly supported as ingroup taxa within subg. *Curcuma* clade. According to cytologic data (genome sizes, number of chromosomes, ploidy levels) both taxa fit well to genome size group II, as both have 1Cx values in range 0.36 – 0.39 pg and 42 chromosomes. It correspond with hexaploids from subgenus *Curcuma.* Morphological differences that lead to classification of these species to separate genera are adaptations to specific environment coupled with specific mode of pollination. As inclusion of these taxa into genus *Curcuma* was justified (see below in section"Taxonomic position of *Curcuma*-like species“) these will be regarded as a members of genome group II species of subg. *Curcuma.*

AFLP data outlined some general relationships among species from group II and III, while ITS data revealed relationships complicated by presence of multiple haplotypes within individuals and their linked relationships. Clear and well supported relationships based on AFLPs were observed between sympatrically occuring *C. pseudomontana* and *C. inodora* (NJ 82% bs, Fig. 8) and between morphologically similar taxa *C. sulcata/C. reclinata/ C. cannanorensis/C. mutabilis* (83% bs in NJ and 58% in MP, Fig. 8). Remaining taxa of genome group II and III, which in general differed from other group II and III species by unique morphological characters and/or by high number of AFLP species specific bands (EB/DB, see Table 2), were supported as separated clusters with 100% bootstrap and their relationships to other group II species remains unresolved (Fig. 8, 14).

According to ITS results C. *inodora*, C. *sulcata* and *C. reclinata* formed monophyletic cluster (subcluster of I e, 55% bs, PP 91%, Fig. 12), where all of the sequence types of these three species are included. Basal to this cluster is one of two sequence types of *C. pseudomontana* that probably shared common ancestor with these species (cluster II e, 82%, Fig. 12). According to Škorničková (pers. comm.) C. *inodora*, C. *sulcata* and *C. reclinata* could be regarded as a single taxon, due to highly similar morphology. According to AFLP results, at least populations of *C. inodora* and *C. pseudomontana* that occurs sympatricaly in Maharashtra, Western Ghats (Santapau, 1952) were probably under introgression or hybridization processes in recent time. Hybridization between these two species could also explain poor bootstrap support for *C. pseudomontana/C. inodora* and

C. sulcata/*C. reclinata* clusters in NJ analyses of AFLP data (Fig. 8, 9), as well as basal position of *C. pseudomontana* to C. *inodora*/C. *sulcata*/*C. reclinata* clade in ITS analysis (Fig. 12, 13A). In addition, *C. pseudomontana* is probably of paleo-allopolyploid origin, as shared one of the two ITS types with *M. scaposus a H. caulina* (clade II a, 74% bs, PP 100%, Fig. 12). Morphologically these species are totaly different.

Sequence types of two species, *C. neilgherensis* and *C. mutabilis*, appeared in the clades IIc (66% bs) and IId (95% bs) in ITS phylogenetic tree (Fig. 12). Many sequence types of these species that had no supported relationships to other sequence types in dataset were, however, removed from final dataset. According to AFLP results relationships of *C. mutabilis* and *C. cannanorensis* was strongly supported, but not supported in ITS analysis (although occuring in common lineage in ITS hexaploid dataset, Fig. 13A). Both species, *C. mutabilis* and *C. cannanorensis,* are endemic species of Western Ghats from northern part of Kerala. Even though their sympatric occurence have not been encountered so far, they are distributed in adjacent districts of Kerala and introgression or hybridization processes could happen in the past. It can be hypothesised that different homoploid genome size of *C. mutabilis* (in comparison with group II species) could be consequence of genome reorganization after hybridization followed by hybrid speciation (e.g. Rieseberg, 1997). For unambiguous detection of relationships between *C. neilgherensis*/ *C. mutabilis* / *C. cannanorensis* further studies are necesarry.

Two species, *C. bhatii* and *C. oligantha*, formed monophyletic clusters of all their sequence types and seems to be basal to other genome size group II and III species (in ITS study, Fig. 12). Seed-setting stenoendemit of Western Ghats, *C. bhatii* is different from other seed-setting Indian *Curcuma* species by its highly specific habitat (exposed lateritic slopes, see Fig. 1) afecting its overall morphology (the smallest *Curcuma* species not exceeding 15 cm, with small, but branched rhizomes, producing only central inflorescence with green and sparsely connated bracts), which previously resulted in treating this species as a member of separate monotypic genus *Paracautleya*. This is congruent with the highest number of diagnostic bands observed by AFLP fingerprinting (DB = 5). Regardless, these species was strongly supported as ingroup species within subg. *Curcuma* in ITS analysis and its morphological differences that lead to classification of these species to separate genera could be assumed as adaptations to specific environment (simmilarly as in case of *M. scaposus* and *H. caulina*).

C. oligantha is so far the only species with 77 chromosomes (and 11 ploid level). This species is Sri Lankan endemit, capable of setting seeds. Origin of *C. oligantha* is still enigmatic, as no plausible ancestors within our dataset were detected. Contrary to high ploidy level, within-individual ITS polymorphism is rather low, probably due to higher rate of concerted evolution, facilitated by sexual mode of reproduction (Baldwin et al., 1995).

C. aurantiaca (= C. ecalcarata) belonging to genome size group II has some unique morphological characters (shape and structure of the anther without anther spur) within the

nominate subg. *Curcuma* and exhibit 3 diagnostic AFLP bands (which is second highest number of DB in AFLP dataset, see Table 2). Apart of India, this species is found in other Asian countries including Indonesia, Thailand and Burma. In AFLP analysis relationship of this species to other *Curcuma* species was unresolved. In ITS analysis all sequence types of this species were removed from the final dataset, as these did not form any supported relationships with any of the sequence types from other species included in present study. Relatives of this species are likely to be native in other parts of Asia rather than India and further studies are necessary.

5.1.4 Hypothesis about evolution of *Curcuma* species

Hypothesis about formation of polyploids in *Curcuma* do not differ much from ideas presented in recent study (Leong-Škorničková et al., 2007). As unreduced gametes are believed to be a major mechanism of polyploid formation (Bretagnolle and Thompson 1995), it is plausible that major of nonaploid cytotypes probably originated by a fusion of reduced and unreduced gametes of hexaploids (in most of cases from different species) giving rise to allopolyploids. Formation of autopolyploids cannot be ruled out, but these cases would be probably less frequent as suggested by ITS data in this study. Both nonaploid (unreduced gamete) and hexaploid (reduced and unreduced gamete) plants were probably involved in the genesis of rare dodecaploid and pentadecaploid species. Although the vast majority of nonaploid species undergo asexual propagation via rhizome branching and produce a high percentage of aborted pollen grains (Prana, 1977; Nasir Uddin, 2000), it is plausible, that even a small fraction of viable pollen may eventually lead to the formation of higher polyploids (Leong-Škorničková et al., 2007). Moreover, ITS analysis suggested strong reticulate pattern in evolution of *Curcuma* species where relationship among *Curcuma* species are complicated by paleopolyploid character of many hexaploid species and by hybridization events.

As many of nonaploids (higher polyploids) are economicaly used and thus cultivated, human factor plays a big part in their current wide distribution. Secondary contacts among previously isolated populations could play role in formation of the polyploids as well. How many geographic centres of origin of higher polyploids there were or if recurrent formation of the polyploids was possible still remains a question. Within the sampling, however, at least four localities with sympatric occurrence of taxa with different ploidy levels and thus potential centres of polyploid formation were detected. In general, the results supports relationships of taxa with different ploidy levels just from these localities (e.g. *C. aeruginosa*/*C. raktakanta* in Assam and Kerala, *C. aromatica*/*C. zanthorrhiza* in Kerala), where individuals grow in sympatry.

Majority of hexaploids are exclusivelly sexually reproducing taxa, many of which were suspected to be able to freely hybridize in overlapping areas (Škorničková 2007,

Škorničková and Sabu, 2005b and Škorničková et al., 2007). Presented results suggested plausible hybridization events in at least two pairs of species with sympatric occurence, between *C. inodora* and *C. pseudomontana* and between *C. cannanorensis* and *C. mutabilis*. However, in both cases paleo-allopolyploid origin of these species cannot be ruled out.

Exact pathways of the reticulate evolution in *Curcuma*, caused by polyploidization and hybridization cannot be infered only based on one rDNA region and AFLP fingerprinting. Rather several DNA regions from different parts of genomes (cpDNA, nDNA) should be compared and preferably more single-copy nuclear region should be employed for inference of exact reticulate relationships among *Curcuma* species.

5.2. Taxonomic position of *Curcuma*-like species

The taxonomic position of some species with *Curcuma*-like morphology has long been discussed (summarized in Škorničková, 2007). Traditionally, they have been placed into separate genera such as *Hitchenia*, *Monolophus, Kaempferia*, *Paracautleya* and *Stahlianthus*. However, a phylogenetic analysis of the tribe Zingibereae based on internally transcribed spacers and *trn*L–*trn*F gene sequences (Ngamriabsakul et al., 2004) did not support their independent generic status as these taxa were well nested within the *Curcuma* complex. This situation was in general confirmed by cytological study (Leong-Škorničková et al., 2007), where genome sizes of the three controversial species *Hitchenia caulina*, *Monolophus scaposus (≡ Kaempferia scaposa)* and *Paracautleya bhatii (≡ Curcuma bhatii*) were found to match hexaploid *Curcuma* taxa from the genome size group II. In addition, all had the hexaploid number of chromosomes (2n = 42) and similar geographical distribution in the Western Ghats (India), a biodiversity hotspot for *Curcuma*. By contrast, *Stahlianthus involucratus* possessed unique holoploid and homoploid genome sizes (Cx =1Cx = 1.56 pg), dissimilar to any other species involved in the cytological study. The diploid number of chromosomes (2n = 22) was shared only with south Indian *C. vamana*, but this species has a different Cx-value and overall morphology.

In the present ITS study all of these taxa were involved in phylogenetic analysis to test the hypothesis about their close relationship to genus *Curcuma*. *Paracautleya bhatii*, which was recently classified as a member of genus *Curcuma* (Škorničková and Sabu, 2005a), were involved also in AFLP study. According to results from both molecular markers, this taxon is unambiguosly member of genus *Curcuma* subg. *Curcuma*, belonging to genome size group II. *Hitchenia caulina*, and *Monolophus scaposus* were involved in ITS study as a members of related genera. Their position within genus *Curcuma* as a members of genome size group II („in molecular sense“) were well supported, similarly as in case of *C. bhatii*, therefore inclusion of these species into genus *Curcuma* is justified by this work.

Two species of *Stahlianthus* (*S. involucratus* and *S. campanulatus*) were another representants of related species in ITS study. Their position in phylogenetic reconstruction

indicated close relationship to taxa belonging to subg. *Hitcheniopsis*. Also *Curcuma vamana* that shared the same diploid number of chromosomes with *Stahlianthus involucratus*, but differ by genome size was classified as a member of subg. *Hitcheniopsis*. Further cytological study of the genus *Curcuma* subg. *Hitcheniopsis* is in progress by Šída et al.

6. Conclusions

Analysis of selected species of genus *Curcuma* by two independent molecular markers (AFLP fingerprinting and cloning and sequencing of ITS region nrDNA) revealed very similar pattern. Species of subgenus *Curcuma* were split into two lineages, that probably evolved independently. The evolution force in the first lineage was mainly allopolyploidization traceable in both, higher ploidy levels as well as hexaploid taxa. Second lineage was most plausibly influenced mainly by paleo-allopolyploidization on hexaploid level or hybridization/introgression between hexaploid species with sympatric occurence.

These two lineages are formed by three groups of species (genome size group I, II, III) previously defined by Leong-Škorničková et al. (2007) that differ by homoploid genome size. First lineage covers all species belonging to genome size group I with lowest homoploid genome sizes, but includes also single species *C. angustifolia* with higher homoploid genome size perfectly matching group II species. Second lineage covers all other group II species and two species belonging to genome group III that have the highest range of homoploid genome size. None of phylogenetic analyses supported genome size group III as independent evolutionary lineage. Although there is no detailed knowledge about evolution of the genome in the genus *Curcuma*, deviations in genome sizes could be e.g. consequences of genome reorganisation after hybridization, followed by hybrid speciation (Rieseberg, 1997) as all above mentioned species are most likely either of allopolyploid or hybrid origin. These „exceptions“ have caused, that genome size groups defined in Leong-Škorničková et al. (2007) cannot be regarded as a monopyletic group in „molecular sense“ outlined by this study.

Yet the determination of cytological characteristics (including homoploid genome sizes) provides primary insights to *Curcuma* species relationships and should be extended to the rest of the genus *Curcuma* s.l. in its complete distribution range. These primary insights should be subsequently verified by molecular studies. Similarly, AFLP fingerprinting and ITS sequencing differed slightly in outcomes, that were further interpreted as phylogenetic hypothesis. Therefore, combination of these markers, rather than using them separately, can provide more reliable phylogenetic hypothesis.

However, exact pattern of the complex relationship network of *Curcuma* species cannot be infered based on one rDNA region and AFLP analyses only. Several DNA regions from different parts of genomes (cpDNA, nDNA) should be compared (similarly as in Guo et

al, 2004; Hodkinson et al., 2002 and many others) and preferably more single-copy nuclear region should be employed for inference exact relationships (Strand et al. 1997, Small et al. 2004) among *Curcuma* species.

Sequencing of ITS region provided unambiguous hypothesis that species belonging to *Curcuma*-like genera, covered in this study have evolved either in lineage of subgenus *Curcuma* (*Hitchenia caulina, Monolophus scaposus*) or within the lineage of subgenus *Hitcheniopsis* (*Stahlianthus involucratus, Stahlianthus campanulatus*). This result is in agreement with previous hypothesis based on molecular markers (Ngambriabsakul et al., 2004) as well as cytological study by Leong-Škorničková et al. (2007). Generic delimitation of the genus *Curcuma* and independent generic status of *Hitchenia*, *Monolophus* and *Stahlianthus* have to be further studied, with inclusion of more species from several other closely related genera including *Smithatris*, *Laosanthus*, and several *Kaempferia* species.

7. References

Alvarez I., Wendel J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29: 417–434.

Adams K.L., Wendel J.F. 2005. Polyploidy and genome evolution in plants. Current Opinion in Plant Biology 8: 135–141.

Apavajtrut P., Anuntalabhochai S., Sirirugsa P., Alisi C. 1999. Molecular markers in the identification of some early flowering *Curcuma* L. (Zingiberaceae) Species. Ann. Bot. (London) 84: 529-534.

Apavatjrut P., Sirisawad T., Sirirugsa P., Voraurai P., Suwanthada C. 1996. Studies on chromosome number of seventeen Thai *Curcuma* species. *Proceedings of 2nd National Conference on Flower and Ornamental Plant* 2: 86-99

Ardiyani M. 2002. Systematic study of *Curcuma* L.: Turmeric and its allies. PhD thesis, University of Edinburgh. (Unpublished).

Arnold M.L., Bennett B.D., Zimmer E.A. 1990. Natural hybridization between *Iris jidva* and *Iris hexagona:* pattern of ribosomal DNA variation. Evolution 44: 1512-1521.

Bailey C.D., Carr T.G., Harris S.A., Hughes C.E. 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol. Phylogenet. Evol. 29: 435-455.

Baker J.G. 1890. *Curcuma.* The Flora of British India. Vol. 6. L. Reeve & Co., London. Pp. 209-216.

Baldwin B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. Mol. Phylogenet. Evol. 1: 3–16.

Baldwin B.G., Robichaux R.H. 1995. Historical biogeography and ecology of the Hawaiian silversword alliance (Asteraceae): new molecular phylogenetic perspectives. In: W. L. WAGNER and V. A. FUNK, eds. Hawaiian biogeography: evolution on a hot spot archipelago. Smithsonian Institution Press, Washington, DC. Pp. 259-287

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 on angiosperm phylogeny. Ann. Missouri Bot. Gard. 82: 247–277.

Beardsley P.M.A., Yen R., Olmstead G. 2003. AFLP phylogeny of *Mimulus* section *Erythranthe* and the evolution of hummingbird pollination. Evolution 57: 1397-1410

Benham J., Jeung J.U., Jasieniuk M., Kanazin V., Blake T.K. 1999. Genographer: A Graphical Tool for Automated Fluorescent AFLP and Microsatellite Analysis. Journal of Agricultural Genomics 4:3

Bobola M.S., Smith D.E., Klein A.S. 1992. Five major nuclear ribosomal repeats represent a large and variable fraction of the genomic DNA of *Picea rubens* and *P. mariana.* Mol. Biol. Evol. 9: 125–137.

Brandizzi F., Caiola M.G. 1998. Flow cytometric analysis of nuclear DNA in *Crocus sativus* and allies (Iridaceae). Pl. Syst. Evol. 211: 149-154.

Branney T.M.E. 2005. Hardy Gingers. Royal Horticultural Society. Timber Press. Portland, Cambridge. Pp. 267

Bretagnolle F., Thompson J.D. 1995. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolutin of autopolyploid plants. New Phytologist 129: 1–22.

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

Buckler E. S. IV, Holtsford T.P. 1996b. *Zea* ribosomal evolution and substitution patterns. Mol. Biol. Evol. 13: 623-632.

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

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. Biological & Pharmaceutical Bulletin 24: 1389-1394.

Cao H., Komatsu K. 2003. Molecular identification of six medicinal *Curcuma* plants produced in Sichuan: evidence from plastid trnK gene sequences. Yao Xue Xue Bao 38: 871-875.

Cluster P.D., Calderini O., Pupilli F., Crea F., Damiani F., Arcioni S. 1996. The fate of ribosomal genes in three interspecific somatic hybrids of *Medicago sativa*: Three different outcomes including the rapid amplification of new spacer-length variants. *Theoretical and Applied Genetics* 93: 801–808.

Crease T.J., Lynch M. 1991. Ribosomal DNA variation in *Daphnia pulex.* Mol. Biol. Evol. 8: 620-640.

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 Phytologist 174: 658–668.

Denduangboripant J., Cronk Q.C.B. 2000. High intraindividual variation in internal transcribed spacer sequences in *Aeschynanthus* (Gesneriaceae): implications for phylogenetics. Proceedings of the Royal Society of London 267: 1407-1415

Delseny M., McGrath J.M, This P., Chevre A.M, Quiros C.E. 1990. Ribosomal RNA genes in diploid and amphidiploid *Brassica* and related species: organization, polymorphism, and evolution. Genome 33: 733-744.

Despres L., Gielly L., Redoutet B., Taberlet P. 2003. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. Mol. Biol. Evol. 27: 185–196.

Doležel J., Greilhuber J., Suda J. 2007a. Flow cytometry with plants: an overview. In: Doležel J., Greilhuber J., Suda J., eds. Flow cytometry with plant cells. Weinheim: Wiley-VCH Verlag, 41–65.

Ekenäs C., Rosén J., Wagner S., Merfort I., Backlund A., Andreasen K. 2009. Secondary chemistry and ribosomal DNA data congruencies in *Arnica* (Asteraceae). Cladistics 25: 78-92

Endress P.K. 1994. Diversity and evolutionary biology of tropical flowers. Cambridge University Press, Cambridge, UK.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. Annual Review of Genetics 22: 521-256

Goldblatt P. 1980. Polyploidy in angiosperms: Monocotyledons. In: Lewis WH, ed. Polyploidy: biological relevance. New York: Plenum Press, 219–239.

Gottlieb A.M., Giberti G.C., Poggio L. 2005. Molecular analyses of the genus *Ilex* (Aquifoliaceae) in southern South America, evidence from AFLP and ITS sequence data. American journal of botany 92: 352–369.

Grimm G.W., Denk T. 2008. ITS Evolution in *Platanus* (Platanaceae): Homoeologues, Pseudogenes and Ancient Hybridization. Ann Bot. 101: 403-419.

Guo Y.P., Ehrendorfer F., Samuel R. 2004. Phylogeny and Systematics of *Achillea* (Asteraceae-Anthemideae) Inferred from nrITS and Plastid *trn*L-F DNA Sequences. Taxon 53: 657-672

Guo Y.P., Saukel J., Mittermayr R., Ehrendorfer F. 2005. AFLP analyses demonstrate genetic divergence, hybridization, and multiple polyploidization in the evolution of *Achillea* (Asteraceae-Anthemideae). New Phytologist 166: 273–290

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

Hodkinson T.R., Chase M.W., Takahashi Ch., Leitch I.J., Bennett M.D., Renevoize S.A. 2002. The use of DNA sequencing (ITS and *trn*L-F), AFLP, and fluorescent in situ hybridization to study allopolyploid *Miscanthus* (Poaceae). Am. J. Bot. 89: 279–286.

Hodkinson T. R., Renvoize S.A., Chonghaile G.N., Stapleton C., Chase M.W. 2000. A comparison of ITS Nuclear rDNA sequence data and AFLP markers for phylogenetic studies in *Phyllostachys* (Bambusoideae, Poaceae). Journal of Plant Research 113: 259–269.

Horaninow P. 1862. Prodromus Monographiae Scitaminearum. Typis Academiae Caesareae Scientiarum. Petropoli. Pp. 1-45.

Hsiao C., Chatterton N.J., Asay K.H., Jensen K.B. 1994. Phylogenetic relationships of 10 grass species: an assessment of phylogenetic utility of the internal transcribed spacer region in nuclear ribosomal DNA in monocots. Genome 37: 112-120.

Huber K.T., Moulton V. 2005. Phylogenetic networks. In: Gascuel O, editor. Mathematics of evolution and phylogeny. Oxford: Oxford University Press. p 178–200.

Huelsenbeck J.P. 2000. MrBayes: Bayesian inference of phylogeny. Distributed by the author. Department of Biology, University of Rochester.

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

Chen Y., Bai S., Cheng K., Zhang S., Nian L. 1999. Random amplified polymorphic DNA analysis on *Curcuma wenyujin* and *C. sichuanensis*. Zhongguo Zhong Yao Za Zhi. 24: 131-133.

Chen Z.Y., Chen S.J. 1984. A report on chromosome numbers of Chinese Zingiberaceae. Guihaia 4: 13-18.

Islam M.A., Meister A., Schubert V., Kloppstech K., Esch E. 2007. Genetic diversity and cytogenetic analyses in *Curcuma zedoaria* (Christm.) Roscoe from Bangladesh. Genetic Research and Crop Evolution 54: 149-156.

Islam M.A, Kloppstech K., Esch E. 2005. Population genetic diversity of *Curcuma zedoaria* (Christm.) Roscoe – a conservation prioritised medicinal plant in Bangladesh. Conservation Genetics 6: 1027-1033.

Islam M.A. 2004. Genetic diversity of the genus *Curcuma* in Bangladesh and further biotechnological approaches for in vitro regeneration and long-term conservation of *C. longa* germplasm. PhD thesis, University of Hannover. (Unpublished).

Jaccard P. 1908. Nouvelles recherches sur la distribution florale. Bulletin de la Société Vaudense des Sciences Naturelles 44: 223-270.

Jellen E. N., Phillips R.L, Hines H.W. 1994. Chromosomal localization and polymorphisms of ribosomal DNA in oat *(Avena* spp.). Genome 37: 23-32.

Jorgensen R.A., Cluster P.D. 1988. Modes and tempo in the evolution of nuclear ribosomal DNA: new perspectives from evolutionary studies and new markers for genetic and population studies. Ann. MO. Bot. Gard. 75: 1238-1247.

Joseph R., Joseph T., Jose J. 1999. Karyomorphological studies in the genus *Curcuma* Linn. Cytologia 64: 313-317.

Kardolus J.P., van Eck H.J., van den Berg R.G. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). Plant Syst. Evol. 210: 87-103.

Karp A., Seberg O., Buiatti M. 1996. Molecular techniques in the assessment of botanical diversity. Annals of Botany (London) 78: 143–149.

Karvonen P., Savolainen O. 1993. Variation and inheritance of ribosomal DNA in *Pinus sylvestris* L. (Scots pine). Heredity 71: 614-622.

Kececioglu J., Gusfield D. 1998. Reconstructing a history of recombinations from a set of sequences. Discrete Applied Mathematics 88: 239-260

Kim K.J., Jansen R.K. 1994. Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelion (*Krigia*): additional information from internal transcribed spacers of nuclear ribosomal DNA. Plant Syst. Evol. 190: 157-185.

Koopman W.J.M. 2005. Phylogenetic Signal in AFLP Data Sets. Systematic Biology 54:2, 197 – 217

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. Am. J. Bot. 92: 167-178.

Kress W.J., Prince L.M., Williams K.J. 2002. The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. Am. J. Bot. 89: 1682-1696.

Kučera J., Tremetsberger K., Vojta J., Marhold K. 2008. Molecular study of the *Cardamine maritima* group (Brassicaceae) from the Balkan and Apennine Peninsulas based on amplified fragment length polymorphism. Plant Syst. Evol. 275:193-207

Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., et al. 2007. ClustalW and ClustalX version 2.0. Bioinformatics 23: 2947–2948.

Larsen K., Lock J.M., Maas H., Maas P.J.M. 1998. Zingiberaceae. In: Kubitzki, K. (Ed.), The families and genera of vascular plants, vol. IV. Pp. 474-495. Springer-Verlag, Berlin, Germany.

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. Botany (London) 100: 505-526.

Lewis P.O., Holder M.T., Holsinger K.E. 2005. Polytomies and Bayesian phylogenetic inference. Systematic Biology 54: 241

Lim S.N. 1972a. Cytogenetics and taxonomy of the genus *Globba* L. (Zingiberaceae) in Malaya. II. Cytogenetics. Notes Roy. Bot. Gard. Edinburgh 31: 271–284.

Lim S.N. 1972b. Cytogenetics and taxonomy of the genus *Globba* L. (Zingiberaceae) in Malaya. IV Distribution in relation to polyploidy. Gard. Bull. Singapore 26: 115–126.

Lin L-S., Ho T.D., Harlan J.R. 1985. Rapid amplification and fixation of new restriction sites in the ribosomal DNA repeats in the derivatives of a cross between maize and *Tripsacum dactyloides*. Developmental Genetics 6: 101-112.

Linares A.R., Bowen T., Dover G.A. 1994. Aspects of nonrandom turnover involved in concerted evolution of intergenic spacers within the ribosomal DNA of *Drosophila melanogaster*. J. Mol. Evol. 39: 151-157.

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

McDade L.A. 1995. Hybridization and phylogenetics. In: Hoch P.C., Stephenson A.G. (Eds.), Experimental and Molecular Approaches to Plant Biosystematics. Missouri Botanical Gardens, St. Louis, Missouri, pp. 305–331

Mukherjee I. 1970. Chromosome studies of some species of *Hedychium*. Botanical Magazine (Tokyo) 83: 237-241.

Nasir Uddin S. 2000. The origin of turmeric: the domestication of an important ginger spice. M. SC. Thesis, University of Edinburgh, Scotland.

Nayak S., Naik P.K., Acharya L.K., Pattnaik A.K. 2006. Detection and evaluation of genetic variation in 17 promising cultivars of turmeric (*Curcuma longa* L.) using 4C nuclear DNA content and RAPD markers. Cytologia 71: 49-55.

Nei M. 1996. Phylogenetic analysis in molecular evolutionary genetics. Annual Review of Genetics 30: 371-403

Nei M., Li W. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Science USA 76: 5269-5273.

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. Edinb. J. Bot. 60: 483–507.

Nylander J.A.A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary BiologyCentre, Uppsala University.

O'Hanlon P.C., Peakall R. 2000. A simple method for the detection of size homoplasy among amplified fragment length polomorphism fragments. Mol. Ecol. 9: 815–816.

Paisooksantivatana Y., Kako S., Seko H. 2001a. Genetic diversity of *Curcuma alismatifolia* Gagnep. (Zingiberaceae) in Thailand as revealed by allozyme polymorphism. Gen. Res. Crop Evol. 48: 459-465.

Paisooksantivatana Y., Kako S., Seko H. 2001b. Isozyme polymorphism in *Curcuma alismatifolia* Gagnep. (Zingiberaceae) populations from Thailand. Sci. Hort. 88: 299-307.

Palmer J.D. 1985. Chloroplast DNA and molecular phylogeny. BioEssays 2: 263-267

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

Podani J. 2001. SYN-TAX 2000, Computer programs for data analysis in ecology and systematics. User's manual. Budapest: Scientia.

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

Popp M., 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 Systematic botany 30: 302-313.

Poulsen A.D. 1993. Two new species of *Boesenbergia* (Zingiberaceae) from Borneo. Nordic Journal of Botany 13: 289–294.

Prana M.S. 1977. Studies on some Indonesian *Curcuma* species. PhD. Thesis, University of Birmingham. (Unpublished)

Rangsiruji A., Newman M.F., Cronk Q.C.B. 2000b. A study of the infrageneric classification of *Alpinia* (Zingiberaceae) based on the ITS region of nuclear rDNA and the *trn*L-F spacer of chloroplast DNA. In: Wilson, K.L. & Morrison, D.A. (Eds.). Monocots: Systematics and Evolution. Pp. 695-709.

Refoufi A., Jahier J., Esnault M.A. 2001. Genome analysis of a natural hybrid with 2n=63 chromosomes in the genus *Elytrigia* Desv. (Poaceae) using the GISH technique. Plant Biology 3: 386-390.

Rieseberg L.R. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): evidence from ribosomal genes. Am. J. Bot. 78: 1218-1237.

Rieseberg L.H. 1997. Hybrid origins of plant species. Annual review of ecology and systematics 28: 359-389

Rieseberg L.R., Carter R., Zona S. 1990. Molecular tests of the origin of two diploid *Heliunthus* species (Asteraceae). Evolution 44: 1498-1511.

Robinson J.P., Harris S.A. 1999. Amplified fragment length polymorphisms and microsatellites: a phylogenetic perspective. In: Which DNA marker for which purpose? Chap. 12. Final Compendium of the Research Project Development, optimization and validation of molecular tools for assessment of biodiversity in forest trees in the European Union. *Edited by* E.M. Gillet. DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. Available from http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm [accessed 1999]

Rogers S.O., Bendich A.J. 1987. Ribosomal DNA genes in plants: variability in copy number and the intergenie spacer. Plant Mol. Biol. 9: 509-520.

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

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. USA 92: 6813–6817.

Santapau H. 1952. On a common species of *Curcuma* of Bombay and Salsette Islands. Journal of the Bombay Natural History Society 51: 135-139.

Searle R.J., Hedderson A.J. 2000. A preliminary phylogeny of the *Hedychieae* tribe (Zingiberaceae) based on ITS sequences of the nuclear rRNA cistron. Pp. 710-718. In: Wilson K.L., Morrison D.A. (Eds.). Monocots: Systematics and Evolution.

Schonswetter P., Tribsch A. 2005. Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). Taxon 54: 725–732.

Schumann K. 1904. Zingiberaceae. In: Engler, A. (Ed.) Das Pflanzenreich IV, vol. 46. Pp. 1-458. Leipzig. Germany

Simmons M.P., Ochoterena H. 2000. Gaps as characters in sequencebased phylogenetic analyses. Syst. Biol. 49: 369-381

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

Sirirugsa P. 1996. The genus *Curcuma* of Thailand. Proceedings of the 2nd Symposium on the Family Zingiberaceae. Guangzhou, China. 39-46.

Sirisawad T., Sirirugsa P., Suwanthada C., Apavatjrut P. 2003. Investigation of chromosome numbers in 20 taxa of *Curcuma.* Proceedings of the 3rd Symposium on the Family Zingiberaceae. Khon Kaen, Thailand. Pp. 54-62.

Sites J.W., Davis S.K. 1989. Phylogenetic relationships and molecular variability within and among six chromosome races of *Sceloporus grammicus* (Saura, Iguanidae) based on nuclear and mitochondrial markers. Evolution 43: 296-317.

Skalická K., Lim K.Y., Matyášek R., Koukalová B., Leitch A., Kovařík A. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. Am. J. Bot. 90: 291–303.

Small R.L., Cronn R.C., Wendel J.F. 2004. Use of nuclear genes for phylogeny reconstruction in plants. Austral. Syst. Bot. 17: 145–170.

Smith R.M. 1981. Synoptic Keys to the genera of Zingiberaceae pro parte. Royal Botanic Garden Edinburgh. Pp. 1-28.

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). Am. J. Bot. 82: 1329–1341.

Strachan T., Webb D., Dover G.A. 1985. Transition stages of molecular drive in multi-copy DNA families in *Drosophila.* EMBO J. 4:1701-1708.

Strand A.E., Leebens-Mack J., Milligan B.G. 1997. Nuclear DNA-based markers for plant evolutionary biology. Molec. Ecol. 6: 113–118

Sun Y., Skinner D.Z, Lang G.H, Hulbert S.H. 1994. Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. 89:26-32.

Swofford D.L. 2002. PAUP*: Phylogenetic Analysis Using Parsimony. ver. 4.0 beta 10. Sinauer, Sunderland.

Škorničková J. 2007. Taxonomic Studies in Indian *Curcuma* L. PhD. Thesis. Charles University, Prague.

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. In: Ravindran, P.N. (ed.) Turmeric: the genus *Curcuma.* CRC Press, Florida Pp. 451-467

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.

Takano A., Okada H. 2002. Multiple occurrence of triploid formation in *Globba* (Zingiberaceae) from molecular evidence. Pl. Syst. Evol. 230: 143-159.

Takano A. 2001. Cytological analyses of 19 taxa in *Globba* (Zingiberaceae). Acta Phytotaxonomica et Geobotanica 52: 65–74.

Tribsch A., Schönswetter P., Stuessy T.F. 2002. *Saponaria pumila* (Caryophyllaceae) and the Ice Age in the Eurépean Alps. Am. J. Bot. 89: 2024-2033.

Valeton T. 1918. New notes on the Zingiberaceae of Java and Malaya. Bulletin du Jardin Botanique de 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.

Vogler A.P., DeSalle R. 1994. Evolution and phylogenetic information content of the ITS-l region in the tiger beetle *Cicindelu dorsalis*. Mol. Biol. Evol. 11: 393-405.

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414.

Weiss H., Maluszynska J. 2000. Chromosomal rearrangement in autotetraploid plants of *Arabidopsis thaliana*. Hereditas 133: 255–261.

Weiss-Schneeweiss H. , Tremetsberger K., Schneeweiss G.M., Parker J.S., Stuessy T.F. 2008. Karyotype diversification and evolution in diploid and polyploid South American *Hypochaeris* (Asteraceae) inferred from rDNA localization and genetic fingerprint data. Ann. Bot. 101: 909-918.

Wendel J.F., Schnabel A., Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proceedings of the National Academy of Sciences, USA 92: 280-284.

White T.J., Bruns T., Lee S., Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M., Gelfand D., Sninsky J., White T. (Eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego, pp. 315–322.

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. Am. J. Bot. 91: 100-114.

Wissemann V. 2002. Molecular evidence for allopolyploid origin of the *Rosa canina* complex (Rosaceae, Rosoideae). J. Appl. Bot. 76: 176–178.

Wood T., Whitten W.M., Williams N.H. 2000. Phylogeny of *Hedychium* and related genera (Zingiberaceae) based on ITS sequenced data. Edinburg J. Bot. 57: 261-270.

Xia Y.M., Kress W.J., Prince L.M. 2004. Phylogenetic analyses of *Amomum* (Alpinioideae: Zingiberaceae) using ITS and *mat*K DNA Sequence Data. Syst. Bot. 29: 334-344.

Young N.D., Healy J. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4:6.

Zonneveld B.J.M., 2001. Nuclear DNA contents of all species of *Helleborus* (Ranunculaceae) discriminate between species and sectional divisions. Pl. Syst. Evol. 229: 125-130.

2. Apendix

Apendix 1. Within-individual ITS polymorphism. For each individual in ITS study (corresponding to collection number, see Table 1 or 3) unique sequences (uppercase letters), duplicate sequences („DUPLICATE“) and recombinant sequences („RECOMBINANT“) were examined by within-individual alignment. Question mark were put, where recombinant sequences were suggested but not unambiguously determinable. Within-individual alignments for outgroup species are not presented. Only variable sites are presented.

HEXAPLOIDS GENOME GROUP I

```
                        10        20        30        40        50
                ....|....|....|....|....|....|....|....|....|....|.
84183arom6x1    ACATAGCTA CA  ATACT    ATAACC    TTG  C ATTGATCCC   A
84183arom6x7    ACATAGCTA CA  ATACT    ATAACC    TTG  C ATTGATCCC   A DUPLICATE
84183arom6x3    ACATAGCTA CA  ATACT    ATAACC    TTGGTT CT    GTT   B (RECOMBINANT A/D)
84183arom6x5    ACATAGCTA CA  ATACT    ATAACC    CCTGTT CC    GCCCG C
84183arom6x6    ACATAGCTA CAACATACT    ATAACC    TTGGTT CT    GTT   D
84183arom6x2    ACAT  TTA CAACATG C    ACGATT    CTG  C ATTGATCCC   E
84183arom6x4    ACAT  TGCC     CG TTTCTACAGTTTGAGCCG  C ATTGATCCC   F
84183arom6x8    CGT   TTA CAACATG C    GCAGCTTGAGCCG  CAAC    GCCCG G

                        10        20        30
                ....|....|....|....|....|....|....|..
84170aromat2      TTA GCCCAACATAG C    AGTGA    T     A
84170aromat4      TTA GCCCAACATAG C    AGTGA    T     A DUPLICATE
84170aromat5      TGCCGCC     CAG TTTCTGGTGA    T     RECOMBINANT A/B
84170aromat1      TGCCGCC     CAG TTTCTAACAGTGAGCTCAA B
84170aromat3    AGCTA GCCCA  ATGACC    AGTGA    T     C

                ....|....|....|....
71484montana1   C  AGTTGGAT    TGCG A
71484montana3   C  AGTTGGAT    TGCG A DUPLICATE
71484montana7   CCAGGTCGGAT         RECOMBINANT B/C
71484montana6   CCAGGTCGGATTGAT     B
71484montana4   CCAGGTCGGATTGAT     B DUPLICATE
71484montana2   T  GATCGGATTGAT     RECOMBINANT C/B
71484montana5   T  GACCTAGC         C
71484montana8   T  GACCTAGCTGGT     RECOMBINANT C/B

                        10        20        30
                ....|....|....|....|....|....|...
73479mon4       CCACGGGCGTCAGTGAT  GCAAC    TGATC A
73479mon10      CCACGGGCGTCAGTGAT  GCAAC    TGATC A DUPLICATE
73479mon13      TC  GGGCGCCAGTAGT  GCGAC    TGATC B
73479mon3       TC  GGGCACCATTAGT  GCGAC    TGATC B DUPLICATE (1 point mutation)
73479mon14      TC  GGGCACCATTAGT  GCGAC    TGATC B DUPLICATE (1 point mutation)
73479mon7       TC      GCCAGTAGT  GCGAT    CGATC C
73479mon8       TC      GCCAGTAGT  GCGAT    CGATC C DUPLICATE
73479mon5       TC      GCCAGTAGT  GCGAT    CGATC C DUPLICATE
73479mon12      TC      GCCAGTAGT  GCGAT    CGATC C DUPLICATE
73479mon15      TC  GGGCACTAGTGAT  ATGAT        C D
73479mon2       TC  GGGCGTCAGTGACCTGTGATTCAATGATG E
73479mon1       CCACGGGCGTCAGTAGT  GCGAT    CGATC RECOMBINANT A/C
73479mon9       CCACGGGCGTCAGTGAT  GCGAT    CGATC RECOMBINANT A/C

                        10        20        30
                ....|....|....|....|....|....|.
73425montana5   GATGTGGCATTGCGTGAGCTCAATGATTGCG A
73425montana2       TT   CCACGGTAGCTCAATGAT     B
73425montana1       CT CACCGTAGGGAT             C
73425montana4       CT CACCGTAGGGAT             C DUPLICATE
73425montana3       CT CACCGTGGGGATTCAATGATTGCG RECOMBINANT C/A
```

```
                   10        20        30        40        50        60
          ....|....|....|....|....|....|....|....|....|....|....|....|.
71443affprak4        AGAGCCGCCG    CGCCTGGGG    GCCCAAGTTTCTGAGGGCCCTGTTT   A
71443affprak6   GCATCATAGAGCCGCCG   CGCCTGGGGTCTTGCCCAAGTTTCTGAGGGCCGCGCCCCG B
71443affprak1   GCATCATGGAGCCGCACATGTCGCCCGGGGTCTTGCCCAAGTTTCTGAGGGCCGCACCC   C
71443affprak5   GCATCATAGAT AGCACATG CGCTCGGGG    GCCCGAGCCTCTGAGGGCCGCGCCC   D
71443affprak8   GCATCATAGAT AGCACATG CGTCTGGGG    GTCCAGATTTC    GGCTGCGCCC   E
71443affprak2   GCA    TATGG       TG CGTCTGGAA    GCTTAAACTCT    AAGTCTGTTT   F
71443affprak7   GCA    TATGG       TG CGTCTGGAA    GCTTAAACTCT    AAGTCTGTTT   F DUPLICATE
71443affprak3   GCA    TATGG       TG CGTCTGGAA    GCTTAAACTCT    AAGTCTGTTT   F DUPLICATE
```

```
                   10        20
          ....|....|....|....|
71476affprak2     C     TTCAAGCCGCA A
71476affprak5     C     TTCAAGCCGCA A DUPLICATE
71476affprak8     C     TTCAAGCCGCA A DUPLICATE
71476affprak1   CAATCTTCTT    T      B
71476affprak3   CAATCTTCTT    T      B DUPLICATE
```

HIGHER POLYPLOIDS GROUP I

```
                  10        20        30        40        50        60        70        80        90       100
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
71431_19     ATACACCGTAGCCAT AAC      ACTGTGGT    GGCTACACTCCCAG    GGTTGCTGT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_20     ATACACCGTAGCCAT AAC      ACTGTGGT    GGCTACACTCCCAG    GGTTGCTGT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_5       CATATTGTA TAGCCAACATG   ACTGTGGT    GGCTACACTCCCAG    GGTTGCTGT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_13      CATATTGTA TAGCCAACATG   ACTGTGGT    GGCTACACTCCCAG    GGTTGCTGT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_10      CATATTGTA TAGCCAACATG   G CGCGGT    GGCCACGTTTCCAG    GGCCACTGC    G                   GCG  CCCC
71431_16      CATATTGTA TAGCCAACATG   G CACGGT    AGCCACGTTTCTAG    GGCCAATGC    G                          CC
71431_18      CATATTGTA TAGCCAACATG   G CGCGGT    GGCTACACTCCCAG    GGCTGCTGC    GGATGTGTGATTGATGATTGCAGACCCCC
71431_7       CATACTGTA TAGCCAACATG   G CGTGGT    GGCCGCATTTCCAG    GGCTGCTGC    GGATGTGTGATTGATGATTGCGGAGCCCC
71431_3       CATACTGTAGCCGC   CGTG   G CGTGGT    GGTTACACCTTCAG    AAGTGCTCT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_15        ACTATGGCCGC   CATG    GCTGTGGT    GGCTACACTCCCAA    GGTTGCTGC    GGATGTGTGATTGATGATTGCAGACCCCC
71431_4      ATATACCGTAGCCAT AAC      ACTGTGGTTCTTGGTTACACCTTCAG    AAGTGCTCTGCAAGGATGCGTGAT    GATTGCGGAGCCCCCG
71431_2      A    CCGTAGCCAT AAC      ACTGTGGT    GGCTACACTCCCAG    GGCTGCTGC    GGATGTGTGAT    GATTGCGGACCCCC
71431_8      ATATACCGTAGCCAT AAC      ACTGTGGT    GGCTACACTCCCAG    GGTTGCTGC    GGATGTGTGAT    GATTGCGGACCCCC
71431_17     ATACACCGTAGCCAT AAC      ACTGTGGT    GGCTACACTCCCAG    GGTTGCCGTTCAAGGATGTGTGAT    GATTGCGGAGTCCT
71431_1      A    CCGCAGCCAT AACTTG  AACTGTAGT    GGCTACACTCCCAG    GGTTGCTGT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_9       CATACTGTAGCCGCCAACATG   G CGTGGT    GGCTACACTCCCAG    GGTTGCTGTA   GGATGTGTGAT    GATTGCGGAGTCCT
71431_6       CATATTGTA TAGCCAACATG   G CGCGGT    GGCCATGTTTCCAGTGAGGGCCGCTGC    GGATGTGTGATTGATGATTGCAGACCCCC
71431_14_BAD  CATACTGTAGCCGCCAACATG   G CGCGAT    GACCACGTTTCCAGTGAGGGCCGCCGTTCAAGGATGTGTGAT    GATTGCGGAGTCTT
71431_11      CATACTGTAGCCGCCAACATG   G TGTGGTTCTTGGCCACGTTTCCAGTGAGGGCCGCTGT    GGATGTGTGAT    GATTGCGGAGTCCT
71431_12      CATACTGTAGCCGCCAACATG   G TGTGGTTCTTGGCCACGTTTCCAG    GGCTGCTGCGCAAGGATGCGTGAT    GATTGCGGAGCCCCCG
```

```
              ....
71431_19           A
71431_20           A DUPLICATE
71431_5            B
71431_13           B DUPLICATE
71431_10           ?
71431_16           ?
71431_18           ?
71431_7       TGCG ?
71431_3            ?
71431_15           ?
71431_4            ?
71431_2            ?
71431_8            ?
71431_17           ?
71431_1            ?
71431_9            ?
71431_6            ?
71431_14           ?
71431_11           ?
71431_12           ?
```

```
                        10        20        30        40        50        60        70
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|...
84142_aerugin8    CAT  TTTCATCAGCA  ACATGTCCG CCCACCAGTTCCAG  CA                    CCCGCC A
84142_aerugin4    CAT  CTTCGTCAGCG CCCGTGTCTG TTAGTCGATTCTGA  CAGACCCTGTGTGATTGATGATTGGACCC B
84142_aerugin6    CAT  CTTCGTCAGCG CCCGTGTCCG CTCACCGATTCTGG  CAGACCCTGTG    TGATGATTGGAGTT RECOM. B/A/D/E
84142_aerugin3    CATA TTTCA   GCAGCCCATGGCTG TTAGTCAGTTCCAG  TAGACCCTGTG    TGATGATTGGAGTT C
84142_aerugin7         CTTCAGCCCGC    GTGTTCG CTCACCGACTCTGG  CAGACCCTGTGTGATTGATGATTGGACCC D
84142_aerugin1    CAT  TTTCATCAGCA  ACATGTCCG CCCACCAGTTCTGGGTTCGACCCTGTG    TGATGATTGGAGTT E
84142_aerugin2    CATAGCC    TCAGCA  TCAACTCCACTTCACTAACCTTGG  TAGACCCTGTG    TGATGATTGGAGTT F
```

```
                     10        20        30        40
               ....|....|....|....|....|....|....|....|.
71492arom9x1      AGGC          TTAAACTTACCTTAAGAGC  GCT    GTT A
71492arom9x5      AGGC          TTAAACTTACCTTAAGAGC  GCT    GTT A DUPLICATE
71492arom9x2   CATAGGC          TTAAACTTACCTTAAGAGC  GGC    CCC B
71492arom9x3   CATAGGC          TTAAACTTACCTTAAGAGC  GGC    CCC B DUPLICATE
71492arom9x7   CATAGGC          TTAAACTTACCTTAAGAGC  GGC    CCC B DUPLICATE
71492arom9x4   CATGATAGCCCAACACCGGGTCCGTTTCAGGGCC  GGC    CCC C
71492arom9x8   CATGATAGCCCAACACCGGGTCCGTTTCAGGGCC  GGC    CCC C DUPLICATE
71492arom9x6   CATAATAGCCCA  ACCGGGTCCGTTCCGGAGTTGTAGTTCAAGTT D
71492arom9x9   CATAATAGCCCA  ATCGGGCCCGTTTCAGGGCC  GGC    CCC RECOMBINANT D/B/C
```

```
                   10        20        30        40
              ....|....|....|....|....|....|....|....|....|..
73222longa7   CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A
73222longa15  CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A DUPLICATE
73222longa16  CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A DUPLICATE
73222longa11  CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A DUPLICATE
73222longa13  CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A DUPLICATE
73222longa9   CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A DUPLICATE
73222longa1   CATAACGTAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATCTGATCC A DUPLICATE
73222longa5   CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATATGATCC B
73222longa8   CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCCTCAATATGATCC B DUPLICATE
73222longa6   CATAACATAGCCCGCCCCGAGTTATACTCA  GGTTC        TT C
73222longa4   CATAACATAGCCCGCCCCGGATTGCACTTA  GGTTC        TT C MUT?
73222longa14  CATAACATAGCCCGCCCCGGGTTGCGTCCGCTGCTTC        TT RECOMBINANT A/C
73222longa2   CATGACATAGCCCGCCCCGGGTTGCGTTCA  AGTTC        TT D
73222longa10     GACGGC        TTAGGTTGTACTCA  GGT    TC    TT E
73222longa3   CATAACATAGCCCGCCCCGGGTTGTACTCA  GGCTCAATCTGATCC RECOMBINANT
73222longa12     GACGGC        TTAGGTCGTACTCA  GCCTCAATCTGATCC RECOMBINANT E/A

                    10        20
               ....|....|....|....|....|.
84182xanthor5  TATAGCCCAAGAGATCTGT C    G A
84182xanthor2  TATAGCCCAAGAGATCTGT C    G A DUPLICATE
84182xanthor4  TGGC       GG GCCCGT CTGATC B
84182xanthor1  GGGC       GG GCCC  AA    G C
84182xanthor3  TGGC       GG GCCC  AA    G C DUPLICATE (1 mutation)

                   10
              ....|....|.
71485ranchi1      CTT     A
71485ranchi8      CTT     A DUPLICATE
71485ranchi9      CTT     A DUPLICATE
71485ranchi5    ACTCT     B
71485ranchi6    ACTCT     B DUPLICATE
71485ranchi2    ACTCT     B DUPLICATE
71485ranchi4  AGACCCCTGAT C
71485ranchi3    ACCCTTGAT D (or RECOMBINANT B/C)

                     10        20        30        40        50        60        70        80
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
71432_rakta14   TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   A
71432_rakta22   TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   A DUPLICATE
71432_rakta5    TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   A DUPLICATE
71432_rakta18   TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   A DUPLICATE
71432_rakta3   GCTTCCT   TCG   GCATCCCATGTGCG CTGCAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   B
71432_rakta8   GCTTCCT   TCG   GCATCCCATGTGCG CTGCAGGCGGGTCTGAGTACAGC  CCGTGTTGC  GTT    TAAT   AC   B DUPLICATE
71432_rakta23  GCTTCCT   TCG   GCATCCCATGTGCG CTGCAGGCGGGTCTAAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   B DUPLICATE
71432_rakta7   GCTTCCT   TCG   GCATCCCATGTGCGCTTGCAGACGGGTCTGAGCACGGCGCCCGCGTCGCGTGTTTCA AAACTGATACCGC C
71432_rakta13  GCTTCCT   TCG   GCATCCCATGTGCG CTGCAGGCGGGTCTGAGTACGGC  CCGTGTTGCGTGCCTCAATAAT   AC   D
71432_rakta16  GCTTCCT   TCG   GCATCCCATGTGCG CTGCAGGCGGGTCTGAGTACGGC  CCGTGTCGCGTGTT   TCAT   GC   E
71432_rakta20  GCTTCCT   TCG   GCATCCCATGTGCGCTTGCAGACGGGTCTGAGTACGGC  CCGTGCTGCGTGTT   TAAT   TCCGT RECOMB.(1mut)
71432_rakta1    TTTTTTTCA      GCAGCCCATGGGTG TTGAGGGTGGGTCTGAGCGTGGT  CCACATTGT  ATT    TAAT   TCCGT F
71432_rakta4    TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   TCCGT ?
71432_rakta21   TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGCGTGTTTCA AAACTGATACCGC ?
71432_rakta12   TTTTCC   TCA   GCA  TCAACTGCACTTACAGGCGGGTCTGAGTACGGC  CCGTGTTGC  GTT    TCAT   AT   ?
71432_rakta11   TCTTCC   TCA   GCA  TCAACTGCACTTGCAGGCAGGTCTGAGTACGGC  CCGTGTTGC  GTT    TAAT   AC   ?
71432_rakta15   TCTTCC   TCA   GCA  TCAACTGCACTTGCAGGCGGGTCTGAGCGTGGT  CCGTGTTGC  GTT    TAAT   AC   ?
71432_rakta19   TCTCCCTTCGTCCGCCCGCATGTTGGTGCGCTCGAGGGCGCCTCTGAGTACGGC  CCGCGTTGC  GTT    TAAT   AC   ?
71432_rakta6    CTTTTTTCA      GCAGCCCATGGCTG TTGAGGGCGGGTCTGAGCGTGGT  CCGTGTTGC  GTT    TAAT   AC   ?
71432_rakta2    CTTTTTTCA      GCAGCCCATGGGTG TTGAGAGCGGGCAAGGACGTGTT  ATGCGTTAC  GTT    TAAT   TGCGT ?
71432_rakta17   CTTTTTTCA      GCAGCCCATGGGTG TTGAGGGTGGGTCTGAGCGTGGT  CCGCGTTGC  GTT    TAAT   TCCGT ?
```

HEXAPLOIDS GROUP II

```
                   10        20
              ....|....|....|....|..
73453angust1   C  GT  GCCCCAATTAA C A
73453angust3   C  GT  GCCCCAATTAA C A DUPLICATE
73453angust2   C  GCAACCCCCAACTAA C B
73453angust5   C  GCAACCCCCAATTAACT C (or RECOMBINANT B/A/D ?)
73453angust4  ATGCCAGTGGTTAAGGTCGGCT D

                  10        20
             ....|....|....|....|....|...
73480angu3   ATGCCAGTGGCTTAAGAGGCGGATCTAT A
73480angu4   ATGCCAGTGGCTTAAGAGGCGGATCTAT A DUPLICATE
73480angu10   C    GCAACCCCCCAAG   GGC TGC B
73480angu14   C    GCAACCCCCCAAG   GGC TGC B DUPLICATE
73480angu2    C    GCAACCCCCCAAG   GGC TGC B DUPLICATE
73480angu7    C    GCAACCCCCCAAG   GGC TGC B DUPLICATE
73480angu8    C    GCAACCCCCCAAG   GGC TGC B DUPLICATE
73480angu6    C    GCAACTCCCCAAG   GGC TGC B DUPLICATE (1 point mutation)
73480angu1    C    GCAACCCCCCAAGGCGGGC CGC C
73480angu12   C    GCAACCCCCCAAGGCGGGC CGC C DUPLICATE
73480angu13   C    GCAACCCCCCAAGGCGGGC CGC C DUPLICATE
73480angu16   C    GCAACCCCCCAAGGCGGGC TGC RECOMBINANT C/B
73480angu9   AC    GCAACCCCCCAAGGCGGATCTAT RECOMBINANT A/C/A
73480angu11   C    GCAACCCCCCAAGGCGGATCTAT RECOMBINANT C/A
73480angu5    C    GCAACCCTAAGAAGCGAGC CGT RECOMBINANT C/A/C/A
73480angu15   C    GCAACCCTCAGAGGCGGGTCTAT RECOMBINANT C/A/C/A  (1 mutation)
```

```
                  10        20
              ...|....|....|....|....|....
84143_canna2   G ACTACTCG                   A
84143_canna7   G ACTACTCG                   A DUPLICATE
84143_canna8   G ACTACTCG                   A DUPLICATE
84143_canna3   G ACTACTTCGCGTGAAAGTACCGTGTC B
84143_canna4   G ACTACTTCGCGTGAAAGTACCGTGTC B DUPLICATE
84143_canna12  G ACTACTTCGCGTGAAAGTACCGTGTC B DUPLICATE
84143_canna16  G ACTACTTCGCGTGAAAGTACCGTGTC B DUPLICATE
84143_canna5   G GCTACTTCGCGTGAAAGTACCGTGTC B DUPLICATE
84143_canna10  G GCTACTTCGCGTGAAAGTACCGTGTC B DUPLICATE
84143_canna6   G GTTACCCGGCGTGAAAGTACCGTGTC B DUPLICATE
84143_canna11  G GTTACCCGGCGTGAAAGTACCGTGTC B DUPLICATE (1 mutation)
84143_canna1   G ACTACTTC                   RECOMBINANT B/A
84143_canna13  A ACATCTCG                   C
84143_canna15  G GCTATTCG                   D
84143_canna9   G GCTATTCG                   D DUPLICATE
```

```
                 10        20
              ....|....|....|....|..
73403_inod10  TT GGA TGGAGTCAATGTAGC A
73403_inod2   TT GGA TGGAGTCAATGTAGC A DUPLICATE
73403_inod13  TTAGGA TGGAGTCAATGTAGC A DUPLICATE (1 indel)
73403_inod4   TT GGACTAGGGCCAATGCGGC B
73403_inod3   TT GGACTAGGGCCAATGCGGT B DUPLICATE (1 mutation)
73403_inod11  TT GGA TGGAGCCAATGCGGC RECOMBINANT A/B
73403_inod5   TT AAA TGGAGTCAATGCAGC C
73403_inod9   TT AAA TGGAGTCAATGCAGC C DUPLICATE
73403_inod12  TT GGA TGGAGTCAATATGGC D
73403_inod6   TT GGA TGGAGTCAGCATAGC D DUPLICATE (1 mutation)
73403_inod1   TC AAA CGGAGTCAATGTAAC E
73403_inod7   TT AAA TGGAGCCAATGTAGC RECOMBINANT C/B/A
73403_inod14  TT AAA TGGAGCCAATGCGGC RECOMBINANT C/B
73403_inod8   TT GGG TGTATTCGATGTAGC E
```

```
                 10
              ....|....|..
73402pseudo7  TAAACTGGGTAC A
73402pseudo5  TAAACTGGGTAC A DUPLICATE
73402pseudo2  TAAGCTGGGTAC A DUPLICATE (1 mutation)
73402pseudo3  CTTATCAAACGT B
73402pseudo4  CTTATCAAACGT B DUPLICATE
73402pseudo6  CTTATCAAACGT B DUPLICATE
```

```
                 10
             ....|....|...
73477recli4  ACGGA GATGCGG A
73477recli5  ACGGT GATGCGG A DUPLICATE (1 mutation)
73477recli1  GCGGA GATGTAG B
73477recli3  GCGGA GATGTAG B DUPLICATE
73477recli6  GAAAACAGCATAA C
73477recli7  GCAAACAGTGCGG RECOMBINANT A/C/A
```

```
                 ....|
73467sulcata6    GACCG A
73467sulcata7    GACCG A DUPLICATE
73467sulcata9    GACCG A DUPLICATE
73467sulcata5    GATCG B
73467sulcata10   GATCG B DUPLICATE
73467sulcata3    GATCG B DUPLICATE
73467sulcata2    GATCG B DUPLICATE
73467sulcata4    AGTTA C
73467sulcata8    GATTA RECOMBINANT A/C
```

```
              ....|..
84155_ecal_1  AGGGTAC A
84155_ecal_3  AGAATAC A DUPLICATE (2 mutation)
84155_ecal_5  AAGGTAC RECOMBINANT A/B/A
84155_ecal_2  GAGGTGC B
84155_ecal_4  AGGGCGT C
```

```
              ....|.
73446bhatt6   GTGTAC A
73446bhatt9   GTGTAC A DUPLICATE
73446bhatt3   GTGTAC A DUPLICATE
73446bhatt5   GTGTAC A DUPLICATE
73446bhatt2   GTGTAC A DUPLICATE
73446bhatt1   GTGTAT A DUPLICATE (1 mutation)
73446bhatt8   GCGTAC A DUPLICATE (1 mutation)
73446bhatt4   GTTGGC B
73446bhatt7   GTTGGC B DUPLICATE
73446bhatt10  ATTGGC B DUPLICATE (1 mutation)
```

```
                       10        20        30        40
              ....|....|....|....|....|....|....|....|....|....
84157neilgh1     CAGC      AGTGTTAACGTGCGTGTCGCGTGAAAGTGCCGTGTC A
84157neilgh2     TAGC      GATATTAGTGTAGACGTCGCGTGAAAGTGCCGTGTC B
84157neilgh5     CGGC      AGTGTTAACGTAGGTGTC                   C
84157neilgh3     TAGC      GATATTAGTATAGGTGTC                   D
84157neilgh4  CATTATAGCCCAAGGCGCCCGTGCAGGT                  GTC E
```

HEXAPLOID GROUP III

```
             10        20        30
              ....|....|....|....|....|....|....
84145_mutab6  TATGA CGTTACAGAT                   A
84145_mutab7  TATGA CGTTACAGGT                   A DUPLICATE
84145_mutab4  TATGAGCGTTACAGGT                   A DUPLICATE (1 indel)
84145_mutab1  TACGA CGTTACAGAT                   B
84145_mutab5  TACGA CGTTACAGGT                   B DUPLICATE
84145_mutab2  TGCGA CGTCGCAGGT                   C
84145_mutab3  CACAG TACTACAGGT                   D
84145_mutab8  TATGA CGTTGTGAGCGCGTGAAAGTGCCGTGTC E
```

ENDECAPLOID GROUP III

```
               10
                ....|....|
73223_olig11x1  TTTGCGGAGC A
73223_olig11x3  TTTGCGGAGC A DUPLICATE
73223_olig11x2  TTTGCGGAGC A DUPLICATE
73223_olig11x8  TCTGCGGAGC A DUPLICATE
73223_olig11x7  CTTGTGGAGC A DUPLICATE (2 mutations)
73223_olig11x5  TTCACGGAGC B
73223_olig11x6  TTTACGGAGC B DUPLICATE
73223_olig11x4  TTTGCTATAT C
```

SUBGENUS HITCHENIOPSIS

```
               ....|....
84156vamana1   ATCT A
84156vamana3   ATCT A DUPLICATE
84156vamana4   ATCT A DUPLICATE
84156vamana5   GTCT A DUPLICATE (1 mutation)
84156vamana6   ATCT A DUPLICATE
84156vamana7   ATCT A DUPLICATE
84156vamana8   ATCT A DUPLICATE
84156vamana9   ATCT A DUPLICATE
84156vamana12  ATCT A DUPLICATE
84156vamana10  AACT B
84156vamana11  AACT B DUPLICATE
84156vamana13  ATTC B DUPLICATE (2 mutations)
84156vamana2   AACT B DUPLICATE
```

```
              ....
20_alismat9   CTTT A
20_alismat10  CTTT A DUPLICATE
20_alismat5   CTTT A DUPLICATE
20_alismat7   CTTT A DUPLICATE
20_alismat3   CTTT A DUPLICATE
20_alismat4   CCTT B
20_alismat6   CCTT B DUPLICATE
20_alismat1   TTCC C
20_alismat2   TTCC C DUPLICATE
20_alismat8   TTCC C DUPLICATE
```

```
               ...
19_pierreana5  GGG A
19_pierreana7  GGG A DUPLICATE
19_pierreana3  GGT B
19_pierreana8  GGT B DUPLICATE
19_pierreana2  GGT B DUPLICATE
19_pierreana1  ACT C
19_pierreana6  ACT C DUPLICATE
19_pierreana4  AGT C DUPLICATE (1 mutation or recombinant C/B)
```

RELATED GENERA I

```
                ...
86407Mscaposu1  GCA A
86407Mscaposu4  GCA A DUPLICATE
86407Mscaposu2  A G B
86407Mscaposu3  A G B DUPLICATE
```

```
              ....|...
84178Hcauli3  TAGC AGA A
84178Hcauli6  TAGC AGA A DUPLICATE
84178Hcauli2  TAGC AGA A DUPLICATE
84178Hcauli1  CTA GGAG B
84178Hcauli4  CTA GGAG B DUPLICATE
```

RELATED GENERA II

```
             ....|.
C086_Scamp1  TGGGAC A
C086_Scamp3  TGGGAC A DUPLICATE
C086_Scamp4  TGGGAC A DUPLICATE
C086_Scamp7  TGGGAC A DUPLICATE
C086_Scamp8  TGGGAC A DUPLICATE
C086_Scamp6  TGGGGC A DUPLICATE (1 mutation)
C086_Scamp5  CAAGAT B
C086_Scamp2  CAAGAT B DUPLICATE
```

Stahlianthus involucatus – no variable sites

Apendix 2. Multiple alignment of final dataset. Total of 102 unique ITS sequence types from 34 individuals. Total of 354 variable characters is presented.

```
                                 10        20        30        40        50        60        70        80        90       100
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
C073_Hcarey_A           TTAAAGAG  CATAAACGTGTGGTTTGGTTCC CTTTGTTGCCCCCCAATCCCAATGTTCGTGTTGGTGTTGG    TGATGCCGCGCCGTAATA TGCA
C073_Hcarey_B           TT.AAGAG   ....A.C.....G.TT..T.C. CTTTGTT....CC AATCCCAATGTTCGTGTTGG......   T............A.T. T..A
C086_Stahlianthus_A     .......    ........C..ATT...... TT.    .....                     .....     ................. T..
C086_Stahlianthus_B     .......    ........C..ATT...... TT.    .....                     .....     ................. T..
20_alismatifolia_A      .......    .......C...AT...... .. TT.  .....                     .....     ................. T..
20_alismatifolia_B      .......    .......C...AT...... .. TT.  .....                     .....     ................. T..
20_alismatifolia_C      .......    .......C...AT...... .. TT.  .....                     .....     T................ T..
71449_Sinvol            .....T     ........C..ATT...... TT.    .....                     .....     ................. T..
84156_vamana_A          .....      .............TT....G. .TT  .....                     .....     .......A......... T..
84156_vamana_B          .....      .............TT....G. .TT  .....                     .....     .......A......... T..
84170_aromat_C          ......    ..............T.... .. ...GC.C.....C                   .....     ................. ...
84183arom6x4            ......    ..............T.... .. ...GC.C.....C                   .....     ................. ...
84183arom6x8            ......    ..............T....... ... .........           CA   ........     ................. ...
71443affprak5           ......    ..............T....... ... .........                ........     ................. ...
71431_16                ......    ..............TT...... ... .........                ........     ................. ...
84142_8                 ......    ..............TT...... ... .........                ........     ................. ...
71431_10                ......    ..............TT...... ... .........                ........     ................. ...
71484_montana_C         ......    ..............TT...... ... .........                ........     ................. ...
73479_mon_B             ......    ..............TT...... ... .........                ........     ................. ...
73425_montana_B         ......    ..............TT...... ... .........                ........     ................. ...
73222longaSRI7          ......    ....A.........T....... ... ......GC            C    ........     ................. ...
73222longaSRI5          ......    ....A.........T....... ... ......GC            C    ........     ................. ...
84183arom6x3            ....... ........................ ... .........                ........     ................. ...
84183arom6x6            ....... ........................ ... .........           CA   ........     ................. ...
84182_xanthor_A         ....... ........................ ... .........                ........     ................. ...
84183arom6x5            ....... ........................ ... .........                ........     ................. ...
84183arom6x1            ....... ........................ ... .........                ........     ................. ...
73453angust2            .......   .............T...... .. ...    .....                .G.CAAC.     ................. ...
73453angust5            .......   .............T...... .. ...    .....                .G.CAAC.     ................. ...
73480_angu_B            .......   .............T...... .. ...    .....                .G.CAAC.     ................. ...
73480_angu_A            .......   .............T...... .. ...    .....                .G.CAAC.     ................. ...
73453angust4            ........  ..............TT...... ...     .....G          CC   .....G..     ................. T..
73480_angu_C            ........  ..............TT...... ...     .....G          CC   .....G..     ................. T..
73453angust1            .......   ..............T....... ...     .....C                  .....     ................. ...
71443affprak2           .......       .....T....T....... ..G     .....T                  .....     ................. T..
71492_arom9x_A          .......       ..........T....... ..G     .....T                  .....     ................. T..
71431_4                 ........T  ..............T....C.. ...    .....T                ..AC....     ................. ...
71431_3                 .......C   ..............T....... ...    ....GC          C    .G.....      ................. ...
71492_arom9x_B          .......C   ..............T....... ..G    .....T                  .....     ................. T..
73222longaSRI10         .......       ..........T....... ..G    .....T                  .....     ................. T..
73222longaSRI6          .......C   ...A.........T....... ... .......GC           C    ........     ................. ...
73222longaSRI2          .......C   ..............T...... ... .......GC           C    ........     ................. ...
71443affprak4           ...          ..........T....... ...GC.C.....C                   .....     ................. ...
71431_11                .......C   ..............T....... ...GC.C......               .......     ................. ...
71431_14_BAD            .......C   ..............T....... ...GC.C......               .......     ................. ...
71431_A                 ........  T.C...........T....C.. ...GC.CAT.A.C                   ....     ................. ...
71431_17                ........  T.C...........T....C.. ...GC.CAT.A.C                   ....     ................. ...
71432_I__11             ........  T.C...........T....C.. ...GC.CAT.A.C                   ....     ................. ...
71432_A                 ........  T.............T....C.. ...GC.CAT.A.C                   ....     ................. ...
71432_I__12             ........  T.............T....C.. ...GC.CAT.A.C                   ....     ................. ...
71431_2                 ........     ...........T....C.. ...GC.CAT.A.C                   ....     ................. ...
71431_8                 ........  T.............T....C.. ...GC.CAT.A.C                   ....     ................. ...
71432_B                 ........G ...................... ..GGC.CAT..C.                 ......     ................. ...
71432_X__16             ........G ...................... ..GGC.CAT..C.                 ......     ................. ...
71432_X__20             ........G ...................... ..GGC.CAT..C.                 ......     ................. ...
71432_I__13             ........G ...................... ..GGC.CAT..C.                 ......     ................. ...
84182_xanthor_B         ........G ...................... ..GGC.CAT                      .....     ................. ...
84182_xanthor_C         ........G ...................... ..GGC.CAT                      .....     ................. ...
71432_7                 ........G ...................... ..GGC.CAT..C.                 ......     ................. ...
71432_1                 ........  T.............TT...... ...     .....G          CC   .....G..     ................. T..
71432_X__17             ........  ..............TT...... ...     .....G          CC   .....G..     ................. T..
71432_2                 ........  ..............TT...... ...     .....G          CC   .....G..     ................. T..
71432_6                 ........  ..............TT...... ...     .....G          CC   .....G..     ................. T..
84142_3                 ........  ..............TT...... ...     .....G          CC   .....G..     ................. T..
86407_Mscaposu_A        .......      ...........TT...... ...    .....                  ....A.     ...A............. T..
86407_Mscaposu_B        .......      ...........TT...... ...    .....                  ....A.     ...A............. T..
84178_Hcauli_A          .......      ...........TT...... ...    .....                  ....A.     ...A............. T..
73402_pseudo_A          .......      ...........TT...... ...    .....                  ....A.     ...A............. T..
73446_bhatt_A           .......      ...........TT...... ...    .....                  ....A.     ................. T..
73446_bhatt_B           .......      ...........TT...... ...    .....                  ....A.     ................. T..
84145_mutabilis_C       ......C      ...........TT...... ...    .....                  ....A.     ................. T..
84145_mutabilis_E       .......      ...........TT.....T ...    .....                  ......     .........A....... T..
73223_oligan11x_A       .......      ...........TT...... ...    .....                  ......     .....T........... T..
73223_oligan11x_C       .......      ...........TT...... ...    .....                  ......     .T..A.T.......... T..
73223_oligan11x_B       .......      ...........TT...... A..    .....                  ......     .....T........... T..
84157neilgh2            .......      ...........TT...... ...    .....                  ......     .........A....... T..
84143_cann_C            .......      ...........TT...... ...    .....                  ......     .........A....... T..
84143_cann_D            .......      ...........TT...... ...    .....                  ......     .........A....... T..
73403_inod5             ......C      ...........TTA..... ...    .....                  ....A.     ............A.... T..
73403_inod1             ......C      .....C.....TTA..... ...    .....                  ....A.     ............A.... T..
73403_inod7             ......C      ...........TTA..... ...    .....                  ....A.     ............A.... T..
73403_inod10            ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73403_inod6             ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73403_inod8             ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73403_inod3             ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73467_sulcata_C         ......C      ....A......TT...... ...    .....                  ....A.     ................. T..
73477_recli_B           ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73467_sulcata_A         ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73467_sulcata_B         ......C      ...........TT...... ...    .....                  ....A.     ................. T..
73477_recli_A           ......C      ....A......TT...... ...    .....                  ....A.     ................. T..
73477_recli_C           ......C      ...........TT.....A ...    .....                  ....A.     ............A.... T..
73402_pseudo_B          ......C      ...........TT...... ...    .....                  ....T.     ................. T..
84157neilgh1            ......C      ...........TT...... ...    .....                  ....A.     ................. T..
84157neilgh5            ......C      ...........TT...... ...    .....                  ....A.     ................. T..
19_pierreana_A          .......      ...........T....... .T.GC.CA....                  ......     ................. T..
19_pierreana_B          .......      ...........T....... .T.GC.CA....                  ......     ................. T..
19_pierreana_C          .......      ...........T....... .T.GC.CA....                  ......     ............A.... T..
C056_Globba_A           .......      G....C...G.TT....C.CCTG    C....                  ...A..   T.G..T......TC.CAG ..T
C056_Globba_B           .......      G....C..AG.TT....C. CTG    C....                  ......   T....T..A.T.TC.CAG ..T
C071_Alpinia_A          ..T....      T........G..T.C..C. CTG    C....                  .....T   .A...T.......C..T. TAT
C071_Alpinia_B          ..T....      T.G......G..T.C..C. CTG    C....                  .....T   .A...T.....T.C..T. TAT
C058_zingiber_B         .......CA T.......C...GC....T.C. C.G    C...C.CCCACCCACAC   CCGTG...C..GGGC..G...............AT..
C058_zingiber_A         .......CA T.......C...GC....T.T. C.G    C...C.CCCACCCACAC   CCGTG...C..GGGC..G...........  ....AT..

                               110       120       130       140       150       160       170       180       190       200
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
C073_Hcarey_A           ACCGGAGAGGC TGGCATGTCGG A  GGCATCGGGGTTC    CGGTCCTGGAGCGTTGTCGTGCTGTCCCTAGAC GTCTTTGTAGTTGTA
C073_Hcarey_B           ACC......G. .G..A..T... .  .G.........C    ...........................C..G..  .........G.T..A
C086_Stahlianthus_A       ..A....G. .G......... .  ..A....AA....    ...............C................  ....A.....T..T
C086_Stahlianthus_B       ..A....G. .G.......A. .  ..A....AA....    ...............C................ A...A.....T..T
20_alismatifolia_A        ..A....G. .G......... .  ..A.....A....    .............. .C...............T ....A..C...T..T
20_alismatifolia_B        ..A....G. .G......... .  ..A.....A.C..    ...............C................T ....A..C...T..T
20_alismatifolia_C        ..A....G. .G......... .  ..A.....A....    .............C..C...............T ....A..C...T..T
71449_Sinvol               .A....G. .G......... .  ........A..    ...............C.. .. ......G.T ....A.....T..T
```

```
84156_vamana_A        ......G. G. . ..... .   ..A. ...A..       ..........T. ..C.... ..T.  ...TT. .C .A. .C.T. .T
84156_vamana_B        ......G. .G......... .   ..A. ...A.. .     ..........T. ..C.......T. ....TT. .C. .A. .C.T. .T
84170_aromat_C        ......G. .... ..... .   ............TCTT............A........C....C...GT ....... . .T. .
84183arom6x4          ......G. .... ..... .   ............TCTT... ................C...GT ....... . .T. .
84183arom6x8          ......G. C.... ..... .   ............      .........................C...G. . .. .....T. .
71443affprak5         ......G. C.......... .   ............      .........................C.G.G. .C.   .....T. .
71431_16              ......G. C...AC..... .   ............      .......A...............C...GT . ... .....T. .
84142_8               ......G. C....C..... .   ............      .........................C...GT ..........T...
71431_10              ......G. C....C..... .   ............      .........................C...GT ..........T...
71484_montana_C       ......G. C.A..C..... .   ............      .....................T...C...GT ..........T...
73479_mon_B           ......G. C.A..C..... .   ............      .....................T...C...GT ..........T...
73425_montana_B       ......G. C.A..C..... .   ............      .....................T...C...GT ..........T...
73222longaSRI7        ......G. C.......... .   ............      .........................C..G.. ..........T...
73222longaSRI5        ......G. C.......... .   ............      .........................C..G.. ..........T. .
84183arom6x3          ................... .   ............      ............................. . ..........
84183arom6x6          ................... .   ............      ............................. . ..........
84182_xanthor_A       ............. .....G.   ............      ...............A........ ......... .....T...
84183arom6x5          ................... .   ............      ............................. . ..........
84183arom6x1          ................... .   ............      ............................. . .... ....
73453angust2          ......G. C.... ..... .   ............      .........................C..G.T ..........T...
73453angust5          ......G. C.......... .   ............      .........................C..G.T .... .....T...
73480_angu_B          ......G. C.......... .   ............      .........................C..G.T ..........T...
73480_angu_A          ......G. C.......... .   ............      .........................C..G.T ..........T. .
73453angust4          ......G. ..... ..A.. .   ..AG........      ..........G..............CC.G.T ..........T...
73480_angu_C          ......G. ..... ..A.. .   ..AG.......       .... ....................C..G.T ..........T...
73453angust1          ......G. C.......... .   ............      .........................C..G.T ..........T...
71443affprak2         ......G. ........... .   ......AA....      ......................T...... ..........CT..T
71492_arom9x_A        ......G. ........... .   ......AA....      ..................A......T...... ..........CT..T
71431_4               .....AG............. .   ............TCTT.....................T...... ..........CT..T
71431_3               ......G. C.......... .   ............      ......................T...... ..........CT..T
71492_arom9x_B        ......G. ........... .   ......AA....      ..................A......T...... ..........CT..T
73222longaSRI10       ......G. ........... .   .......A....      ............................. ..........CT..T
73222longaSRI6        ......G. C.......... .   ............      ............................. ..........CT..T
73222longaSRI2        ......G. C.......... .   ............      .........................C..G.. ..........T..T
71443affprak4         ......G. ........... .   ............      .........................C...GT ..........T...
71431_11              ......G. ........... .   ............TCTT.....................C...GT ..........T...
71431_14_BAD          ......G. C....C..... .   ..........A...    .........A...............C...GT ..........T...
71431_A               .....AG............. .   ............      ............................. ..............
71431_17              .....AG............. .   ............      ............................. ..............
71432_I__11           .....AG............. .   ............      ............................. ..............
71432_A               .....AG........A.... .   ............      ............................. ..............
71432_I__12           .....AG........A.... .   ............      ............................. ..............
71431_2               .....AG............. .   ............      ............................. ..............
71431_8               .....AG............. .   ............      ............................. ..............
71432_B               ......G. C.......... .   ............      ............................. ..............
71432_X__16           ......G. C.......... .   ............      ............................. ..............
71432_X__20           ......G............. .   ..........A...    ............................. ..............
71432_I__13           ......G. C.......... .   ............      ............................. ..............
84182_xanthor_B       ......G............. .   ............      ............................. ..............
84182_xanthor_C       ......G............. .   ............      ............................. ..............
71432_7               ......G............. .   .........A...     .........................C.... ..............
71432_1               ......G. ........... .   ..AG........      T........................C..G.T ..........T...
71432_X__17           ......G. ........... .   ..AG........      T........................C..G.T ..........T...
71432_2               ......G. ........... .   ..AG........      .........................C..G.T ..........T...
71432_6               ......G. ........... .   ..AG........      .........................C..G.T ..........T...
84142_3               ......G. ........... .   ..AG........      T........................C...GT ..........T...
86407_Mscaposu_A      ......G. ..........AG   .A....A....       .............................T ..........CT..T
86407_Mscaposu_B      ......G. ..........AG   .A....A....       .............................T ..........CT..T
84178_Ecauli_A        ......G.............AG   .A.........       .............................T ...........T..T
73402_pseudo_A        ......G. ..........AG   .A.........       .............................T ..........CT..T
73446_bhatt_A         ......G. ..........AG   .A.........       .............................T .........G.T..T
73446_bhatt_B         ......G. ..........AG   .A.........       .............................T ...........T..T
84145_mutabilis_C     ......G. ..........AG   .A..T.A....       ......C......................T ...........T..T
84145_mutabilis_E     ......G. ..........AG   .A.........       .............................T ...........T..T
73223_oligan11x_A     ......G. ..........AG   .A.........       .............................T ...........T..T
73223_oligan11x_C     ......G. .T........AG   .A....A....       T............................T ...........T..T
73223_oligan11x_B     ......G. ..........AG   .A.........       .............................T ...........T..T
84157neilgh2          .A.....G. ..........AG   .A.........       .............................T ...........T..T
84143_cann_C          ......G. ..........AG   .A.........       .............................T ...........T..T
84143_cann_D          ......G. ..........AG   .A...A.....       .............................T ...........T..T
73403_inod5           ......G. ..........GG   .A....A....       .....T.......................T ...........T..T
73403_inod1           ......G. .....C....GG   .A....A....       .....T.......................T ...........T..T
73403_inod7           ......G. ..........GG   .A....A....       .............................T ...........T..T
73403_inod10          ......G. ..........GG   .A....A....       .....T.......................T ...........T..T
73403_inod6           ......G. ..........GG   .A....A....       .....T.......................T ........GC.TA.T
73403_inod8           ......G. ..........GG   .A....TAT...      .....T.......................T ...........T..T
73403_inod3           ......G. ..........GG   .A....A....       .............................T ...........T..T
73467_sulcata_C       ......G. ..........GG   .A....A....       .............................T ...........T..T
73477_recli_B         ......G. ..........GG   .A....A....       .............................T ...........T..T
73467_sulcata_A       ......G. ..........GG   .A....A....       ..........................C...T ...........T..T
73467_sulcata_B       ......G. ..........GG   .A....A....       .............................T ...........T..T
73477_recli_A         ......G. ..........GG   .A....A....       .............................T ...........T..T
73477_recli_C         .A.....G.............AG   .A.........       ..............C...A..........T ...........T..T
73402_pseudo_B        ......G. ..........AG   .A.........       .............................T ...........T..T
84157neilgh1          ......G. ..........AG   .A....A....       ......C......................T ...........T..T
84157neilgh5          ......G. ..........AG   .A....A....       ......C......................T ...........T..T
19_pierreana_A        C......G. CGC.......AG   .A...A.A...C      ................C............G.. ..........TT.T
19_pierreana_B        C......G. CGC.......AG   .A...A.A...C      ................C............G.. ..........TT.T
19_pierreana_C        C......C. CGC.......AG   .A...A.A...C      ................C............G.. ..........TT.T
C056_Globba_A         C....G.GT .G....CT...G  GT...........C      .A..................T..C..C..G.. ..........T.CT
C056_Globba_B         C...A..GT .G..A.CTT..GTGGT...........C      .A..G...A..........CAT..C..C..G.. ..T.A.....T..T
C071_Alpinia_A        GA.....G  CG.T...T..AG     ..........GC      ..A..............T......C..C..G.. ...C....T..T.CT     G
C071_Alpinia_B        GA.....G  CG.T...T..AG     ..........GC      ..A..............T......C..C..G.. ...C....T..T.CT     G
C058_zingiber_B       C..G...G. .G.....C.A.G.   ...GC.......C      ...C....................C..C..G.. ...CCA...CCT...AGTCGG
C058_zingiber_A       C..G...G. .G.....C.A.G.   ...GC.......C      ...C....................C..C..G.. ...CCA. .CCT. ..AGTCGG

                          210       220       230       240       250       260       270       280       290       300
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
C073_Ecarey_A         GTGGCG    CGCGTAGCTC    GGCTCCTGAGTGGCA GTTCGCACATGGTGCTATGAGGGCAGACGTGGTCCTGTTTTGGTGCTATCTCAAGA
C073_Ecarey_B         .TGGC.    ....TA..T.    .....C..AGT..C. G...G.....GT.CTAT...G..A........C.T.........C.A.CTCAA..
C086_Stahlianthus_A   ......    T....A....    ......CC   ..C. ...........G.....A............   ...A.........     A..
C086_Stahlianthus_B   ......    T....A....    ......CC   ..C. ...........G.....A............   ...A.........     A..
20_alismatifolia_A    ....C.    T....A....    ..T....C   ..C. ...........G.....A............   ...A.........     A..
20_alismatifolia_B    ....C.    T....A....    ..T....C   ..C. ...........G.....A............   ...A.........     A..
20_alismatifolia_C    ....C.    .....A....    ..T....C   ..C. ...........G.....A............   ...A.........     A..
71449_Sinvol          ......    T....A....    ..T...CC   ..C. ...........G.....A............   ...A.........     A..
84156_vamana_A        A.....    A....A....    ........   .... .....G......T...........A..A..   ........... .A...  A..
84156_vamana_B        A.....    A....A....    ........   .... .....G......T...........A..AA.   ........... .A...  A..
84170_aromat_C        ......TGAG..........    .....C..   ..C. ...............................   ..............C..AATCA..
84183arom6x4          ......TGAG..........    .....C..   ..C. ...............................   ..............C..   A..
84183arom6x8          ......TGAG..........    .....C..   ..C. ...............................   ..............C..   A..
71443affprak5         ......TGAG..........    .....C..   ..C. ...............................   ..............C..   A..
71431_16              .T....    ..........    .....C..   ..C. ..........A...............A...   ........... ..C..   A..
84142_8               ......    ..........    .....C..   ..C. ..........A....................   ........... ..C..   A..
71431_10              ......    ..........    .....C..   ..C. ..........A....................   ..............C..   A..
71484_montana_C       ......    ..........    .....C..   ..C. ...............................   ...........C..C..   A..
73479_mon_B           ......    ..........    .....C..   ..C. ...............................   ...........C..C..   A..
73425_montana_B       ......    ..........    .....C..   ..C. ...............................   ..............C..AATCA..
73222longaSRI7        ......    ..........TC  .....C..   .... ...............................   ..............C..AATCA..
73222longaSRI5        ......    ..........TC  .....C..   .... ...............................   ..............C..AATAA..
84183arom6x3          ......    ..........    ........   .... ...............................   ............        ...
84183arom6x6          ......    ..........    ........   .... ...............................   ............        ...
84182_xanthor_A       ......    ..........    ........   .... ...............................   ............        ...
```

```
84183arom6x5          .... .   ..........    ....C .   ..C. .   ...... ...T........ .. . ..   ...  .. ....     ...
84183arom6x1          ......   ..........    ........  .... ........................... .... .  ............C..   A..
73453angust2          ......CGC G.........   .....C..  .... ........................... .... .  ..............AAGCA..
73453angust5          ......CGC G.........   .....C..  .... ........................... .... .  ..............AAGCA..
73480_angu_B          ......CGC G.........   .....C..  .... ........................... .... .  ...........C..AAGCA..
73480_angu_A          ......   ..........    .....C..  .... ........................... .... .  ...........C..AAGCA..
73453angust4          ......CGC G.........   .....C..  ...G ...........................G.. .  ..............AAGCA..
73480_angu_C          ......CGC G.........   .....C..  .... ........................... .... .  ..............AAGCA..
73453angust1          ......CGC G.........   .....C..  .... ........................... .... .  ..............AAGCA..
71443affprak2         ......   .....A....    A....G..  .... ........................... .... .  ........C. .....    A..
71492_arom9x_A        ......   .....A....    A....G..  .... ........................... .... .  ........C. .....    A..
71431_4               ......   .....A....    A....G..  .... ........................... .... .  ........C......AAGCA..
71431_3               ......   .....A....    A....G..  .... ........................... .... .  ........C......     A..
71492_arom9x_B        ......   .....A....    A....G..  .... ........................... .... .  ...........C..      A..
73222longaSRI10       ......   .....A....    .....G..  .... ........................... .... .  ...............      A..
73222longaSRI6        ......   .....A....    .....G..  .... ........................... .... .  ...............      A..
73222longaSRI2        ......   .....A....    .A...G..  .... ........................... .... .  ...............      A..
71443affprak4         ......TGAG.........    .....C..  ..C. ........................... .... .  ........C. .....    A..
71431_11              ......TGAG.........    .....C..  ..C. ........................... .... .  ...............      A..
71431_14_BAD          ......TGAG.........    .....C..  ..C. ........................... .... .  ......C.......AATCA..
71431_A               ......   ..........    ........  .... ........................... .... .  ...............      A..
71431_17              ......   ..........    ........  .... ........................... .... .  ......C.. ....AATCA..
71432_I__11           ......   ..........    ........  .... ........................... .... .  ...............      A..
71432_A               ......   ..........    ........  .... ........................... .... .  ...............      A..
71432_I__12           ......   ..........    ........  .... ........................... .... .  ...............      ...
71431_2               ......   ..........    .....C..  .... ........................... .... .  ...........C..      A..
71431_8               ......   ..........    ........  .... ........................... .... .  ...........C..      A..
71432_B               ......   ..........    ........  .... ........................... .... .  ...............      A..
71432_X__16           ......   ..........    ........  .... ...................................................    ...
71432_X__20           ......   ..........    ........  .... ...................................................    A..
71432_I__13           ......   ..........    ........  .... ........................... .... .  ...........C..AATCA..
84182_xanthor_B       ......   ..........    .....C..  .... ...................................................    ...
84182_xanthor_C       ......   ..........    ........  .... ........................... .... .  ...............     AA..
71432_7               ......CG ..........    .....C..  .... ...................................................AA CA..
71432_1               ......   ..........    .....C..  .... ........................... .... .  ...............      A..
71432_X__17           ......   ..........    .....C..  .... ........................... .... .  ...............      A..
71432_2               ......   ..........    .....C..  .... ........................... .... .  ...............      A..
71432_6               ......   ..........    ........  .... ........................... .... .  ...............      A..
84142_3               ......   ..........    .....C..  ..C. ..........A................ .... .  ...............      A..
86407_Mscaposu_A      ......   .....A....    ........  .... ........................... .... .  ...............     AA..
86407_Mscaposu_B      ......   .....A....    ........  .... ........................... .... .  ...............     AA..
84178_Hcauli_A        ......   ...A.A....    ........  .... ........................... .... .  ...............     AA..
73402_pseudo_A        ......   .....A....    ........  .... ........................... .... .  ...............     AA..
73446_bhatt_A         ......   .....AA...    ........  .... ...............A........... .... .  ...............     AA..
73446_bhatt_B         ....G.   .....AA...    ........  .... ........................... .... .  ...............     AA..
84145_mutabilis_C     A...C.   .....A....    ........  .... ..................A........ .... .  ...............     AA..
84145_mutabilis_E     ......   .....A....    ........  .... ........G.........A........ .... .  ...............     AA..
73223_oligan11x_A     ......   .....A....    ........  .... ..................A........ .... .  ..........C....     AA..
73223_oligan11x_C     ......   .....A....    ........  .... ..................A........ .... .  ..........C....     AA..
73223_oligan11x_B     ......   .....A....    ........  .... ..................A........ .... .  ..........C....     AA..
84157neilgh2          A.....   .....A....    ........  .... ..................A........ .... .  ...............     AA..
84143_cann_C          ......   .....A....    ........  .... ..................A........ .... .  ...............     AAC.
84143_cann_D          ......   .....A....    ........  .... ..................A........ .... .  ...............     AAC.
73403_inod5           ......   .....A....    .....C..  .... ..............A...A........ .... .  ...............     AA..
73403_inod1           ......   .....A....    ..T..C..  .... ..............A...A.........A       ...............     AA..
73403_inod7           ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73403_inod10          ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73403_inod6           ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73403_inod8           ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73403_inod3           ......   .....A....    .....C..  .... ..................A........ .... .  ...............     AA..
73467_sulcata_C       ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73477_recli_B         ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73467_sulcata_A       ......   .....A....    .....C..  .... ..................A........ .... .  ...............     AA..
73467_sulcata_B       ......   .....A....    .....C..  .... ..................A........ .... .  ...............     AA..
73477_recli_A         ......   .....A....    .....C..  .... ..................A........ .... .  ...............     AA..
73477_recli_C         ......   .....A....    ..T..C..  .... ..............A...A........ .... .  ...............     AA..
73402_pseudo_B        ......   .....A....    .....C..  .... ..................A........ .... .  ............A...    AA..
84157neilgh1          A.....   .....A....    ........  .... ..................A........ .... .  ...............     AA..
84157neilgh5          A.....   .....A....    ........  .... ..................A........ .... .  ...............     AA..
19_pierreana_A        ......CGCG.....A....   .A......  .... G.......C..T.C.AT...G........      .....C.C.ACA...G     A..
19_pierreana_B        ......CGCG.....A....   .A......  .... G.......C..T.C.AT...G........      .....C.C.ACA....     A..
19_pierreana_C        ......CGCG.....A....   .A......  .... G.......C..T.C.AT...G........      .....C.C.ACA....     A..
C056_Globba_A         .TGCCT   .....A...T    ....T...  A.C. G..TG..T...TT....T...G.A.......    CT..CG.......C..     A.G
C056_Globba_B         .TGCCT   .AT..A...T    ....T...  AAC. G..TG..T...TT..T.T..AG.A..G....    CT..CG.......C..     A.G
C071_Alpinia_A        CTATGC.A T....A.T..GAGCGA....T...  .... GA..G.G....GT.CTCTAT.T.TCA..A..  CT...AC.........      A..
C071_Alpinia_B        CTATGC.A T....A.T..GAGCGA....T...  .... GA..G.G....GT.CTCTAT.T.TCA..A..  CT...AC.........      A..
C058_zingiber_B       CA.TGC.. .....A....    ...C.C..  ....AG.C.G..T.C.GT.C..T...G..A......  C.......C.C..C..     A..
C058_zingiber_A       CA.TGC.. .....A....    ...C.C..  ....AG.C.G....C.GT.C..T...GA.A.....  C.......C.C..C..     A..
```

```
                           310       320       330       340       350
                      ....|....|....|....|....|....|....|....|....|....|...
C073_Hcarey_A         GGCC TGTGTG                                         A
C073_Hcarey_B         .... ......                                         .
C086_Stahlianthus_A   .G.. .A...ATG   .......A.....GA.C.......T.T      ....
C086_Stahlianthus_B   .G.. .A...ATG   .......A.....GA.C.......T.T      ....
20_alismatifolia_A    .... .AC..      .......A.....GA.........T.T      ....
20_alismatifolia_B    .... .AC..      .......A.....GA.........T.T      ....
20_alismatifolia_C    .... .AC..      .......A.....GA.........T.T      ....
71449_Sinvol          .... .A...      .......A.....GAC.....T....T      ....
84156_vamana_A        .... .A...      .......A.....G.......T..T.T      ....
84156_vamana_B        .... .A...      .......A.....G.......T..T.T      ....
84170_aromat_C        .... .... .ATTG. .........CC....C..........      ....
84183arom6x4          .... .....ATTG. .........CC....C..........      ....
84183arom6x8          .... ..C.. .    ..........C....C...........GC    ....
71443affprak5         .... ..C.. .    ..........C....C..........       ....
71431_16               . C.C.                        C..........       ....
84142_8                . C.C.                   C....C..........       ....
71431_10               . C.C.                   C....C..........       ....
71484_montana_C       .... .... .ATTGG .........CC....C..........      ....
73479_mon_B           .... .....ATTG. .........CC....C..........       ....
73425_montana_B       .... .....ATTG. .........CC....C..........       ....
73222longaSRI7        .... .....ATTG. ..........C....C..........       ....
73222longaSRI5        .... .....ATTG. ..........C....C..........       ....
84183arom6x3          .... .....      . ..........................     ....
84183arom6x6          .... .....      ...........................      ....
84182_xanthor_A       .... .....      ...........................      ....
84183arom6x5          .... ..C.. .    ..........C....C...........GC    ....
84183arom6x1          .... .....ATTG. .........CC....C..........       ....
73453angust2          .... .....      ...............C..........CG   T....
73453angust5          ....C.... .     .........................CG    T....
73480_angu_B          .... .....      ...............C..........CG   C....
73480_angu_A          .... .....      ..........................CG   C....
73453angust4          ....C.... .     ..........................CG   T....
73480_angu_C          ....C.... .     ..........................CG   T....
73453angust1          .... .....      ...............C..........CG   T....
71443affprak2         .... .....      ...........................      ....
71492_arom9x_A        .... .....      ...........................      ....
71431_4               .... ..C.. .    ..........C....C...........GC    ....
71431_3               .... .....      ...........................      ....
71492_arom9x_B        .... . ...      .........CC....C..........       ....
73222longaSRI10       .... .....      ...........................      ....
73222longaSRI6        .... .....      ...........................      ....
73222longaSRI2        .... .....      ... .... ..................      ....
```

```
71443affprak4         .... ....  .       ..........................      . ..
71431_11              .... .... .       ..........................      ....
71431_14_BAD          .... ... .       ..............T...........      ....
71431_A               .... .. ..       ..........................      ....
71431_17              .... ....       ..........................      . ..
71432_I__11           .... ....       ..........................      ....
71432_A               .... ... .       .........................       . ..
71432_I__12           .... ... .       ..........................      . ..
71431_2               .... .....       ..........CC....C..........      ....
71431_8               .... .....       ..........CC....C..........      ....
71432_B               .... .....       ..........................      ....
71432_X__16           .... .....       ..........................      ....
71432_X__20           .... .....       ....................T.......CG  T...
71432_I__13           .... .....       .........................       ....
84182_xanthor_B       ....  ....ATTG..C.........................       ....
84182_xanthor_C       .... .....       .........................       ....
71432_7               .... C....ATTG...........................CG  C...
71432_1               .... .....       ....................T.......CG  T...
71432_X__17           .... .....       ....................T.......CG  T...
71432_2               .... .....       ....................T.......CG  T...
71432_6               .... .....       .........................       ....
84142_3               .... .....       .........................       ....
86407_Mscaposu_A      ...  .....       .........................T      ....
86407_Mscaposu_B      .... .....       .........................T      ....
84178_Ecauli_A        .... .....       .........................T      ....
73402_pseudo_A        .... .....       ......................A....T      ....
73446_bhatt_A         ...T .....       .......................T...T      ....
73446_bhatt_B         ...T .....       ................ .....T...T      .  .
84145_mutabilis_C     .... .....       .............                       .
84145_mutabilis_E     ....  ..A..       ....C..................T...T      . ..
73223_oligan11x_A     ....  ....       .......................T...T      . ..
73223_oligan11x_C     .... .....       .......................T...T      ....
73223_oligan11x_B     .... .... .       .......................T...T      ....
84157neilgh2          .... ..A..       ....C..................T...T      ....
84143_cann_C          .... .....       .......................TA..T      ....
84143_cann_D          .... .....       .......................TA..T      ....
73403_inod5           .... ...A..       ....C..................T...T      ....
73403_inod1           .... ...A..       ....C..................T...T      ....
73403_inod7           .... ...A..       ....C..................T...T      ....
73403_inod10          .... ...A..       ....C..................T...T      ....
73403_inod6           .... ...A..       ....C..................T...T      ....
73403_inod8           .... ...A..       ....C..................T...T      ....
73403_inod3           .... ...A..       ....C..................T...T      ....
73467_sulcata_C       .... ...A..       ....C..................T...T      ....
73477_recli_B         .... ...A..       ....C..................T...T      ....
73467_sulcata_A       .... ...A..       ....C..................T...T      ....
73467_sulcata_B       .... ...A..       ....C..................T...T      ....
73477_recli_A         .... ...A..       ....C..................T...T      ....
73477_recli_C         .... ...A..       ....C..................T..AT      ....
73402_pseudo_B        .... ...A..       ....C..................T...T      ....
84157neilgh1          .... ......       .......................T...T      ....
84157neilgh5          .... ......       .............                       .
19_pierreana_A        .... .A....          ...........GAC.....T...T      ....
19_pierreana_B        .... .A....          ...........GAC.....T...T      ....
19_pierreana_C        .... .A....          ...........GAC.....T...T      ....
C056_Globba_A         .... ......        AT......C....G.......T....T      .C.C
C056_Globba_B         .... ......        AT....A.C....A.......T.T..T      ...C
C071_Alpinia_A        ...T ......        AT. ..A.CA....A......T....T      ..TC
C071_Alpinia_B        ...T ......        AT. ..A.CA....A......T....T      ..TC
C058_zingiber_B       .... ..C...        AC. ....CAC..G.C.....AAAG .GCCGT.C.C
C058_zingiber_A       .... ......        AC. ....CAC..G.C.....AAAG .GCCGT.C.C
```

Apendix 3. Poster presentations in 2009;
A: Záveská E., Fér T., Marhold K., Šída O., Sabu M., Leong-Škorničková J. 2009. Phylogenetic studies in Indian polyploid *Curcuma* species using AFLP marker. *International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, FRANCE, May 17-20 2009.*

B: Záveská E., Fér T., Marhold K., Šída O., Sabu M., Leong-Škorničková J. 2009. Phylogenetic studies in Indian polyploid *Curcuma* species using AFLP´s and ITS sequencing. *The Fifth International Symposium on the Family Zingiberaceae, Xishuangbanna, CHINA, July 6-9 2009.*

International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, FRANCE, May 17-20 2009

Phylogenetic studies in Indian polyploid *Curcuma* species using AFLP marker

Eliška Záveská (1), Tomáš Fér (1), Karol Marhold (1), Otakar Šída (2), Mamiyil Sabu (3) & Jana Leong-Škorničková (4)

(1) Department of Botany, Charles University in Prague, Prague, CZ; (2) Department of Botany, National Museum in Prague, 252 43 Pruhonice, CZ; (3) Department of Botany, Calicut University P.O., 673635, Kerala, India;(4) Singapore Botanic Gardens, 1 Cluny Road, 259569 Singapore

Background

Economically important genus *Curcuma* is taxonomically critical polyploid complex. Study of Indian *Curcuma* by Leong-Škorničková et al., 2007 revealed six ploidy levels (2x, 6x, 9x, 11x, 12x and 15x) and four groups of taxa – „*Genome size groups*" - with significantly different 1Cx-values and probably with different evolutionary history. Various ploidy levels developed different modes of reproduction and exhibit uneven levels of morphological variability within and between populations, blurring specific boundaries. Complicated nomenclature further contributes to taxonomic complexity of the group (Leong-Škorničková et al., 2008). So far there is neither suitable infrageneric classifications of genus *Curcuma* (Škorničková, 2007), nor major published works focusing on **phylogeny of *Curcuma*.**

Material & methods

Plant material: A total of 19 Indian *Curcuma* species from 25 populations (total 115 plants) were analysed. We sampled 13 hexaploid (2n = 42), 7 nonaploid (2n = 63), one endecaploid (2n = 77), one dodecaploid (2n = 84) and two pentadecaploid populations (2n = 105) from nominate subgenus *Curcuma* and one population of diploid (2n = 22) from subgenus *Hitcheniopsis*. Selected species covered all four genome size groups with different 1Cx values (Leong-Škorničková et al. 2007) (see Table 1). We analyzed two to ten individuals per population and one to three populations per taxon.

Fig. 1: Two regions of main *Curcuma* diversity - SW and NE of India (red line), where majority of species were sampled.

AFLP analysis: AFLP analysis was carried out using AFLP Core Reagent Kit I (Invitrogen) and AFLP Pre-Amp Primer Mix I (Invitrogen), with three primer combinations used for selective amplification. Fragments were resolved on an ABI 3100 Avant DNA sequencer (Applied Biosystems) together with GeneScan-ROX-500 size standard and scored with Genographer 2.0 (Benham, 1999; http://hordeum.msu.montana.edu/genographer/) in the range of 100-500 bp.

Data analysis: Neighbor-joining analysis based on Nei-Li distances as well as heuristic parsimony analysis were performed using PAUP* v. 4.0b10 (Swofford, 2003) to reveal phylogenetic relationships among species. Robustness of branches was estimated using the bootstrap approach with 5000 replicates in case of NJ and 500 replicates for MP analysis. Principal coordinate analysis (PCoA) using Jaccard's coefficient performed with SYN-TAX 2000 (Podani, 2001) was computed to visualize non-hierarchical structure in complete AFLP data set. Phylogenetic relationships among diploid and hexaploid taxa (12 species), where complex evolutionary history are expected, were depicted by NeighborNet network based on Nei-Li distances in SplitsTree 4 (Huson and Bryant, 2006).

Results

The AFLP analysis of 115 individuals resulted in a total of 364 scorable bands, where 96% were polymorphic. All methods used for data analysis (NJ, MP, PCoA, NN) revealed similar pattern of forming two main clades - first clade covered all taxa belonging to genome size group G1 (hexaploids and higher polyploids), while second group comprised [illegible] taxa belonging to [illegible] and all of the remaining samples. Position of *C. angustifolia* (G2) was on the border of these two groups. The basal relationships among the genome size groups, were poorly resolved, on the other hand, some particular species relationships within genome size groups G1 and G2 were highly supported by bootstrap analysis.

Fig. 2: Neighbor-Network of diploid and hexaploid species only. 55 individuals, 350 AFLP characters. Bootstrap values were obtained in NJ analysis of the same dataset (5000 replicates).

Aims

→ to check the monophyletic character of „Genome size groups" sensu Leong-Škorničková et al. (2007)

→ to outline the basic phylogenetic relationships between species within and also among this groups.

Fig. 3: Neighbor-joining tree of 115 *Curcuma* individuals; 350 informative characters; NJ as well as MP bootstrap value higher than 50% are presented (NJ / MP).

Genome size group	2n	1Cx value				
Species population	Sample size	Mode of reproduct.	Origin of population	BT	BA	% *PPL*
Hexaploid Group II	2n=42	1Cx (0.38 – 0.39 pg)				
C. angustifolia U	5	sexual	Uttaranchal	108	62.8	63
C. angustifolia Ch	4	sexual	Chhattisgarh	117	75	50
C. cannanorensis K	4	sexual	Kerala	105	69.5	58
C. [illegible] M	5	sexual	Maharashtra	77	39.6	71
C. pseudomontana M	5	sexual	Maharashtra	85	47	65
C. [illegible] Ch	2	sexual	Chhattisgarh	95	77	37
C. [illegible] MP	4	sexual	Mad. Prad.	82	44.5	75
C. [illegible] K	4	sexual	Kerala	52	34	48
C. [illegible] Ka	5	sexual	Karnataka	95	60.25	33
Hexaploid Group III	2n=42	1Cx (0.41 – 0.48 pg)				
C. mutabilis K	4	sexual	Kerala	128	98.78	63
11x Group III	2n=77	1Cx (0.41 – 0.43 pg)				
C. [illegible] SL	4	sex/veg	Sri Lanka	96	73.5	47
TOTAL	115	-	-	350	72.43	96.15

Table 1: Species in study and AFLP data characteristics; Group 0, I, II, III - genome size groups sensu Leong-Škorničková et al. (2007); 1Cx - intervals of homoploid genome sizes of individual genome size groups; (BT) – total number of bands present in a given population; bands average (BA) – mean number of bands present per individual; % PPL - percentage of polymorphic bands (countered for 4 random chosen individuals per population, 1000 times repeated and median value presented)

Discussion

AFLP data partially supports former hypothesis designed in Leong-Škorničková et al. (2007), that at least two genome size groups (G1 and G2) have different evolutionary histories. It is, however, not reliable to determine if Genome size groups are of monophyletic origin based on AFLP data only.

Genome size Group G1

Within Genome size group G1 we outline relationships mainly between taxa of different ploidy levels: between hexaploids and nonaploids or between nonaploids and 12-ploids/ 15-ploids. We hypothesize that majority of nonaploid cytotypes plausibly originated by a fusion of reduced and unreduced gametes of hexaploids and that both nonaploid and hexaploid taxa could be involved in formation of higher polyploids (12x, 15x). Existence of common ancestor of all these species is thus very likely.

C. *montana* Roxb. | C. *xanthorrhiza* Roxb. | C. *aeruginosa* Roxb.

Genome size Group G2 and G3

Among taxa of G2 group, which are mainly sexually reproducing hexaploids, well supported clades were detected, and these were also well separated from taxa of G1 groups. Interestingly while no support was detected for relationship between the two G3 taxa in this study, there was observed high bootstrap support for cluster of G3 taxon *C. mutabilis* and G2 taxon *C. cannanorensis*. We speculate, that *C. mutabilis* either evolved from G2 group lineage with subsequent increase of its genome size, or some hybridization event(s) occured between these two hexaploids with sympatric occurence.

Hybridization?

As hybridizations in genus *Curcuma* were observed in the field (Škorničková and Sabu, 2005) and experimentally proved in pilot molecular studies (PCR-RFLP) between G1 and G2 species (*C. montana* x *angustifolia*) (Fér unpubl.), we explain position of G2 species *C. angustifolia* within subclade of G1 group by plausible former or simultaneous hybridization events. Exact ways of polyploid formation, parental linkage of particular higher polyploids and detailed investigation of hybridization events between *Curcuma* hexaploids are subject of our further studies.

Fig. 4: Principal Coordinate Analysis based on AFLP data of 115 individuals. First three coordinates cumulatively explaining 19.9% of the total variation. Combination of the first and third coordinate show obvious separation of taxa from genome size group G1 (black shapes) and taxa mainly from genome size group G2 (white dots). Individuals from genome size group G3 (grey dots) and G0 (stars) were integrated into G2 individuals.

References

Benham J, Jeung JU, Jasieniuk M, Kanazin V, Blake TK. 1999. Genographer: A Graphical Tool for Automated Fluorescent AFLP and Microsatellite Analysis. Journal of Agricultural Genomics 4:3

Huson DH, Bryant D, 2006. Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23. 254-267.

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* (Zingiberaceae). Annals of Botany 100: 527-536.

Leong-Škorničková J, Šída O, Sabu M, Marhold K. 2008. Taxonomic and nomenclatural puzzles in Indian Curcuma: the identity and nomenclatural history of *C. zedoaria* (Christm.) Roscoe and *C. zerumbet* Roxb. (Zingiberaceae). Taxon 57(3): 949-962.

Podani J. 2001.SYN-TAX 2000. Computer programs for data analysis in ecology and systematics, User's manual, Scientia, Budapest

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (* and other methods). version 4.0b10. Sinauer Associates, Sunderland, Massachusetts, USA.

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

Škorničková J, Sabu M, 2005. *Curcuma roscoeana* Wall (Zingiberaceae) in India. Gardens' Bulletin Singapore 57: 187-198.

Acknowledgments

This study was financially supported by the Grant Agency of Charles University (GA UK No. 112808, Eliška Záveská, Tomáš Fér). Long-term stay of JL-Š at the Calicut University, Kerala, India was supported by the fellowships from the Indian Council of Cultural Relationships and the Ministry of Education, Youth and Sport of the Czech Republic. We thank the authorities of the Ministry of Environment and Forests, Government of India and the Botanical Survey of India for granting collecting permits and permits to access Indian herbaria.

A.

Phylogenetic Studies In Indian Polyploid Curcuma Species Using AFLP Marker And ITS Sequencing

Eliška Záveská (1), Tomáš Fér (1), Karol Marhold (1), Otakar Šída (2), Mamiyil Sabu (3) & Jana Leong-Škorničková (4)

(1) Department of Botany, Charles University in Prague, Prague, CZ; (2) Department of Botany, National Museum in Prague, 252 43 Pruhonice, CZ; (3) Department of Botany, Calicut University P.O., 673635, Kerala, India; (4) Singapore Botanic Gardens, 1 Cluny Road, 259569 Singapore

Background

Economically important genus *Curcuma* (Zingiberaceae) is taxonomically critical polyploid complex. Study of Indian *Curcuma* by Leong-Škorničková et al., 2007 revealed six ploidy levels (2x, 6x, 9x, 11x, 12x and 15x) and four groups of taxa – „Genome size groups" - with significantly different 1Cx-values, distribution and probably with different evolutionary history. Various ploidy levels developed different modes of reproduction and exhibit uneven levels of morphological variability within and between populations, blurring specific boundaries. Unresolved nomenclature further complicates overall situation (Leong-Škorničková et al., 2008). So far there is neither suitable infrageneric classifications of genus *Curcuma* (Škorničková, 2007), nor major published works focusing on phylogeny of *Curcuma*.

Materials and Methods

Plant material: Total of 20 Indian *Curcuma* species (overall 120 individuals) were under the study. Using AFLP's we analysed 13 hexaploid (2n = 42), 7 nonaploid (2n = 63), one endecaploid (2n = 77), one dodecaploid (2n = 84) and two pentadecaploid populations (2n = 105) from nominate subgenus *Curcuma* and one population of diploid (2n = 22) from subgenus *Hitcheniopsis*. Selected species covered all four genome size groups with different 1Cx values (Leong-Škorničková et al. 2007) (see Table 1). We analyzed two to ten individuals per population and one to three populations per taxon (in total 115 plants). In ITS study we followed in general sampling for AFLP's, but we were more interested in G1 and G2 group representatives. Thus 6 taxa (8 individuals) from G1 group and 11 taxa (12 individuals) from G2 group were analysed.

AFLP analysis: AFLP analysis was carried out using AFLP Core Reagent Kit I (Invitrogen) and AFLP Pre-Amp Primer Mix I (Invitrogen), with three primer combinations used for selective amplification. Fragments were resolved on an ABI 3100 Avant DNA sequencer (Applied Biosystems) together with GeneScan-ROX-500 size standard and scored with Genographer.2.0 (Benham, 1999; http://hordeum.msu.montana.edu/genographer/) in the range of 100-500 bp. Neighbor - joining analysis based on Nei - li distances as well as heuristic parsimony analysis were performed using PAUP* v. 4.0b10 (Swofford, 2003) to reveal phylogenetic relationships among species. Robustness of branches was estimated using the bootstrap approach with 5000 replicates in case of NJ and 500 replicates for MP analysis.

ITS clonning and sequencing: The internal transcribed spacers (ITS1 and ITS2) of the rDNA were sequenced with universal primers ITS4 and ITS5 (White et al., 1990). All of the PCR products were cloned using pGEM-T Easy Vector System (Promega). Five to ten clones from each individual were directly sequenced with the same primers as for initial PCR. The sequences were aligned in ClustalX, then corrected manually in BioEdit. Prior the phylogenetic analyses, suspected recombinant sequences were excluded (see Popp and Oxelman, 2001). Parsimony analyses and bootstrap analysis (1000 replicates) were conducted with PAUP*. As complex evolutionary history of species were expected, phylogenetic relationships were depicted by NeighborNet network in SplitsTree 4 (Huson and Bryant, 2006) based on distance matrix calculated in PAUP.

Aims

By AFLP marker and cloning analysis of ITS region:

- to check the monophyletic character of „genome size groups" sensu Leong-Škorničková et al. (2007)
- to outline the basic phylogenetic relationships between species within and also among this groups.

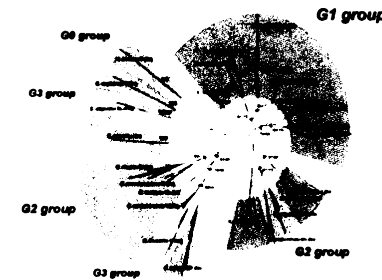


Fig. 1. AFLP analysis; Neighbor - joining tree of 115 *Curcuma* individuals; 350 informative characters; NJ as well as MP bootstrap value higher than 50% are presented (NJ / MP). Genome size groups are marked by colours as in Table 1.

Results

The AFLP analysis of 115 individuals resulted in a total of 364 scorable bands, of which 96% were polymorphic. All methods used for data analysis (NJ, MP, PCoA) revealed similar pattern of forming two main clades. First clade covered all taxa belonging to genome size group G1 (hexaploids and higher polyploids), while second group comprised hexaploid taxa belonging to G2 and G3 group and all of the remaining samples. *C. angustifolia* (G2) appeared on the border of these two groups. The basal relationships among the genome size groups were poorly resolved, on the other hand, particular species relationships within genome size groups G1 and G2 were highly supported by bootstrap analysis.

Analysis of ITS resulted in overall same pattern as observed by AFLP's. High diversity of haplotypes within individuals were detected by ITS cloning. Moreover, within G1 group ITS haplotypes of two or more different taxa belonging to various ploidy levels clustered together.

Discussion

Both AFLP and ITS data partially supports former hypothesis designed in Leong-Škorničková et al. (2007), that at least two genome size groups (G1 and G2) have different evolutionary histories. Cloning of ITS region additionally showed tremendous, but not unexpected, complexity of phylogenetic relationships between NE Indian *Curcuma* species and shed some light into parental linkage between hexaploids and higher polyploids.

Genome size Group G1

AFLP data outlined and supported relationships between G1 hexaploids and higher polyploids in general. ITS data additionally suggest repeated introgression/hybridization within hexaploids or paleo-allopolyploid origin of hexaploids. Based on proportion of ploidy levels in the genus we suggest that higher polyploids arise by fusion of reduced and unreduced gametes of hexaploids and both, one or more parental species, can participate in their formation.

Genome size Group G2 and G3

Majority of taxa belonging to genome size group G2, which are mainly sexually reproducing hexaploids, were well supported as a clade by AFLP as well as by ITS data. Although there were detected different ITS haplotypes within individuals, these clustered together and therefore we assume, that this variability was caused by within-genome mutations. Our data supported formerly proposed inclusion of *Hitchenia caulina* and *Monolophus scaposus* into genus *Curcuma*. *Curcuma angustifolia* previously classified as member of G2 group by its 1Cx value is well nested into clade of G1 species.

Genome size group G3 was confirmed as polyphyletic based on this study. *C. mutabilis* is overall similar to G2 group species and its higher 1Cx value likely evolved by recent increase of genome size.

Conclusions

Both methods, AFLP and sequencing of rDNA ITS region, reveal similar pattern in Indian *Curcuma* species and supported the hypothesis outlined in our former cytological study (Leong-Škorničková et al. 2007). Genome groups G1 and G2 have independent evolutionary histories, while G3 seem to be artificial group of polyphyletic origin. Contrary to previous study, border-line cases *C. angustifolia* and *C. mutabilis*, are members of G1 and G2 group respectively.

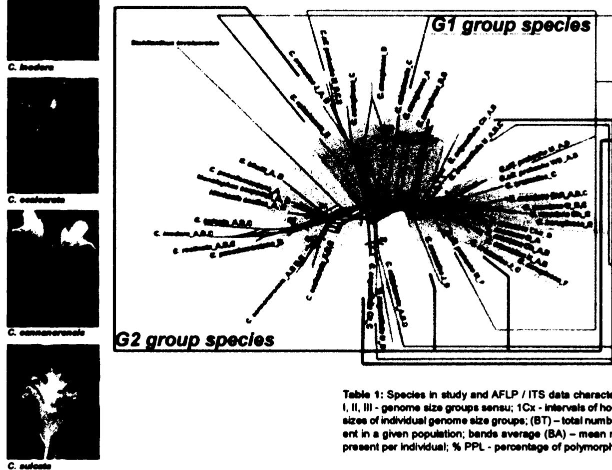
C. inodora

C. ecalcarata

C. cannanorensis

C. sulcata

Fig. 2. ITS analysis; Neighbor - Network of 20 *Curcuma* individuals, based on genetic distances calculated in PAUP. Bootstrap values (1000 replicates) higher than 50% are presented; Letters A - G behind each sample represent different ITS haplotypes



Genome size group	2n	1Cx value		AFLP	ITS
Species in study	Sample size AFLP / ITS	Mode of reproduct.	Origin of population	BT/BA/%PPL	No. of clones/no. of haplotypes
[illegible]					
C. vamana K	3/ -	sex/veg	Kerala	30/ 30/ 32	-
[illegible]					
C. aromatica K	3 / -	sex/veg	Kerala	98/ 88/ 17	-
C. aromatica KH	- / 1	sex/veg	Kerala	-	5/ 3
C. montana J	4/ -	sex/veg	Jharkhand	119/ 84/ 49	-
C. montana Ch	3/ 1	sex/veg	Chhattisgarh	84/ 58/ 49	8/ 3
C. montana ChII	-/ 1	sex/veg	Chhattisgarh	-	8/ 3
C. montana O	-/ 1	sex/veg	Orissa	-	5/ 3
C. aff. prakasha M	-/ 1	sex/veg	Meghalaya	-	8/ 6
[illegible]					
C. aeruginosa As	9/ 1	vegetative	Assam	154/ 108/ 34	8/ 6
C. aeruginosa K1	10/ -	vegetative	Kerala	125/ 93/ 44	-
C. aeruginosa K2	7/ -	vegetative	Kerala	47/ 34/ 42	-
C. aromatica J	-/ 1	vegetative	Jharkhand	-	9/ 4
C. leucorhiza J	4/ -	vegetative	Jharkhand	150/ 85/ 54	-
C. longa K	4/ -	vegetative	Kerala	91/ 76/ 74	-
C. longa SL	5/ -	vegetative	Sri Lanka	93/ 73/ 24	-
C. zanthorrhiza K	3/ -	vegetative	Kerala	46/ 31/ 38	-
[illegible]					
C. sp. '[illegible]' J	4/ -	vegetative	Jharkhand	93/ 74/ 35	-
[illegible]					
C. raktakanta As	5/ 1	vegetative	Assam	115/ 103/ 25	7/ 5
C. raktakanta K	5/ -	vegetative	Kerala	107/ 81/ 12	-
[illegible]					
C. angustifolia U	5/ 1	sexual	Uttaranchal	108/ 62/ 63	5/ 4
C. angustifolia Ch	4/ 1	sexual	Chhattisgarh	117/ 75/ 59	8/ 3
C. cannanorensis K	4/ 1	sexual	Kerala	105/ 69/ 50	8/ 4
C. inodora M	5/ 1	sexual	Maharashtra	77/ 39/ 71	7/ 3
C. pseudomontana M	5/ 1	sexual	Maharashtra	85/ 47/ 63	7/ 3
C. reclinata Ch	2/ 1	sexual	Chhattisgarh	95/ 77/ 37	7/ 3
C. sulcata MP	4/ 1	sexual	Mad. Prad.	82/ 44/ 75	10/ 3
C. ecalcarata K	4/ 1	sexual	Kerala	52/ 34/ 48	5/ 3
C. bhatii Ka	5/ 1	sexual	Karnataka	95/ 60/ 33	10/ 2
Hitchenia caulina	-/ 1	sexual	Maharashtra	-	6/ 2
Monolophus scaposus	-/ 1	sexual		-	5/ 2
Hexaploid Group III	2n=42	1Cx (0. 41 – 0. 43 pg)			
C. mutabilis K	4/ -	sexual	Kerala	128/ 98/ 63	-
C. aff. prakasha WB	-/ 1	sexual	W. Bengal	-	8/ 2
11x Group III	2n=77	1Cx (0. 41 – 0. 43 pg)			
C. oligantha SL	4/ -	sex/veg	Sri Lanka	96/ 75/ 47	-

Table 1: Species in study and AFLP / ITS data characteristics; Group 0, I, II, III - genome size groups sensu; 1Cx - intervals of homoploid genome sizes of individual genome size groups; (BT) – total number of bands present in a given population; bands average (BA) – mean number of bands present per individual; % PPL - percentage of polymorphic bands;

C. aromatica

C. aeruginosa

C. raktakanta

C. montana

Acknowledgement

This study was financially supported by the Grant Agency of Charles University (GA UK No. 112608, Eliška Záveská, Tomáš Fér), Long-term stay of JL-Š at the Calicut University, Kerala, India was supported by the fellowships from the Indian Council of Cultural Relationships and the Ministry of Education, Youth and Sport of the Czech Republic. We thank the authorities of the Ministry of Environment and Forests, Government of India and the Botanical Survey of India for granting collecting permits and permits to access Indian herbaria

References

Benham J, Jeung JU, Jasieniuk M, Kanazin V, Blake TK. 1999. Genographer: A Graphical Tool for Automated Fluorescent AFLP and Microsatellite Analysis. Journal of Agricultural Genomics 4:3

Huson DH, Bryant D, 2006. Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution. 23: 254-267.

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* (Zingiberaceae). Annals of Botany 100: 527-536.

Leong-Škorničková J, Šída O, Sabu M, Marhold K. 2008. Taxonomic and nomenclatural puzzles in Indian *Curcuma*: the identity and nomenclatural history of *C. zedoaria* (Christm.) Roscoe and *C. zerumbet* Roxb. (Zingiberaceae). Taxon 57(3): 949-962

Podani J. 2001.SYN-TAX 2000. Computer programs for data analysis in ecology and systematics, User's manual, Scientia, Budapest

Popp M, Oxelman B. 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophylaceae) using plastid and homoeologous nuclear DNA sequences. Molec. Phylogenet. Evol. 20 (3): 471-481.

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer Associates, Sunderland, Massachusetts, USA.

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

Škorničková J, Sabu M. 2005. *Curcuma roscoeana* Wall (Zingiberaceae) in India.Gar. Bull. Singapore 67: 187-198