Phylogenetics and historical biogeography of Southeast Asian *Begonia* L. (Begoniaceae)

by

Daniel Caspar Thomas

This thesis is submitted to the University of Glasgow for the degree of Doctor of Philosophy

Institute of Biomedical and Life Sciences

Division of Environmental and Evolutionary Biology

June, 2010

Abstract

The *Begonia* flora of Southeast Asia comprises more than 540 species. This exceptional species diversity and the wide distribution of the genus in tropical rainforests offers the opportunity to address biogeographical questions and to investigate the processes which underlie modern patterns of biodiversity, but also poses major taxonomic challenges. Only few apomorphies characterising infrageneric taxa in this large genus have been identified and delimitation of Asian *Begonia* sections is highly problematic. A robust phylogenetic framework of Asian *Begonia* informing taxonomic monographs and facilitating biogeographical and evolutionary studies is currently lacking.

Maximum parsimony, maximum likelihood and Bayesian analyses of plastid (*ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer; 115 taxa) and nuclear ribosomal (ITS; 89 taxa) sequence data were used to reconstruct the phylogeny of Southeast Asian *Begonia* and to determine whether major Asian sections are monophyletic. Morphological characters which are crucial in current sectional circumscriptions were mapped on the phylogeny to determine their degree of homoplasy and to assess their suitability in infrageneric classifications. Relaxed molecular clock analyses of a Cucurbitales-Fagales datatset (cpDNA: *mat*K gene, *rbc*L gene, *trn*L intron, *trn*L-F spacer; 92 taxa; five fossil calibrations) and a Begoniaceae datatset (cpDNA: *ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer; 110 taxa; two alternative secondary calibrations), as well as ancestral area reconstructions were employed to elucidate temporal and spatial diversification patterns in Asian *Begonia*.

The results indicate that Asian and Socotran *Begonia* species form a well supported clade. Most major Asian sections are not supported as monophyletic and the strong systematic emphasis placed on single, homoplasious characters such as undivided placenta lamellae (section *Reichenheimia*) and fleshy pericarps (section *Sphenanthera*), and the recognition of sections primarily based on a plesiomorphic fruit syndrome and the absence of characteristic features of other taxa (section *Diploclinium*) has resulted in the circumscription of several highly polyphyletic sections. Ovary and fruit characters have traditionally played a major role in sectional delimitation, however the high level of homoplasy associated with these has obscured systematic relationships in Asian *Begonia*. Gene trees derived from separate analyses of the plastid and nuclear ribosomal data show congruent support for several major clades, but there is hard incongruence within the clades comprising species of the species-rich sections *Platycentrum* s.l. (including section *Sphenanthera*) and *Petermannia* s.l. (including section *Symbegonia*), indicating that hybridization might have had a significant impact on the evolution of the genus.

The molecular divergence ages and the biogeographical analyses indicate an initial

diversification of Asian *Begonia* on the Indian subcontinent and in continental Southeast Asia in the Middle Miocene, and subsequent colonization of Malesia by multiple lineages. The predominant directional trend of the reconstructed dispersals between continental Asia and Malesia and within Malesia is from west to east including four independent dispersal events from continental Southeast Asia and the Malesian Sunda Shelf region to Wallacea dating from the Late Miocene to the Pleistocene. Dispersal across the ancient deep water channels separating intervening islands of the Sunda Shelf and Wallacea and subsequent successful colonisation of Wallacean islands seem to have been infrequent events during this period. This suggests that the water bodies which have separated the Sunda Shelf region from Wallacea have been distinct, yet porous barriers to the predominantly anemochorous dispersal in *Begonia*. The inferred timing of dispersals from the Sunda Shelf region to Wallacea is generally concordant with hypotheses about the geological history of the region, which indicate that the period from the Late Miocene onwards offered opportunities for dispersal to Wallacea and across Wallacea to New Guinea as substantial land masses emerged in Sulawesi and New Guinea, and newly emergent volcanic islands along the Sunda Arc, the Banda Arc and the Halmahera Arc formed potential routes for dispersal by island hopping.

The results further suggest that *Begonia* section *Petermannia* (>270 spp.) originated in the Malesian Sunda Shelf region, and subsequently dispersed to Wallacea, New Guinea and the Philippines. Lineages within this section diversified rapidly since the Pliocene with diversification peaking in the Pleistocene. The timing of diversifications coincides with orogenesis on Sulawesi and New Guinea, as well as pronounced glacioeustatic sea-level and climate fluctuations. It can be hypothesised that a complex interplay of extrinsic and intrinsic factors including the presence and formation of suitable microhabitats by orogenesis, cyclical vicariance by frequent habitat fragmentations and amalgamations caused by sea-level and climate fluctuations, as well as only weakly developed mechanisms to maintain species cohesion in fragmented habitats in *Begonia* could have driven speciation in allopatry and could have resulted in the remarkable *Begonia* species diversity found in Southeast Asia today.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgements

First and foremost I thank my advisers Mark Hughes, James Richardson and Rod Page for their support in all of its various forms, and for giving me the opportunity to join the Tropical Group at the Royal Botanic Garden Edinburgh to carry out research on *Begonia*. It was a genuine privilege to work in such a dynamic and international environment, and I am grateful to all members of the Tropical Group and the molecular biology laboratories for sharing their expertise.

The funding of this PhD project by the M. L. MacIntyre Trust, and the support by the trustees, are gratefully acknowledged.

I would like to thank several colleagues who have constantly been willing to discuss *Begonia* and to impart their expert knowledge on the subject: Mobina Shaukat Ali, Wisnu H. Ardi, Deden Girmansyah, Nikki Harrison, Catherine Kidner, Thamarat Phutthai, Sangeeta Rajbhandary, Alex Twyford, and Harry Wiriadinata.

Four months of productive field work in Indonesia would not have been possible without the help and support of many people and organizations. Firstly, I would like to thank RISTEK for allowing me to collect and conduct research in Indonesia; secondly, the Kebun Raya Bali and the Kebun Raya Bogor and staff, particularly former director Dr Irawati, the current director Ir. Mustaid Siregar and his wife Mrs Hartutiningsih, who were always supportive. My counterpart Mr Wisnu H. Ardi was invaluable in helping and advising me at all stages.. The Herbarium Bogoriense provided us with export permits for the herbarium material and the director Dr Eko Baroto Walujo could not have been more helpful. I could not have gone without the money provided by my sponsors, the Royal Horticultural Society, the Stanley Smith Horticultural Trust and the Systematics Association, whose support is gratefully acknowledged. Expeditions to remote parts of Sulawesi depend ultimately for their success on the strength and knowledge of the local people involved, and I am aware of a debt, that can never be repaid to these people who gave so freely of their time, knowledge and strength.

I am grateful to everyone who has helped with the collection of plant material, provided silica dried samples, DNA isolations and sequence data during the course of my research. Notable mentions must go to Marlina Ardiyani, Kate Armstrong, Daniele Cicuzza, Janet Gagul, Deden Girmansyah, Hartutiningsih, David Middleton, Mark Newman, Axel Poulsen, Sangeeta Rajbhandary, Thamarat Phutthai, Peter Wilkie, and Harry Wiriadinata.

Living collections were essential for my research, and I would like to thank the horticulture staff at the RBGE and Glasgow Botanic Gardens, and particularly Steve Scott, Neil

Watherson and Sadie Barber, for the excellent maintenance of the collections.

I would also like to acknowledge Ruth Kiew and Mark Tebbitt for their constructive reviews of the alpha-taxonomic work, and Kate Armstrong, Barbara Nobmann, and Alex Twyford for most helpful corrections of an earlier version of the manuscript.

Thanks also go to my fellow PhD students for their moral support and for imparting occasional pearls of wisdom: Bhaskar Adhikari, Mobina Shaukat Ali, Kate Armstrong, Rhiannon Chrichton, Jane Droop, Emma Goodyer, Nikki Harrison, Tobias Marczweski, Carmen Puglisi, Sumudu Rubasinghe, Tiina Sarkinen, and Alex Twyford.

Finally, I am forever grateful to my parents, family and Barbara. Without their love, encouragement, understanding, and selflessness, none of this would have been possible.

Declaration

I declare that the work described in this thesis has been carried out by myself unless otherwise acknowledged. It is entirely of my own composition and has not, in whole or part, been submitted for any other degree.

Daniel C. Thomas

Edinburgh

June, 2010

CHAPTER 1. Introduction

1.1 Introduction to *Begonia*

1.1.1 The current state of Begoniaceae phylogenetics

Begoniaceae are a pantropically distributed family, which comprises more than 1550 species of herbs and soft-wooded shrubs (Carlquist, 1985; Doorenbos et al., 1998; Forrest and Hollingsworth, 2003; Hughes, 2008). Two genera are currently recognised in the family: the monotypic genus *Hillebrandia* Oliv., and the species-rich and morphologically diverse genus *Begonia* L. (Doorenbos et al., 1998; Forrest and Hollingsworth, 2003). Diagnostic characters of the family include asymmetrical leaves, unisexual flowers and monoecy, twisted, papillose stigmas, and dry, three-winged capsules (Doorenbos et al., 1998; Judd et al., 2008). These characters are present in the majority of species, but there are numerous deviations from this typical syndrome. The minute seeds of Begoniaceae species exhibit an autapomorphic syndrome with a transverse ring of elongated collar cells bordering the micropylar-hilar part, whose rupture results in an operculate opening of the seed (Bouman and de Lange, 1983; de Lange and Bouman, 1992, 1999). *Hillebrandia sandwicensis* Oliv., which is endemic to Hawaii (Clement et al., 2004), can be differentiated from *Begonia* by a suite of characters including more differentiated segments of the perianth (a perigone is present *Begonia*), semi-inferior ovaries (inferior in *Begonia*), and fruit dehiscence between the styles (usually loculicidal dehiscence with lines of dehiscence along most of the length of the ovary in *Begonia*) (Clement et al., 2004; Forrest et al., 2005). The Begoniaceae are placed within the order Cucurbitales in the Eurosids I (Angiosperm Phylogeny Group, 2003, 2009). A molecular phylogenetic study by Zhang et al. (2006) indicates a close relationship of Datiscaceae, Tetramelaceae and Begoniaceae, and a sister group relationship of Datiscaceae and Begoniaceae, but the latter relationship is only weakly supported.

Begonia is one of the largest genera of vascular plants (Frodin, 2004), with more than 1550 species divided into 68 sections (de Wilde and Plana, 2003; Doorenbos et al., 1998; Forrest and Hollingsworth, 2003; Ku et al., 2007; Ku, 1999; Shui et al., 2002). The genus has a pantropical distribution, and is absent only from the Australian tropical forests and the Pacific region from the east of Fiji to the Galapagos Islands (Heywood, 2007; Tebbitt, 2005). A single species, *Begonia grandis* Dryand., extends the distribution of the genus into the temperate zone growing as far north as Hebei Province in China (Heywood, 2007; Tebbitt, 2005). The African *Begonia* flora is with *c.* 160 species relatively species poor (Doorenbos et al., 1998; Plana, 2003; Sosef, 1994), while the bulk of the species diversity is relatively equally distributed between the Neotropics and Asia (Doorenbos et

al., 1998). Based on the *Annotated Checklist of Southeast Asian Begonia* (Hughes, 2008) and subsequent new species descriptions (Girmansyah, 2009; Girmansyah et al., 2009; Hughes and Coyle, 2009; Hughes et al., 2010; Hughes et al., 2009; Kiew and Sang, 2009; Thomas et al., 2009a; Thomas et al., 2009b; Thomas and Hughes, 2008), a conspicuous hotspot of species diversity can be identified in Southeast Asia, whose *Begonia* flora comprises more than 540 species.

Several molecular phylogenetic trees that include low-density worldwide taxon samplings of *Begonia* have been published (Clement et al., 2004; Forrest, 2001; Forrest and Hollingsworth, 2003; Forrest et al., 2005; Goodall-Copestake, 2005; Goodall-Copestake et al., 2009; Plana, 2002, 2003; Plana et al., 2004). The work by Goodall-Copestake (2005), who analysed *c.* 13 kb of sequence data from eleven regions of nuclear, mitochondrial and chloroplast DNA of 31 Begoniaceae species, provides a well resolved phylogenetic framework for *Begonia* (Fig. 1.1). His analyses suggest that both Asian and American *Begonia* lineages are derived from African ancestors, that Socotran and Asian *Begonia* form a well supported monophyletic group, and that the split up of the ancient continent of Gondwana long preceded the generation of the pantropical distribution of the genus (Goodall-Copestake, 2005).

The phylogenetic relationships within the relatively small group of African Begonias (*c.* 160 species) are relatively well understood. Revisions exist for the majority of the 17 African sections (see references in Plana, 2003), and their intersectional relationships have been studied in some detail using molecular systematic approaches (Plana, 2002, 2003; Plana et al., 2004). In contrast to this, there is only very limited knowledge of the phylogenetics of the extensive American and Asian *Begonia* radiations. Morphological circumscriptions of several Neotropical and Asian sections are highly problematic. Doorenbos et al. (1998: 181), who revised sectional circumscriptions in *Begonia*, emphasized that several neotropical sections "shade off into each other," i.e. sectional boundaries are often inconsistent and transitional species are present. Moreover, several crucial fruit and ovary characters traditionally used in infrageneric classifications of *Begonia* were identified as highly homoplasious in phylogenetic trees based on low-density, world-wide taxon sampling of the genus (Forrest and Hollingsworth, 2003; Forrest et al., 2005). Doorenbos et al. (1998) recognised 18 Asian sections, and another four sections were subsequently proposed (Forrest and Hollingsworth, 2003; Ku et al., 2007; Ku, 1999; Shui et al., 2002). However, the circumscriptions of several of these sections are questionable. Analyses of sequence data of the nuclear ribosomal DNA provided strong evidence for the polyphyly of sections *Sphenanthera* (Hassk.) Warb., *Platycentrum* (Klotzsch) A.DC. and *Leprosae* (T.C. Ku) Y.M. Shui and hinted at the polyphyly of section *Diploclinium* (Lindl.) A.DC. (Forrest, 2001; Tebbitt et al., 2006). Moreover, the New Guinean genus *Symbegonia* Warb., which was traditionally separated from *Begonia* based on floral characters including a syntepalous perigone and a characteristic monadelphous androecium, was

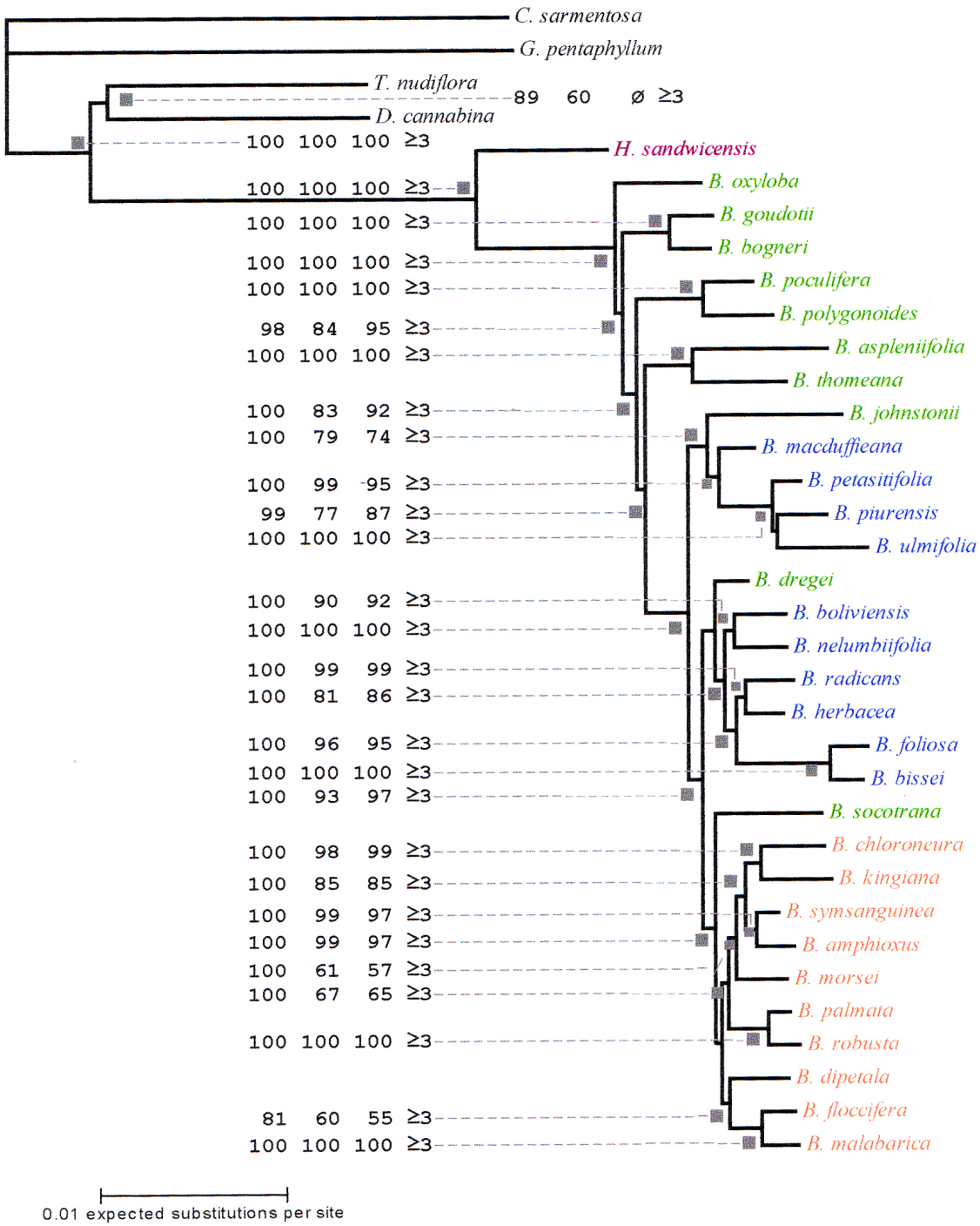


Fig. 1.1. Bayesian majority rule phylogram from Goodall-Copestake (2005: 56). Captions from Goodall-Copestake (2005): "Bayesian majority rule consensus phylogram for eleven DNA fragment dataset, see methods for the model and partitioning strategy. Begoniaceae taxa are colored: *H. sandwicensis* is purple, African *Begonia* are green, Asian *Begonia* are red, and American *Begonia* are blue. The phylogeny confidence measures are: Bayesian posterior probability, maximum likelihood bootstrap, maximum parsimony bootstrap, and if the inferred branch length is ≥/<3 substitutions long. Branches lengths ≥3 substitutions are in bold type. No support measures are given for poorly supported nodes that were only resolved in the Bayesian consensus topology and nor resolved in either of the maximum likelihood and maximum parsimony consensus topologies."

shown to be nested within *Begonia* section *Petermannia* (Klotzsch) A.DC. (Forrest and Hollingsworth, 2003). However, because of limited taxon sampling and unresolved or only poorly statistically supported deeper internal relationships of published phylogenetic trees, the intersectional relationships within Asian *Begonia* have remained only very fragmentarily understood.

1.1.2 The ecology of Asian *Begonia*

1.1.2.1 Habitats and life strategies of Asian Begonia

The majority of Begoniaceae species are perennial herbs or soft-wooded shrubs adapted to shady, moist microhabitats in primary rainforests, often associated with rocky slopes along small streams and waterfalls (Goodall-Copestake, 2005; Kiew, 2005; Phutthai et al., 2009). In areas with a limestone karst topography Begonias can frequently be found on shaded, moist rocks at the base of limestone hills or growing directly on moist rock faces (Kiew, 1998, 2001a, 2001b; Kiew, 2005; Kiew and Sang, 2009). Lithophytic Asian Begonias often grow on granite or limestone, but shale, sandstone and quartzite substrates have also been reported for some species (Kiew, 2005; Phutthai et al., 2009; locality data in Hughes & Pullan, 2007). Many Asian species can grow both lithophytically and on soil in the ground layer of primary rainforest, and Kiew (2005) pointed out that most forest floor dwelling Begonias are found in habitats with steep inclinations where leaf litter does not collect, which would otherwise inhibit the establishment and growth of the minute seedlings. Some Asian species can also be found growing epiphytically, either on wet bark of tree trunks near the ground (Phutthai et al., 2009), or rarely as climbing epiphytes (*Begonia oxysperma* A.DC. in the Philippines and *B. kaniensis* Irmsch. in New Guinea). Many *Begonia* species are highly sensitive towards disturbances which significantly alter light exposure and humidity intensities in the habitat, and only a few Asian species such as the widespread *B. longifolia* Blume and *B. alicida* C.B.Clarke, show wider ecological tolerances towards habitat disturbances and tolerate high levels of insolation (Phutthai et al., 2009; Tebbitt, 2003).

Adaptations to seasonal climates can be found in several *Begonia* lineages. Many continental Asian and a few Malesian species in sections *Parvibegonia* A.DC., *Diploclinium* and *Sphenanthera* can resist drought by dying down during the dry season and by resprouting from tuberous or rhizomatous perennating organs in the next rainy season (Kiew, 2005; Phutthai et al., 2009). Although they are perennials, some tuberous Asian species, like *Begonia sibthorpioides* Ridl. and *B. sinuata* Wall. ex Meissner, show a similar life strategy as drought avoiding annual plants: they are small plants, i.e. they only produce a small amount of biomass in each growth period, mature flowers are often developed relatively soon after resprouting and germination, and they profusely regenerate from the seed bank in the rainy season (Kiew, 2005). These adaptations seem to be essential to survive dry seasonal extremes or pronounced seasonal monsoonal climates with several

dry months, not just in continental Southeast Asia (Phutthai et al., 2009) and parts of the Himalaya (Sangeeta Rajbhandary, Tribhuvan University, Kathmandu, Nepal, *pers. com.*), but also in the North of the Malay Peninsula (Kiew, 2005), and eastern Java and the Lesser Sunda Isles (*pers. obs.*). Adaptations to seasonally dry climates are also known from African *Begonia* species in the sections *Augustia* (Klotzsch) A.DC., *Peltaugustia* (Warb.) Barkley, *Rostrobegonia* Warb., and *Sexlaria* A.DC. Adaptations in these sections include drought avoidance through an annual life cycle (sections *Rostrobegonia* and *Sexlaria*), drought resistance by resprouting from swollen stem bases (sections *Augustia*, *Peltaugustia*, and *Rostrobegonia*), and the production of bubils, i.e. compressed, fleshy lateral branches with reduced, scaly leaves, which serve as perennating organs (section *Peltaugustia*) (Goodall-Copestake, 2005; Hughes and Miller, 2002; Plana, 2002, 2003).

1.1.2.2 Reproduction and dispersal in Asian Begonia

The artificial propagation of Begonias by the use of leaf and shoot cuttings is a common horticultural practise (Tebbitt, 2005)(Tebbitt, 2005), and clonal reproduction seems to be a common phenomenon in the wild. Regeneration of new plantlets on the veins of fallen leaves or from shoot or rhizome fragments has been observed in several species and natural hybrids (Kiew, 2005; Peng and Chiang, 2000; Peng and Ku, 2009). Other species show specialised vegetative organs, which likely serve for vegetative reproduction. Examples are the Himalayan species *Begonia gemmipara* Hook.f. & Thomson which produces clusters of bulbils in modified inflorescences (Clarke, 1879; Grierson, 1991), and *B. sinuata*, which develops bulbils at the base of the leaf blades (Kiew, 2005). In other species, roots are developed at the apices of their leaves, where new plantlets are developed (*Begonia elisabethae* Kiew, *B. vagans* Craib), and some species develop plantlets on the veins ending in the leaf margins (*B. elisabethae*) (Craib, 1930; Kiew, 2005).

Most Begonias seem to be zoophilous and pollinated by generalist insect pollinators. Stingless bees (*Trigona* species), honey bees (*Apis cerana*) and bumble bees (*Bombus ephippiatus*) have been reported as flower visitors and likely pollinators in *Begonia* (Ågren and Schemske, 1991; Burt-Utley, 1985; Hughes and Hollingsworth, 2008; Kiew, 2005; Schemske et al., 1996). A few species like the neotropical *Begonia boliviensis* A.DC., and *B. ferruginea* L.f., are thought to be bird-pollinated (Vogel, 1975, 1993, 1998), which may also be the case in species in the New Guinean *Begonia* section *Symbegonia* (Warb.) L.L.Forrest & Hollingsw., which is characterised by syntepalous, tubular perigones (Hughes, 2002). Moreover, Hughes (2002) suggested that some species which grow in wind exposed habitats and exhibit large inflorescences which produce copious amounts of pollen may be wind pollinated (e.g. *Begonia glabra* Aubl.). Female *Begonia* flowers do not offer nectar or other rewards to pollinators. Pollinators apparently mistake the rewardless female flowers, which signal reward by the same yellow or orange colouration of their styles and stigmas as the androecia of the male flowers, for male flowers, which offer a reward in the form of pollen (Renner, 2006; Schemske et al., 1996).

The resemblance of male and female flowers seems to be essential for the effectiveness of this deceit pollination and effective mimicry may be selected for (Le Corff et al., 1998). An exception to the general rewardlessness of female *Begonia* flowers can be observed in the bird-pollinated *B. ferruginea*, in which nectar secreting tissue is developed in the female, but not in the male flowers (Vogel, 1998). Other putatively bird-pollinated *Begonia* species are completely rewardless and pollination is by double deceit, as nectar is lacking and pollen is not taken as reward by the bird pollinators (Hughes, 2002; Renner, 2006).

The majority of Asian species exhibit winged capsules with loculicidal deshiscence, and a single capsule produces hundreds to thousands of minute seeds. It has been estimated that 30000-70000 begonia seeds are needed to weigh one gram, and it can be assumed that these dust seeds, which are usually about 300-600 μm in length, can be effectively distributed by wind (de Lange and Bouman, 1992, 1999; Kiew, 2005). Anemochorously dispersed taxa usually exhibit capsules with a dry, membranous pericarp and well developed wings (Fig. 2.2 A-C). After dehiscence seeds are gradually released through slits along the wing attachments when the winged capsules are shaken by wind. However, numerous anemochorous species are narrow endemics, and anemochory is unlikely to result in long dispersal distances in the sheltered conditions of the moist ground layer habitats preferred by the majority of Asian *Begonia* species (Burt-Utley, 1985; Hughes, 2002; Hughes and Hollingsworth, 2008). This hypothesis is corroborated by the strong genetic and morphological differentiation of subpopulations of some *Begonia* species at very local scales indicating very limited dispersal capabilities and limited gene flow between populations (Hughes and Hollingsworth, 2008; Hughes et al., 2003; Matolweni et al., 2000).

Alternatives to the predominant anemochorous dispersal mechanisms can be found in several Asian sections. The majority of species in section *Platycentrum* exhibit a rain-ballist syndrome characterised by capsules with a robust pericarp and two smaller and one distinctly larger wing (Fig. 2.2 D). These fruits are held by stiff pedicels, which curve at maturity so that parts of the pericarp and the two smaller wings form an upwards facing splash cup, while the longer wing points downwards. When rain drops hit the splash cup the capsule moves up and down on the stiff pedicel releasing the seeds through lines of dehiscence along the bases of the wings (Kiew, 2005; Savile and Hayhoe, 1978; Tebbitt et al., 2006). Several species in section *Parvibegonia* exhibit a similar splash cup syndrome. However, the capsules of other species in the section are held on thin pedicels, which wilt and hang downward at maturity, releasing their seeds when the capsule is moved by wind or rain drops.

While epi- and endozoochorous dispersal seem to be common in African *Begonia*, zoochory has not been observed in Asian *Begonia* (de Lange and Bouman, 1992, 1999;

Tebbitt et al., 2006). However, the development of fleshy pericarps in some Asian *Begonia* sections has been interpreted as indicative of endozoochory (Kiew, 2005; Tebbitt et al., 2006). Indehiscent fruits with a thick, fleshy pericarp characterize *Begonia* section *Sphenanthera* (Fig. 2.2 E-F), and have also been reported from section *Leprosa* (Shui et al., 2002) and a few species in the large section *Petermannia* (Thomas et al., 2009a). Fleshy pericarps have generally been interpreted as related to zoochory, but dispersal of some fleshy fruited species may be mainly by rain-wash from the decomposing fruit or a combination of rain-wash and epizoochory, which has been proposed as predominant dispersal mechanism of *c*. 40% of the African *Begonia* flora (de Lange and Bouman, 1992, 1999; Tebbitt et al., 2006; Thomas et al., 2009a). Tebbitt et al. (2006) hypothesised that some fleshy-fruited species in the Asian section *Sphenanthera* are dispersed by bats and other animals, and that zoochory may be a factor contributing to the unusually wide distributions of some taxa in the section.

1.2 Introduction to Southeast Asia's palaeogeography

1.2.1 Palaeogeography of Malesia

The geological history of Southeast Asia is remarkably complex, and has been discussed in detail in a large body of geological and biogeographical literature (Hall, 1998, 2001, 2002, 2009; Metcalfe, 2002; Michaux, 1991; Morley, 2000; Ridder-Numan, 1996). This brief overview of the palaeogeography and geological history is mainly based on the palaeogeographic and tectonic reconstructions by Robert Hall (Hall, 1998, 2001, 2002, 2009), which integrated various data sources including GPS measurements of plate motions, observations of seismic and volcanic activity, geophysical imaging, palaeomagnetic data, isotope data and stratigraphy. These reconstructions provide a valuable framework for addressing biogeographical questions, but it is important to note that the underlying geological data is often fragmentary and its interpretation sometimes contentious (Hall, 2009).

Southeast Asia is located at a meeting point of several converging, major tectonic plates. The Australian continent is moving northwards and is colliding with the eastern Eurasian Plate, and the Pacific and Philippine Plates are moving westwards subducting below the Eurasian Plate. The present configuration of Southeast Asia has broadly resulted from the amalgamation of fragments, which have rifted from Gondwana, with Eurasia, as well as convergence of the Australian and the Eurasian Plates closing a large water barrier between the land masses, and subsequent subduction, collision and arc volcanism at the plate margins (Hall, 2009). These geological processes in combination with sea-level fluctuations between glacial and interglacial periods have produced a complex mosaic of terrestrial and marine areas evolving throughout the Cenozoic (Hall, 1998, 2001, 2002,

2009). Some biogeographically important aspects of the geological genesis of the region are elaborated below.

During the Mesozoic the western Malesian region, including Peninsula Malaysia, Sumatra, Borneo, Java and the western part of Sulawesi, was formed by the accretion of continental fragments which had broken off the northern margin of Gondwana and rifted northwards with the Eurasian plate margin (Hall, 2001, 2009; Metcalfe, 2002). By the Late Cretaceous, *c.* 80 Ma ago, these western Malesian areas had sutured to Eurasia and palaeogeographic reconstructions indicate that substantial parts of western Malesia were terrestrial throughout the Cenozoic (65.5 Ma onwards) (Fig. 1.2) (Hall, 2001, 2009). In the Middle Eocene the Makassar Strait region extended by block faulting and subsidence, separating a western Sulawesian fragment from Borneo and forming a deep water channel, which purportedly hindered biotic exchange between western Malesia and areas to the east throughout the Cenozoic (Hall, 2002, 2009; Morley, 1998; Moss and Wilson, 1998; Voris, 2000). Most of this western Sulawesian fragment was submerged in the Miocene (Hall, 2009), but palynological data indicates that it carried a Laurasian flora, which might have survived on smaller emergent parts or neighboring islands acting as a source of these Laurasian floristic elements for the eastern Malesian areas (Morley, 2000, 2003). Nevertheless, the emergent parts of the Sunda Shelf region remained separated from terrestrial areas of the Australian plate by a major water barrier throughout the Eocene and Oligocene (Hall, 2009).

Eastern Malesia including parts of eastern Sulawesi, the Moluccas, the Lesser Sunda Islands and New Guinea is derived from continental fragments of the Australian Plate. These fragments drifted to their current position only during the Cenozoic, when western Malesian areas had largely been in place already (Fig. 1.2) (Hall, 2002, 2009; Metcalfe, 2001; van Welzen et al., 2005). The movement of these fragments closed the gap between the converging Australian Continent and the Eurasian Plate, but during most of their migration the Eastern Malesian fragments were submerged and substantial land in Wallacea and New Guinea only emerged from the Late Miocene onwards. However, smaller ephemeral islands were present before that and there is some evidence that the southeastern parts of Sulawesi had already emerged at 20 Ma (Hall, 2002, 2009; van Welzen et al., 2005). By 5 Ma Sulawesi was largely emergent and high mountains were present in West and Central Sulawesi (Hall, 2009). The New Guinean orogenesis was initiated at around 10 Ma, but the rapid rise of the major mountain ranges of New Guinea reaching a height of almost 5000 m probably occurred only since 5 Ma (Hall, 2009; Pigram and Davies, 1987).

Despite the elimination of the wider deep water barrier between the Sunda Shelf region and Australia, there was never one clear overland track which allowed biotic exchange by overland migration, even during periods of lower sea-level during the Pleistocene, and

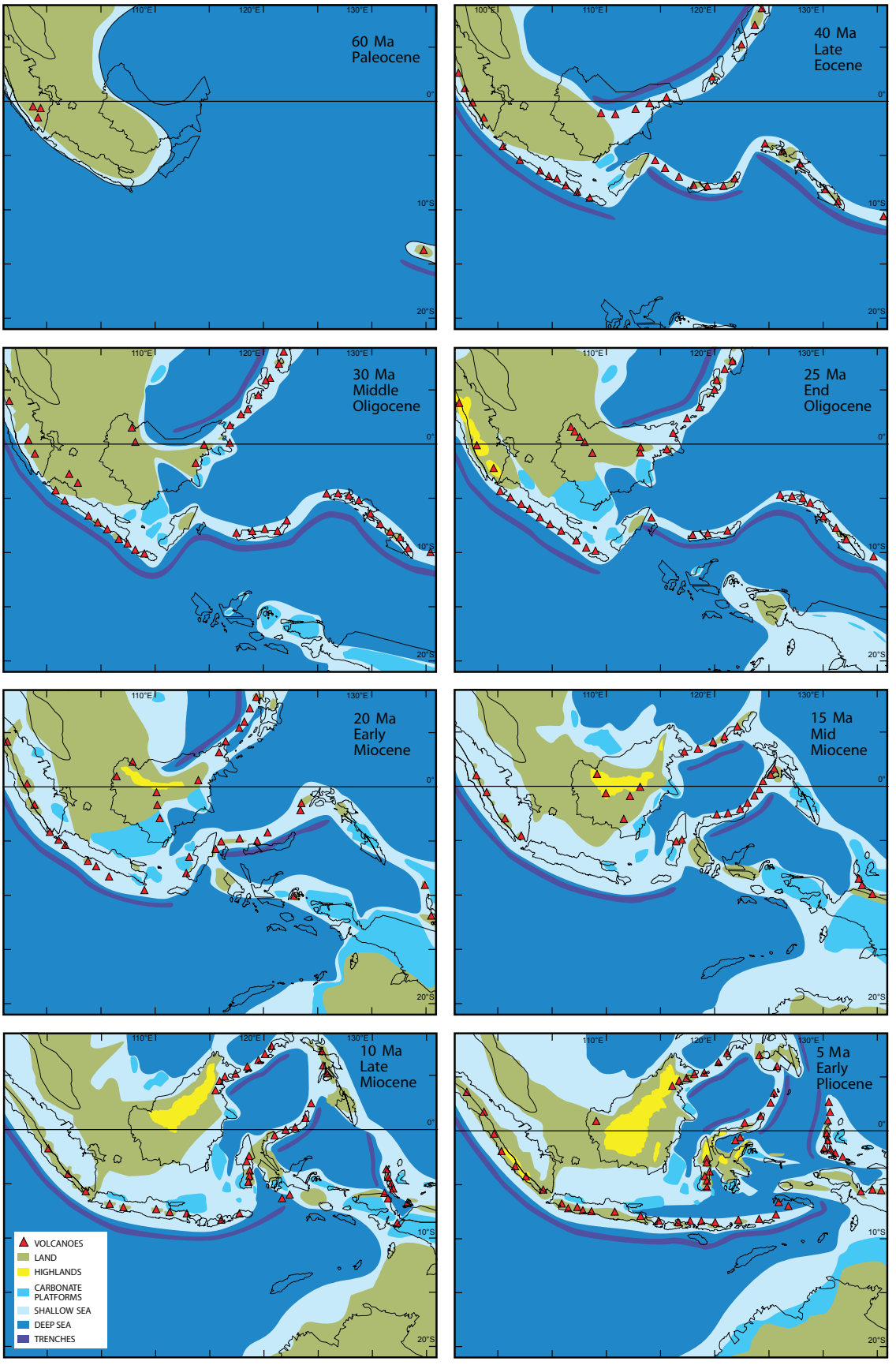


Fig. 1.2. Postulated distribution of land and sea in the Wallacea region. Figures from Hall (2009: 155-156; position and size of the legend, and position of age and epoch specifications slightly modified). Ma: Millions of years before present.

a variety of scenarios involving island hopping and rafting on tectonic fragments could be hypothesized for biotic exchange between the two regions (Hall, 2001). Adding to the complexity of these scenarios is the fact that the major Malesian islands and island assemblages have a composite nature. Sulawesi is an amalgamation of a continental fragment, which rifted from the eastern margin of Borneo during the Eocene, a northern volcanic arc, which was part of the subduction margin beneath which the northwards moving Indian-Australian Plate subducted in the past, extensive ophiolitic masses, i.e. uplifted oceanic crust and associated tectonic melanges, and Australian continental fragments such as Banggai-Sula and Buton (Hall, 2001; Morley, 2000; Moss and Wilson, 1998; Ridder-Numan, 1996). New Guinea consists of a larger southern part, the Australian continental craton, which collided and coalesced with about 32 smaller terranes of different origin (Daly et al., 1991; Pigram and Davies, 1987). Finally, the Philippines are an assemblage of blocks comprising, for the main part, ancient and more recent volcanic island arcs, but Palawan and Mindoro are Eurasian continental fragments (Hall, 2002).

Larger-scale patterns of biodiversity are strongly influenced by the physical environment (Benton, 2009), and the geological processes which contributed to the geological evolution of Southeast Asia certainly had a profound impact on the diversification and distribution of the Malesian biota. Physical barriers like large water bodies were reduced, e.g. by the convergence of the Australian and the Eurasian Plates, or eliminated, e.g. by the creation of land bridges connecting islands on the Sunda Shelf by glacioeustatic sea-level fluctuations, which potentially facilitated biotic exchange by overland migration and rafting on fragments. Other processes like the creation of the Makassar Straits by rifting of the western Sulawesi fragment from Borneo, and the genesis of new basins and deep water trenches created water barriers which likely acted as biogeographic filters or barriers, hindering range expansion of taxa with poor dispersal capabilities. However, the creation of physical barriers by fluctuating sea-levels and rifting of tectonic fragments could also have lead to speciation by vicariance, when previously wider distributed and connected ancestral populations were split and geographically separated. Moreover, the creation of topographical heterogeneity as in the relatively recent massive orogenesis on New Guinea is likely to promote microallopatric speciation and rapid divergence (Lomolino et al., 2006), and the emergence of formerly submerged tectonic fragments, terranes and volcanic islands offered opportunities for colonization of new land. Extrinsic factors like tectonic processes, changes in the palaeoclimate, and sea-level fluctuations have certainly played key roles in shaping biotic distributions in Malesia over millions of years.

1.2.2 Palaeoclimate and sea-level changes in Southeast Asia

The powerful impact of Pleistocene climate fluctuations and refugial dynamics on biogeographic patterns is well documented for Europe and other northern latitude areas,

and distinctive genetic patterns and subdivisions in various biota can be linked with isolation in cold-stage refugia (Hewitt, 2004; Lomolino et al., 2006; Petit et al., 2002). The impact of Pleistocene climate fluctuations on distribution patterns in tropical areas is much more contentious (Willis and Whittaker, 2000). The palynological record indicates that in tropical areas cooler and drier conditions during the cold stages resulted in the expansion of seasonal open forest and savannahs replacing and fragmenting tropical rainforest (Morley, 2000). However, the extent of this fragmentation may have differed dramatically between the Neotropics, Africa and Asia (Cannon et al., 2009; Colinvaux et al., 2000; Willis and Whittaker, 2000). Cyclical vicariance by recurrent isolation of once more widespread species in separate rainforest refugia during glacial maxima has been proposed as a model to explain the generation of high diversity and centres of endemism in the Neotropics (Haffer, 1969, 1997) and recent studies evoke cyclical vicariance as a speciation mechanism which may have contributed to large radiations in the Neotropics and African tropics (Harris et al., 2000; Janssens et al., 2009; Richardson et al., 2001; Sosef, 1994). However, Morley (2000) pointed out that the extent of aridification and rainforest contraction may have been overestimated in some cases, e.g. Brandon-Jones (2001) proposed almost complete reduction of Southeast-Asian rainforests by glacial drought, only leaving small refugia in north Sumatra, the Mentawi islands, north Borneo, west Java, northeast Indochina and the Western Ghats of south India. Recent modelling of rainforest distribution in Southeast Asia during the last million years incorporating geographic, geological and palaeoclimatic data indicate that the current rainforest distribution represents a refugial state and that Southeast Asian lowland rainforest had its widest extent at the time of Pleistocene glacial maxima (Cannon et al., 2009). This expansion might have been possible because the decrease of temperatures lowered the elevational zonation between upland and lowland rainforest, while eustatic sea-level changes resulted in the exposure of vast areas of the Sunda Shelf, which potentially provided suitable habitats for lowland rainforest (Cannon et al., 2009). The presence of more widespread rainforests during most of the Pleistocene is also supported by phylogeographic data which indicates that the main vicariance events in the region occurred before the Pleistocene (Cannon and Manos, 2003; Gorog et al., 2004). However, the palynological record also shows indicators for seasonal vegetation and there may have been a seasonal climate corridor with open woodland vegetation across the Sunda Shelf, and savannah vegetation on Java during substantial phases of the Pleistocene (Cannon et al., 2009; Morley, 2000, 2007).

Sequence stratigraphy (Woodruff, 2003, 2010), isotope data (Zachos et al., 2001), and the palynological record (Morley, 2000) indicate drastic eustatic sea-level fluctuations, and climate and vegetation changes in Southeast Asia long before the Pleistocene. From the early Miocene to the Pliocene the climate was predominantly moist, with a warm phase peaking around 15-17 Ma (Zachos et al., 2001). During this Mid Miocene Climatic Optimum Asian rainforest distribution reached as far north as Japan (Morley,

2000, 2007). However, the Miocene-Pliocene climate was not uniformly moist, but there were short-lived phases of glaciations, dryer, cooler climates, e.g. a short, deep glaciation at the Oligocene/Miocene boundary at around 23 Ma (the Mi-1 glaciation), which was followed by several intermittent but smaller glaciations (Zachos et al., 2001). The Mid Miocene Climatic Optimum was followed by gradual cooling until the early Pliocene, which showed a slight warming trend until ca. 3.2 Ma, when the Northern Hemisphere Glaciation commenced (Zachos et al., 2001). Moreover, eustatic sea-level curves indicate that during a *c.* 11 Ma period beginning at 24 Ma in the early Miocene, and during another *c.* 1.0-1.4 Ma period beginning at 5.5 Ma in the early Pliocene, sea-levels were distinctly higher than today, and multiple Miocene and Pliocene sea-level changes had amplitudes of *c.* 90 m (Woodruff, 2003, 2010).

Fragmentation and replacement of rainforests by seasonal or savannah vegetation caused by drier and cooler climates may have been less pronounced in Southeast Asia and may have followed different patterns than in other tropical areas during the Pleistocene (Cannon et al., 2009; Morley, 2000, 2007). However, Southeast Asian rainforest distributions did show major expansions and contractions based on pronounced cycles of submergence and exposure of land and shifts in the elevational zonation during the Quaternary and earlier. These fluctuations in combination with the rapid geological reconfiguration of the region likely had a profound impact on biotic distributions and diversification patterns (Cannon et al., 2009; Morley, 2000, 2007; Woodruff, 2010).

1.3 Aims of the doctoral research and structure of the thesis

The need for a robust phylogenetic framework of Southeast Asian *Begonia* which has the power to inform taxonomic monographs and to facilitate biogeographical and evolutionary studies was the motivation for this thesis. The first part of the thesis therefore aimed to determine suitable DNA markers for phylogenetic analyses of Southeast Asian *Begonia* species, to reconstruct the phylogeny of Southeast Asian *Begonia*, to determine whether major Asian sections are monophyletic, and to assess the degree of homoplasy of morphological characters which are crucial in current sectional delimitations. The first empirical research chapter of this thesis (Chapter 2) problematises current classifications of Asian *Begonia*, describes the methods used for DNA sequence generation, and presents the results of phylogenetic analyses and ancestral character state reconstructions based on this data.

The second part of this study aimed to contribute to a greater understanding of the historical biogeography of Southeast Asian *Begonia* in the context of global climate changes and the complex palaeogeography of Southeast Asia. The aims of this part of the research were to infer spatial and temporal patterns of *Begonia* diversification in Southeast Asia

and to detect potential geological or climatic correlates, as well as to investigate specific hypotheses about the origin and dispersal routes of Malesian *Begonia* lineages. The second empirical research chapter (Chapter 3) gives an introduction to research questions and hypotheses about the biogeography of Southeast Asian *Begonia*, describes the methods used to infer molecular divergence age estimates and to reconstruct ancestral areas of distribution, and presents biogeographical inferences derived from these analyses.

The final chapter (Chapter 4) summarises the main findings of the research described in the two empirical research chapters (Chapters 2 and 3). Research questions raised by the results are discussed and recommendations for further research strategies are made.

Finally, Appendix 1 presents three recently published papers, which represent alpha-taxonomic contributions to the knowledge of the Malesian *Begonia* flora. These papers provide detailed descriptions and IUCN conservation assessments of five *Begonia* species from the Indonesian island of Sulawesi. Some aspects of the character evolution in Sulawesi *Begonia* are discussed.

CHAPTER 2. Phylogenetics and character evolution of Southeast Asian *Begonia* L. (Begoniaceae)

Chapter Summary

The *Begonia* flora of Southeast Asia comprises more than 540 species. This exceptional species diversity in combination with a wide distribution of the genus in tropical rainforests offers the opportunity to address biogeographical questions and to investigate the processes which underlie modern patterns of biodiversity, but also poses major taxonomic challenges. Only few apomorphies characterising infrageneric taxa in this large genus have been identified and delimitation of Asian *Begonia* sections is highly problematic. A robust phylogenetic framework of Asian *Begonia* informing taxonomic monographs and facilitating biogeographical and evolutionary studies is currently lacking.

Maximum parsimony, maximum likelihood and Bayesian analyses of plastid (*ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer, 115 taxa) and nuclear ribosomal (ITS, 89 taxa) sequence data were used to reconstruct the phylogeny of Southeast Asian *Begonia* and to determine whether major Asian sections (sections *Coelocentrum*, *Diploclinium*, *Parvibegonia*, *Petermannia*, *Platycentrum*, *Reichenheimia*, *Sphenanthera*, *Symbegonia*) are monophyletic. Morphological characters which are crucial in current sectional circumscriptions were mapped on the phylogenetic trees to determine their degree of homoplasy and to assess their suitability in infrageneric classifications.

The results indicate that Asian and Socotran *Begonia* species form a well supported clade. The basal relationships within this clade involve two small subclades including five species placed in sections *Haagea*, *Reichenheimia* and *Peltaugustia*, whose relationships are unresolved or only poorly supported. The vast majority of Asian *Begonia* species fall into two major clades: Clade A and Clade B. Clade A includes species of section *Parvibegonia*, continental Asian species of section *Diploclinium*, and species in sections *Platycentrum* and *Sphenanthera*. Clade B includes species of section *Coelocentrum*, *Ridleyella*, *Bracteibegonia*, *Petermannia*, *Symbegonia* and Malesian species of sections *Diploclinium* and *Reichenheimia*. The cpDNA phylogenetic trees indicate that a base chromosme number of n = 15 is ancestral within Asian *Begonia*, and chromosome counts of 30 or 44, likely a triploid derivate from the base chromosome number of n = 15, are dominant within Clade B, while Clade A seems to be characterised by a primary base chromosome number of n = 11. Most major Asian sections are not supported as monophyletic and the strong systematic emphasis placed on single, homoplasious characters like undivided placenta lamellae (section *Reichenheimia*), fleshy pericarps (section *Sphenanthera*), and the recognition of sections primarily based on a plesiomorphic fruit syndrome and the

absence of characteristic features of other taxa (section *Diploclinium*) has resulted in the circumscription of several highly polyphyletic sections. Ovary and fruit characters have traditionally played a major role in sectional delimitation, however the high level of homoplasy associated with these has obscured systematic relationships in Asian *Begonia*. The presence or absence and type of stem metamorphoses and perennating organs like tubers and rhizomes is of more systematic importance than has been assumed in the past.

Several major clades are congruently supported in gene trees derived from the analyses of the plastid and nuclear ribosomal data, but there is hard incongruence within the clades comprising species of the species-rich sections *Platycentrum* s.l. (including section *Sphenanthera*) and *Petermannia* s.l. (including section *Symbegonia*), indicating that hybridization might have had a significant impact on the evolution of the genus, and highlighting the importance of using multiple independent sources of phylogenetic data to detect discrepancies between gene and species trees in *Begonia*.

2.1 Introduction

The pantropically distributed genus *Begonia* is with more than 1550 species among the ten largest genera of vascular plants (Frodin, 2004; Hughes, 2008). One hotspot of species diversity lies in Southeast Asia, whose *Begonia* flora comprises more than 540 species (Girmansyah, 2009; Girmansyah et al., 2009; Hughes and Coyle, 2009; Hughes et al., 2010; Hughes et al., 2009; Kiew and Sang, 2009; Thomas et al., 2009a; Thomas et al., 2009b; Thomas and Hughes, 2008). This exceptional species diversity in combination with a wide distribution of the genus in tropical rainforests offers opportunities to address biogeographical questions and to investigate phenomena such as rapid radiations, shifts in diversification rates, character evolution and the evolution of key innovations, but also poses major taxonomic challenges. Although Southeast Asian Begonias are morphologically diverse, especially with regards to their growth habits, stem metamorphoses, perennating organs, leaf shapes, inflorescence architectures and fruit types, only few apomorphies characterising infrageneric taxa have been identified and delimitation of Asian sections is often problematic (Doorenbos et al., 1998). A robust phylogenetic framework of Asian *Begonia* informing taxonomic monographs and facilitating biogeographical and evolutionary studies is currently lacking.

Doorenbos et al. (1998) revised sectional circumscriptions in *Begonia* recognizing 18 Asian sections, and another four Asian sections were subsequently proposed (Forrest and Hollingsworth, 2003; Ku et al., 2007; Ku, 1999; Shui et al., 2002). These 22 sections are highly unbalanced with regards to species numbers with the largest eight, sections *Petermannia*, *Platycentrum*, *Diploclinium*, *Reichenheimia* (Klotzsch) A.DC., *Coelocentrum* Irmsch., *Parvibegonia, Sphenanthera* and *Symbegonia*, accounting for more

than 95% of the ca. 750 species in Asia, while 14 sections, five of which are monotypic, have less than five species. Morphological and anatomical fruit and ovary characters have traditionally played an essential role as diagnostic characters and for infrageneric taxon delimitation in *Begonia* (Irmscher, 1925; Warburg, 1894), and modern classifications still strongly rely on these characters (Doorenbos et al., 1998). One example is the recent *Begonia* treatise in the *Flora of China*, in which a new monotypic *Begonia* section was proposed and sections present in China were re-circumscribed solely based on character combinations of carpel number, ovary locule number, placentation type, and the type of placenta divisions (Ku et al., 2007). This allowed clear sectional placement for all of the 173 described Chinese species. To a lesser extent, vegetative characters like the presence and type of perennating organs like tubers and rhizomes, and floral and inflorescence characters like the numbers of tepals, the number of style branches and the distribution of male and female flowers in the inflorescences have contributed to the delimitation of infrageneric taxa in *Begonia* (Doorenbos et al., 1998). However, recent molecular phylogenetic studies challenge the strong emphasis on few easily observable morphological and anatomical characters in infrageneric *Begonia* classifications, and most crucial characters used in sectional circumscription were identified as highly homoplasious in phylogenetic trees based on low-density, world-wide taxon sampling of the genus (Forrest, 2001; Forrest and Hollingsworth, 2003; Tebbitt et al., 2006). The results of a study by Tebbitt et al. (2006), who analysed sequence data of the nrDNA internal transcribed spacer region (ITS) of 46 Asian *Begonia* species, indicate that fruit syndromes associated with rain and animal dispersal evolved multiple times within Asian *Begonia*, and the sections *Sphenanthera*, *Platycentrum* and *Leprosae* were identified as polyphyletic. Moreover, Forrest and Hollingsworth (2003) showed, based on nrDNA ITS and 26S sequence data, that the New Guinean genus *Symbegonia*, which was traditionally separated from *Begonia* based on floral characters (a syntepalous perianth and a characteristic monadelphous androecium), is nested within *Begonia* section *Petermannia*. They proposed to recognize *Symbegonia* at sectional level rendering the large section *Petermannia* paraphyletic, but retaining a morphologically easily recognizable taxon (Forrest and Hollingsworth, 2003). Thus, the monophyly of some sections was tested and their phylogenetic positions were clarified in molecular phylogenetic studies including a wider sampling of Asian *Begonia* (Forrest, 2001; Forrest and Hollingsworth, 2003; Tebbitt et al., 2006). However, these analyses of nrDNA sequence data have largely failed to resolve deeper relationships within Asian *Begonia*. One reason for the limited utility of the ITS region for phylogenetic analyses of Asian *Begonia* is the extensive nucleotide and sequence length variation, which can be observed in ITS alignments (Forrest and Hollingsworth, 2003; Forrest et al., 2005; Tebbitt et al., 2006), e.g. the alignment of 52 *Begonia* accessions representing 46 Asian and two African species in the study by Tebbitt et al. (2006) exhibited *c.* 56% of variable sites and *c.* 39% of potentially parsimony informative sites. Despite high percentages of potentially parsimony informative characters, ITS phylogenetic trees of Asian *Begonia* show only poorly resolved and/or poorly supported backbones, which indicates that the analyses

are likely confounded by extensive nucleotide variation and associated high levels of alignment ambiguity and homoplasy. Because of limited taxon sampling and the poorly resolved phylogenetic trees of previous studies, intersectional relationships within Asian *Begonia* are still only very fragmentarily understood, and the extent to which parallel evolution of crucial characters has obscured systematic relationships in Asian *Begonia* requires further investigation.

The aims of this study are:

- to determine suitable markers for phylogenetic analyses of Southeast Asian *Begonia* species;
- to reconstruct the phylogeny of Southeast Asian *Begonia*;
- to determine whether major Asian sections (sections *Coelocentrum*, *Diploclinium, Parvibegonia*, *Petermannia, Platycentrum, Reichenheimia, Sphenanthera, Symbegonia*) are monophyletic;
- to map morphological characters which are crucial in current sectional delimitations on the phylogenetic trees to determine their degree of homoplasy and to assess their suitability in infrageneric classifications.

2.2 Material and Methods

2.2.1 DNA region sampling

The potential utility of ten coding and non-coding plastid regions (3'*trn*V-*ndh*C spacer, *ndh*A intron, *ndh*F-*rpl*32 spacer, *mat*K gene, *pet*D gene and intron, *psb*B gene, *psb*D-*trn*T spacer, *rpl*32-*trn*L spacer, *trn*L intron, *trn*Q-5'*rps*16 spacer) for phylogenetic reconstruction in Asian *Begonia* were assessed in a trial study. This trial compared the ease of amplification of the different regions and the variability of sequences of six Asian species (*Begonia malabarica* Lam., *B. kingiana* Irmsch., *B. morsei* Irmsch., *B. palmata* D.Don, *B. robusta* Blume, *B. symsanguinea* L.L.Forrest & Hollingsw.) representing a wide taxonomic and geographic range in Asia. Based on the results (Table 2.1), the *ndh*A intron, the *ndh*F-*rpl*32 spacer and the *rpl*32-*trn*L spacer, which are located in the small single copy unit of the cpDNA (Fig. 2.1 A) (Shaw et al., 2007), were selected for this study as they exhibited high percentages of variable and parsimony informative characters and complex potentially phylogenetically informative indel structures, while posing only few amplification and alignment difficulties.

In addition, sequences of the ITS region including the fast evolving internal transcribed spacers and the conservatively evolving 5.8S-rRNA coding gene of the 18S-5.8S-26S nuclear ribosomal DNA cistron (Fig. 2.1 B) (Álvarez and Wendel, 2003; Hughes et al., 2006; Small et al., 2004) were generated. Nuclear ribosomal genes exist in several hundred

to several thousand copies in each plant genome, and phenomena like incomplete concerted evolution and the evolution of pseudogenes can obscure phylogenetic relationships (reviewed in Álvarez and Wendel, 2003). Despite these drawbacks, ITS sequences and phylogenetic trees can provide useful insights when analysed carefully, and ITS has continued to be the most popular non-plastid region for species-level phylogenetic studies of plant groups, as the identification and development of phylogenetically useful low-copy nuclear markers in non-model organisms is often time and cost prohibitive (Feliner and Rossello, 2007). For this study, the ITS region was chosen, because it allowed the

Table 2.1. Six-taxon comparison of ITS and plastid region variability.

DNA Region	PCR success [# of 6 taxa]	Aligned positions [#]	Fragment length [bp]	Variable sites [# (%)]	Parsimony informative sites [# (%)]	SIC Indel codes [#]*	Amplification & sequencing protocols
ITS	6	698	610-655	178 (25.5)	55 (7.9)	39	Clement et al., 2004
*rpl*32-*trn*L spacer	6	1001	862-987	87 (8.7)	27 (2.7)	28	Shaw et al., 2007
*ndh*F-*rpl*32 spacer	6	1177	1011-1144	105 (8.9)	23 (2.0)	32	Shaw et al., 2007
*ndh*A intron	6	1223	1168-1192	73 (6.0)	16 (1.3)	21	Shaw et al., 2007
*pet*D gene and intron**	6	629	617-624	28 (4.5)	6 (1.0)	5	Copestake, 2005
*psb*B gene**	6	1320	1320	19 (1.4)	4 (0.3)	0	Copestake, 2005
*trn*L intron**	6	544	536-544	15 (2.8)	1 (0.2)	3	Copestake, 2005
*mat*K gene**	6	1195	1195	11 (0.9)	1 (0.1)	0	Copestake, 2005
*trn*Q-5'*rps*16 spacer	5	n/a	n/a	n/a	n/a	n/a	Shaw et al., 2007
*psb*D-*trn*T spacer	4	n/a	n/a	n/a	n/a	n/a	Shaw et al., 2007
3'*trn*V-*ndh*C spacer	0	n/a	n/a	n/a	n/a	n/a	Shaw et al., 2007

* Indels scores calculated with the simple indel coding method (SIC; Ochoterena and Simmons, 2000) as implemented in *SeqSate* (Müller, 2005)

** Sequence data from Goodall-Copestake, 2005

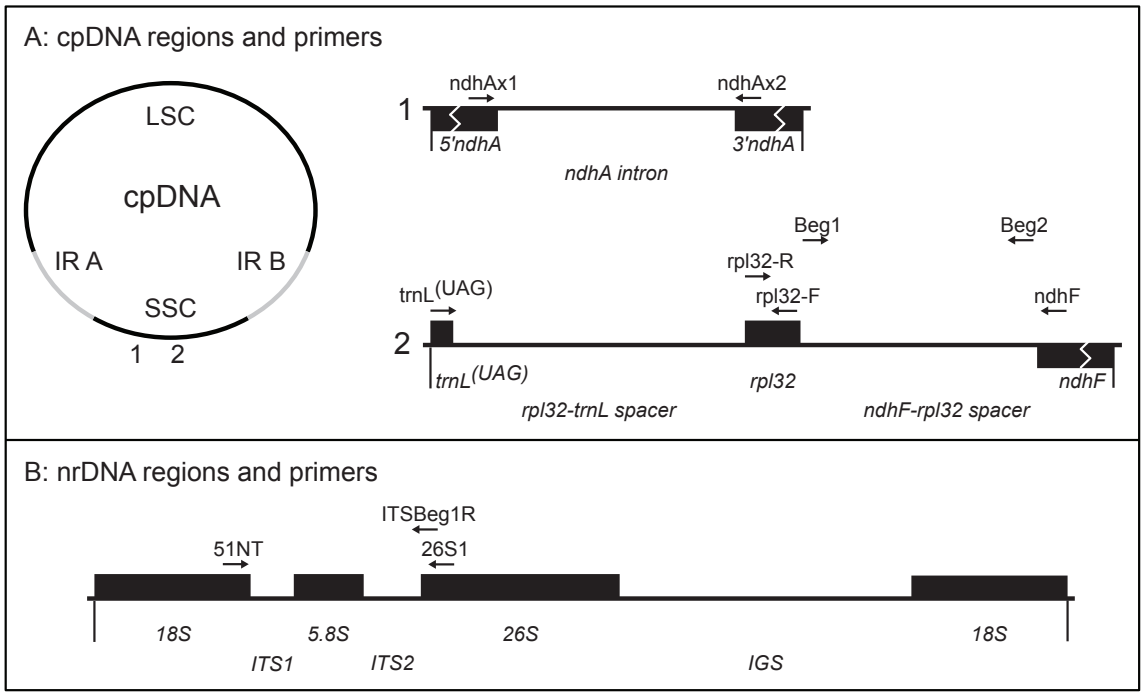


Fig. 2.1. cpDNA and nrDNA regions. **A**: Schematic diagram of the cpDNA and the nrDNA regions used in this study (adapted from Shaw et al., 2007). The relative position of used DNA regions is indicated by the numbers 1 and 2 (based on the position in the *Nicotiana* chloroplast genome). DNA region names are italicized below and primer names are in regular typeface above with directional arrows. IR A and B: inverted repeats; LSC: large single copy region; SSC: small single copy region. **B**: Schematic diagramm of a nrDNA repeat in angiosperms (adapted from Soltis and Soltis, 1998). DNA region names are italicized below and primer names are in regular typeface above with directional arrows. IGS: Non-coding intergenic spacer; ITS1 and 2: internal transcribed spacers; 5.8S, 18S, 26S: coding genes for rRNA.

inclusion of published ITS sequences of species of some monotypic or small Asian sections and mainland lineages of sections *Reichenheimia* and *Diploclinium*, for which no samples were available. Moreover, the ITS region is of nuclear origin and hence independent from the plastid DNA dataset. This allowed the assessment of conflict between the different gene trees, and assessment of potential disparity between gene and species trees, which would otherwise go undetected.

2.2.2 Taxon sampling

Non-coding cpDNA dataset: The cpDNA sequence dataset (*ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer) comprised 115 taxa sampled broadly from all major Asian *Begonia* sections (sections *Coelocentrum*, *Diploclinium, Parvibegonia*, *Petermannia, Platycentrum, Reichenheimia, Sphenanthera, Symbegonia*) with a focus on Malesian lineages. Moreover, samples of the small Asian sections *Bracteibegonia* A.DC., *Haagea* (Klotzsch) A.DC. and *Ridleyella* Irmsch., and samples of the only two *Begonia* species known from Socotra, *B. socotrana* Hook.f. and *B. samhaensis* M.Hughes & A.G.Mill., were included. Four African species were chosen to form the outgroup based on molecular phylogenetic studies by Goodall-Copestake (2005) and Plana et al. (2004) which indicate that *Begonia* initially diversified in Africa and that the Asian *Begonia* lineage is derived from African *Begonia*. Because of the unavailability of samples, several monotypic or small Asian sections including sections *Baryandra* A.DC. (one species), *Leprosae* (three species), *Monopteron* (A.DC.) Warb. (two species), *Pleiothece* T.C.Ku (one species), and *Putzeysia* (Klotzsch) A.DC. (one species) were only included in the ITS datasets using sequences provided by collaborators or downloaded from the nucleotide database of the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), and the sections *Alicida* C.B.Clarke (three species), *Apterobegonia* Warb. (one species), *Heeringia* Irmsch. (one species), *Lauchea* (Klotzsch) A.DC. (two species), *Monolobium* T.C.Ku (one species), and *Monophyllon* A.DC. (one species), were not included in the analyses. However, this likely had only a minor impact on the ability of this study to identify major clades and their relationships. Most of these sections, which together represent less than three percent of the species diversity of Asian *Begonia*, are defined by single or few apomorphic morphological characters and morphological observations indicate that they are likely to be closely related or nested within the major Asian sections (see discussion of phylogenetic relationships of sections *Parvibegonia*, *Diploclinium*, and *Platycentrum*). All 345 sequences of the cpDNA regions were newly generated for this study. Voucher information is listed in Appendix 2.

ITS dataset: The ITS dataset included sequences of 89 taxa. The ingroup comprised taxa of all major Asian sections and samples of the monotypic or small Asian sections *Baryandra*, *Bracteibegonia*, *Leprosae*, *Monopteron*, *Ridleyella* and *Putzeysia*. Two African taxa were chosen as an outgroup based on studies by Plana et al. (2004) and

Goodall-Copestake (2005). Twenty-four sequences were downloaded from the nucleotide database of the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), an unpublished sequence of *Begonia gemmipara* was provided by Sangeeta Rajbhandary (Tribhuvan University, Kathmandu, Nepal), and 64 sequences were newly generated. Voucher information and Genbank accession numbers are listed in Appendices 2 and 3, respectively.

Reduced cpDNA and ITS datasets: To test for congruence between cpDNA and nrDNA gene trees two reduced datasets, each comprising 64 taxa for which both cpDNA and rDNA data had been generated, were analysed. For ancestral character state reconstructions, six Bornean taxa in *Begonia* section *Petermannia*, which showed conflicting positions in analyses of the cpDNA and the nrDNA datasets (see results and discussion of gene tree incongruence) were excluded from the 115 taxa cpDNA dataset.

Table 2.2 gives an overview of all datasets and analyses.

2.2.3 DNA extraction, amplification and sequencing

Total genomic data was extracted from living material or silica gel dried material using the DNeasy Plant Mini Kit (Qiagen, UK) according to the manufacturer's protocols. For amplification of both the cpDNA regions and the ITS region, each 25 µl PCR reaction

Table 2.2. Overview of datasets and analyses. BI: Bayesian inference (*MrBayes*); BI-An: Ancestral character state reconstruction (*MrBayes, Mesquite*); BI-Ge: ITS and cpDNA gene tree comparison (*MrBayes*); ML: Maximum likelihood analysis (*RAxML*); ML-Ge: ITS and cpDNA gene tree comparison (*RAxML*); MP: Maximum parsimony analysis (*PAUP*).

Dataset	Data	Taxa [#]	Partitions in BI and ML analyses [# (Name)]	Analyses
Cp1A	cpDNA	115	0 (concatenated cpDNA)	BI, MP
Cp1B	cpDNA	115	3 (*ndh*A, *ndh*F-*rpl*32, *rpl*32-*trn*L)	BI, ML
Cp2A	cpDNA + indel codes	115	2 (concatenated cpDNA; indel codes)	BI, MP
Cp2B	cpDNA + indel codes	115	4 (*ndh*A, *ndh*F-*rpl*32, *rpl*32-*trn*L; indel codes)	BI
Cp3A	cpDNA	64	0 (concatenated cpDNA)	BI-Ge
Cp3B	cpDNA	64	3 (*ndh*A, *ndh*F-*rpl*32, *rpl*32-*trn*L)	BI-Ge, ML-Ge
Cp4A	cpDNA + indel codes	64	2 (concatenated cpDNA; indel codes)	BI-Ge
Cp4B	cpDNA + indel codes	64	4 (*ndh*A, *ndh*F-*rpl*32, *rpl*32-*trn*L; indel codes)	BI-Ge, ML-Ge
Cp5A	cpDNA + indel codes	109	2 (concatenated cpDNA; indel codes)	BI-An
Cp5B	cpDNA + indel codes	109	4 (*ndh*A, *ndh*F-*rpl*32, *rpl*32-*trn*L; indel codes)	BI-An
ITS1A	ITS	89	0 (concatenated ITS)	BI, MP
ITS1B	ITS	89	3 (ITS1, 5.8S, ITS2)	BI, ML
ITS2A	ITS + indel codes	89	2 (concatenated ITS; indel codes)	BI, MP
ITS2B	ITS + indel codes	89	4 (ITS1, 5.8S, ITS2; indel codes)	BI
ITS3A	ITS	64	0 (concatenated ITS)	BI-Ge
ITS3B	ITS	64	3 (ITS1, 5.8S, ITS2)	BI-Ge
ITS4A	ITS + indel codes	64	2 (concatenated ITS; indel codes)	BI-Ge
ITS4B	ITS + indel codes	64	4 (ITS1, 5.8S, ITS2; indel codes)	BI-Ge

contained 15.25 µl of ddH_2O, 2.5 µl of 10× reaction buffer, 1.25 µl of 25mM $MgCl_2$, 2.5 µl dNTPs (2mM), 0.75 µl of each forward and reverse primer (10µM), 0.8 µl bovine serum albumin (BSA, 0.4%), 0.2 µl of Biotaq DNA polymerase (Bioline, UK) and 1 µl of DNA template. Table 2.3 shows all primers used in this study. Amplification of the cpDNA *ndh*A intron, and the *ndh*F-*rpl*32 and *rpl*32-*trn*L spacers was carried out using primers designed by Shaw et al. (2007). The amplification of the *ndh*F-*rpl*32 spacer of some samples in *Begonia* section *Reichenheimia*, *Coelocentrum*, and *Petermannia* failed with these primers and required the design of specific internal primers, Beg1F and Beg2R (Fig. 2.1 A, Tab. 2.3). The PCR temperature profile used was the same as in Shaw et al. (2007): template denaturation at 80°C for 5 min followed by 30 cycles of denaturation at 95°C for 1 min, primer annealing at 50°C for 1 min, followed by a ramp of 0.3°C/s to 65°C, and primer extension at 65°C for 4 min; followed by a final extension step at 65°C for 5 min. Poly A/T homonucleotide strands composed of eight or more nucleotides, which are present in most samples, can cause PCR artefacts by slipped-strand mis-pairing (Shinde et al., 2003). To mitigate this problem an alternative amplification protocol was applied for some problematic samples using a *Pfu*-based DNA polymerase which is attached to nonspecific DNA binding proteins (Phusion by Finnzymes, Finland). Phusion polymerase was shown to reduce slipped-strand mis-pairing, at least for homonucleotide strands of up to 15 bp length, possibly because of increased contact surface between enzyme and DNA in comparison to *Taq* polymerases (Fazekas et al., 2010). Each 25 µl PCR reaction contained 13.0-14.0 µl of ddH_2O, 5.0 µl of 5× Phusion HF buffer, 2.5 µl dNTPs (2mM), 1.25 µl of each forward and reverse primer (10µM), 0.25 µl of Phusion DNA Polymerase, 1.0-2.0 µl of DNA template. The PCR temperature profile included template denaturation at 98°C for 30 s followed by 32 cycles of denaturation at 98°C for 10 s, primer annealing at 62°C (*ndh*F-*rpl*32) or 63°C (*ndh*A intron, *rpl*32-*trn*L) for 30 s, primer extension at 72°C for 30 s; followed by a final extension step at 72°C for 10 min. The ITS region was amplified using the primers 51NT and 26S1R (Clement et al., 2004). The amplification of several samples, which showed no or poor quantity amplifications products using these primers, required the design of an internal *Begonia* specific reverse primer, ITSBeg1R, which is located at the transition of the ITS2 spacer and the 26S rRNA gene and was used in combination with the 51NT primer (Fig. 2.1 B, Tab. 2.3). The PCR temperature profile used is the same as in Clement et al. (2004): denaturation at 94°C for 4 min followed by 35 cycles of denaturation at 94°C for 1 min, primer annealing at 60°C for 1 min, and primer extension at 72°C for 1 min; followed by a final extension step at 72°C for 4 min. Amplification products were visualized under UV light after electrophoretic separation on a 1% agarose gel stained with Sybr Safe (Invitrogen, USA). To remove superfluous dNTPs and primers, PCR products were subsequently purified using illustra GFX PCR purification spin columns (GE Healthcare, UK) or an enzymatic Shrimp Alkaline Phosphatase-based cleanup using ExoSAP-IT (Affymetrix, UK) according to the manufacturer's protocols. Sequencing PCRs were quarter reactions using the BigDye Terminator Cycle Sequencing ready Reaction Kit (Applied Biosystems, UK) and involved an initial denaturation at

Table 2.3. Primers used in this study.

DNA Region	Primer	Primer Sequence (5'–3')	Source
ITS	51NT	AGG TGA ACC TGC CGA AGG ATC ATT G	Clement et al., 2004
	26S1Rev	CGC CTG ACC TGG GGT CG	Kuzoff et al., 1998
	ITSBeg1R	GGG GTC GCT TYG AYA ACG	this study
3'*trn*V-*ndh*C	trnV$^{(UAC)}$x2	GTC TAC GGT TCG ART CCG TA	Shaw et al., 2007
	ndhC	TAT TAT TAG AAA TGY CCA RAA AAT ATC ATA TTC	Shaw et al., 2007
*ndh*A intron	ndhAx1	GCY CAA TCW ATT AGT TAT GAA ATA CC	Shaw et al., 2007
	ndhAx2	GGT TGA CGC CAM ARA TTC CA	Shaw et al., 2007
*ndh*F-*rpl*32	rpL32-R	CCA ATA TCC CTT YYT TTT CCA A	Shaw et al., 2007
	ndhF	GAA AGG TAT KAT CCA YGM ATA TT	Shaw et al., 2007
	Beg1F	TGG ATG TGA AAG ACA TAT TTT GCT	this study
	Beg2R	TTT GAA AAG GGT CAG TTA ATA ACA A	this study
*psb*D-*trn*T	psbD	CTC CGT ARC CAG TCA TCC ATA	Shaw et al., 2007
	trnT$^{(GGU)}$-R	CCC TTT TAA CTC AGT GGT AG	Shaw et al., 2007
*rpl*32-*trn*L	trnL$^{(UAG)}$	CTG CTT CCT AAG AGC AGC GT	Shaw et al., 2007
	rpL32-F	CAG TTC CAA AAA AAC GTA CTT C	Shaw et al., 2007
*trn*Q-5'*rps*16	trnQ$^{(UUG)}$	GCG TGG CCA AGY GGT AAG GC	Shaw et al., 2007
	rpS16x1	GTT GCT TTY TAC CAC ATC GTT T	Shaw et al., 2007

95°C for 1 min, followed by 25 cycles of denaturation at 96°C for 10 s, primer annealing at 60°C for 5 s, and primer extension at 60°C for 4 min. Sequencing PCR products were sent to the GenePool facilities at the University of Edinburgh (GenePool, UK), were purified using Shrimp Alkaline Phosphatase (Amersham, UK) and Exonuclease I (New England Biolabs, USA), and subsequently analysed on an AB 3730 DNA Analyser (Applied Biosystems, UK).

2.2.4 Alignment and gap coding

Sequences were assembled and edited using *GeneiousPro* v4.8.4 (Drummond et al., 2010). Plastid DNA sequences (*ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer) were aligned using the multiple sequence alignment software *MUSCLE* (Edgar, 2004) implemented in *GeneiousPro* using default settings, and subsequently manually modified in *GeneiousPro*. An inversion of 355 bp, or due to a deletion within this inversion of 309 bp, flanked on both sides by A/T homonucleotide strands, was identified in the *ndh*F-*rpl*32 spacer region of all Philippine samples of *Begonia* section *Diploclinium*. Sequences of the *rpl*32-*trn*L spacer of two Bornean species in *Begonia* section *Petermannia* showed single inversions of 27 bp and 37 bp, respectively, which are flanked by complementary regions indicating hairpin or stem folding secondary structures. These three inversions were reverse-complemented, thereby retaining substitution information in the fragments in the matrix (Borsch and Quandt, 2009; Graham et al., 2000; Löhne and Borsch, 2005). A fourth inversion, flanked by homonucleotide repeats of four G/Cs on each side, was identified in the *rpl*32-*trn*L spacer. This 11 bp inversion was present in six Bornean species in *Begonia* section *Petermannia* as well as in *Begonia masoniana* Irmsch. ex Ziesenh. (section *Coelocentrum*), and *B. roxburghii* A.DC. (section *Sphenanthera*). These species are only distantly related in phylogenetic trees resulting from the analysis of the cpDNA

nucleotide data, and the region of this homoplasious inversion was excluded from the analyses. Twenty-four mutational hotspots, most of which were length differences of T/A homonucleotide strands, together *c.* 5.8% of the aligned positions, were excluded from the final matrix because of uncertain homology (Table 2.4) (Borsch and Quandt, 2009; Kelchner, 2000, 2002). A 13 bp region of ambiguous alignment caused by poor sequence reads at the beginning of sequences of the *ndh*F-*rpl*32 intergenic spacer was removed. Phylogenetically informative indels were coded as presence/absence characters using the simple indel coding method (SIC) (Simmons and Ochoterena, 2000) implemented in *SeqState* (Müller, 2005).

The ITS dataset was partitioned into the 5.8S gene and the ITS1 and the ITS2, which were aligned separately. The 5.8S-rRNA gene showed low levels of sequence variation and was manually aligned without difficulty. Alignments of the variable ITS1 and ITS2 regions were performed using MAFFT v6.717 (Multiple Alignment using Fast Fourier Transform) (Katoh et al., 2009) applying the iterative refinement method (FFT-NS-i) and using default parameter settings (gap opening penalty: 1.53, offset-value: 0.0). The resulting alignments were investigated for the presence of the conserved domains C1-C6 and the variable domains V1-V6 (Hershkovitz and Zimmer, 1996) and other structural motifs (Jobes and Thien, 1997; Liu and Schardl, 1994) and manually modified guided by the position of conserved motifs and secondary structure predictions of ITS RNA transcripts (Forrest, 2001). A highly variable region of 1-16 bp in the ITS2, corresponding to a part of the variable domain V1 (Hershkovitz and Zimmer, 1996) was excluded from the analyses. Short regions directly adjacent to the primer sites at the end and beginning

Table 2.4. Excluded alignment positions and inversions. Alignment positions refer to the reference alignments in Appendix 4 (cpDNA data:115 taxa; ITS: 89 taxa).

Region	Aligned sites [#]	Excluded fragments [#]	Excluded aligned sites [#]	Excluded fragment (Position)	Excluded aligned sites [% aligned]	Inversions [#]	Inversion [bp (# taxa)]	Inversion (Position)
*ndh*A	1402	6	39	167-170, 305-307, 679-686, 715-725, 773-777, 1110-1117	2.8	0	n/a	n/a
*ndh*F	1210	8	62	1489-1494, 1940-1951, 2084-2089, 2158-2164, 2227-2232, 2347-2357, 2451-2454, 2556-2565	5.1	1	309 (2) 345 (3)	1501-1933 1501-1933
*rpl*32	1504	10	138	2720-2722, 2738-2742, 3133-3136, 3177-3209, 3249-3259, 3306, 3404-3410, 3784-3793, 3844-3898, 4102-4110	9.2	3	11 (8) 27 (1) 37 (1)	3628-3638, 3939-3975, 3906-3985
cpDNA combined	4116	24	239	See above	5.8	4	See above	See above
ITS	842	1	17	483-499	2.1	0	n/a	n/a

of the ITS were removed because of poor sequence reads. Phylogenetically informative indels were coded as presence/absence characters using the same approach as for the cpDNA data.

2.2.5 Phylogenetic analyses

The different datasets were analysed using three commonly applied methods for phylogenetic analyses: Bayesian inference (BI), maximum likelihood (ML) and maximum parsimony (MP). Knoop and Müller (2009) provide an overview of the relative strengths and weaknesses of the three methods and the following brief method comparison is primarily based on this overview.

Maximum parsimony analyses are based on the minimalistic principle of Ockham's razor, i.e. that the simplest, most parsimonious explanation of a problem should be preferred to more complex explanations. In phylogenetic applications of this principle, the tree or the trees which require the fewest character changes, provide the most parsimonious solution. Advocates of this method see the intuitive simplicity of this approach as a major strength in comparison to model-based methods which include numerous underlying assumptions. Moreover, software that apply the principle of parsimony for phylogenetic analyses, e.g. in *Paup** (Swofford, 2002), are computationally very efficient. However, the lack of many explicit assumptions does not automatically lead to a good performance of the method. Under certain conditions such as a high degree of sequence divergence and homoplasy, distinctly unequal frequencies of different kinds of nucleotide substitutions and considerably faster evolution of some sequences in an alignment in comparison to others, the use of the parsimony optimality criterion can produce misleading results. For example, MP analyses are particularly prone to long-branch attraction, i.e. the erroneous reconstruction of clades comprising taxa which show long branches in comparison to other taxa in the phylogenetic tree and a high degree of random congruence of homoplasious characters along these long branches (Bergsten, 2005; Felsenstein, 1978).

ML analyses search for the tree topology that maximizes the likelihood that the observed data have occurred under a given model of sequence evolution. In contrast to MP analyses, ML analyses can incorporate assumptions about sequence evolution in the form of nucleotide substitutions models, can account for among site rate heterogeneity, and do not ignore branch length in the evaluation of a tree, but account for the higher probability of nucleotide change along longer branches than along shorter branches. Therefore, variable but parsimony uninformative characters can be ML informative, and the method is less prone to long-branch attraction. However, if incorporated assumptions about sequence evolution strongly deviate from the real sequence evolution, the results will be flawed. Moreover, the ability to incorporate complex assumptions into ML analyses comes with a high computational cost, although the recent and ongoing development of new algorithms

implemented in software like *GARLI* (Zwickl, 2006) and *RAxML* (Stamatakis, 2006) allows the analysis of large datasets without losing much precision (Morrison, 2007).

BI also uses a likelihood function and an explicit model of sequence evolution, and, therefore, shares most strengths of the ML approach. BI uses a stochastic search algorithm, an explicit model of sequence evolution and the observed data to estimate the posterior probability distribution of trees (Huelsenbeck et al., 2002). This posterior probability distribution describes the probability of trees considering the probability of trees based on prior knowledge, which is usually very limited, the model of sequence evolution, and most importantly the data. Because of the properties of the stochastic search algorithm, BI is computationally efficient and can deal with complex, parameter-rich models. Moreover, BI is able to avoid getting stuck in local suboptimal solutions, and by sampling a set of plausible trees, BI directly produces estimates of the uncertainty of any branching event, while ML and MP analyses need an extra step, like bootstrapping, which is often computationally expensive, to estimate confidence intervals (Huelsenbeck et al., 2002). BI also allows the incorporation of prior information in the analyses, e.g. by specifying a prior probability distribution of trees. The specification of informative priors offers great flexibility, which can be seen as strength of the approach, but the inherent subjectivity can also be interpreted as weakness and the sensitivity of the results to priors has to be examined carefully (Huelsenbeck et al., 2002). Several specific methodological issues have been raised since the introduction of Bayesian methods in phylogenetics, e.g. with regards to the apparent overestimation of clade support in BI (Erixon et al., 2003; Simmons et al., 2004), the impact of supposedly uninformative priors on posterior clade probabilities (Pickett and Randle, 2005), and failure of analyses of complex, partitioned datasets (Brown et al., 2010; Marshall, 2010). Despite these drawbacks, BI has become one of the most commonly employed methods for phylogenetic reconstruction.

BI analyses were performed in *MrBayes* v3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Partitions were defined a priori based on spacer, intron and coding region identity. For the cpDNA dataset three partitions based on spacer and intron identity (*ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L spacer) were chosen, or alternatively, the three regions were concatenated and analysed as an unpartitioned matrix. For the ITS datasets three partitions were selected based on spacer and coding region identity (ITS1, ITS2, 5.8S gene), or alternatively, the three regions were concatenated and analysed as an unpartitioned matrix. One additional partition was selected for the binary indel code matrices. Table 2.2 gives an overview of the different partition strategies. Models of sequence evolution for each nucleotide sequence partition were determined using *jModelTest* (Posada, 2008). Maximum Likelihood topologies were used to estimate the optimal evolutionary model comparing 88 distinct models (11 substitution schemes, with equal or unequal base frequencies, a proportion of invariable sites, and rate variation among sites). Log-likelihoods of different models of substitution under ML

tree topologies were compared using the corrected version of the Akaike Information Criterion for small samples (AICc) as the model selection criterion (Posada and Buckley, 2004). The AICc converges towards the AIC, when larger sampling sizes are used, and should therefore always be used regardless of the sample size (Burnham and Anderson, 2004). Selected models which are not implemented in *MrBayes* were substituted by the closest overparameterized implemented model (Huelsenbeck and Rannala, 2004). For the indel code partition a simple F81-like binary model with rate variation among sites accommodated using four discrete gamma categories was selected as suggested by Ronquist et al. (2005). Overall performance of analyses of unpartitioned, concatenated nucleotide datasets and partitioned nucleotide datasets were assessed with comparison of the mean -ln*L* of all trees sampled from the posterior distribution at stationarity for each strategy, and with Bayes Factor comparison implemented in *Tracer* v1.5 (Rambaut and Drummond, 2009), which is based on smoothed estimates of marginal likelihoods (Newton and Raftery, 1994; Suchard et al., 2001). The criterion of 2ln Bayes Factor of ≥10 was used as a benchmark indicating very strong evidence in favour of one strategy over another (Kass and Raftery, 1995). For all datasets, four independent Metropolis-coupled MCMC analyses were run. Each search used three incrementally heated and one cold Markov chain, a temperature constant setting of 0.2, and was run for 2×10^7 generations and sampled every 1000 generations. If convergence diagnostics indicated poor convergence, mixing in these analyses was optimized by using eight incrementally heated and one cold Markov chains and applying different heating schemes lowering the temperature constant value from the default to minimally 0.05 (temperature constant settings of 0.2, 0.15, 0.1, 0.05), so that acceptance rates of attempted swaps between adjacent chains were in the range of *c.* 20-60%. Independent analyses of the partitioned datasets showed convergence of the topology and other parameters except for the rate multiplier and the gamma shape parameters for some partitions which showed poor convergence and different estimates including nonsensically high estimates of the rate multiplier. Analyses using low temperature constants (0.05) resulted in considerably diffuse posterior distributions of these parameters with samples repeatedly jumping between extremely small and extremely large values. This is consistent with observations that partitioned *MrBayes* analyses are prone to become trapped in regions of the parameter space characterized by distorted partition rate multipliers and unrealistically long trees, which has been demonstrated for both multiple published datasets and simulated data (Brown et al., 2010; Marshall, 2010). Following recommendations by Marshall (2010), the mean branch length prior was set from the default mean (0.1) to 0.01, which reduces the likelihood of stochastic entrapment in local tree length optima, and resulted in good convergence and realistic rate multiplier estimates. All parameters except topology and branch lengths were unlinked across partitions. Convergence was assessed by using the standard deviation of split frequencies as convergence index with values < 0.005 interpreted as indicating good convergence. *Tracer* v1.5 (Rambaut and Drummond, 2009) was used to determine whether the MCMC parameter samples were drawn from

a stationary, unimodal distribution, whether adequate effective sample sizes for each parameter (ESS > 200) were reached, and to compare mean -ln*L* of all trees sampled from the posterior distribution at stationarity for the different runs. Topological convergence of tree topologies resulting from different runs was visually checked using the online version of *AWTY* (Nylander et al., 2008). The initial 25% of samples of each Metropolis-coupled MCMC run were discarded as a conservative burnin, and the post burnin samples were summarized as 50% majority rule consensus phylograms with nodal support expressed as posterior probabilities.

Maximum likelihood analyses were performed using *RAxML* v7.0.4 (Stamatakis, 2006). Only nucleotide datasets were analysed. Chloroplast DNA datasets were divided into three partitions based on spacer and intron identity. The ITS datasets were split into three partitions based on spacer and coding region identity. One thousand inferences were run from distinct random stepwise addition sequence MP starting trees under a general time reversible nucleotide substitution model with among-site rate variation modelled with a gamma distribution. Subsequently, 1000 non-parametric bootstraps were performed under the partition data mode, and bootstrap support values were drawn on the best-scoring ML tree.

Maximum parsimony searches were performed using *PAUP** v4.0b10 (Swofford, 2002) with all characters treated as unordered, independent and of equal weight, and gaps treated as missing data. A two-stage heuristic search strategy was applied. In the first stage, heuristic tree searches were performed using the following specifications: 1000 replicates with random taxon sequence addition, tree bisection-reconnection branch-swapping (TBR), keeping multiple shortest trees found during branch swapping (MulTrees=on), saving no more than 10 trees per replicate to achieve reasonable computing times spent on branch swapping, and all other search settings at default values. The shortest trees found in the first stage were used as starting trees in the second heuristic search stage using the following specifications: tree bisection-reconnection branch-swapping (TBR), keeping multiple shortest trees (MulTrees=on), analysing all input trees in subsequent swapping rounds (steepest=yes), with a maximum of 10000 trees saved. Non-parametric bootstrapping was used to estimate confidence values (Felsenstein, 1985) with 10000 replicates of simple sequence addition, TBR branch-swapping, saving no more than 10 trees per replicate.

All BI and MP analyses were performed using the Titan computer cluster at the University of Oslo (Oslo Bioportal, www.bioportal.uio.no). ML analyses were performed using the CIPRES Cluster at the San Diego Supercomputer Center (http://phylo.bio.ku.edu:8080/portal2).

2.2.6 Ancestral character state reconstructions

2.2.6.1 Characters

Ancestral character states of five characters that have traditionally been used to define infrageneric taxa including the presence and type of perennating organs and stem metamorphoses, fruit types, locule numbers, placentation types, and placenta divisions were reconstructed.

Perennating organs and stem metamorphoses: The presence and type of specialised perennating organs and stem metamorphoses have traditionally provided important diagnostic characters for taxon circumscription in *Begonia* (Doorenbos et al., 1998), and the terms "tuberous" and "rhizomatous" define major *Begonia* groups in horticultural classifications (Tebbitt, 2005).

Rhizomes in *Begonia* can be defined as stem-homologous organs, which grow plagiotropically under- and/or above ground, exhibit compressed internodes, are thickened by primary growth, bear adventitious roots, and produce leaves and inflorescences or vertical leafy shoots. Rhizomes occur in the majority of major Asian sections of *Begonia*.

Numerous Asian species in sections *Diploclinium* and *Parvibegonia*, as well as species in the small Asian sections *Alicida*, *Lauchea*, *Heeringia*, *Monophyllon* and *Putzeysia* exhibit small but distinctly thickened storage organs directly below the substrate surface. These tubers seem to be stem-homologous (Badcock, 1998), although without investigations of the ontogeny, histology and vascularisation patterns it is difficult to assess whether tubers in all tuberous Asian species are homologous and do not have different histological origins (hypocotyl, epicotyl, root or a composite of different tissues). Tubers can be differentiated from rhizomes as they do not show a conspicuous plagiotropic length growth. However, tubers can occur in clusters, and sometimes short moniliform structures are developed, which differ from most rhizomes by the distinct constrictions between the tuberous units. For several species in sections *Diploclinium* and *Parvibegonia* tubers are essential adaptations to survive dry seasons during which the aboveground parts die down (Kiew, 2005; Phutthai et al., 2009). Some *Begonia* species in the African sections *Augustia* and *Peltaugustia*, like *B. dregei* Otto & Dietr., *B. socotrana* and *B. samhaensis*, exhibit thickened stem bases as perennating organs, which in the literature are variously referred to as swollen stem base, caudex, swollen rootstock or tuber (Hughes and Miller, 2002; Irmscher, 1961; Tebbitt, 2005; Warburg, 1894). The Indian *Begonia dipetala* Graham (section *Haagea*) exhibits fleshy stems and similar, though less distinct, thickened stem bases. In *Begonia socotrana* and *B. samhaensis* the aboveground surfaces of the thickened stem bases are covered by bulbils, which are strongly compressed, fleshy shoots bearing scaly bracts (Hughes and Miller, 2002). The reserves in the swollen stems and bulbils allow these species to re-sprout after prolonged dry seasons. Four character states are

differentiated for the ancestral character reconstructions: 1. No specialisation; 2. Rhizome; 3. Tuber; 4. Thickened stem base.

Fruit types: Several fruit syndromes have been described in Asian *Begonia* including dry capsules, rain-ballist capsules and indehiscent fleshy fruits (Fig. 2.2) (Doorenbos et al., 1998; Kiew, 2005; Tebbitt et al., 2006). Fruit types and dispersal mechanisms in Asian *Begonia* are described in the introduction (see 1.1.2.2) and are illustrated in Fig. 2.2. Three character states are differentiated for the ancestral character reconstructions: 1. Dry capsule; 2. Rain-ballist capsule; 3. Fleshy fruit.

Ovary locule number: The majority of Asian species have three-locular ovaries, but several fleshy-fruited species in section *Sphenanthera* exhibit four-locular fruits (Fig. 2.2 L), and sections *Parvibegonia*, *Platycentrum* and *Ridleyella* are characterized by two-locular ovaries (Fig. 2.2 D, J) (Doorenbos et al., 1998; Tebbitt et al., 2006). Section *Coelocentrum* is characterised by unilocular ovaries (Fig. 2.2. C, I). Four character states are differentiated for the ancestral character reconstructions based on locule numbers (one to four).

Placenta configuration: The placenta configuration is usually axillary in Asian *Begonia*, with the exception of section *Coelocentrum*, which exhibits a parietal placentation in the middle part of the ovary (Fig. 2.2 I). This character is discussed in more detail in the discussion of the phylogenetic relationships of section *Coelocentrum* (see 2.4.2.7).
Two character states are differentiated for the ancestral character reconstructions: 1. Axillary; 2. Parietal.

Placenta division: Types of placenta division have traditionally provided important characters in sectional delimitation, which is reflected in names of infrageneric *Begonia* taxa like section *Uniplacentales* (Clarke, 1879) or *Monolobium* (Ku et al., 2007). The majority of Asian species exhibit placentas which are bilamellate, i.e. each placenta is divided into two main lamellae (Fig. 2.2 H-L). These lamellae are usually distinct in cross-section, although they can exhibit further branching or folds which enlarge the placenta surface. Very rarely more than two main placenta lamellae are developed per locule, e.g. in *Begonia sizemoreae* Kiew (Kiew, 2004). Section *Reichenheimia* is characterized by undivided placentae, which is the only clear differential character against section *Diploclinium* (Doorenbos et al., 1998), and undivided placentae are also found in the small or monotypic Asian sections *Haagea* (Fig. 2.2. G) and *Ridleyella*, and in the Socotran section *Peltaugustia*. Three character states are differentiated for the ancestral character reconstructions: 1. Unilamellate; 2. Bilamellate; 3. More than two primary placenta lamellae.

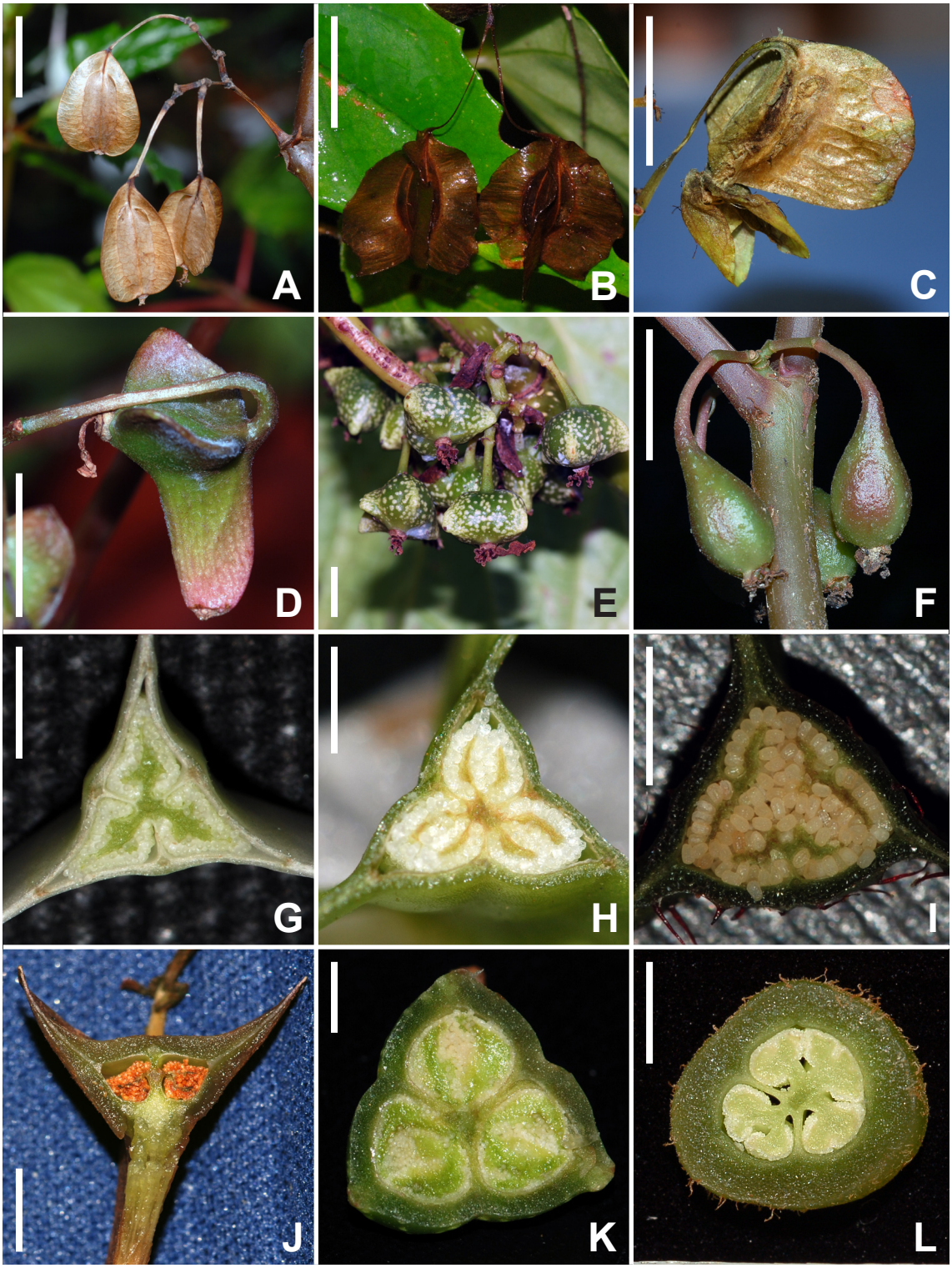


Fig. 2.2. Overview of fruit and ovary morphology and anatomy in Asian *Begonia*. A: *Begonia dipetala*: capsule, dry pericarp, equal wings. Scale bar = 16 mm. **B:** *Begonia varipeltata:* capsule, dry pericarp, equal wings. Scale bar = 12 mm. **C:** *Begonia masoniana*: capsule, dry pericarp, unequal wings. Scale bar = 10 mm. **D:** *Begonia pavonina*: rain-ballist capsule, coriaceous pericarp, unequal wings. Scale bar = 8 mm. **E:** *Begonia aptera*: berry, fleshy pericarp, wings reduced. Scale bar = 10 mm. **F:** *Begonia obovoidea*: berry, fleshy pericarp, wings absent. Scale bar = 15 mm. **G:** *Begonia dipetala*: ovary cross-section, three-locular ovary with axillary, undivided placentae. Scale bar = 3 mm. **H:** *Begonia varipeltata*: ovary cross-section, three-locular ovary with axillary, bilamellate placentae. Scale bar = 3 mm. **I:** *Begonia masoniana*: ovary cross-section, unilocular ovary with parietal, bilamellate placentae. Scale bar = 3 mm. **J:** *Begonia pavonina*: ovary cross-section, two-locular ovary with axillary, bilamellate placentae. Scale bar = 3 mm. **K:** *Begonia aptera*: ovary cross-section, three-locular ovary with axillary, bilamellate placentae. Scale bar = 3 mm. **L:** *Begonia obovoidea*: ovary cross-section, four-locular ovary with axillary, bilamellate placentae. Scale bar = 3 mm. Pictures of F and L taken by T. Phutthai.

2.2.6.2 Parsimony and likelihood ancestral character state reconstructions

Trees derived from the analyses of the plastid sequence data were chosen as input for the ancestral area reconstructions, as they provided much better resolved topologies than the trees derived from the analyses of the ITS sequence data (see results and discussion).

Ancestral character states were reconstructed using parsimony and likelihood methods implemented in *Mesquite* v2.7.2 (Maddison and Maddison, 2009). To account for phylogenetic uncertainty the "Trace over trees" option was selected, and 1000 randomly chosen trees from the stabilized part of the MC^3 of the Bayesian analysis of the 109 taxa dataset (combined cpDNA regions + indel codes, four data partitions) were included as input trees. Ancestral character reconstructions were mapped on the majority rule consensus tree obtained from the Bayesian analysis. Parsimony reconstruction searches for the ancestral character states which minimize the number of required steps of character change given a tree and observed extant character distributions at the terminals. Character-state changes were modelled as unordered for all characters. The "trace over trees" option summarizes for every node the percentage of the ancestral character state which was reconstructed for the same clade in the 1000 input trees. Likelihood reconstructions optimize the character states at each node which maximize the probability of arriving at the observed extant character states of the terminals, given a model of evolution. The Mk1 model (Markov k-state 1 parameter model) (Lewis, 2001) was selected. Under this model any particular change is equally probable, and the rate of change is the only parameter. The "trace over trees" option summarizes for every node the proportion of the average likelihood received by each character state as the ancestral character of a given clade.

2.3 Results

2.3.1 Comparison of DNA region variability

The six-taxon comparison of ITS and plastid region sequence variability showed that the three non-coding plastid regions used in this study (*ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer) exhibited distinctly higher percentages of variable and parsimony informative sites than several other chloroplast regions which have been used in molecular phylogenetic studies of *Begonia* (*mat*K gene, *pet*D gene and intron, *psb*B gene, *trn*L intron) (Table 2.1). The percentage of potentially parsimony informative sites for the *ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer datasets ranged from *c.* 1.3-2.7 percent, the range of potentially parsimony informative sites of the datasets of the other analysed plastid regions ranged from *c.* 0.1-0.95 percent, while the ITS dataset exhibited 7.9 percent of potentially parsimony informative sites. Parsimony informative indel scores were also much higher in the *ndh*A intron, and *ndh*F-*rpl*32 and *rpl*32-*trn*L spacer datasets (range 16-27) than in other analysed plastid region datasets (range 1-6), while the ITS dataset

exhibited the highest number of potentially parsimony informative indels (39).

2.3.2 Dataset descriptive statistics and model selection

Descriptive statistics for 115-taxon and 64-taxon cpDNA datasets and the 89-taxon and 64-taxon ITS datasets and their nucleotide partitions including the number of aligned positions, the length of the analysed fragments, the number and percentage of variable sites, the number and percentage of parsimony informative sites, and the number of indel codes are given in Table 2.5.

The number and percentage of variable and parsimony informative characters were calculated as crude indicators of variability and phylogenetic utility of the different analysed DNA regions. Of the three non-coding chloroplast regions used for this study, the *ndh*A intron was the most conservatively evolving region exhibiting distinctly lower percentages of variable and potentially parsimony informative sites than the *ndh*F-*rpl*32 and *rpl*32-*trn*L spacers.

Nucleotide model selection under the AIC and its corrected version for small sample sizes (AICc) did not differ for most partitions (Table 2.6), with the exception of the smallest and least variable partitions, the 5.8S coding regions of the ITS datasets. For the 5.8S

Table 2.5. Dataset descriptive statistics.

Dataset	Partition	Aligned positions [#]	Fragment length [bp]	Variable sites [# (%)]	Parsimony informative sites [# (%)]	SIC indel codes* [#]
cpDNA, 115 taxa	*ndh*A intron	1363	1077-1186	280 (20.5)	137 (10.1)	79
	*ndh*F-*rpl*32	1133	772-991	350 (30.9)	195 (17.2)	87
	*rpl*32-*trn*L	1353	620-1097	382 (28.2)	204 (15.1)	116
	Combined	3849	2745-3250	1012 (26.3)	536 (13.9)	282
cpDNA, 64 taxa	*ndh*A intron	1314	1090-1189	193 (14.7)	82 (6.2)	59
	*ndh*F-*rpl*32	1060	772-968	268 (25.3)	125 (11.8)	57
	*rpl*32-*trn*L	1246	620-1094	271 (21.8)	127 (10.2)	77
	Combined	3620	2745-3233	732 (20.2)	334 (9.2)	193
ITS, 89 taxa	ITS1	292	204-231	197 (67.5)	144 (49.3)	80
	ITS2	367	212-296	228 (62.1)	169 (46.1)	115
	5.8S	162	160-162	20 (12.4)	10 (6.2)	0
	Combined	821	587-686	445 (54.2)	323 (39.3)	195
ITS, 64 taxa	ITS1	273	211-241	177 (64.8)	120 (44.0)	55
	ITS2	342	220-290	208 (60.8)	137 (40.1)	84
	5.8S	162	160-162	16 (9.9)	6 (3.7)	0
	Combined	777	600-669	401 (51.6)	263 (33.9)	139

* Indel scores calculated with the simple indel coding method (SIC; Simmons and Ochoterena, 2000) implemented in *SeqState* (Müller, 2005)

partitions the relatively complex transitional model (TIM2ef+G; Posada, 2003) was chosen under the AIC, while the least parameter-rich model (JC model; Jukes and Cantor, 1969) was selected under the AICc.

2.3.3 Phylogenetic analyses

In the Bayesian analyses partitioning improved mean –ln*L* values considerably and the analyses using more complex partition strategies provided distinctly better explanations of the data than all other analyses according to Bayes Factor comparison (Fig. 2.3). The subsequent presentation of the results of the Bayesian analyses will be limited to the trees derived from the analyses which showed the best mean –*ln*L values and were chosen based on Bayes Factor comparison over the other analyses. For the results of the Bayesian analyses a 90% posterior probability (PP) lower threshold was considered to indicate moderate support, and a 95% lower threshold to indicate well supported relationships. For the results of the MP and ML analyses a 70% bootstrap support value lower threshold was considered to indicate moderate support, and an 85% lower threshold to indicate well supported relationships.

Table 2.6. Model selection using *jModelTest.* AIC: Akaike Information Criterion; AICc: Akaike Information Criterion corrected for small samples; n/a: not applicable. Models: JC (Jukes and Cantor, 1969), F81 (Felsenstein, 1981), TrN (Tamura and Nei, 1993), TPM (''3-parameter model'' = K81) (Kimura, 1981), TIM (''transitional model'') (Posada, 2003), TVM (''transversional model'') (Posada, 2003), SYM (Zharkikh, 1994), and GTR (Tavaré, 1986). +G: among-site rate variation modelled with a gamma distribution; +I: proportion of invariable sites; ef: equal base frequencies; uf: unequal base frequencies. For different types of the TIM and TPM models see Posada (2008).

Dataset	Partition	Aligned characters [#]	Model selected (AIC)	Model selected (AICc)	Model applied (*MrBayes*)	Model applied (*RAxML*)
cpDNA, 115 taxa	*ndh*A intron	1363	TVM+G	TVM+G	GTR+G	GTR+G
	*ndh*F-*rpl*32	1133	TVM+G	TVM+G	GTR+G	GTR+G
	*rpl*32-*trn*L	1353	TVM+G	TVM+G	GTR+G	GTR+G
	Combined	3849	TVM+G	TVM+G	GTR+G	n/a
	Indel codes	282	n/a	n/a	F81-like	n/a
cpDNA, 64 taxa	*ndh*A intron	1314	TVM+G	TVM+G	GTR+G	GTR+G
	*ndh*F-*rpl*32	1060	TVM+G	TVM+G	GTR+G	GTR+G
	*rpl*32-*trn*L	1246	TVM+G	TVM+G	GTR+G	GTR+G
	Combined	3620	TVM+G	TVM+G	GTR+G	n/a
	Indel codes	193	n/a	n/a	F81-like	n/a
ITS, 89 taxa	ITS1	292	TIM3ef+G	TrNef+G	SYM+G	GTR+G
	ITS2	367	TrN+G	TIM1ef+I+G	SYM+I+G	GTR+G
	5.8S	162	TIM2ef+G	JC	JC	GTR+G
	Combined	821	TIM3ef+G	TIM3ef+G	SYM+G	n/a
	Indel codes	195	n/a	n/a	F81-like	n/a
ITS, 64 taxa	ITS1	273	TIM3ef+G	TrNef+G	SYM+G	GTR+G
	ITS2	342	TPM2uf+G	TPM2uf+G	GTR+G	GTR+G
	5.8S	162	TIM2ef+G	JC	JC	GTR+G
	Combined	777	TIM3ef+G	TIM3ef+G	SYM+G	n/a
	Indel codes	139	n/a	n/a	F81-like	n/a

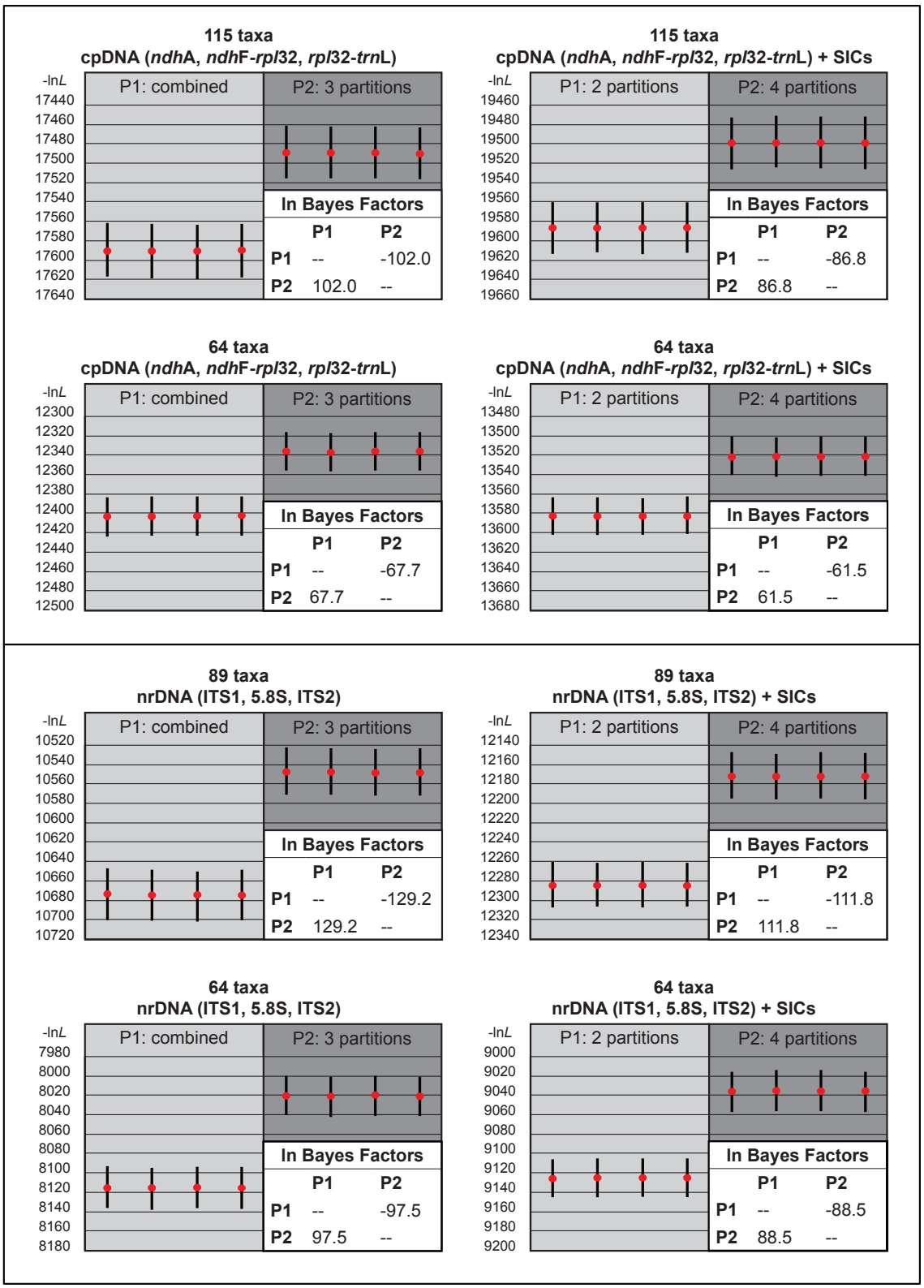


Fig. 2.3. Bayesian analysis and Bayes Factor comparison of different partition strategies. Charts show mean −ln*L* values of all trees sampled from the posterior distribution at stationarity of four independent runs for each partitioning strategy. The bars indicate the 95% interval, and the red dot indicates the mean of the distribution. Embedded Base Factor matrices show lnBase Factors calculated in *Tracer* v1.5 (Rambaut and Drummond, 2009). A positive value >5 indicates strong evidence against alternative hypotheses (partition strategies indicated in the first column are compared with partition strategies indicated in subsequent columns). SICs: Indels derived from simple indel coding (Simmons & Ochoterena, 2000).

Sectional placement of taxa shown on Figs 2.4-11 primarily follows Doorenbos et al. (1998) or more recent sectional placement corrections and new species descriptions (Hughes, 2008; Kiew, 2001; Kiew, 2005; Tebbitt and Dickson, 2000). Unidentified species are placed to section based on sectional circumscriptions in Doorenbos et al. (1998).

2.3.3.1 Non-coding cpDNA phylogenetic trees

Trees derived from the analyses of non-coding cpDNA sequences of 115 taxa are presented in Figs. 2.4-6. Figure 2.4 shows a phylogram based on the majority rule consensus trees of the Bayesian analysis, Fig. 2.5 presents a phylogram based on the best-scoring ML tree, and Fig. 2.6 shows the strict consensus tree from the MP analysis.

Asian and Socotran taxa form a strongly supported clade. The relationships of two clades which diverge at two of the deepest nodes within the Asian-Socotran crown group are only poorly supported. The first of these two clades is well supported and composed of *Begonia floccifera* Bedd. (section *Reichenheimia*) and *B. malabarica* (unplaced to section). The second includes *Begonia dipetala* (section *Haagea*) as sister to a strongly supported clade including two species of section *Peltaugustia*. Apart from these two lineages, two main clades can be differentiated: Clade A and Clade B.

Clade A is strongly supported in the Bayesian analyses, moderately supported in the ML analysis, but receives only weak support in the MP analyses. It comprises species of sections *Parvibegonia*, *Diploclinium*, *Platycentrum* and *Sphenanthera*. Within Clade A, species of section *Parvibegonia* form a well supported sister clade to the rest of the clade whose topology exhibits a grade of species of section *Diploclinium*, and a well supported clade comprising species of sections *Platycentrum* and *Sphenanthera*. Within the *Platycentrum-Sphenanthera* clade species of both sections form intermixed assemblages.

Clade B is well supported in all analyses. It comprises species of section *Coelocentrum*, as well as clades consisting of species of sections *Diploclinium*, *Reichenheimia*, *Ridleyella*, *Bracteibegonia*, *Petermannia* and *Symbegonia*. Section *Coelocentrum* is sister to the moderately to strongly supported rest of the clade, which comprises four main subclades: 1. A strongly supported clade comprising species of section *Diploclinium*; 2. A strongly supported clade comprising species of section *Reichenheimia*; 3. A strongly supported clade of six species of section *Petermannia*; 4. A strongly supported clade which comprises a clade of species of section *Bracteibegonia* as sister to a well supported section *Petermannia* clade, in a subclade of which the New Guinean section *Symbegonia* is nested. The relationships among these four main clades and the monotypic section *Ridleyella* are unresolved or only poorly supported.

Most major Asian sections are not supported as monophyletic (Fig. 2.7). Section *Diploclinium* is polyphyletic with Asian mainland species found in Clade A, while

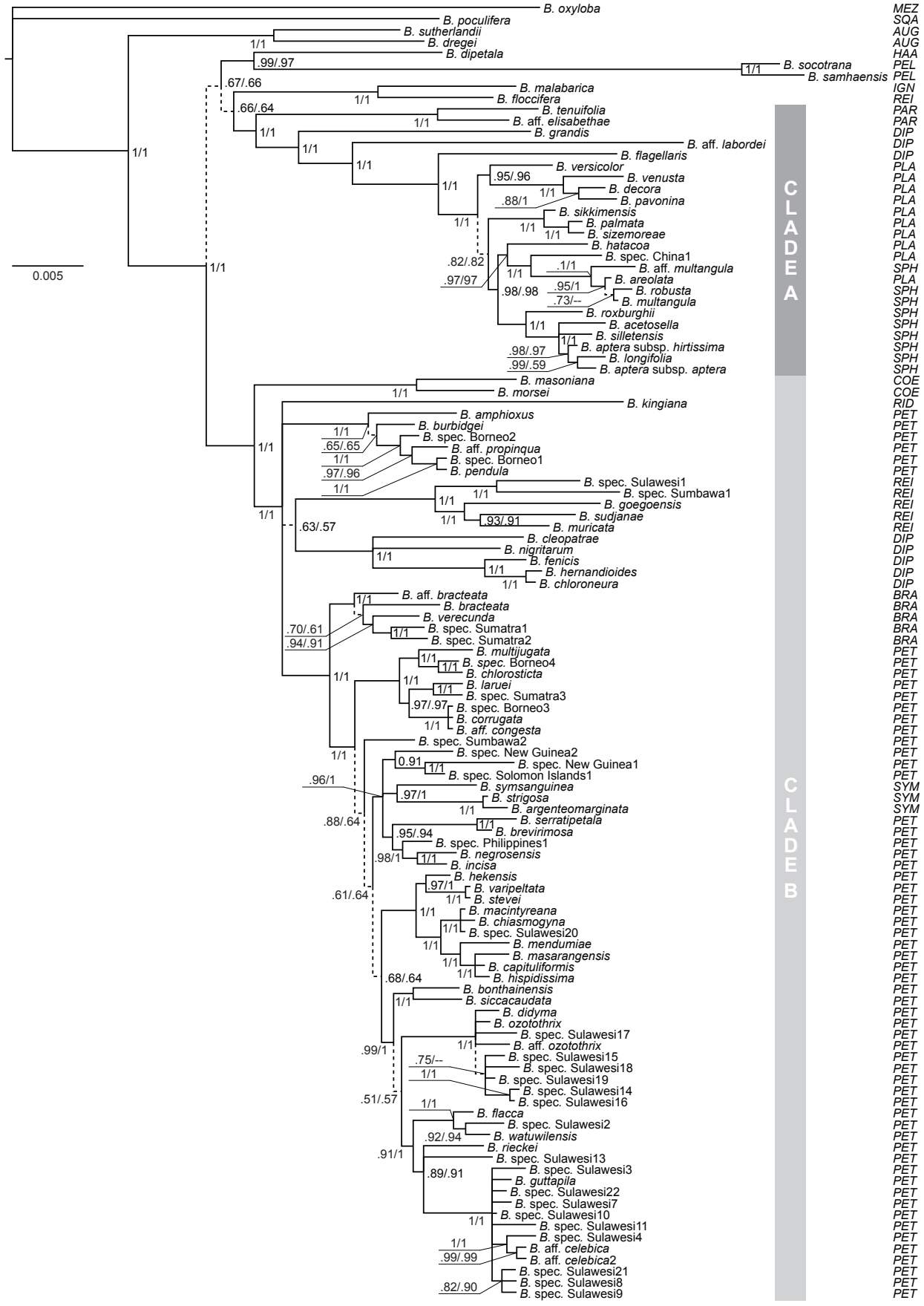


Fig. 2.4. Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L; 3 data partitions; 115 taxa). Bayesian posterior probability (PP) support values > 0.5 are indicated next to the nodes, and PPs of corresponding clades of an analysis additionally including 282 indel codes are mapped on the tree: PP (analysis without indel codes)/PP (analysis with additional indel code partition). Broken lines indicate branches which lead to nodes with a PP < 0.9. The scale bar indicates substitutions per site. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, HAA: *Haagea*, MEZ: *Mezierea*, PAR: *Parvibegonia*, PEL: *Peltaugustia*, PET: *Petermannia*, PLA: *Platycentrum*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQA: *Squamibegonia*, SYM: *Symbegonia.*

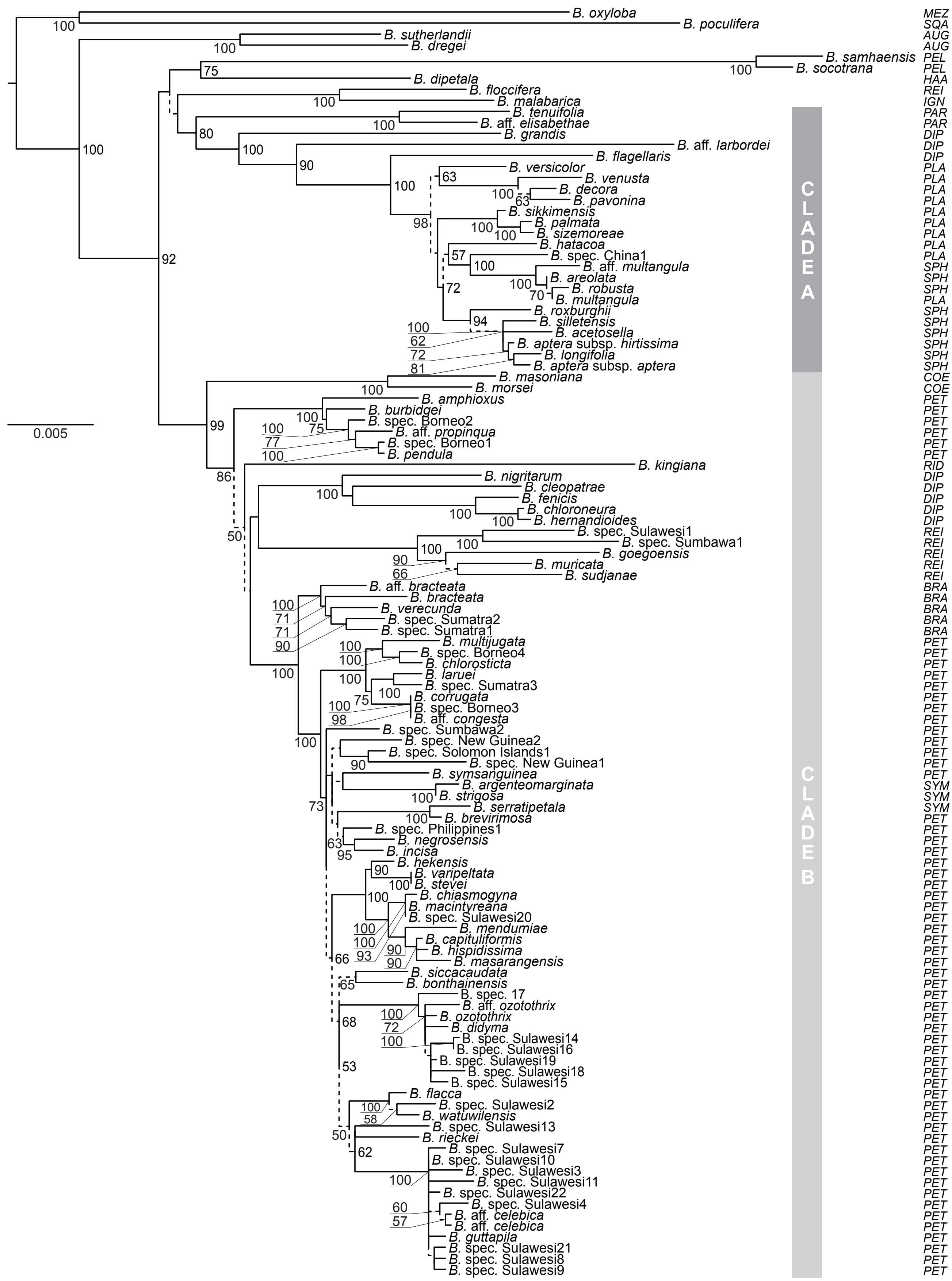


Fig. 2.5 Best scoring maximimum likelihood phylogram (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L; 3 data partitions; 115 taxa). Bootstrap support (BS) values > 50 are indicated next to the nodes. Broken lines indicate branches which lead to nodes with a BS < 70. The scale bar indicates substitutions per site. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, HAA: *Haagea*, MEZ: *Mezierea*, PAR: *Parvibegonia*, PEL: *Peltaugustia*, PET: *Petermannia*, PLA: *Platycentrum*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQA: *Squamibegonia*, SYM: *Symbegonia.*

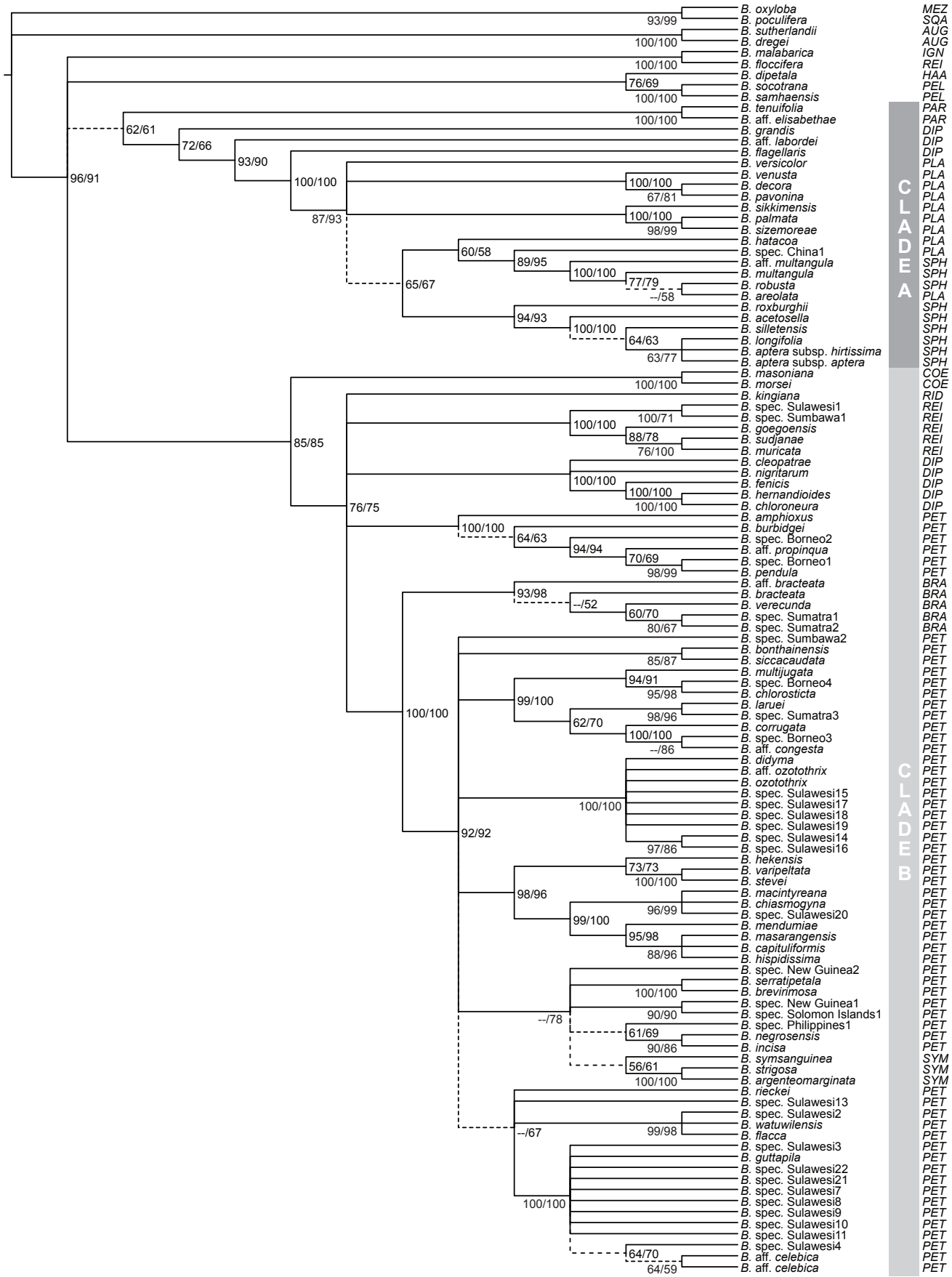


Fig. 2.6. Maximum parsimony strict consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L; 3 data partitions; 115 taxa). Bootstrap support values (BS) > 50 are indicated next to the nodes, and bootstrap support of corresponding clades of an analysis additionally including 282 indel codes are mapped on the tree: BS (analysis without indel codes)/BS (analysis including indel codes). Broken lines indicate branches which lead to nodes with a BS < 70. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, HAA: *Haagea*, MEZ: *Mezierea*, PAR: *Parvibegonia*, PEL: *Peltaugustia*, PET: *Petermannia*, PLA: *Platycentrum*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQA: *Squamibegonia*, SYM: *Symbegonia.*

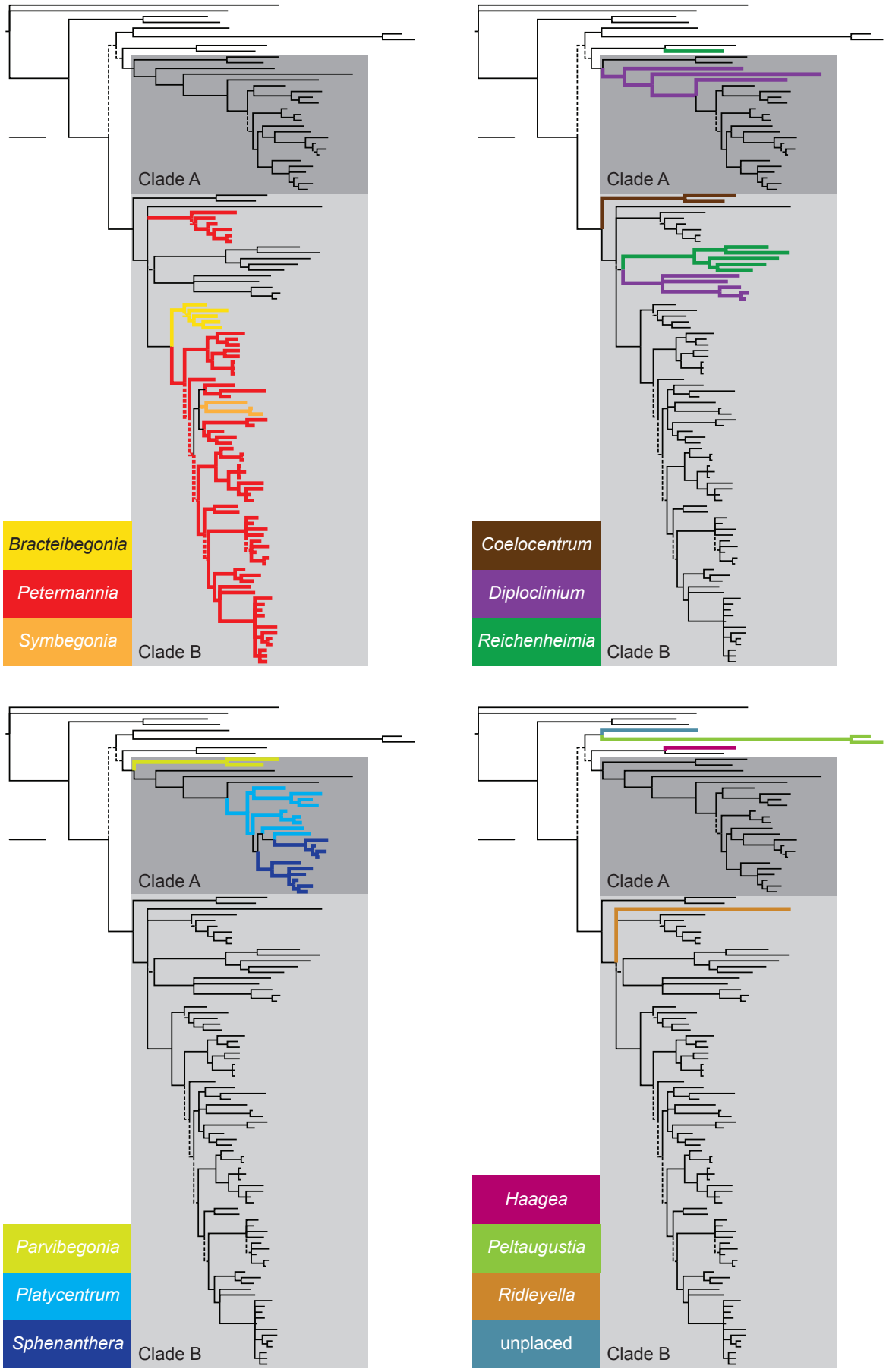


Fig. 2.7. Overview of monophyly, paraphyly and polyphyly of Asian *Begonia* sections based on cpDNA data. Tree topologies are based on the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L; 3 data partitions; 115 taxa) (see Fig. 2.4). Broken lines indicate branches which lead to nodes with a PP < 0.9. The scale bar indicates 0.005 substitutions/site.

Philippine species fall into Clade B. Malesian taxa of section *Reichenheimia* fall in a strongly supported clade in Clade B, to which the Indian species *Begonia floccifera*, which has also been placed in section *Reichenheimia*, is only distantly related. Sections *Platycentrum* and *Sphenanthera* are also not monophyletic and form interdigitated assemblages in Clade A. Species of section *Petermannia* are found in two apparently only distantly related subclades in Clade B, and species in section *Symbegonia* are nested within a subclade of the larger *Petermannia* clade.

2.3.3.2 ITS phylogenetic trees

Trees derived from the analyses of ITS sequences of 89 taxa are presented in Figs. 2.8-10. Figure 2.8 shows a phylogram based on the majority rule consensus trees of the Bayesian analyses, Fig. 2.9 present a phylogram showing the best-scoring ML tree, and Fig. 2.10 shows a strict consensus tree from the MP analyses.

Asian taxa form a strongly supported clade in the Bayesian analyses, but receive only weak support in the MP and ML analyses. The backbone of the Asian clade is largely unresolved or only poorly supported, and the relationships of several species including *Begonia boisiana* Gagnep. and *B. malabarica*, both of which are unplaced to section, *B. kingiana* (section *Ridleyella*), two species placed in section *Reichenheimia*, and *B. grandis* (section *Diploclinium*), are unresolved or get only poor support. Apart from these, several well supported clades can be differentiated: Clades A-F. All sampled species of section *Parvibegonia* constitute a strongly supported clade (Clade A). Four species of section *Reichenheimia* compose a strongly supported clade (Clade B), but two species assigned to this section are not included within this clade and their relationships are only poorly supported. A well supported clade contains all sampled species of sections *Platycentrum* and *Sphenanthera*, but also *Begonia nepalensis* Warb. (section *Monopteron*), *B. longicarpa* K.Y.Guan & D.K.Tian (section *Leprosae*), *B. balansana* Gagnep. (section *Pleiothece*) as well as two species placed in section *Diploclinium* (Clade C). Species of section *Coelocentrum* fall into a strongly supported clade with *Begonia cavaleriei* H.Lév. (section *Diploclinium*) and *B. leprosa* Hance (section *Leprosae*) (Clade D). The relationships among these species are largely unresolved. Four species of section *Diploclinum* form a clade together with *Begonia oxysperma*, which constitutes the monotypic section *Baryandra* (Clade E). Other species of section *Diploclinium* are found in several other well supported clades or show unresolved or only poorly supported relationships. All samples species of sections *Bracteibegonia*, *Petermannia*, and *Symbegonia* fall into a strongly supported clade (Clade F). Within this clades section *Bracteibegonia* is well supported as monophyletic, and section *Symbegonia*is weakly to moderately supported as monophyletic, and nested within a clade also including species assigned to section *Petermannia*.

Most Asian sections are not supported as monophyletic (Fig. 2.11). Section *Diploclinum*

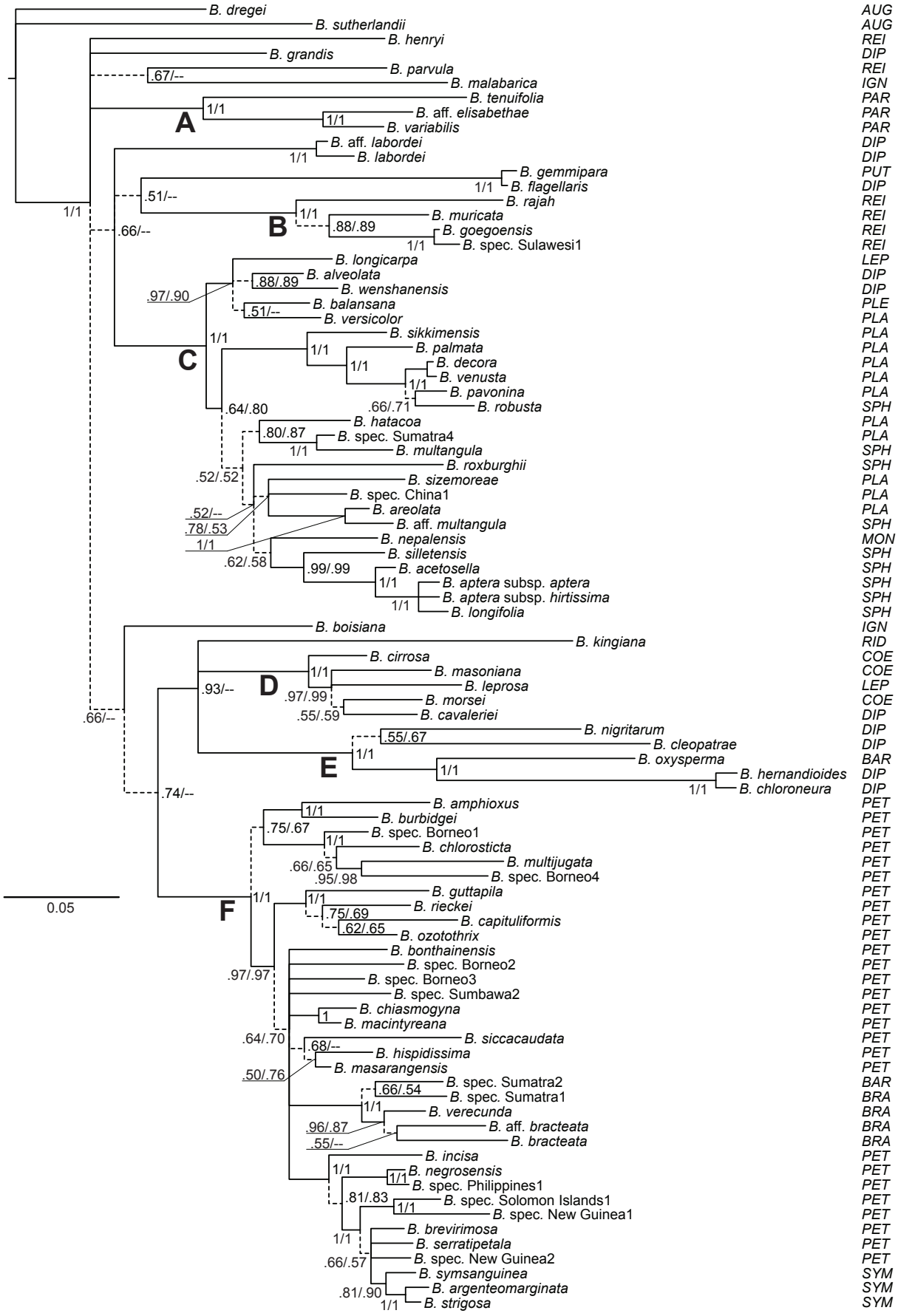


Fig. 2.8. Bayesian majority rule consensus tree (ITS data; 3 data partitions; 89 taxa). Bayesian posterior probability (PP) support values > 0.5 are indicated next to the nodes, and PPs of corresponding clades of an analysis additionally including 195 indel codes are mapped on the tree: PP (analysis without indel codes)/PP (analysis with additional indel code partition). Broken lines indicate branches which lead to nodes with a PP < 0.9. The scale bar indicates substitutions per site. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BAR: *Baryandra*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, LEP: *Leprosae*, MEZ: *Mezierea*, PAR: *Parvibegonia*, PET: *Petermannia*, PLA: *Platycentrum*, PLE: *Pleiothece*, PUT: *Putzeysia*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQA: *Squamibegonia*, SYM: *Symbegonia.*

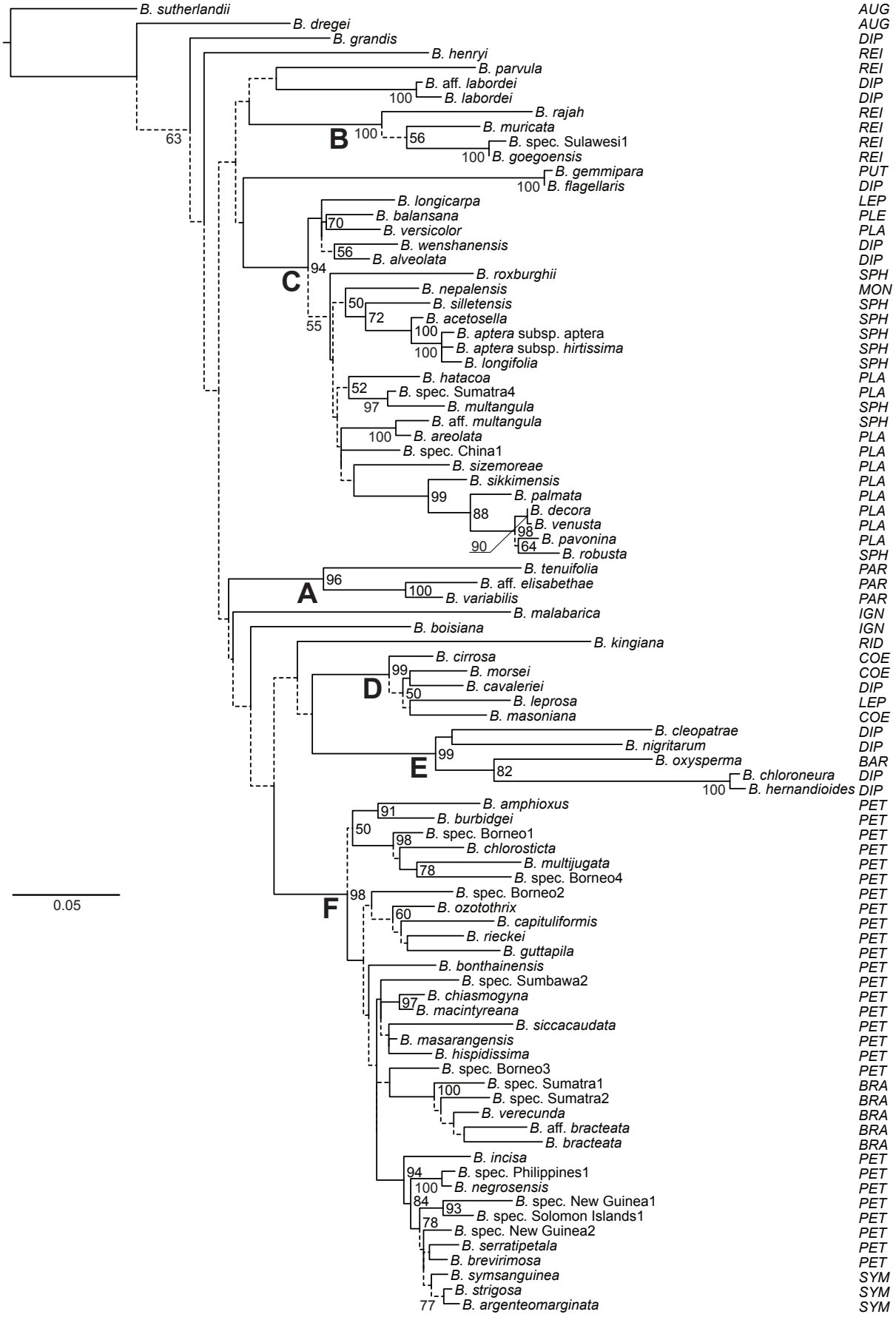


Fig. 2.9 Best scoring maximimum likelihood phylogram (ITS data; 3 data partitions; 89 taxa). Bootstrap support (BS) values > 50 are indicated next to the nodes. Broken lines indicate branches which lead to nodes with a BS < 70. The scale bar indicates substitutions per site. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BAR: *Baryandra*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, LEP: *Leprosae*, MEZ: *Mezierea*, PAR: *Parvibegonia*, PET: *Petermannia*, PLA: *Platycentrum*, PLE: *Pleiothece*, PUT: *Putzeysia*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQA: *Squamibegonia*, SYM: *Symbegonia.*

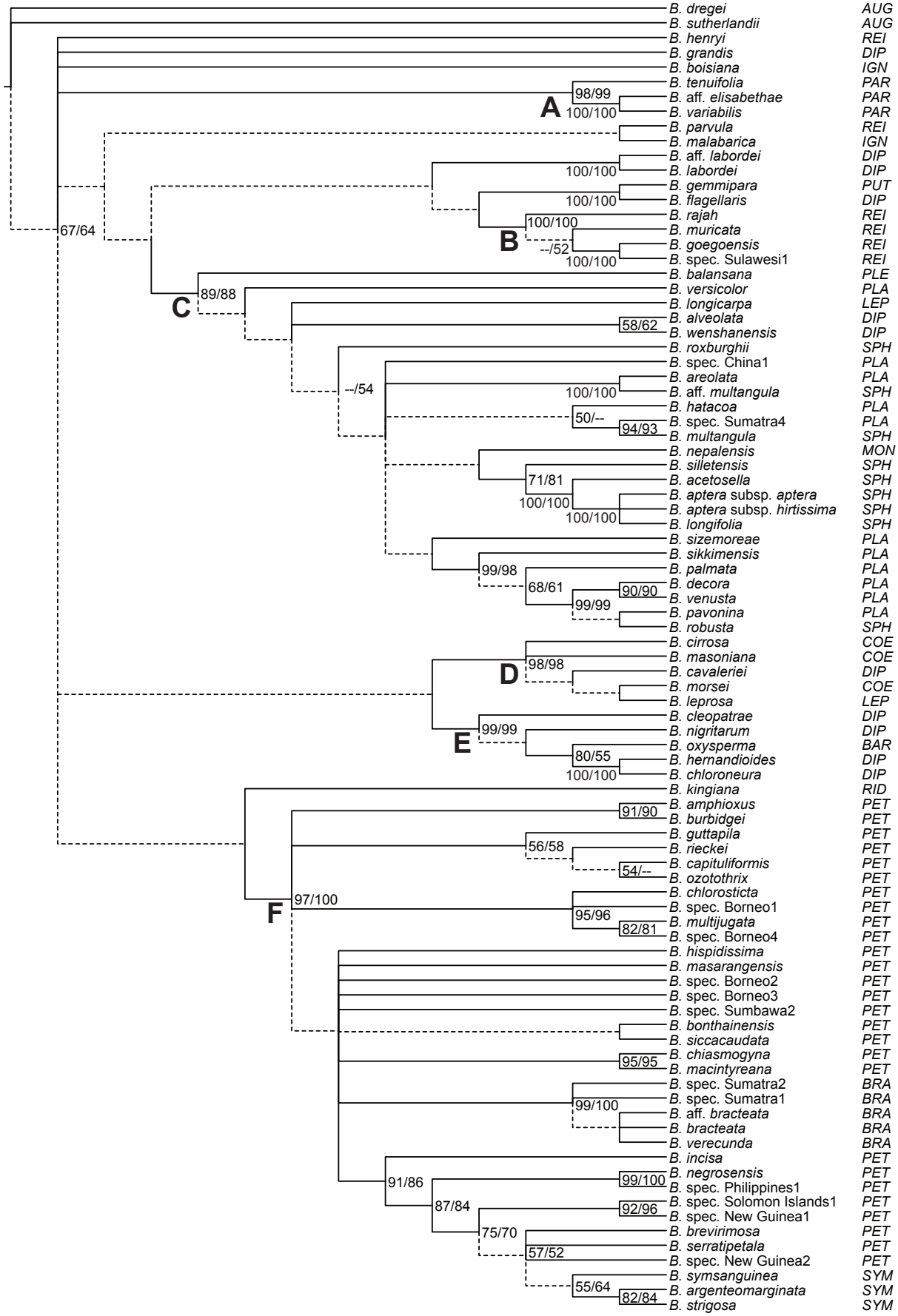


Fig. 2.10. Maximum parsimony strict consensus tree (ITS data; 3 data partitions; 89 taxa). Bootstrap support values (BS) > 50 are indicated next to the nodes, and bootstrap support of corresponding clades of an analysis additionally including 282 indel codes are mapped on the tree: BS (analysis without indel codes)/BS (analysis including indel codes). Broken lines indicate branches which lead to nodes with a BS < 70. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BAR: *Baryandra*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, LEP: *Leprosae*, MEZ: *Mezierea*, PAR: *Parvibegonia*, PET: *Petermannia*, PLA: *Platycentrum*, PLE: *Pleiothece*, PUT: *Putzeysia*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQA: *Squamibegonia*, SYM: *Symbegonia.*

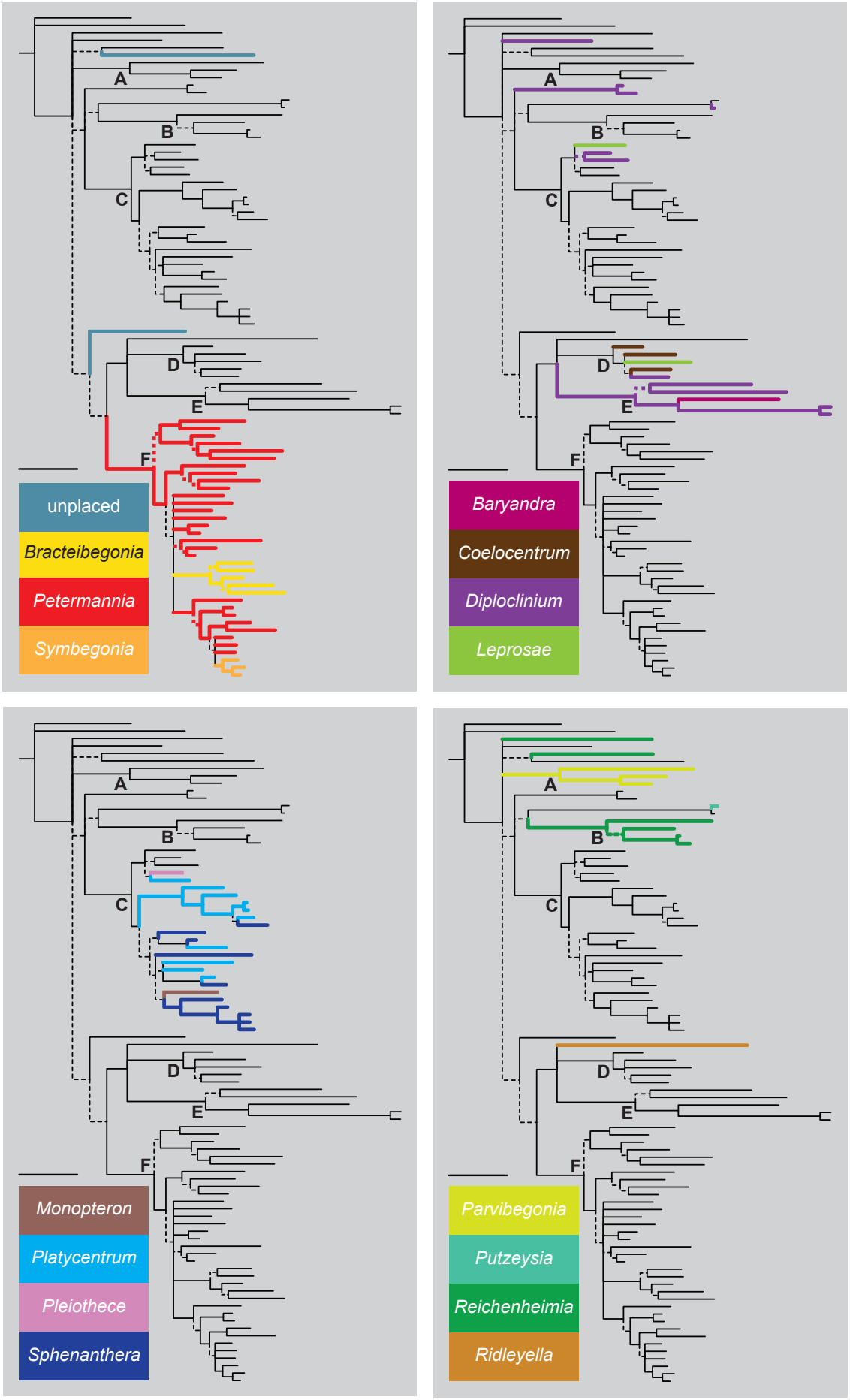


Fig. 2.11. Overview of monophyly, paraphyly and polyphyly of Asian *Begonia* sections based on ITS data. Tree topologies are based on the Bayesian majority rule consensus tree (ITS data; 3 data partitions; 89 taxa) (see Fig. 2.8). Broken lines indicate branches which lead to nodes with a PP < 0.9. The scale bar indicates 0.05 substitutions/site.

is highly polyphyletic and present in several major clades (Clades C, D, E) as well as in unresolved or poorly supported positions. The monotypic section *Baryandra* is nested within Philippine *Diploclinium* (Clade E). Section *Leprosae* is polyphyletic with *Begonia leprosa* closely related to section *Coelocentrum* (Clade D), while *B. longicarpa* is more closely related to section *Platycentrum* and *Sphenanthera* (Clade C). Samples from the Asian mainland of section *Reichenheimia* are apparently only distantly related to a well supported Malesian *Reichenheimia* clade (Clade B). Section *Symbegonia* is nested within section *Petermannia* (Clade F). Species of section *Platycentrum* and *Sphenanthera* form interdigitated species assemblages, and *Begonia nepalensis* (section *Monopteron*) is nested within this *Platycentrum-Sphenanthera* clade (Clade C).

2.3.3.3 Comparison of ITS and cpDNA phylogenetic trees

Two phylograms based on the majority rule consensus trees of the Bayesian analyses of the 64 taxa ITS and cpDNA datasets are presented in Fig. 2.12. Support values of the ML analyses of these datasets are mapped on the trees.

Direct comparison of the two gene tree topologies is problematic as the backbone of the ITS phylogenetic tree is only poorly supported. However, several major clades are congruently strongly supported in both the cpDNA and in the nrDNA gene trees: Clades I-VI. Clade I is composed of two species of section *Parvibegonia*. Clade II includes all sampled species of sections *Platycentrum* and *Sphenanthera*. Clade III includes species of section *Coelocentrum*. Clade IV includes three species of section *Diploclinium*. Clade V comprises species in section *Bracteibegonia*. Clade VI comprises nine species of sections *Petermannia* and *Symbegonia*.

However, the two gene trees exhibit hard incongruence, i.e. well supported, conflicting positions of several taxa. Multiple conflicting positions can be found within the *Playcentrum-Sphenanthera* clade (Clade II). In the cpDNA phylogenetic tree *Begonia robusta* is part of a strongly supported subclade of clade II also comprising *B. multangula* Blume, *B.* aff. *multangula* and *B. areolata* Miq. In the ITS phylogenetic tree *Begonia robusta* falls into a well supported clade with *Begonia pavonina* Ridl., *B. venusta* King, and *B. decora* Stapf. In addition, *Begonia palmata* is strongly supported as sister to *B. sizemoreae* in the cpDNA phylogenetic tree, but is the sister to the clade comprising *Begonia robusta*, *Begonia pavonina*, *B. venusta*, and *B. decora* in the ITS phylogenetic tree. Several taxa placed in section *Petermannia* show conflicting positions in the cpDNA and ITS phylogenetic trees. Within the cpDNA phylogenetic tree *Begonia amphioxus* Sands*, B. burbidgei* Stapf and two unidentified species form a well supported clade, which is not included in a strongly supported clade comprising the majority of species of section *Petermannia* as well as sections *Bracteibegonia* and *Symbegonia*. In the ITS phylogenetic tree these four species are nested within a moderately to strongly supported

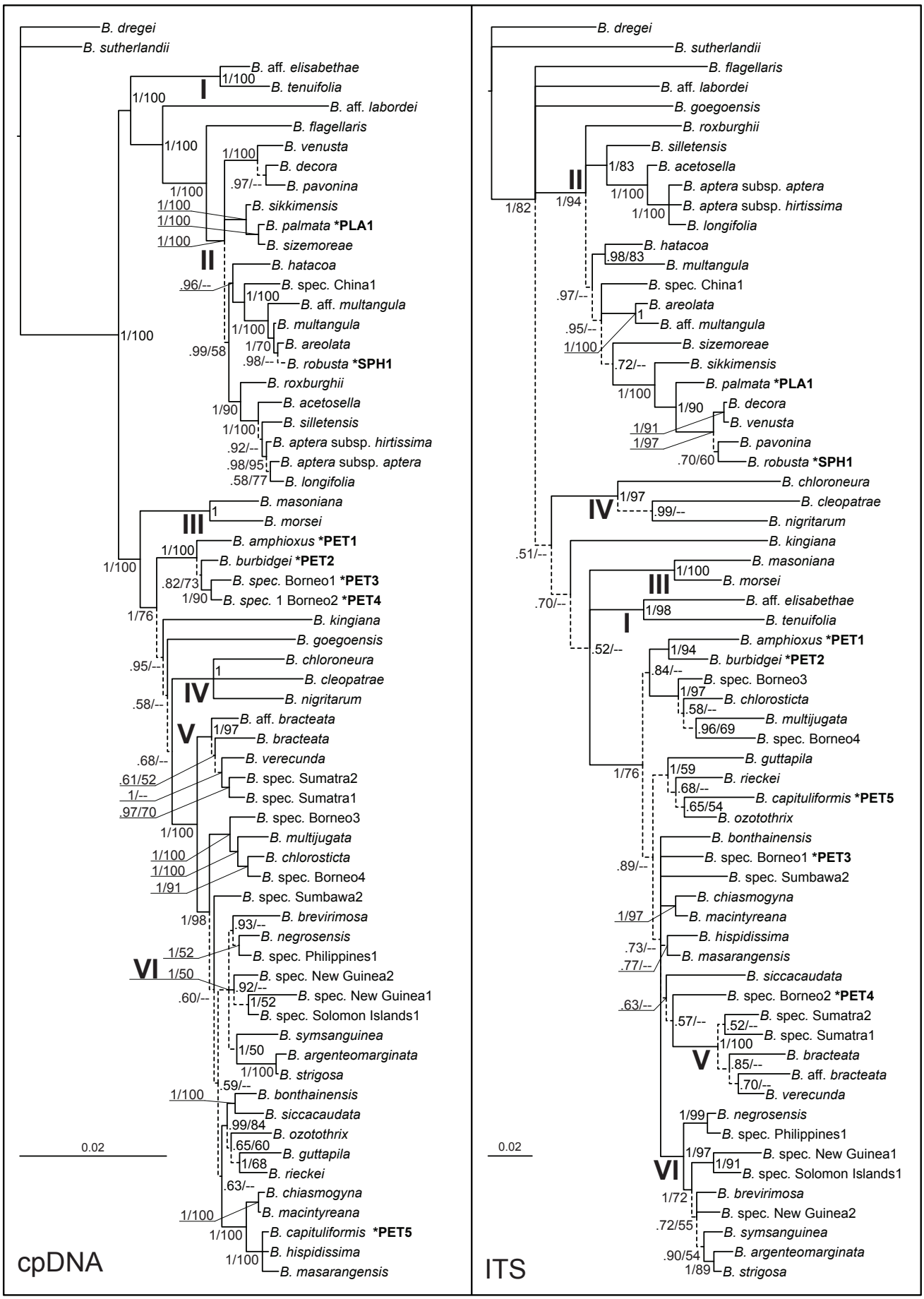


Fig. 2.12. Comparison of nrDNA and cpDNA gene trees. Tree topologies are based on the Bayesian majority rule consensus trees of analyses of nrDNA (ITS, 3 data partitions, 64 taxa) and cpDNA (*ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L; 3 data partitions; 64 taxa). Bayesian posterior probability (PP) support values > 0.5 are indicated next to the nodes, and boostrap support values > 50 of corresponding clades of ML analyses are mapped on the tree: PP/BS. Broken lines indicate branches which lead to nodes which did not receive at least moderate support in both Bayesian and ML analyses (PP ≥ 0.90 and BS ≥ 70). *PLA1, *SPH1 and *PET1-5 indicates taxa with conflicting positions in the cpDNA and ITS phylogenies.

Bracteibegonia/*Petermannia*/*Symbegonia* clade. This clade is only poorly resolved, and while *B. amphioxus* and *B. burbidgei* are supported as monophyletic, the two unidentified Bornean species show only poorly supported relationships. Finally, *Begonia capituliformis* Irmsch. falls into a strongly supported clade with *B. hispidissima* Zipp. ex Koord. and *B. masarangensis* Irmsch. in the cpDNA phylogenetic tree, but falls into a well supported clade with *B. guttapila* D.C.Thomas & Ardi, *B. rieckei* Warb., and *B. ozotothrix* D.C.Thomas in the ITS phylogenetic tree.

2.3.4 Ancestral character state reconstructions

Parsimony and likelihood ancestral character state reconstructions of five characters including the presence and type of specialised perennating organs, fruit types, locule numbers, placentation types, and placenta divisions are presented in Figs. 2.13-17.

Presence of specialised perennating organs (Fig. 2.13): Reconstructions of the ancestral character states at the deepest nodes are equivocal. Within Clade A, species of section *Parvibegonia*, which are the sister to the rest of the clade, and species of section *Diploclinium* exhibit tubers. The reconstructions indicate a character transition from the tuberous habit found in species placed in section *Diploclinium* to the rhizomatous habit exhibited by species in the *Platycentrum-Sphenanthera* clade. Within the rhizomatous *Platycentrum-Sphenanthera* clade, rhizomes were lost in a lineage comprising *Begonia acetosella* Craib, *Begonia longifolia* and *Begonia aptera* Blume, which exhibit erect stems and fibrous root systems. Within Clade B, species of section *Coelocentrum*, which form the sister clade to the rest of the clade, as well as species in clades comprising Malesian taxa of sections *Diploclinium*, *Reichenheimia* and the monotypic section *Ridleyella*, are characterized by rhizomatous habits. Rhizomes were lost in the lineage comprising sections *Bracteibegonia*, *Petermannia* and section *Symbegonia*. Within this lineage, only two species, *Begonia siccacaudata* J.Door. and *B. mendumiae* M.Hughes show tuberous or rhizomatous organs, respectively, which evolved independently from the rhizomes and tubers of other lineages of Asian *Begonia*.

Fruit type, ovary locule numbers and placenta configuration (Figs. 2.14-16): Character state reconstructions indicate that three-locular ovaries developing into dry capsules at maturity are ancestral in Asian *Begonia*. Within Clade A, two-locular ovaries and rain-ballist fruits likely evolved independently in both sections *Parvibegonia* and *Platycentrum*. The reconstructions indicate two independent character transitions from two-locular, rain-ballist fruits to fleshy fruits within the *Platycentrum-Sphenanthera* clade. Within Clade B, two-locular ovaries are present in section *Ridleyella* and in a single species in the *Diploclinium* clade, and fleshy fruits evolved in some species of section *Petermannia*. Two-locular ovaries and fleshy fruits evolved independently in both clades A and B. The unilocular ovaries with parietal placentation in section *Coelocentrum* are likely derived

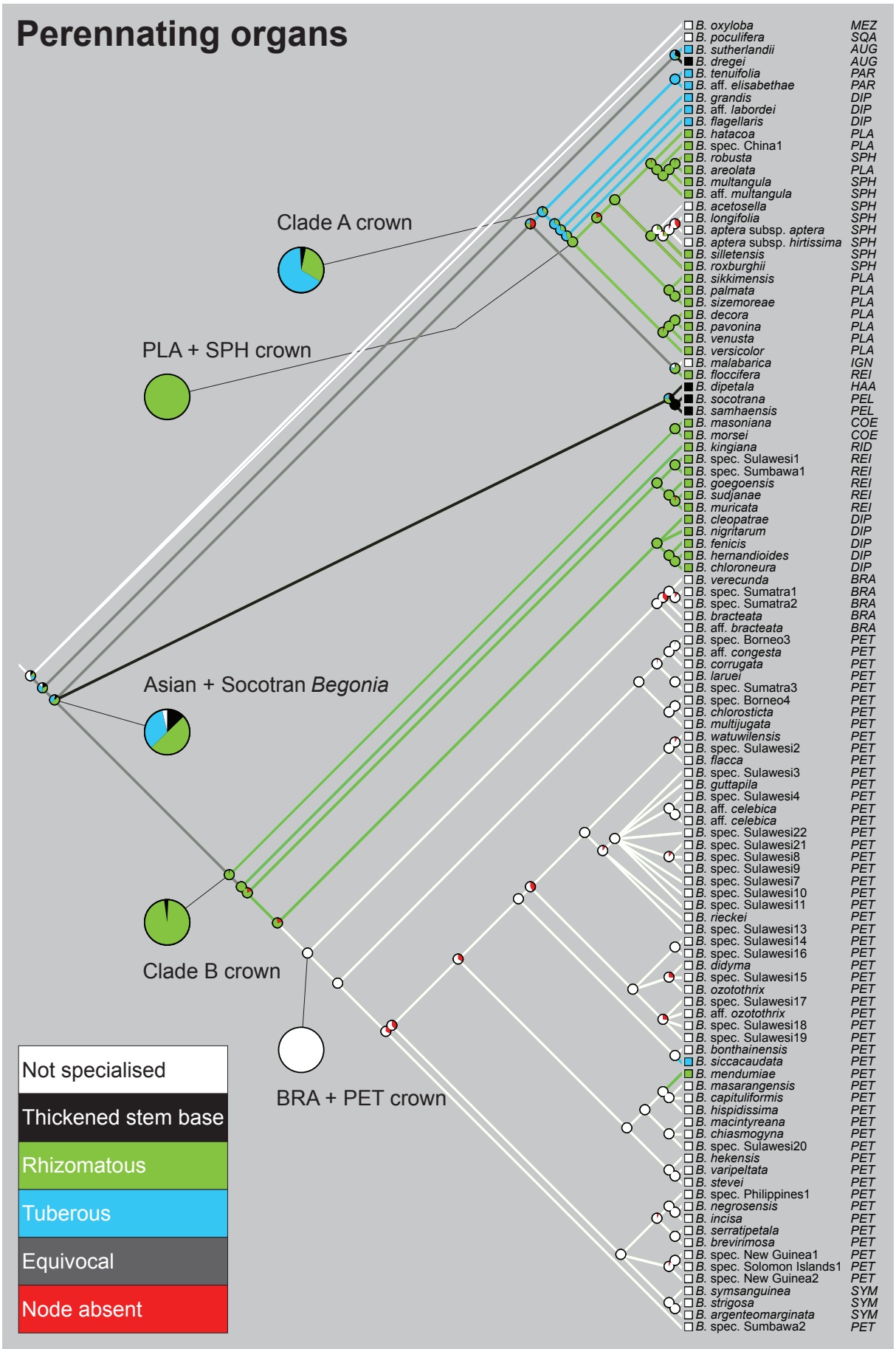


Fig. 2.13. Parsimony and likelihood ancestral character reconstruction: Perennating organs and stem metamorphoses. Character reconstructions across 1000 Bayesian input trees are shown on the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L + indel codes; 4 data partitions; 109 taxa). Branch colour indicates parsimony optimization of ancestral character states. Pie charts at each node illustrate the likelihood reconstructions and show the proportion of the average likelihood received by each character state as the ancestral character of a given clade.

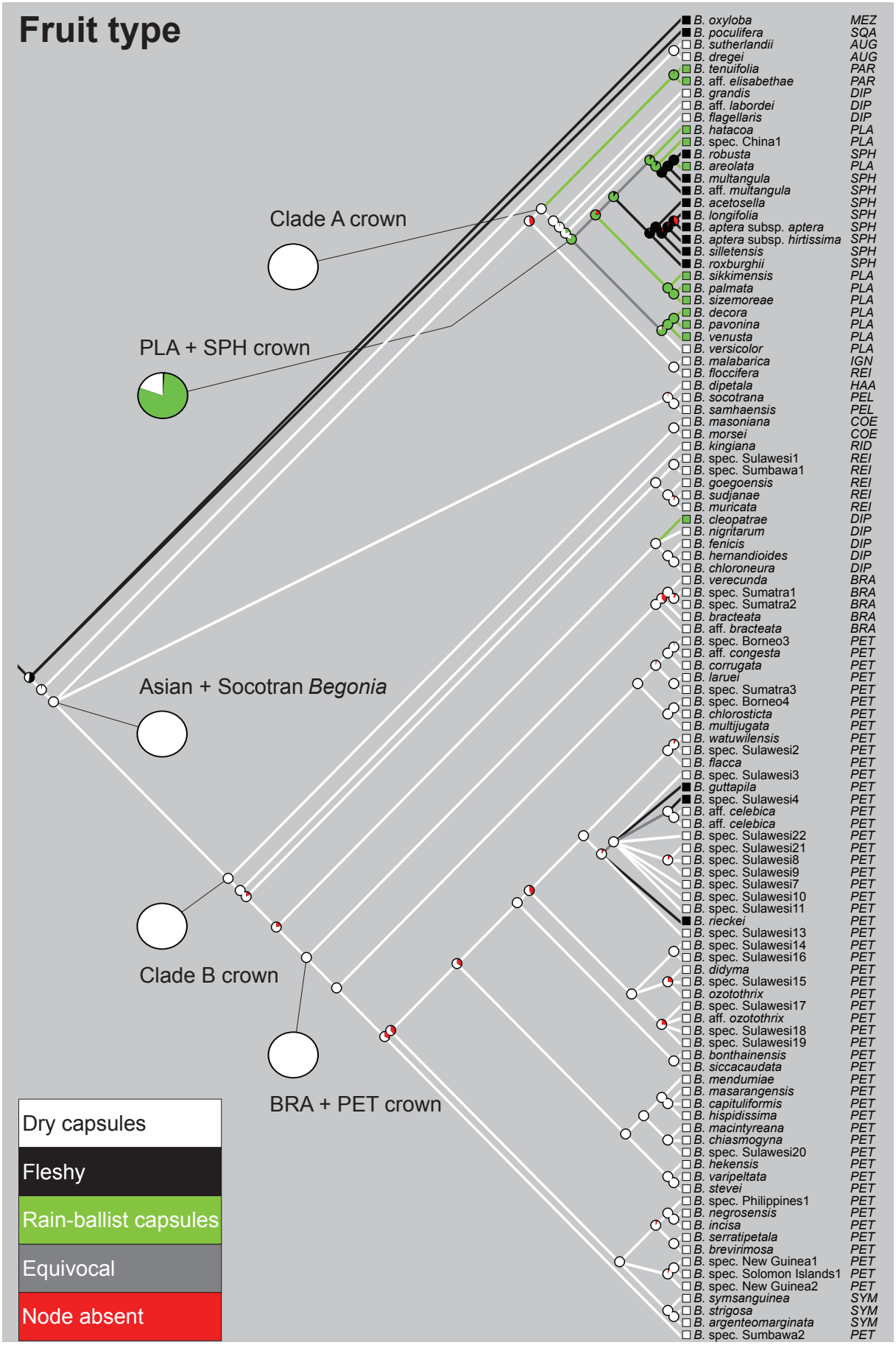


Fig. 2.14. Parsimony and likelihood ancestral character reconstruction: Fruit types. Character reconstructions across 1000 Bayesian input trees are shown on the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L + indel codes; 4 data partitions; 109 taxa). Branch colour indicates parsimony optimization of ancestral character states. Pie charts at each node illustrate the likelihood reconstructions and show the proportion of the average likelihood received by each character state as the ancestral character of a given clade.

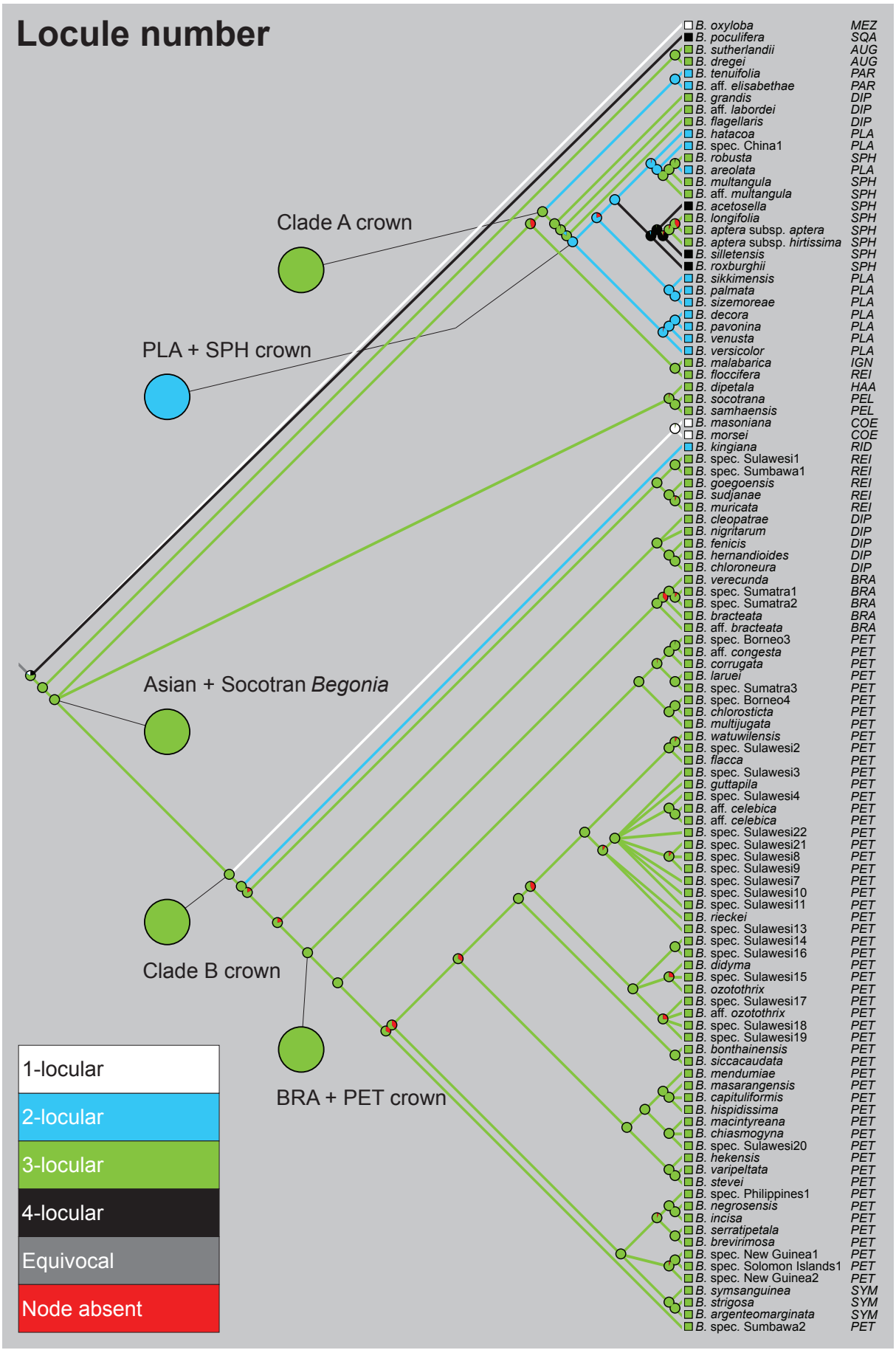


Fig. 2.15. Parsimony and likelihood ancestral character reconstruction: Ovary locule numbers. Character reconstructions across 1000 Bayesian input trees are shown on the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L + indel codes; 4 data partitions; 109 taxa). Branch colour indicates parsimony optimization of ancestral character states. Pie charts at each node illustrate the likelihood reconstructions and show the proportion of the average likelihood received by each character state as the ancestral character of a given clade.

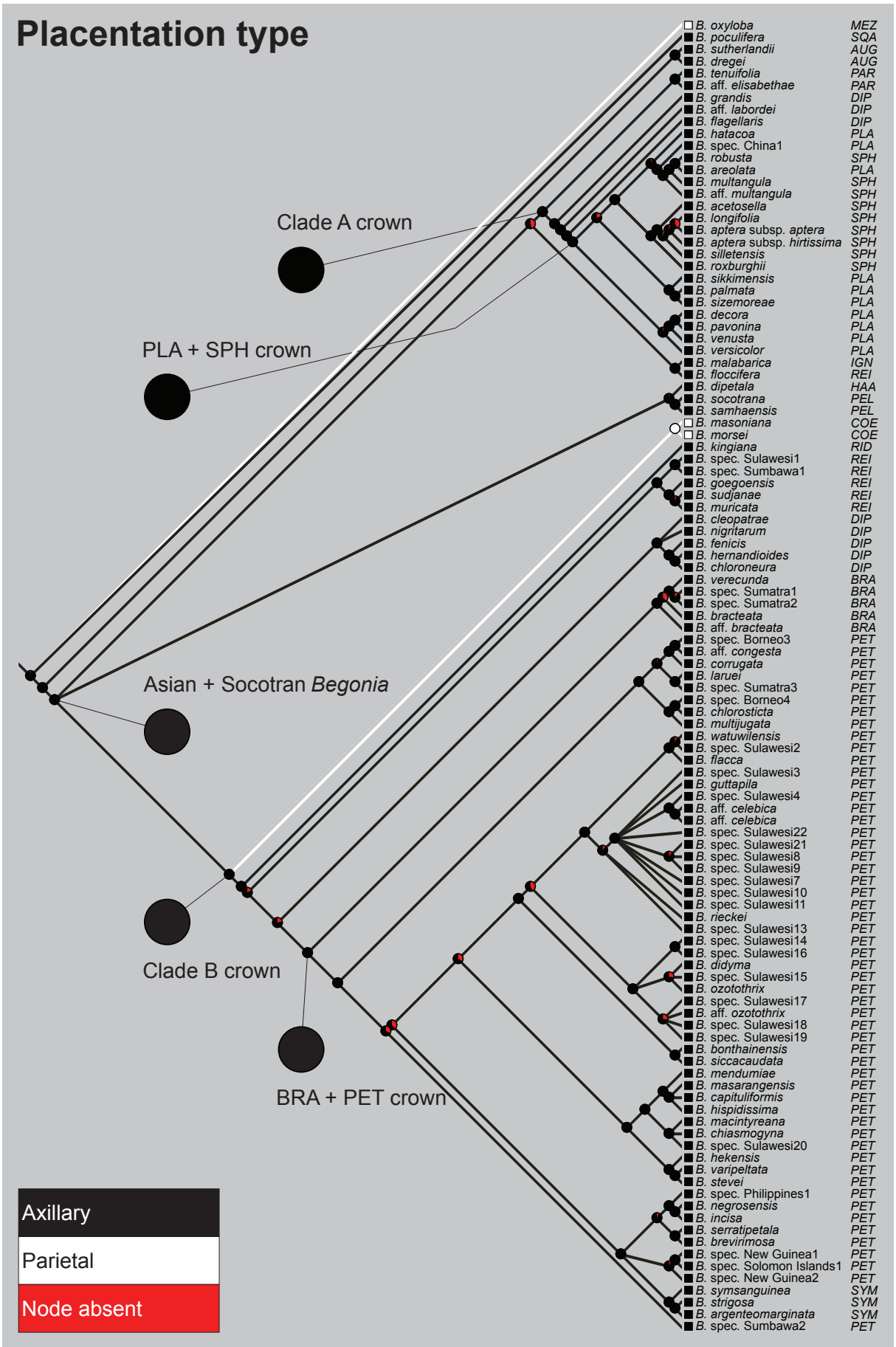


Fig. 2.16. Parsimony and likelihood ancestral character reconstruction: Placentation type. Character reconstructions across 1000 Bayesian input trees are shown on the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L + indel codes; 4 data partitions; 109 taxa). Branch colour indicates parsimony optimization of ancestral character states. Pie charts at each node illustrate the likelihood reconstructions and show the proportion of the average likelihood received by each character state as the ancestral character of a given clade.

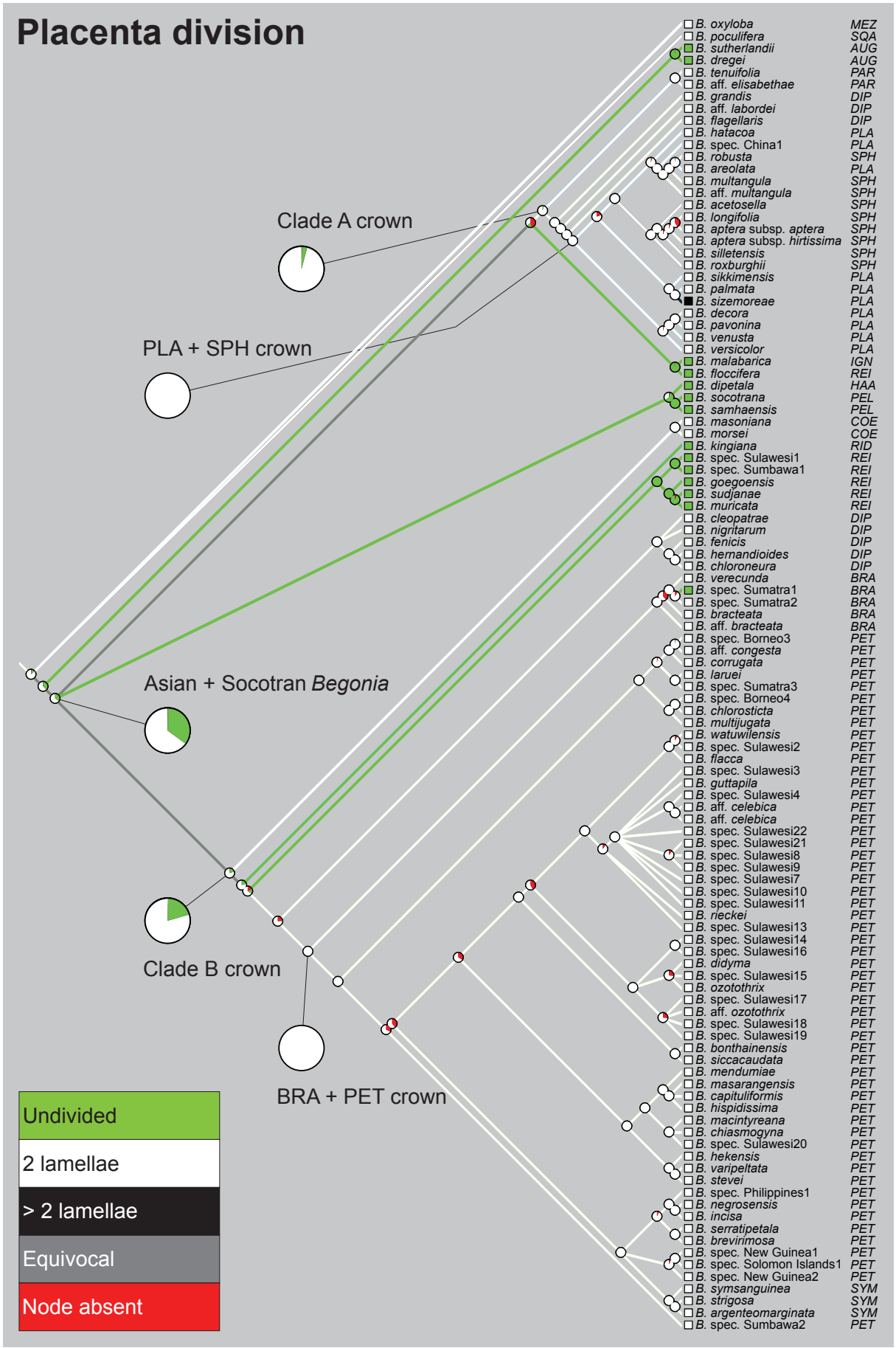


Fig. 2.17. Parsimony and likelihood ancestral character reconstruction: Placenta division. Character reconstructions across 1000 Bayesian input trees are shown on the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L + indel codes; 4 data partitions; 109 taxa). Branch colour indicates parsimony optimization of ancestral character states. Pie charts at each node illustrate the likelihood reconstructions and show the proportion of the average likelihood received by each character state as the ancestral character of a given clade.

from a three-locular ovary with axile placentation.

Placenta division (Fig. 2.17): The reconstructions indicate that placentae with undivided lamellae are likely to be homoplasious and evolved at least twice independently within Clade B (Sections *Reichenheimia* and *Ridleyella*, as well as one species in section *Bracteibegonia*). Undivided placentae can also be found in early divergent lineages comprising species assigned to sections *Haagea*, *Reichenheimia*, *Peltaugustia* and species unplaced to section, but reconstructions are equivocal at the deepest nodes.

2.4 Discussion

2.4.1 Utility of the ITS and non-coding cpDNA markers for phylogenetic analyses of Southeast Asian *Begonia*

Phylogenetic analyses of ITS sequence data have largely failed to resolve deeper relationships within Asian *Begonia* (this study; Forrest et al., 2001; Forrest et al., 2005; Tebbitt et al., 2006). One reason for the limited utility of the ITS region for phylogenetic analyses of Asian *Begonia* is the extensive nucleotide and sequence length variation which can be observed in ITS datasets (Tables 2.1, 2.5). Sequence variability of the ITS1 and ITS2 regions is over 60% in the 89 taxon dataset resulting in high levels of ambiguous sequence alignment. Despite high percentages of potentially parsimony informative characters, ITS phylogenetic trees of Asian *Begonia* show only poorly resolved backbones, which indicates that the analyses are likely confounded by extensive nucleotide variation and associated high levels of alignment ambiguity and homoplasy. Non-coding cpDNA regions, i.e. introns and intergenic spacers, are less functionally constrained and usually exhibit greater average sequence variation than coding regions, but distinctly less average variation than the ITS. However, non-coding cpDNA regions, especially introns, can have well-conserved secondary structures resulting in mosaics of highly conserved and variable parts (Borsch and Quandt, 2009). While some commonly employed markers, like the *trn*L intron, were shown to have relatively low levels of average sequence variation in *Begonia* (Goodall-Copestake, 2005; Plana, 2003; Plana et al., 2004), the three non-coding cpDNA markers employed in the analyses presented here (*ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer) exhibit intermediate levels of average variability in comparison to coding cpDNA regions and the ITS region (Tab. 2.1). The alignments of these non-coding regions show little ambiguity and the analyses resulted in well resolved, informative phylogenetic trees of Asian *Begonia*. However, one problematic aspect of employing non-coding cpDNA regions for phylogenetic analyses is the frequent occurrence of poly A/T homonucleotide strands. Homology assessment of length differences in these plastid microsatellite regions is usually not possible, as they show fast mutational dynamics which may involve overlapping insertions and deletions

of one to several nucleotides often resulting in highly homoplasious indel structures (Borsch and Quandt, 2009; Tesfaye et al., 2007). Consequently, these mutational hotspots were excluded from the analyses. Moreover, poly A/T homonucleotide strands composed of eight or more nucleotides often cause PCR artefacts due to slipped-strand mis-pairing during PCR-mediated DNA replication (Shinde et al., 2003). The sequence chromatograms of most samples with poly A/T homonucleotide strands of 8 to 10 bases exhibited noise in the sequence chromatogram, while the correct signal was usually still distinct, but the quality of the sequence data following longer homonucleotide strands were greatly reduced, sometimes to the point of being completely obliterated. The application of *Pfu*-based DNA polymerases which are attached to nonspecific DNA binding proteins mitigates this problem, at least for homonucleotide strands of up to 15 bp length, possibly because of increased contact surface between enzyme and DNA in comparison to *Taq* polymerases (Fazekas et al., 2010). Moreover, the cpDNA alignments presented here show levels of conservation which allow unproblematic sequencing primer design for major lineages in Asian *Begonia*. The negatives of additional costs and time requirements to sequence these markers by using high-quality *Pfu*-based DNA polymerases and additional internal primers are outweighed by the presence of intermediate rates of sequence evolution, which were appropriate for the questions addressed in this study, and which resulted in little alignment ambiguity and highly informative phylogenetic trees.

2.4.2 Phylogenetics and character evolution of Southeast Asian *Begonia*

2.4.2.1 Major subclades of the monophyletic Socotran-Asian clade

The results from the analyses of non-coding cpDNA sequence data show that Asian and Socotran *Begonia* species form a well supported clade. This confirms, with greater sampling, the results of former phylogenetic studies (Forrest and Hollingsworth, 2003; Forrest et al., 2005; Goodall-Copestake, 2005). The majority of species within the Socotran-Asian clade fall into two major subclades in the cpDNA phylogenetic trees: Clade A and Clade B (Figs. 2.4-6).

Clade A exhibits subclades of species of taxa which are most diverse on the Asian mainland including sections *Parvibegonia*, continental Asian species placed in section *Diploclinium*, and species of section *Platycentrum* s.l. (inclusive section *Sphenanthera*). Species within this clade exhibit diverse fruit morphologies and anatomies including dry capsules, rain-ballist capsules and fleshy fruits and the vast majority of species have tubers or rhizomes. The predominant somatic chromosome number within this clade seems to be 2n = 22 (Fig. 2.18), which has been reported from species in all sections within the clade, but there are also polyploid series and numerous aneuploid derivates (Doorenbos et al., 1998; Ku et al., 2007; Legro and Doorenbos, 1969, 1971, 1973; Oginuma and Peng, 2002). Further chromosome counts in sections *Parvibegonia* and mainland *Diploclinium* are needed to test whether a base chromosome number of n = 11 was likely present in

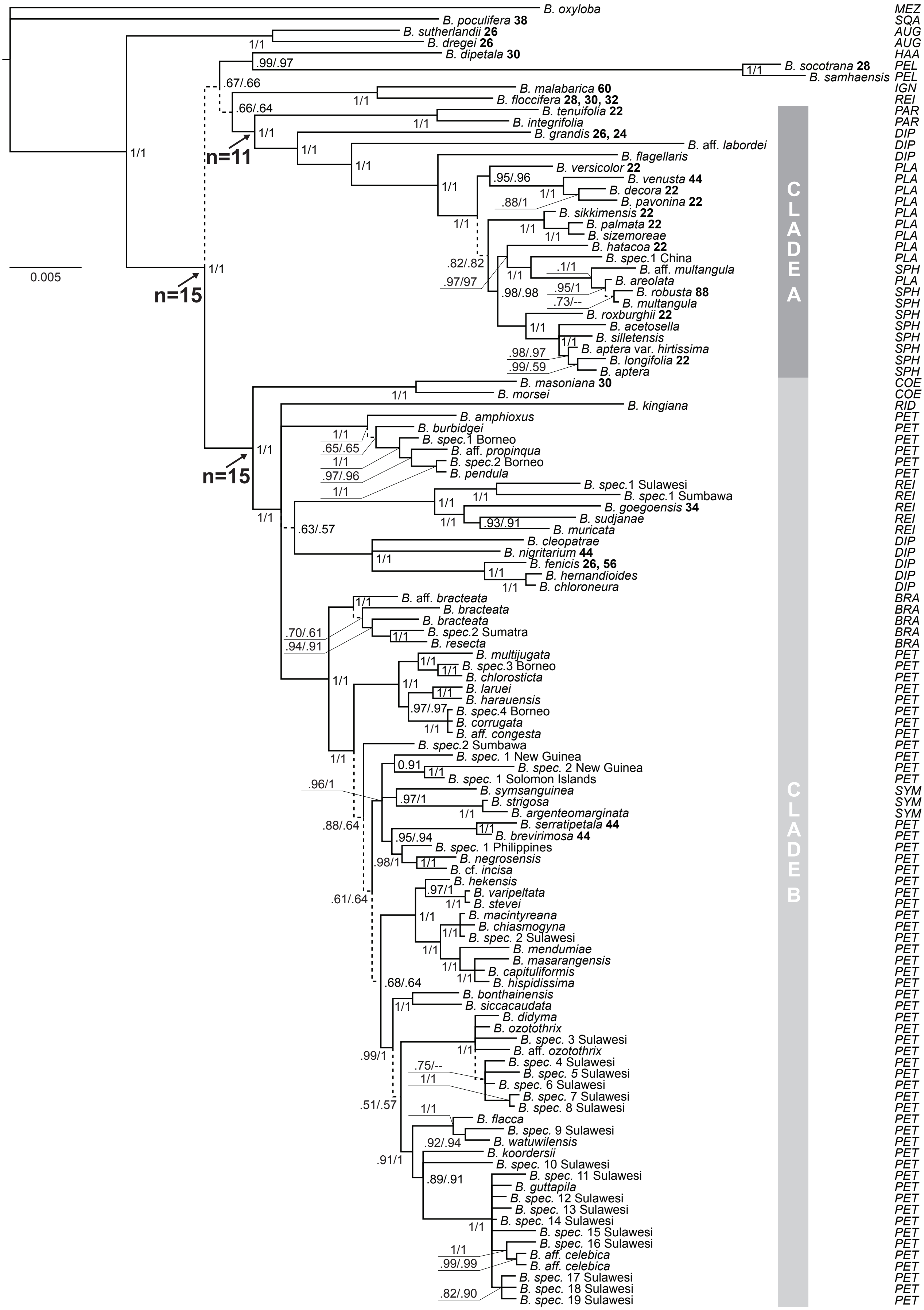
the most recent common ancestor of Clade A or whether it is characteristic for a subclade including sections *Platycentrum* s.l. and closely related, continental Asian species in section *Diploclinium* (see 2.4.2.5).

Clade B includes species of the predominantly Chinese section *Coelocentrum*, species in the predominantly or exclusively Malesian sections *Ridleyella*, *Bracteibegonia*, *Petermannia*, *Symbegonia*, and Malesian species placed in sections *Diploclinium* and *Reichenheimia*. Species in most of these taxa exhibit a rhizomatous habit, but rhizomes were lost in the large section *Petermannia* s.l. (inclusive *Symbegonia*) and the closely related section *Bracteibegonia*. The vast majority of taxa within Clade B exhibit dry capsules, but fleshy fruits evolved independently in a few Sulawesian species in section *Petermannia* and, as indicated by ITS data, in two Chinese species in the polyphyletic section *Leprosae*. The predominant somatic chromosome number in this clade seems to be 2n = 30 (Fig. 2.18). Chromosome counts are sparse for the Malesian taxa, but the vast majority of species in section *Coelocentrum*, which is the sister to the rest of the clade, have a chromosome number of 2n = 30 (Ku et al., 2007; Ku, 2006), somatic chromosome numbers of 30 and 44 have been reported from some Malesian species in sections *Reichenheimia*, *Diploclinium* and *Petermannia* (Doorenbos et al., 1998; Legro and Doorenbos, 1969, 1971, 1973), and Legro and Doorenbos (Legro and Doorenbos, 1971, 1973) hypothesised that the somatic chromosome numbers of 44 in some species in sections *Reichenheimia* and *Petermannia* likely arose from triploids of species with 30 somatic chromosomes.

The basal relationships in the phylogeny, which involve two subclades comprising five Indian, Sri Lankan and Socotran species assigned to sections *Haagea*, *Reichenheimia*, and *Peltaugustia*, are unresolved or only weakly supported. Somatic chromosome numbers are 30 for the three Indian and Sri Lankan species, and 28 for one of the two Socotran species, indicating that a primary base chromosome number of n = 15 might be ancestral in Asian *Begonia*.

2.4.2.2 Basal relationships in Asian Begonia

The unresolved or poorly supported basal relationships within the cpDNA phylogenetic trees involve two clades which comprise Indian, Sri Lankan and Socotran taxa assigned to sections *Reichenheimia* (*Begonia floccifera*), *Haagea* (*B. dipetala*), *Peltaugustia* (*B. samhaensis, B. socotrana*) and one species unplaced to section (*B. malabarica*). This corroborates the phylogenetic analyses and molecular divergence age estimates of Goodall-Copestake (2005), which indicated that these taxa form the earliest divergent subclades within the Socotran-Asian clade. The sequence data supports a close relationship of the South Indian and Sri Lankan species *Begonia dipetala* and the only two species known from the Socotra Archipelago, which has not been suggested before. Species in both of these basal clades exhibit morphologically heterogeneous growth habits and



2.18 Karyotype evolution in Asian *Begonia*. Somatic chromosome counts (Doorenbos et al., 1998; Ku et al., 2007; Legro and Doorenbos, 1969, 1971, 1973; Oginuma and Peng, 2002) are indicated as bold numbers after the taxon names of the Bayesian majority rule consensus tree (cpDNA data: *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L; 3 data partitions; 115 taxa). Putative base chromosome numbers of Socotran-Asian *Begonia* and the major subclades A and B are indicated as n=15 and n=11. See captions of Fig. 2.4 for explanations of support values and abbreviations.

floral characters, but, interestingly, most of them show adaptations to survive seasonally dry conditions (see 2.2.5.1.1) and the ovaries and fruits of all of these taxa exhibit undivided placenta lamellae, while the majority of Asian species exhibit bilamellate placentae. However, the character reconstructions remain largely equivocal with regards to the question whether undivided or bilamellate placentae are ancestral within Asian *Begonia*. Of the seventeen *Begonia* species known from the Ghats of India and from Sri Lanka (Jayasuriya, 1983; Uddin, 2007), eight, including the lectotype species of section *Reichenheimia* (*Begonia thwaitesii* Hook., a heterotypic synonym of *B. tenera* Dryand.; designated by Barkley and Baranov, 1972), exhibit undivided placentae and have been placed in section *Reichenheimia* (Doorenbos et al., 1998). Of these eight species, *Begonia albococcinea* Hook., *B. phrixophylla* Blatt. & McCann, *B. subpeltata* Wight, *B. tenera,* and *B. trichocarpa* Dalzell were not included in the analyses, and further phylogenetic and detailed morphological studies are needed to investigate their relationships and to clarify the morphological circumscriptions of the early divergent lineages within the Socotran-Asian clade.

2.4.2.3 Polyphyly of section Reichenheimia *and homoplasy of undivided placenta lamellae in Asian* Begonia

Since Klotzsch (1854) erected the genus *Reichenheimia* to accommodate two species from Sri Lanka and South India, which are characterised by a tuberous, acaulescent habit, and three-locular ovaries with undivided placentae, almost all Asian *Begonia* species which exhibit ovaries with undivided placenta lamellae were placed in section *Reichenheimia*. Exceptions are two species placed in section *Ridleyella*, which Irmscher (1929) described to accommodate species with two-locular ovaries and undivided placentae, and the erect, non-tuberous *Begonia dipetala*, which Klotzsch (1854) placed in a separate monotypic genus *Haagea*. Clarke (1879) placed Indian and Indo-Chinese species with undivided placenta lamellae, including the Indian and Sri Lankan *Begonia floccifera* and *B. malabarica*, which fall into the unresolved basal clades of the cpDNA phylogenetic trees, in section *Uniplacentales*, and subsequently most of the species in Clarke's section *Uniplacentales* were placed in section *Reichenheimia* (Doorenbos et al., 1998). However, Irmscher (1939) already emphasised that the practise of pooling all Asian species with three-locular ovaries and undivided placentae in section *Reichenheimia* had resulted in a morphologically heterogeneous group. The analyses of the cpDNA sequence data indicate that this morphological heterogeneity is correlated with the polyphyly of the section, and the Indian and Sri Lankan *Begonia floccifera* and *B. malabarica* form a strongly supported clade, which is only distantly related to Malesian species assigned to section *Reichenheimia*. Irmscher (1939) placed four Chinese species in section *Reichenheimia*, and the recent *Flora of China* treatise (Ku et al., 2007) followed these placements, and another four species were added to the list. The relationships of two Chinese species, *Begonia henryi* Hemsl. and *B. parvula* H.Lév. & Vaniot, which were assigned to section *Reichenheimia*,

are largely unresolved in the ITS phylogenetic trees, but they are not included in a strongly supported clade of Malesian species placed in section *Reichenheimia*. The eight Chinese species assigned to section *Reichenheimia* in the *Flora of China* (Ku et al., 2007) form a morphologically heterogeneous assemblage: six species, including *Begonia henryi* and *B. parvula* show a similar growth habit as Chinese species placed in section *Diploclinium*, which exhibit a tuberous, acaulescent habit. The development of variably undivided and bifid placenta lamellae has been described for some Chinese species placed in section *Diploclinium* (*Begonia labordei* H.Lév., *B. fimbristipula* Hance, *B. wilsonii* Gagnep.), and these species seem to morphologically link the tuberous Chinese *Reichenheimia* species with tuberous Chinese species assigned to section *Diploclinium* (Irmscher, 1939; Shui et al., 2002). However, the unresolved or poorly supported early divergent positions in the ITS phylogenetic trees, a chromosome count of 2n = 30 for *Begonia henryi* (Ku et al., 2007), and the undivided placentae lamellae indicate that *B. henryi* and *B. parvula* may belong to early divergent lineages within Asian *Begonia*. The other two Chinese species assigned to *Reichenheimia*, *Begonia cylindrica* D.R.Liang & X.X.Chen and *B. filiformis* Irmsch., are not tuberous, but exhibit well-developed rhizomes. *Begonia cylindrica* is morphologically aberrant for the section and exhibits rhizomes, fleshy, wingless fruits and uni- or bilamellate placentae (Ku et al., 2007; Tebbitt, 2005; Tebbitt et al., 2006). Tebbitt et al. (2006) pointed out that *Begonia cylindrica* is morphologically most similar and maybe conspecific with *B. leprosa*, which, based on ITS data, falls into a well supported clade with species of section *Coelocentrum* (Tebbitt et al., 2006). *Begonia filiformis* also shows close morphological affinities to section *Coelocentrum*, in which it was placed by Shui et al. (2002). The character combination of rhizomes, uniloculate ovaries and the yellowish-greenish tepals strongly support this placement.

Ten species which are currently placed in section *Reichenheimia* have been described from Burma, Thailand, Laos, and Vietnam (Hughes, 2008). Most of these species, like *Begonia brandiana*, exhibit tubers and an acaulescent habit morphologically similar to Continental Asian *Dipoclinium* lineages, but some of these species maybe more closely related to other tuberous Continental lineages like section *Parvibegonia* or the early divergent South Asian and Socotran lineages, and further phylogenetic and morphological studies are needed to clarify their relationships.

Twenty-eight Malesian species are currently placed in section *Reichenheimia* (Hughes, 2008; Hughes et al., 2009; Kiew and Sang, 2009). The samples of this group which were included in the molecular analyses form a clade which is well supported as monophyletic in Clade B of the cpDNA phylogenetic trees, and there is no indication for a close relationship to either Sri Lankan, Indian, or Chinese species placed in this section. The well developed rhizomes in the Malesian species clearly separate them from most Continental species assigned to section *Reichenheimia*, which are predominantly tuberous and either acaulescent or erect. Malesian *Reichenheimia* species share a suite of vegetative and

generative characters with section *Ridleyella*, which can be separated by two-locular ovaries (Irmscher, 1929), and Malesian species assigned to section *Diploclinium*, which can be separated by ovaries with bilamellate placentae (Doorenbos et al., 1998). The close relationships between these taxa is indicated by the cpDNA phylogenetic trees, although the relationships between section *Ridleyella*, Malesian *Diploclinium* and Malesian *Reichenheimia* remain unresolved or only poorly supported. Thus, rhizomatous species placed in section *Reichenheimia* seem to form a natural group, if you exclude both the only distantly related rhizomatous Indian species *Begonia floccifera* and the rhizomatous Chinese species which are misplaced in *Reichenheimia* and belong to section *Coelocentrum*. This Malesian group is mainly distributed in the predominantly everwet Sunda Shelf region, and has its centre of diversity on Sumatra and the Malay Peninsula. Only few species extend the distributional range to eastern Malesia including the Lesser Sunda Islands, Southeast Sulawesi and the Malukku Islands (Fig. 2.19) (Hughes, 2008), and it is apparently absent from continental Asian regions north of the Thai-Malay Peninsula, which show a monsoonal seasonal climate with pronounced dry seasons. This group has been thoroughly revised for the Malay Peninsula (Kiew, 2005) and an ongoing revision of section *Reichenheimia* at the Royal Botanic Garden Edinburgh (Mark Hughes, Royal Botanic Garden Edinburgh, Edinburgh, UK, *pers. com.*) will provide the necessary morphological detail for the circumscription and formal description of this taxon.

Begonia section *Reichenheimia*, in its current circumscription, includes almost all Asian species with undivided placentae (Doorenbos et al., 1998), and this section is a prime example illustrating how the strong systematic importance associated with a single, homoplasious character resulted in the circumscription of a polyphyletic, morphologically heterogeneous taxon.

2.4.2.4 Phylogenetic relationships of section Parvibegonia *and homoplasy of rainballist capsules in Asian* Begonia

Sections *Parvibegonia* and *Platycentrum* share a character syndrome of two-locular fruits with bifid placenta lamellae, and De Candolle (1864) and Clarke (1879) lumped several species currently placed in section *Parvibegonia* together with species currently placed in section *Platycentrum.* However, Irmscher (1929) and Doorenbos et al. (1998) treated section *Parvibegonia* as distinct from section *Platycentrum*. Species in section *Parvibegonia* are predominantly tuberous, small plants, while species in section *Platycentrum* usually exhibit well developed rhizomes and a much more robust growth habit. Moreover, species in section *Platycentrum* exhibit characteristic anthers with elongated connectives, which are not present in section *Parvibegonia*. The phylogenetic relationships of section *Parvibegonia* remained unclear in a recent phylogenetic study based on ITS data (Tebbitt et al., 2006), because of poor statistical support of the relationships of the only included species, and Tebbitt et al. (2006) pointed out that the possibility that the two-locular rainballist capsules in section *Platycentrum* arose from similar dispersed

taxa in section *Parvibegonia* needed further investigation. The results of the analyses of the three non-coding cpDNA regions and the ancestral character reconstruction indicate that section *Parvibegonia*, which is well supported as monophyletic, is the sister to a clade comprising a grade of continental Asian species assigned to section *Diploclinium* and a well supported subclade of species placed in sections *Platycentrum* s.l. The two-locular fruits and the rain-ballist syndrome apparently evolved independently in sections *Parvibegonia* and *Platycentrum*.

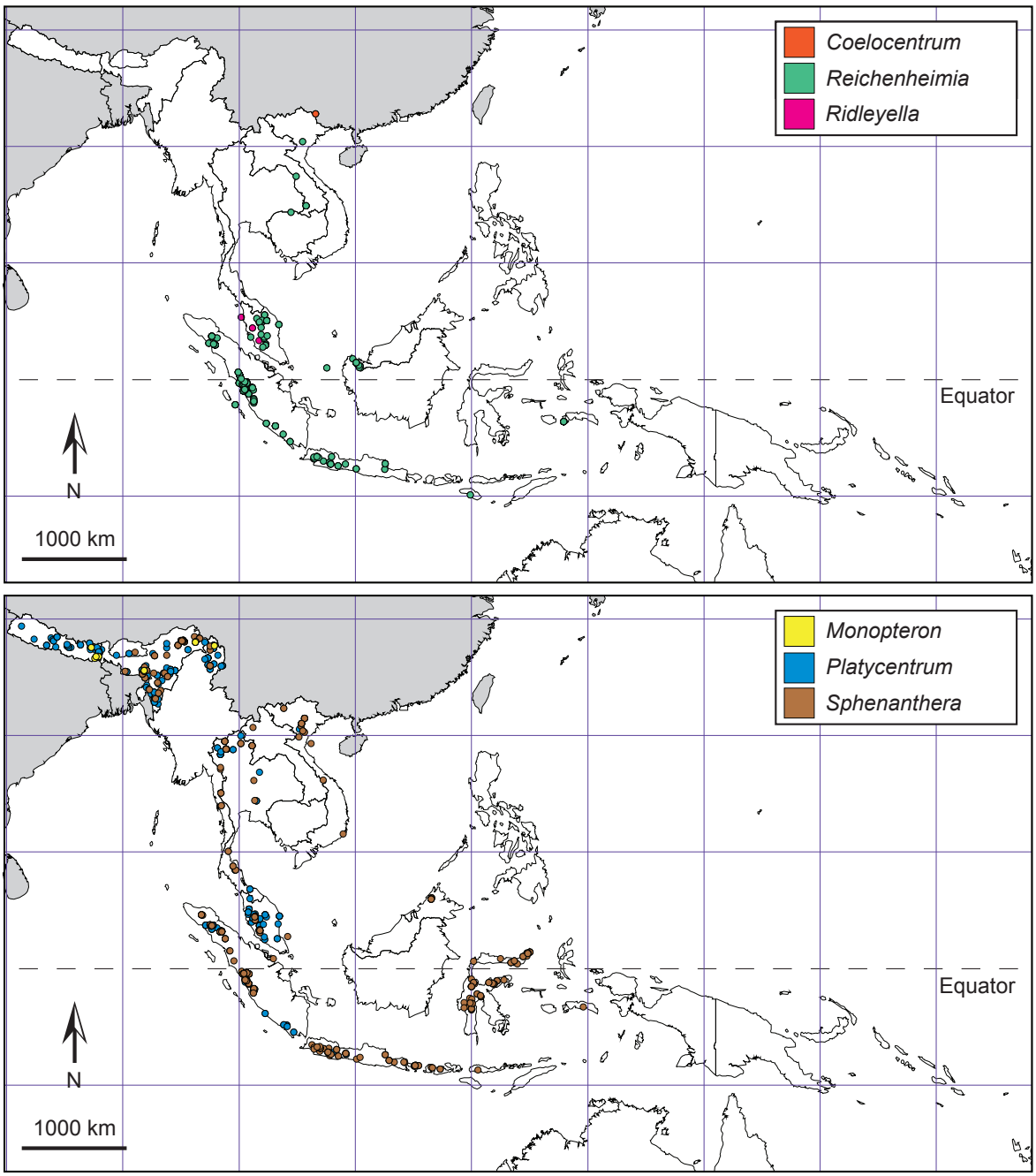


Fig. 2.19. *Begonia* distribution in Southeast Asia: sections *Coelocentrum*, *Monopteron*, *Platycentrum*, *Reichenheimia*, and *Ridleyella*. Distribution data is based on GPS data and georeferenced locations in the *Southeast Asian Begonia Database* (7946 specimens from 57 herbaria; Hughes and Pullan, 2007). Areas shaded in grey (Sri Lanka, India except for the north-eastern states, Bangladesh, China) are not included in the database. GIS data was downloaded from the following sites: DIVA-GIS: http://www.diva-gis.org/Data (administrative boundaries); Natural Earth: http://www.naturalearthdata.com (10° longitude/latitude grid).

Species in several monotypic or small sections which show narrow distributions in Indo-China and the Malay Peninsula (Fig. 2.20) show close morphological affinities with section *Parvibegonia*. Clarke (1879) lumped several species of section *Parvibegonia*, the monotypic section *Monophyllon* and section *Lauchea* in a new section *Papyraceae* C.B.Clarke, and Irmscher (1929) emphasised in his description of the monotypic section *Heeringia* the close relationship to section *Parvibegonia*. Species in sections *Heeringia*, *Monophyllon* and *Lauchea* exhibit a character syndrome with small tubers, two-locular fruits and a well developed filament column, which is characteristic of section *Parvibegonia*.

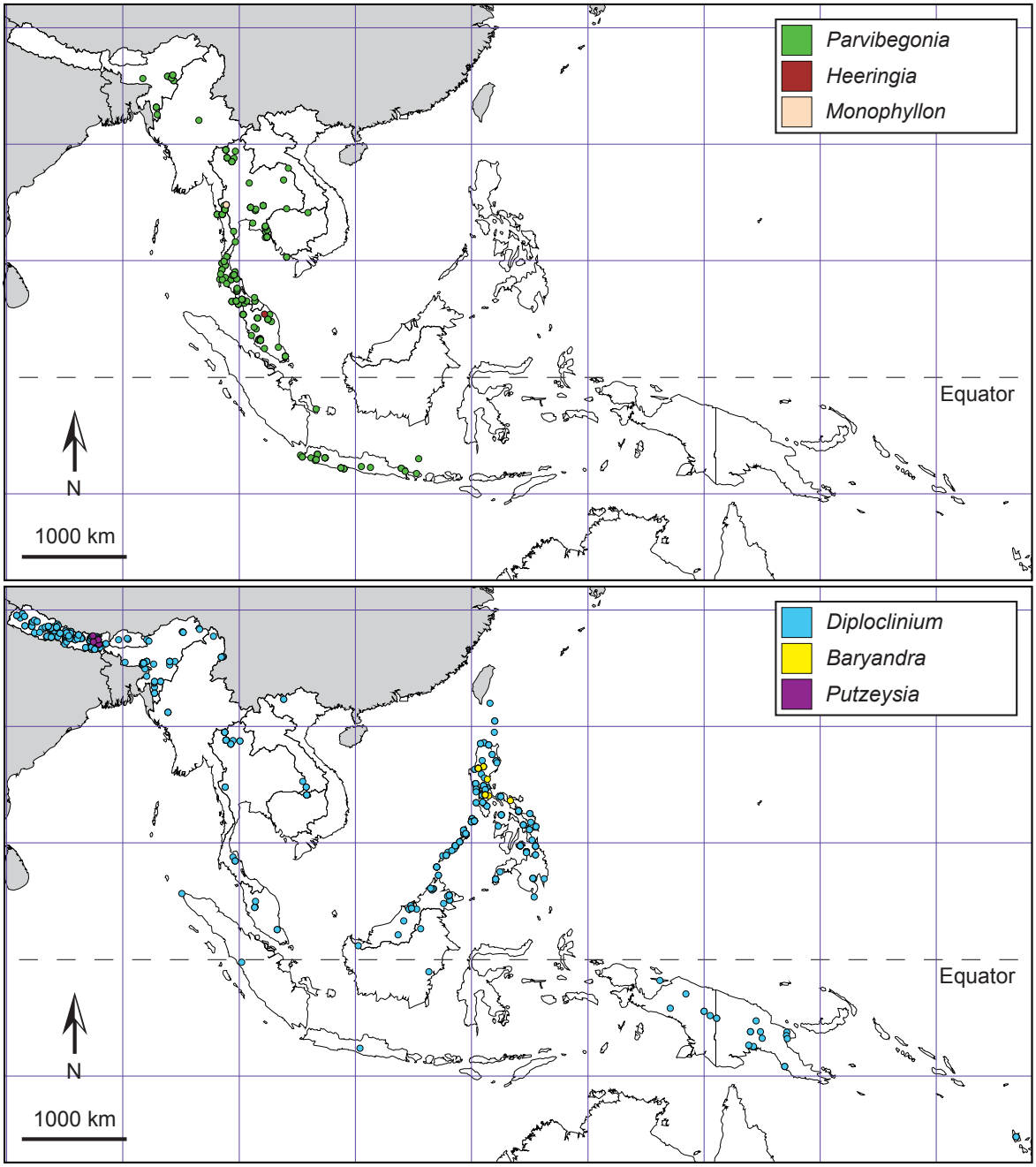


Fig. 2.20. *Begonia* distribution in Southeast Asia: sections *Baryandra, Diploclinium, Heeringia, Monophyllon, Parvibegonia*, and *Putzeysia*. Distribution data is based on GPS data and georeferenced locations in the *Southeast Asian Begonia Database* (7946 specimens from 57 herbaria; Hughes and Pullan, 2007). Areas shaded in grey (Sri Lanka, India except for the north-eastern states, Bangladesh, China) are not included in the database. GIS data was downloaded from the following sites: DIVA-GIS: http://www.diva-gis.org/Data (administrative boundaries); Natural Earth: http://www.naturalearthdata.com (10° longitude/latitude grid).

They are differentiated from section *Parvibegonia* only by few morphological characters like inflorescences which are developed in the axile of a single, sessile leaf (section *Monophyllon*, see Clarke 1879; Doorenbos et al., 1998), verticillate leaves associated with the inflorescence (section *Lauchea*, see Doorenbos et al., 1998), or opposite leaves and anther morphology (section *Heeringia*, see Irmscher, 1929; Kiew, 2005). The poorly known, monotypic section *Apterobegonia* is only based on one specimen described as *Begonia delicatula* Parish ex C.B.Clarke, which exhibits androecia with long filament columns, three-locular ovaries and wingless fruits (Clarke, 1879; Warburg, 1894). Despite the differences in generative characters, Clarke (1879) associated this species with species in section *Parvibegonia*, but morphologically it is also difficult to distinguish from continental Asian *Diploclinium* lineages.

2.4.2.5 Polyphyly of section **Diploclinium**

Doorenbos et al. (1998) lumped Asian species previously variously placed in sections *Diploclinium*, *Knesebeckia*, *Begonia* and *Begoniastrum* (Irmscher, 1939; Ku et al., 2007; Ku, 1999) in section *Diploclinium*. Thereby, they restricted sections *Begonia* and *Knesebeckia* to the New World, and the separation of Asian taxa from these New World sections has subsequently been supported by morphological and molecular data (Badcock, 1998; Forrest, 2001). The circumscription of *Diploclinium* sensu Doorenbos et al. (1998) has been widely accepted, with the exception of the *Flora of China* treatment (Ku et al., 2007), in which Chinese species placed in section *Diploclinium* sensu Doorenbos et al. (1998) were instead placed in section *Begonia*. Doorenbos et al. (1998) emphasized that *Diploclinium* in their wide circumscription is a morphologically heterogeneous taxon, and previous molecular phylogenetic studies indicated the polyphyly of section *Diploclinium*, and showed that Philippine species form a well supported clade, while some Asian mainland species like *Begonia grandis*, *B. alveolata* Yu and *B. wenshanensis* C.M.Hu ex C.Y.Wu & T.C.Ku, may be more closely related to sections *Platycentrum* and *Sphenanthera* than to other species in section *Diploclinium* (Badcock, 1998; Forrest, 2001; Tebbitt et al., 2006). However, the phylogenetic relationships between the clades which include taxa placed in section *Diploclinium* remained unclear, because of unresolved or poorly supported backbones of the phylogenetic trees. The results of phylogenetic analyses of the cpDNA sequence dataset further clarify the relationships of this morphologically heterogeneous, polyphyletic taxon. Species assigned to section *Diploclinium* fall into a well supported clade of rhizomatous, Malesian species in Clade B, and form a grade of predominantly tuberous, erect or acaulescent continental Asian species in Clade A of the cpDNA phylogenetic trees (Figs. 2.4-7.)

Malesian species placed in section *Diploclinium* include a large radiation on the Philippines (> 40 spp.), but also five species described from Borneo and seven species on New Guinea (Fig. 2.21) (Hughes, 2008; Hughes et al., 2010). The New Guinean species assigned to the section are poorly known and morphologically heterogeneous. Some

species, like the shortly rhizomatous *Begonia sharpeana* F.Muell. and related species, are morphologically similar to Philippine *Diploclinium* species. Others, like the climbing *Begonia kaniensis* and the erect species *B. brassii* Merr. & L.M.Perry and *B. oligandra* Merr. & L.M.Perry seem to be more closely related to New Guinean species of section *Petermannia* than to other species in section *Diploclinium*. Most Malesian species in section *Diploclinium* exhibit a rhizomatous habit, and can thus be differentiated from continental Asian species assigned to the section which predominantly exhibit tubers and erect leafy stems or an acaulescent habit. ITS data indicate that the Philippine *Begonia oxysperma*, the only species in the monotypic section *Baryandra*, is nested within section *Diploclinium*. This corroborates observations of Merrill (1912), who emphasised that the two sections can hardly be differentiated. Moreover, *Begonia oxysperma* and *B. calcicola* Merr., which is placed in section *Diploclinium*, share the synapomorphic character of multicellular hairs which have a broad and flat stalk which divides into few to several thinner branches at the apex (Doorenbos et al., 1998). The cpDNA phylogenetic trees and morphological similarities indicate close phylogenetic relationships to Malesian species placed in section *Reichenheimia* and to the small Malesian section *Ridleyella*. Philippine *Diploclinium* species share a suite of morphological characters with these sections (Doorenbos et al., 1998), but can be differentiated by the bilamellate placentae. Moreover, the five Philippine *Diploclinium* species included in the analyses of the cpDNA dataset are well supported as monophyletic and show a synapomorphic moderately large inversion of 345 or, due to deletion, 309 bp in the *ndh*F-*rpl*32 spacer.

The limited available molecular data indicates that continental Asian taxa placed in section *Diploclinium* do not form a monophyletic group, but instead form a grade within Clade A of the cpDNA phylogenetic trees (Fig. 2.4-7). Some taxa seem to be more closely related to a species in sections *Sphenanthera* and *Platycentrum* than to other mainland *Diploclinium* species, and *Begonia flagellaris* Hara, which is endemic to Nepal, is the sister to the *Platycentrum*-*Sphenanthera* clade in the cpDNA phylogenetic trees. The analyses of the ITS data further indicate that *Begonia gemmipara*, the only species placed in the monotypic section *Putzeysia,* falls into a strongly supported clade with *B. flagellaris* (Figs. 2.8-11). Clarke (1879) placed this species into section *Knesbeckia*, and the Asian species of this section were later transferred to section *Diploclinium* (Badcock, 1998; Doorenbos et al., 1998). Doorenbos et al. (1998) retained section *Putzeysia* as a distinct section because of conspicuous autapomorphic characters including papillate seed testae, and clusters of bulbils produced in modified inflorescences (Clarke, 1879; Grierson, 1991). However, Doorenbos et al. (1998) also emphasised that in other respects section *Putzeysia* could be accommodated in section *Diploclinium*. The ITS phylogenetic trees corroborates the hypothesis of a close relationship of sections *Putzeysia* and continental Asian *Diploclinium* lineages.

Species in the small section *Alicida* were not included in the analyses, but as they exhibit

a tuberous, erect habit, three-locular capsules with bifid placentae and monadelphous androecia, this section cannot be differentiated from the current circumscription of *Begonia* section *Diploclinium*, and species in section *Alicida* are likely closely related to other tuberous continental Asian *Diploclinium* lineages.

The rhizomatous, Chinese species *Begonia cavaleriei* seems to be misplaced in section *Diploclinum*. It falls into a clade with species of section *Coelocentrum* in the ITS phylogenetic trees. The ovaries of this species have been described as three-locular and bilamellate (Ku et al., 2007), but the rhizomatous habit and a chromosome number of 2n = 30 (Ku et al., 2007; Ku, 2006) like in most species in section *Coelocentrum*, support the molecular data.

The current circumscription of section *Diploclinium* is not based on synapomorphic characters, but is primarily based on a plesiomorphic ovary and fruit syndrome with dry, three-locular capsules with bilamellate placentae, and the absence of easily observable morphological characters like unilamellate placentae, fleshy fruits, rain-ballist fruits etc. which are characteristic for other infrageneric taxa. Rhizomatous Malesian species of *Diploclinium* form a well supported clade in the cpDNA phylogenetic trees, and are clearly distinct from tuberous Indian, Sri Lankan and continental Asian species assigned to section *Diploclinium*. An ongoing revision of Philippine *Diploclinium* by Rosario Rubite (Rosario Rubite, University of Manila, Manila, Philippines, *pers. com.*) (Rubite, 2010) will provide the necessary morphological and anatomical detail to circumscribe and formally describe this group. However, molecular analyses based on a geographically robust taxon sampling of the predominantly tuberous Indian, Sri Lankan and continental Asian species placed in section *Diploclinium* and detailed morphological studies are needed to identify major clades and apomorphic characters within this heterogeneous group.

2.4.2.6 Paraphyly of section Platycentrum*, polyphyly of section* Sphenanthera*, and homoplasy of fleshy fruits in Asian* Begonia

Species of sections *Sphenanthera* and *Platycentrum* form a well supported clade in the molecular analyses of the cpDNA dataset. Within this clade, species of section *Sphenanthera* form two derived subclades nested within a paraphyletic section *Platycentrum*. This is largely congruent with findings of former morphological and molecular studies (Doorenbos et al., 1998; Forrest, 2001; Tebbitt et al., 2006). Species of these two sections are predominantly rhizomatous, although rhizomes were secondarily lost in the *Begonia longifolia* complex (Tebbitt, 2003), and they exhibit androecia with characteristic anthers which dehisce via lateral slits and exhibit apically elongated connectives (Tebbitt et al., 2006). Some authors have emphasised that the two sections also share a somatic chromosome number of 2n = 22 (Forrest, 2001; Legro and Doorenbos, 1973; Tebbitt et al.,

2006). However, the phylogenetic framework provided by the cpDNA phylogenetic trees indicates that a primary base chromosome number of n = 11 might be a synapomorphic character for a wider taxon as somatic chromosome numbers of 2n = 22 have also been reported for the continental Asian *B. picta* Sm., *B. rubella* Buch.-Ham. ex D.Don, and *B. fimbristipula* which have been placed in section *Diploclinium* (Doorenbos et al., 1998; Forrest, 2001; Ku et al., 2007; Legro and Doorenbos, 1969, 1971, 1973). Moreover, somatic chromosome counts of 2n = 22 have reported for *Begonia tenuifolia* Dryand. in section *Parvibegonia* (Legro and Doorenbos, 1971), species of which form the sister clade to the rest of the Clade A in the cpDNA phylogenetic trees. This might indicate that a primary base chromosome number of n = 11 is synapomorphic for all taxa in Clade A, but further chromosome counts from species in sections *Parvibegonia* and *Diploclinium* are needed to test this hypothesis.

Within the *Platycentrum-Sphenanthera* clade, three and four-locular, fleshy, indehiscent fruits, which are characteristic of section *Sphenanthera*, apparently evolved multiple times independently from ancestors which had two-locular, rain-ballist capsules, which characterize section *Platycentrum* (Fig. 2.14). Doorenbos et al. (1998) emphasised that the fleshy-fruited *Begonia robusta*, which is placed in section *Sphenanthera*, is morphologically quite divergent from most species in the section, and could be easily accommodated in section *Platycentrum*, if it did not have a three-locular ovary. In the analyses of the cpDNA dataset, *Begonia robusta*, the morphologically similar species *B. multangula*, a species with strong morphological affinities to *B. multangula* from Sulawesi, and *B. areolata* form a well supported clade. These taxa show similarities in their growth habits characterized by short stout rhizomes from which well developed leafy stems arise. However, *Begonia areolata* exhibits two-locular capsules, while the other species have three-locular fleshy fruits. A second *Sphenanthera* clade is formed by other three- or four-locular fleshy-fruited species placed in section *Sphenanthera* including species of the *Begonia longifolia* complex. The analyses by Tebbitt et al. (2006) indicate that fleshy fruits likely evolved not just twice, but multiple times from rain-ballist ancestors within the *Platycentrum-Sphenanthera* clade. However, the backbone of the *Platycentrum-Sphenanthera* clade is only poorly supported in their ITS phylogenetic tree, and the comparison of cpDNA and ITS gene trees presented here shows hard incongruence within the *Platycentrum-Sphenanthera* clade, indicating that hybridisation and/or biological processes like incomplete or differential homogenisation of the ITS may have obscured phylogenetic relationships and character evolution within this group (see discussion of gene tree incongruence in 2.4.3). The analyses of the ITS data elucidate the phylogenetic relationships of the small sections *Monopteron* and *Pleiothece* (Fig. 2.8-11). Section *Monopteron* includes two morphologically similar species, *Begonia nepalensis* and *B. griffithiana* Warb., which are endemic to the Himalaya region (Fig. 2.19). *Begonia nepalensis* is deeply nested within the *Platycentrum-Sphenanthera* clade in the ITS phylogenetic trees, and the rhizomatous habit and the two-locular capsules found in section

Monopteron, and a chromosome count of 2n = 22 for the second species in this section, *B. griffithiana*, are consistent with the same placement for this taxon. The monotypic section *Pleiothece* was erected for *Begonia balansana*, which exhibts an autapomorphic ovary and fruit syndrome with very unusual fleshy, 5-7-locular fruits. In the ITS phylogenetic trees *Begonia balansana* falls into a well supported clade, which also includes all other samples of sections *Platycentrum* and *Sphenanthera*, but also two Chinese species placed in section *Diploclinium* (*B. wenshanensis*, *B. alveolata*), and one species placed in the polyphyletic section *Leprosae* (*B. longicarpa*). Tebbitt et al. (2006) hypothesised that some of the earliest divergent lineages in the *Sphenanthera-Platycentrum* clade, including the capsule-fruited *Begonia wenshanensis*, *B. alveolata*, and *B. versicolor* Irmsch., exhibit fruit syndromes which are somewhat transitional between two-locular rain-ballist capsules and the typical three-locular, dry capsules of wind dispersed taxa as indicated by considerable variation in pericarp thickness, locule numbers, style numbers and fruit wing development within this group. Moreover, Irmscher (1939) observed elongated anther connectives, similar to those which characterize sections *Platycentrum* and *Sphenanthera*, in *Begonia yunannensis* H.Lév. (syn. *Begonia modestiflora* Kurz), which otherwise fit the circumscription of *Begonia* section *Diploclinium* well. The interpretation of this variation as putatively morphologically transitional between sections *Diploclinium* and *Platycentrum* s.l. is plausible, but further molecular data from independent sources like the plastome should be analysed to investigate whether there are indicators of hybridization, which could lead to similar variation, within this group.

Section *Monolobium*, which was not included in the analyses, is differentiated from section *Platycentrum* only by unilamellate placentae (Ku et al., 2007). However, the original description of the only species in the monotypic section, *Begonia wutaiana* C.I.Peng & Y.K.Chen, shows that the placenta and ovary locule development is variable in this species and both bilamellate placentae and three-locular ovaries have been described in this species, which was originally assigned to section *Platycentrum* (Peng et al., 2005). The short, stout rhizome, the coriaceous pericarp and the elongated connectives of the anthers strongly support this placement.

2.4.2.7 Paraphyly of section Coelocentrum *and polyphyly of section* Leprosae

Irmscher (1939) described section *Coelocentrum* to accommodate four Chinese species with uniloculate ovaries and a parietal placentation in the middle part of the ovary (Fig. 2.2 I). Since Irmscher's description numerous species exhibiting these conspicuous characters were discovered in Southeastern China and Northern Vietnam, and the recent *Flora of China* treatise lists 35 species in the section (Ku et al., 2007). Irmscher (1939) interpreted the unilocular ovaries and the parietal placentation in section *Coelocentrum* as a derived character within Asian *Begonia*, but Jin and Wang (1994), after detailed anatomical studies, linked the ovary syndrome in *Coelocentrum* to the parietal placentation known from the African section *Mezierea* (Gaud.) Warb., and interpreted it as ancestral within

Asian *Begonia*. Consecutive cross-sections of ovaries of species in section *Coelocentrum* show that septa are usually developed in the basal part of the ovary, which shows a typical three-locular ovary with axillary placentation, while in the middle part of the ovary a parietal placentation is developed (Jin and Wang, 1994; Ku, 2006; Wu and Raven, 2008). Each of the three parietal placentas consists of a stem, which is equivalent to one of the septa developed in the basal part of the ovary, but which does not radially reach the centre of the ovary. Two placenta branches adhere to this stem, each equivalent with one placenta branch of two adjacent locules developed at the base of the ovary. The ancestral character reconstructions presented here indicate that parietal placentation is not plesiomorphic, but likely derived from three-locular ovaries with axillary placentae. Parietal placentae are likely derived from axillary placentae in many groups of angiosperms by developmental inhibition of the radial growth of the septae of a coenocarp gynoecium, so that the placentae are developed at the periphery of the ovary (Leins, 2000). The results of the analyses of the cpDNA dataset indicate that species in section *Coelocentrum* form a well supported monophyletic group, which is the sister to the moderately to strongly supported clade which includes Philippine species of section *Diploclinium*, Malesian species of section *Reichenheimia*, and species of sections *Ridleyella*, *Bracteibegonia, Petermannia* and *Symbegonia*. Analyses of ITS data and morphological observations indicate further that *Begonia leprosa* and *B. cylindrica*, two of the three species assigned to the small section *Leprosae*, are closely related to section *Coelocentrum*, while a third species placed in section *Leprosae*, *B. longicarpa*, is more closely related to section *Platycentrum* s.l. (Tebbitt et al., 2006). The results of the ITS analyses presented here indicate that *Begonia leprosa* is not the sister taxon to section *Coelocentrum* but likely nested within this section (Fig. 2.8-11). This species shows three-locular ovaries with axillary, uni- or bilamellate placentae, and a thick, fleshy pericarp. The presence of this species nested within the *Coelocentrum* clade indicates that reversals from a unilocular ovary to a three-locular ovary state may have occurred in this group, which seems to be a small developmental step given that the ovaries of most *Coelocentrum* species exhibit a three-locular basal part with axillary placentation. It also indicates that fleshy fruits evolved independently within sections *Petermannia* and *Coelocentrum* (inclusive section *Leprosae pro parte*), species of both of which form subclades in the major Asian Clade B in the cpDNA phylogenetic trees (Figs. 2.4-7).

2.4.2.8 Phylogenetic relationships of section Ridleyella

Section *Ridleyella* includes only two species endemic to the Malay Peninsula (Fig. 2.19), one of which is considered to be extinct (Kiew, 2005). Irmscher (1929) separated *Begonia kingiana* and *B. eiromischa* Ridl. in section *Ridleyella* from other sections because of the unique character combination of a rhizomatous habit, two-locular ovaries and undivided placentae, and further hypothesized that section *Ridleyella* is derived from section *Reichenheimia*, whose species exhibit three-locular ovaries, but share a suite of vegetative and generative characters including a rhizomatous habit and undivided

placentae lamellae with the species assigned to section *Ridleyella*. In the cpDNA phylogenetic trees section *Ridleyella* falls in a clade in which subclades are formed by Malesian species assigned to section *Reichenheimia*, Malesian species assigned to section *Diploclinium*, and species assigned to sections *Bracteibegonia*, *Petermannia* and *Symbegonia*. However, the relationships between these subclades and *Begonia kingiana* are not resolved or only poorly supported in the cpDNA phylogenetic trees.

2.4.2.9 Paraphyly of section Petermannia *and phylogenetic relationships of sections* Bracteibegonia *and* Symbegonia

Begonia section *Petermannia* comprises with *c.* 270 species almost half of the *Begonia* species diversity in Southeast Asia, and the results of recent expeditions to under-collected areas on Borneo, Sulawesi, and New Guinea as well as studies of available herbarium material indicate that there are numerous morphologically distinct species awaiting description (Girmansyah, 2009; Hughes and Coyle, 2009; Hughes et al., 2009; Kiew and Sang, 2009; Thomas et al., 2009a; Thomas et al., 2009b; Thomas and Hughes, 2008). Section *Petermannia* has an almost exclusively Malesian distribution with only a few species extending the geographic range of the section to Thailand and Vietnam (Fig. 2.21) (Hughes, 2008). Shui and Chen (2004) placed three species from Southeast China in section *Petermannia*, but this placement was not followed in the *Flora of China* (Ku et al., 2007), and a combination of characters which are rare or absent in section *Petermannia*, including peltate leaves, distinctly monadelphous androecia, two-tepaled female flowers and apparently bisexual, protandrous inflorescences, indicate that at least the

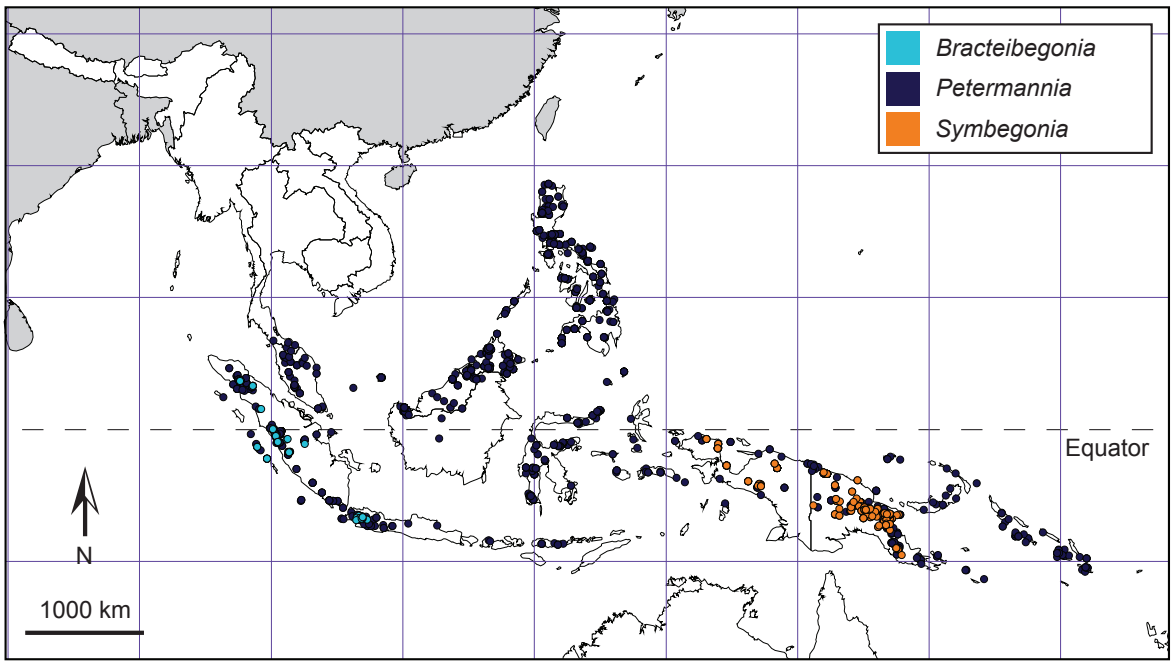


Fig. 2.21. *Begonia* distribution in Southeast Asia: sections *Bracteibegonia*, *Petermannia*, *Symbegonia*. Distribution data is based on GPS data and georeferenced locations in the *Southeast Asian Begonia Database* (7946 specimens from 57 herbaria; Hughes and Pullan, 2007). Areas shaded in grey (Sri Lanka, India except for the north-eastern states, Bangladesh, China) are not included in the database. GIS data was downloaded from the following sites: DIVA-GIS: http://www.diva-gis.org/Data (administrative boundaries); Natural Earth: http://www.naturalearthdata.com (10° longitude/latitude grid).

Chinese species *Begonia sinofloribunda* Dorr. is misplaced in section *Petermannia*.

The majority of species assigned to section *Petermannia* fall into a well supported subclade in Clade B of the cpDNA phylogenetic trees, but a small, well-supported clade of Bornean species including *Begonia amphioxus, B. burbidgei, B. pendula* Ridl., and three unidentified species (Fig. 2.4-6) is not included in this clade and has unresolved or poorly supported relationships within Clade B. However, both groups fall into a well supported, but poorly resolved clade in the ITS phylogenetic trees (Figs. 2.8-11), and this incongruence between the results of cpDNA and nrDNA datasets is discussed below (see 2.4.3). Analyses of nrDNA data (26S, ITS) indicate that species traditionally assigned to *Symbegonia*, a separate genus within the Begoniaceae, are nested within a Philippine and New Guinean clade of species assigned to *Begonia* section *Petermannia* (Forrest and Hollingsworth, 2003). This is corroborated by the results of the analyses of non-coding plastid regions presented here. Species in section *Symbegonia* lack rhizomes or tubers, as do the vast majority of species in section *Petermannia*, the two sections share a suite of generative characters including three-locular ovaries with bifid placentae, and most of the species exhibit protogynous, two- or sometimes one-flowered female inflorescences or partial inflorescences, which are basal to or not directly associated with the male inflorescences. Forrest and Hollingsworth (2003) proposed the recognition of *Symbegonia* at sectional level rendering the large section *Petermannia* paraphyletic, but retaining a morphologically distinct taxon. Species placed in section *Symbegonia* can be easily identified by a floral syndrome with a syntepalous perianth and a characteristic monadelphous androecium, and most species in this section exhibit characteristic endothecial cells of the anthers with faint or lacking endothecial thickenings (Tebbitt and MacIver, 1999). However, the presence of basally fused tepals in some species assigned to section *Petermannia* (Forrest and Hollingsworth, 2003; Sands, 2009) as well as transitions between the endothecial types found in *Symbegonia* and *Petermannia* (Tebbitt and MacIver, 1999) indicate that there are morphologically and anatomically transitional species between the two taxa.

Section *Bracteibegonia* includes only four species from Sumatra and Java, but there are several more species to be described from Sumatra (Hughes, 2008; Hughes et al., 2009). Species assigned to section *Bracteibegonia* form the well supported sister clade to the major *Petermannia* s.l. clade in the cpDNA phylogenetic trees. Some of the species assigned to section *Bracteibegonia* exhibit prostrate or weak stems, but rhizomes are not developed in this section, a derived character within Clade B of the cpDNA phylogenetic trees, which supports a close relationship of the two sections. Doorenbos et al. (1998) emphasised the general morphological similarities of the two sections, but they retained *Bracteibegonia* as a distinct section, because species of this section exhibit bisexual inflorescences which consist of one to few-flowered cymes and lack the clear separation of a basal female part and a distal male part, or separate female and male inflorescences,

which are characteristic for the vast majority of species in section *Petermannia* (Irmscher, 1914; Thomas et al., 2009b). The cpDNA phylogenetic trees clearly support the distinctness of section *Bracteibegonia*.

2.4.3 Incongruence of cpDNA and nrDNA gene trees

Reproductive barriers are notoriously weak in *Begonia*. Thousands of frequently fertile *Begonia* hybrids have been developed artificially in cultivation (Tebbitt, 2005), a considerable number of putative natural *Begonia* hybrids have been observed in the field (Burt-Utley, 1985; Kiew et al., 2003; Sosef, 1994; Teo and Kiew, 1999), and hybridization and allopolyploidy seem to be common in some regional *Begonia* floras such as the extensively studied *Begonia* flora of Taiwan (Chiang et al., 2001; Ku et al., 2007; Oginuma and Peng, 2002; Peng and Chiang, 1998; Peng and Chen, 1991; Peng and Chiang, 2000; Peng and Ku, 2009; Peng and Sue, 2000). However, the genus *Begonia* exhibits very high numbers of local, often morphologically very distinct endemics, widespread species are rare in the genus, and population-genetic studies indicate very strong population structures and limited dispersal capabilities of some species (Hughes and Hollingsworth, 2008; Hughes et al., 2003; Matolweni et al., 2000). This may explain to a certain extent, why natural hybrids have not been reported more frequently.

Phylogenetic incongruence between independent datasets like the nuclear ITS and the non-coding plastid region employed in this study can be indicative of hybridisation (Linder and Rieseberg, 2004), but several other biological processes like incomplete homogenization of the numerous copies in nrDNA arrays, the stochasticity of lineage sorting, recombination, gene paralogy, and pseudogene formation, can result in similar patterns (Álvarez and Wendel, 2003; Feliner and Rossello, 2007; Linder and Rieseberg, 2004; Small et al., 2004). The phylogenetic analyses of the 64 taxon plastid and nuclear datasets indicate considerable phylogenetic incongruence between the plastid phylogenetic tree and the ITS phylogenetic tree. While several major clades are congruently supported, there are several taxa which show conflicting positions in the two phylogenetic trees. However, direct comparison is confounded by the poor resolution, especially of the deeper relationships, of the ITS phylogenetic tree. Major contributing factors causing this poor resolution are high levels of sequence variability and associated alignment ambiguity and high levels of homoplasy in the ITS1 and ITS2 regions. Soft incongruence, i.e. weakly supported conflicting positions in the plastid and nuclear phylogenetic trees, is probably, to a large extent, the result of these confounding factors rather than other biological processes. However, there are also several instances of hard incongruence, i.e. strongly supported conflicting positions of taxa in the phylogenetic trees, especially in the species-rich sections *Platycentrum* s.l. and *Petermannia* s.l., which need further investigation. One instance of incongruence at a deeper level of the phylogeny involves four *Bornean* species placed in *Begonia* section *Petermannia*. In

the cpDNA phylogenetic tree *Begonia amphioxus, B. burbidgei* and two unidentified species (indicated with *PET1-4 in Fig. 2.12) form a well supported clade which is not included within the strongly supported clade comprising all other species placed in section *Petermannia*, as well as sections *Bracteibegonia* and *Symbegonia*. In the ITS phylogenetic tree these four species are not supported as monophyletic and are nested in three different subclades in a moderately to strongly supported but poorly resolved clade comprising all samples of sections *Petermannia*, *Bracteibegonia* and *Symbegonia* (Fig. 2.12). Species in the two separated *Petermannia* clades in the cpDNA phylogenetic tree are morphologically similar: they do not have rhizomes (except for one species from Sulawesi), which is a derived condition within Clade B of the plastid phylogenetic tree, they share characteristic protogynous, two- or sometimes one-flowered female inflorescences or partial inflorescences, which are basal to or separated from the male inflorescences, and species in both groups exhibit characteristic perforate cell wall thickenings in the endothecium layer of the anthers, while the endothecium cells of species of other Asian section predominantly show U-shaped wall thickenings (Tebbitt and MacIver, 1999). The inclusion of the four orphan species in the main *Petermannia* clade, as indicated by the ITS data, is concordant with the morphological and anatomical data, which supports a close relationship with other species placed in section *Petermannia*. The plastid and ITS gene trees seem to reflect different evolutionary histories, and the observed patterns seem to be consistent with an evolutionary scenario which involves the transfer of foreign plastids in a Bornean *Petermannia* lineage via hybridisation, introgression and plastid capture, and subsequent diversification on Borneo. However, only *Begonia amphioxus* and *B. burbidgei* are supported as monophyletic in the ITS phylogenetic tree, while the other two species fall in two other subclades. The relationships within the *Petermannia-Bracteibegonia-Symbegonia* clade are only poorly resolved or poorly supported in the ITS phylogenetic tree making further interpretation difficult, but the non-monophyly of these four taxa might hint at additional confounding factors including alignment ambiguity and homoplasy, incomplete homogenization or further hybridisation and differential homogenization of the ITS copies of Bornean species in section *Petermannia*. A second example of incongruence within the main *Petermannia* clade involves the Sulawesian species *Begonia capituliformis* indicated with *PET5 in Fig. 2. This species is an endemic to Northern Sulawesi and falls in a well supported clade with two other Northern Sulawesian species in the cpDNA phylogenetic tree, but falls in a clade with *B. rieckei* and species from Southwest and central Sulawesi in the ITS phylogeny. *Begonia rieckei* is a fleshy-fruited species which is unusually widespread on Sulawesi, while the vast majority of Sulawesian *Begonia* species in section *Petermannia* are local endemics. The more widespread distribution of *Begonia rieckei* makes sympatry and hence hybridisation with other *Petermannia* species more likely. The observed incongruence in the plastid and nuclear DNA phylogenetic trees might indicate hybridisation, but further analyses including data from multiple accessions, data from other nuclear markers, and cloning of nrDNA regions, which was beyond the scope of this study, are needed to achieve a

better resolved nuclear phylogeny and to investigate whether incomplete homogenization and incompletelineage sorting were factors which contributed to the incongruence found in this group, which is likely the result of a massive, relatively recent diversification on Sulawesi in the Plio- and Pleistocene (see Chapter 3).

Further incongruence can be detected in conflicting positions of taxa placed in section *Platycentrum* s.l. All samples of sections *Platycentrum* and *Sphenanthera* form well supported clades in both the plastid and the ITS phylogenetic trees, but some taxa like *Begonia sizemoreae* and *B. robusta* exhibit conflicting phylogenetic position within these clades. The most conspicuous incongruence in the *Platycentrum-Sphenanthera* clade involves *Begonia robusta*, which is indicated with *SPH1 in the cpDNA and ITS phylogenetic trees in Fig. 2.12. In the analyses of the cpDNA dataset *Begonia robusta* falls into a clade with *B. multangula*, an unidentified species with strong morphological affinities to *B. multangula* from Sulawesi, and *B. areolata*. *Begonia robusta* and *B. multangula* are morphologically most similar and are characterised by fleshy fruits and a short, stout rhizome from which well developed leafy stems arise. Hughes (2008) pointed out that some specimens are morphologically intermediate between the typical forms of these two species, and maybe of hybrid origin. *Begonia areolata* exhibits two-locular capsules, but a somewhat similar growth habit as *B. multangula*. *Begonia robusta* falls into a well supported clade with *B. decora*, *B. pavonina*, and *B. venusta*, all of which show two-locular rain-ballist capsules, in the ITS phylogenetic tree. *Begonia multangula*, the unidentified morphologically similar species from Sulawesi, and *B. areolata* are not included in this clade and can be found in two subclades within the *Platycentrum-Sphenanthera* clade of the ITS phylogenetic tree. Overall morphology is more concordant with the clade retrieved in the cpDNA analyses, and the separation of the morphologically most similar *Begonia robusta*, *B. multangula*, and *B.* aff. *multangula* from Sulawesi in three different clades in the ITS phylogenetic tree indicates that the ITS gene tree may not accurately reflect the species tree. A tentative somatic chromosome count of 2n = 88 for *Begonia robusta* (Doorenbos et al., 1998), while the majority of species in *Platycentrum* s.l. show a somatic chromosome of 2n = 22, indicates that allopolyploidy might have played an important role in the evolution of this group. The incorporation of xenologous loci in the nuclear genome through hybridization, and subsequent incomplete homogenization and/or different directions of homogenization to either parental species can lead to complex patterns as observed in the ITS phylogenetic tree (Álvarez and Wendel, 2003; Cronn et al., 2003; Sang et al., 1995). However, studies using multiple accessions and cloning efforts are needed to test whether different homogenization directions can be detected and whether divergent intragenomic ITS copies are present in the nrDNA arrays.

The results highlight the importance of using multiple independent sources of phylogenetic data to detect discrepancies between gene and species trees in *Begonia*. The considerable incongruence between the plastid and the nuclear marker phylogenetic trees

presented here may be partially explained by hybridisation. However, other biological processes, like incomplete homogenization of the numerous copies in nrDNA arrays, the stochasticity of lineage sorting, recombination, gene paralogy, and pseudogene formation, can result in similar patterns (Álvarez and Wendel, 2003; Feliner and Rossello, 2007; Linder and Rieseberg, 2004; Small et al., 2004), and further research is needed to identify and disentangle the relative impacts of these processes in causing incongruence in phylogenetic reconstructions of *Begonia*.

2.4.4 Conclusions

The species in the mega-diverse genus *Begonia* exhibit an enormous vegetative diversity and even closely related species often show conspicuous differences in growth habits, indumentum characters and leaf morphologies. It is likely that this morphological diversity arose due to both genetic drift and natural selection for adaptations to specific habitat conditions (Kidner and Umbreen, in press; Matolweni et al., 2000; Neale et al., 2006). In contrast to this, generative organs provide easily observable, qualitative and quantitative, and apparently relatively complex characters such as differences in carpel and ovary locule numbers, placentation types, and pericarp types, which have been crucial for the circumscription of sections in Asian *Begonia*. The results of the phylogenetic analyses and ancestral character reconstructions presented here indicate that apparently complex fruit syndromes like three- or four-locular indehiscent fruits with fleshy pericarps, and two-locular rain-ballist capsules evolved multiple times independently in Asian *Begonia*. The genetic-developmental complexity, an essential criterion in character homology assessments, of the gain or loss of carpels, the inhibition of the development of locules in a three-locular ovary, the development of unilamellate or bilamellate placentae, and the development of a fleshy pericarp seem to have been overestimated in the past. The extensive homoplasy of these characters has confused systematic relationships in Asian *Begonia* and most Asian sections are not supported as monophyletic in the phylogenetic analyses. The strong systematic emphasis placed on single, homoplasious characters like undivided placenta lamellae (section *Reichenheimia*), fleshy pericarps (section *Sphenanthera*), and the recognition of sections primarily based on a plesiomorphic fruit syndrome and the absence of characteristic features of other taxa (section *Diploclinium*) has resulted in the circumscription of several highly polyphyletic taxa. The results indicate further that the presence and absence and type of stem metamorphoses and perennating organs like tubers and rhizomes and correlated growth habits are of greater systematic value in Asian *Begonia* than has been assumed in the past. These vegetative characters allow us to differentiate monophyletic predominantly Malesian species groups in both sections *Reichenheimia* and *Diploclinum* from distantly related Indian, Sri Lankan and continental Asian species placed in these polyphyletic sections. The Malesian taxa are characterized by usually plagiotropically growing rhizomes, which directly produce leaves and inflorescences at their nodes, while the distantly related Indian, Sri Lankan and continental Asian species

placed in sections *Diploclinium* and *Reichenheimia* predominantly exhibit tubers which produce erect leafy stem, or, in acaulescent species, directly leaves and inflorescences. Several Chinese rhizomatous species which were placed in sections *Reichenheimia* and *Diploclinium* were shown to be more closely related to the predominantly Chinese, rhizomatous section *Coelocentrum*, while others seem to be more closely related to section *Platycentrum* s.l. However, detailed comparative morphological and anatomical studies are needed to further investigate the homology and systematic importance of these organs in *Begonia*. Moreover, their ecological significance as perennial organs, anchor organs, and in habitat occupation and clonal reproduction are only poorly understood and studies are needed to establish whether there are correlations between precipitation, seasonality, substrate types and other environmental factors and the development of tubers, clusters of tubers and rhizomes.

Previous phylogenetic and population-genetic studies have indicated that most *Begonia* species seem to have low dispersal capabilities and a strong geographic structure has been observed at different geographic and taxonomic levels reaching from the continental, sectional level (Goodall-Copestake, 2005; Plana, 2003; Plana et al., 2004) to the regional, population-genetic level (Hughes and Hollingsworth, 2008; Hughes et al., 2003; Matolweni et al., 2000). The cpDNA phylogenetic trees of Asian Begonias exhibit the same trend, and indicates a stronger geographic structure than previously suggested. Clade A comprises taxa like sections *Parvibegonia*, and *Platycentrum* s.l. (incl. section *Sphenanthera*) and continental Asian lineages in the polyphyletic sections *Diploclinium*, which are most diverse on the Asian mainland, although several smaller Malesian lineages can also be recognized. Clade B comprises the predominantly Chinese section *Coelocentrum*, several predominantly or exclusively Malesian sections (*Ridleyella*, *Bracteibegonia*, *Petermannia* s.l.) and Malesian species groups in both sections *Reichenheimia* and *Diploclinum* that are only distantly related to Indian, Sri Lankan and other continental Asian species which were placed in these two highly polyphyletic sections. The strong geographic structure in cpDNA phylogeny and the general preponderance of narrow endemics in the genus highlight the potential of *Begonia* as model group to study the biogeography of the Southeast Asian tropical region.

The current artificial infrageneric classification of Asian *Begonia* has a certain diagnostic, but only poor predictive value, and has hampered the understanding of the evolution of morphological and anatomical characters, karyotypes and the ecology and biogeography of Southeast Asian *Begonia*. The phylogenetic trees derived from non-coding plastid data provides for the first time a reasonably resolved, supported phylogenetic framework for Asian *Begonia*, which has the power to inform taxonomic work, and evolutionary and biogeographical studies. This cpDNA phylogenetic framework represents an important step towards a natural re-classification of Asian *Begonia*, and it clarifies the phylogenetic relationships, and aspects of the character evolution and karyotype evolution within this

species-rich and morphologically diverse group. However, it also indentifies several problematic groups including Indian, Sri Lankan and continental Asian lineages of species placed in the polyphyletic sections *Diploclinium* and *Reichenheimia*, whose phylogenetic relationships are still only fragmentarily understood. Before the aim of a comprehensive, stable and natural re-classification can be achieved, it will be essential to further elucidate the phylogenetic relationships of these groups using morphological and phylogenetic studies based on a geographically robust taxon sampling and multiple independent sources of molecular data.

CHAPTER 3. Historical biogeography of Southeast Asian *Begonia* L. (Begoniaceae)

Chapter summary

The biogeography of the species-rich and geologically complex phytogeographic region of Malesia has intrigued biologists since Alfred Russel Wallace's seminal zoogeographical studies in the nineteenth century. Wallace identified major faunistic differences between neighbouring central Malesian islands such as Bali and Lombok and emphasized the impact of past geological connections and past biotic migrations on current distribution patterns. Many of Wallace's hypotheses were supported by research in the 20th and 21st century, and the multitude of islands and island assemblages in Malesia differ indeed greatly in their origin, age and their past land connections, as well as in the composition of their biota. However, despite the long-standing interest in the biogeography and geology of the Malay Archipelago, the relative contributions of overland migration, rafting on tectonic fragments, long-distance dispersal, vicariance, and autochthonous speciation to the composition of Malesian island biota are still poorly understood. Few phylogenetic studies have investigated the temporal and spatial diversification patterns of taxa whose distributions span the wider archipelago.

This study focuses on the mega-diverse genus *Begonia* (> 1550 spp.) to explore questions about Malesian biogeography. Bayesian phylogenetic and relaxed molecular clock analyses of a Cucurbitales-Fagales datatset (92 taxa; cpDNA: *mat*K gene, *rbc*L gene, *trn*L intron, *trn*L-F spacer; five fossil calibrations) and a Begoniaceae datatset (110 taxa; cpDNA: *ndh*A intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer; two alternative secondary calibrations), as well as likelihood and Bayesian ancestral area reconstructions were employed to elucidate temporal and spatial diversification patterns in Asian *Begonia*, the temporal and spatial origin of Malesian *Begonia* lineages, and the timing and frequency of crossings of purported dispersal barriers like the ancient deep water channels separating the Sunda Shelf region (Peninsula Malaysia, Borneo, Sumatra, Java, Bali) from Wallacea (Sulawesi, the Maluku Islands, the Lesser Sunda Islands east of Bali).

The analyses suggest an initial diversification of Asian *Begonia* on the Indian subcontinent and in continental Southeast Asia in the Middle Miocene, and subsequent colonization of Malesia by multiple lineages. The predominant directional trend of dispersals between continental Asia and Malesia and within Malesia is from west to east including four independent dispersal events from continental Southeast Asia and the Malesian Sunda Shelf region to Wallacea dating from the Late Miocene to the Pleistocene. Dispersal across the ancient deep water channels separating intervening islands of the Sunda

Shelf and Wallacea and subsequent successful colonisation of Wallacean islands seem to have been infrequent events during this period, suggesting that the water bodies which have separated the Sunda Shelf region from Wallacea have been distinct, yet porous barriers to dispersal in *Begonia*. The inferred timing of dispersals from the Sunda Shelf region to Wallacea is generally concordant with hypotheses about the geological history of the region, which indicate that the period from the Late Miocene onwards offered opportunities for dispersal to Wallacea and across Wallacea to New Guinea as substantial land masses emerged in Sulawesi and New Guinea, and newly emergent volcanic islands along the Sunda Arc, the Banda Arc and the Halmahera Arc formed potential routes for dispersal by island hopping.

The results further suggest that *Begonia* section *Petermannia* (>270 spp.) originated in the Malesian Sunda Shelf region, and subsequently dispersed to Wallacea, New Guinea and the Philippines. Lineages within this section diversified rapidly since the Pliocene with diversification peaking in the Pleistocene. The timing of diversifications coincides with orogenesis on Sulawesi and New Guinea, as well as pronounced glacioeustatic sea-level and climate fluctuations. It can be hypothesised that a complex interplay of extrinsic and intrinsic factors including the presence and formation of suitable microhabitats by orogenesis, cyclical vicariance by frequent habitat fragmentations and amalgamations caused by sea-level and climate fluctuations, as well as only weakly developed mechanisms to maintain species cohesion in fragmented habitats in *Begonia* could have driven speciation in allopatry and could have resulted in the remarkable *Begonia* species diversity found in Southeast Asia today.

3.1 Introduction

The phytogeographic region of Malesia, also known as the Malay Archipelago, extends from southern Thailand through Malaysia, Singapore, Indonesia and East Timor, to the Philippines, Papua New Guinea and the Solomon Islands (Raes and van Welzen, 2009). It is one of the three regions in the world with extensive areas of tropical rainforest and comprises biodiversity hotspots like the Sunda Shelf region, Wallacea, the Philippines, and New Guinea, harbouring an estimated 42,000 species of vascular plants (Brooks et al., 2006; Myers et al., 2000; Roos, 1993). Within the geographic range of Malesia three floristic subregions can be differentiated based on species-level floristic similarities: 1. The everwet Sunda Shelf area in the west comprising Peninsula Malaysia, Sumatra and Borneo; 2. The everwet Sahul Shelf area in the east including New Guinea and adjacent island groups; and 3. Wallacea, a biotic interface region in the heart of the Malay Archipelago comprising the Lesser Sunda Islands, Sulawesi, the Moluccas and the Philippines (Fig. 3.1) (van Welzen and Slik, 2009; van Welzen et al., 2005). These three floristic subregions largely correlate with the geological history of Southeast Asia,

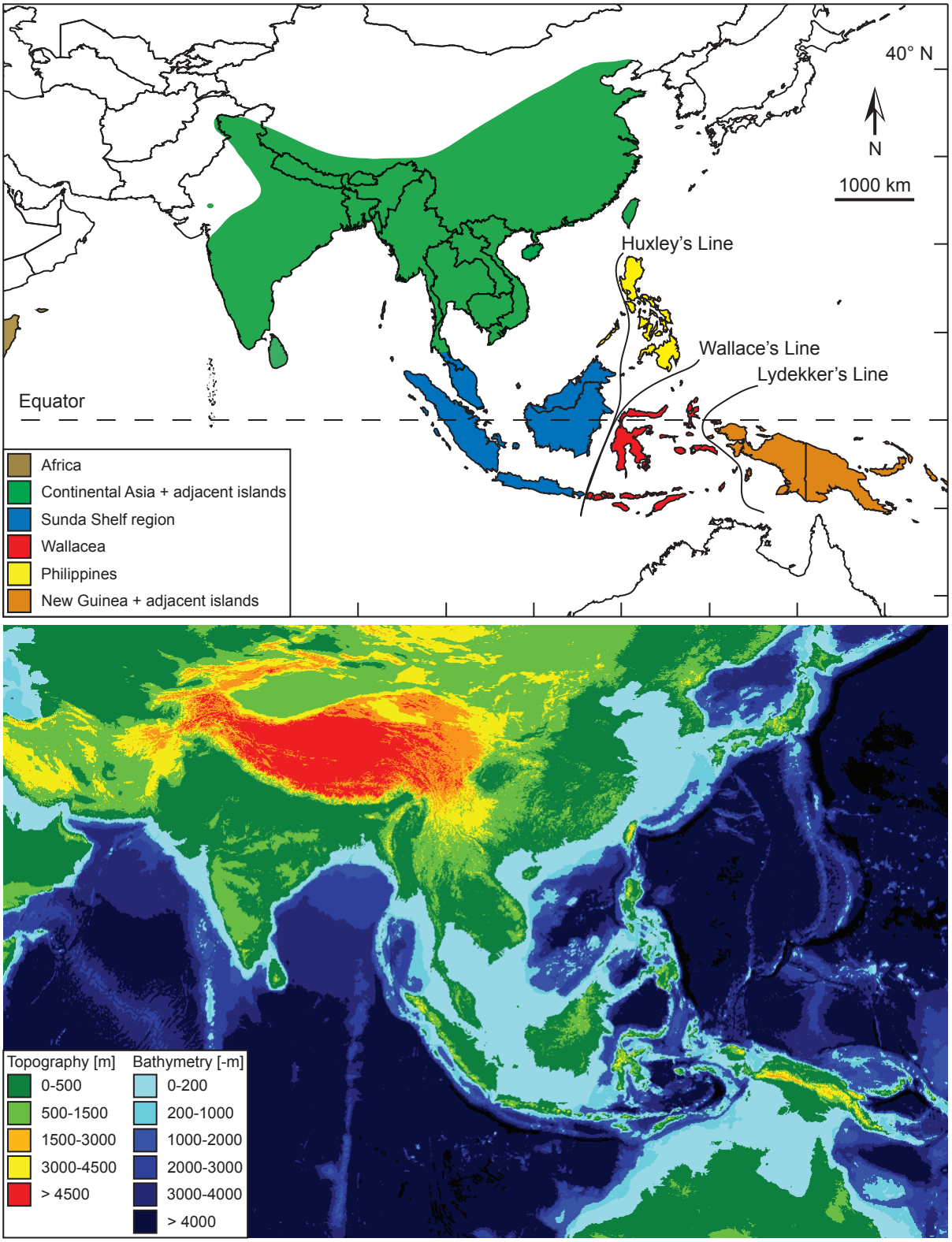


Fig. 3.1. *Begonia* distribution in Asia and delimitation of geographic regions used in ancestral area reconstructions. *Begonia* distribution based on data from the Global Biodiversity Information Facility (http://www.gbif.org/), the South East Asian *Begonia* Database (Hughes and Pullan, 2007), the *eFlora of China* and the *eFlora of Pakistan* (eFloras, 2008), Uddin (2007), Kumar (1993) and Parmar (1987). The upper map shows the geographic regions used in the ancestral area reconstructions: Africa and the Yeminite Socotra Archipelago; Continental Asia and adjacent major islands (Sri Lanka, Taiwan, Hainan); the Sunda Shelf region (southern Thailand, Peninsula Malaysia, Sumatra, Borneo, Java, Bali); Wallacea (Sulawesi and adjacent islands, the Lesser Sunda Islands and the Maluku Islands between Wallace's and Lydekker's Line); the Philippines inclusive Palawan; New Guinea and adjacent islands (islands east of Lydekker's Line, the Bismarck Archipelago, the Solomon Islands). The lower map shows the topography of the region. GIS data was downloaded from the following sites: DIVA-GIS: http://www.diva-gis.org/Data (administrative boundaries); CGIAR Consortium for Spatial Information: http://srtm.csi.cgiar.org/ (topographic data); Natural Earth: http://www.naturalearthdata.com/ (bathymetric data, longitude/latitude grid).

except for the position of the Indonesian island of Java, which geologically belongs to the Sunda Shelf region, but is floristically associated with the islands of Wallacea, large parts of which like most of Java exhibit a dryer monsoonal climate (Hall, 2009; van Welzen et al., 2005).

The biogeography of the remarkably species-rich Malesian region has intrigued biologists since Alfred Russel Wallace's seminal zoogeographical studies in the 19th century. Wallace identified major faunistic differences between neighbouring central Malesian islands such as Bali and Lombok and emphasized the impact of past geological connections and past biotic migrations on current distribution patterns (George, 1981; Lomolino et al., 2006; Wallace, 1860, 1863, 1869, 1876; Whitten et al., 2002). Many of Wallace's ideas were supported by research in the 20th and 21st century, and the multitude of islands and island assemblages in Malesia indeed differ greatly in their origin, age and their past land connections, as well as in the composition of their biota (Hall, 2002, 2009; van Welzen and Slik, 2009; van Welzen et al., 2005). However, despite the long-standing interest in the biogeography and geology of the Malay Archipelago, the origins and biogeographical affinities of its biota have remained enigmatic. The relative contributions of overland migration, rafting on tectonic fragments, long-distance dispersal, vicariance, and autochthonous speciation to the composition of Malesian island biota are still only poorly understood. Few phylogenetic studies have investigated the temporal and spatial diversification patterns of taxa whose distributions span the wider archipelago (Alfaro et al., 2008; Evans et al., 2003; Muellner et al., 2008; Su and Saunders, 2009).

This study focuses on the mega-diverse genus *Begonia* (> 1550 spp.) to explore questions about Malesian biogeography. The great potential of this genus with regards to biogeographical analyses lies in its wide distribution in Asia, which spans the biotic interface of Wallacea as well as the wider Malay Archipelago and large parts of continental Asia; in its good representation with regards to species numbers on all major islands and island groups in Malesia (> 540 spp. in Southeast Asia); in the preponderance of narrow endemics limited to primary habitats and the low dispersal capabilities of most *Begonia* species as indicated by population genetic studies (Hughes and Hollingsworth, 2008); in the high geographic structure observed in studies at different geographic and taxonomic levels reaching from the continental, sectional level (Goodall-Copestake, 2005) to the local, population-genetic level (Hughes and Hollingsworth, 2008; Hughes et al., 2003; Matolweni et al., 2000). This study uses phylogenetic analyses, reconstructions of ancestral areas of distribution and molecular divergence age estimates of Southeast Asian *Begonia* to disentangle the impacts of earth-driven processes which create or eliminate physical barriers and can lead to vicariance and geodispersal (sensu Liebermann, 2005), and stochastic processes such as long-distance dispersal, on current distribution patterns in Malesia. Pattern congruence between the sequence and timing of geological events and the sequence and timing of cladogenetic events indicates a key role of geology in

shaping current distribution patterns. However, if congruence at either the temporal or the spatial level is not evident, key roles of vicariance and geodispersal are not supported, and hypotheses based on other explanatory processes such as long-distance dispersal across physical barriers must be evoked (Rutschmann et al., 2004). By correlating the phylogeny, divergence ages, and reconstructions of ancestral areas of Southeast Asian *Begonia* with knowledge about the palaeogeography of the region, this study investigates the following topics:

1. Colonization of the Malay Archipelago: There is some support for the hypothesis that Southeast Asian *Begonia* first diversified in Continental Southeast Asia and subsequently colonized the Malay Archipelago. A molecular phylogenetic study, analysing eleven DNA regions of a low-density world-wide taxon sample of the genus, indicates that Socotran and Asian Begonias form a strongly supported clade (Goodall-Copestake, 2005). Within this Socotran-Asian clade, the Socotran species *Begonia socotrana* is the sister taxon to the Asian clade, and within the Asian clade South Indian and Sri Lankan species form the sister clade to the rest of the clade. However, the deepest nodes in the Socotran-Asian clade received only weak statistical support. Previous molecular divergence age estimates indicate that migration from Africa to Asia likely occurred in the early Miocene (*c.* 20 Ma), possibly by an overland migration route via the Arabian Peninsula, which might have been hospitable for Begonias during moist and humid phases in the Miocene, or by sweepstake dispersal from Africa to Western Asia (Goodall-Copestake, 2005). The molecular divergence age estimates put the diversification of Asian *Begonia* into a time frame long after initial collision and final suturing of the Indian Plate and the Eurasian Plate *c.* 50 Ma ago (Beck et al., 1995; Briggs, 2003; Morley, 2000). If the Asian *Begonia* lineage initially diversified in Western and South Asia, when land connections between India and the Southeast Asian mainland were already established, progressive overland range expansions may have led to colonization of and diversifications in suitable habitats on the Southeast Asian mainland and subsequent colonization of the Malay Archipelago from west to east. Phylogenetic trees and ancestral area reconstructions which indicate that Malesian clades are derived from Southeast Asian mainland ancestors would corroborate this hypothesis. However, secondary migration from Malesia to continental Southeast Asia, as well as long-distance dispersal from India to Malesia, and subsequent dispersal to continental Asia may have resulted in much more complex scenarios than the simple progression scenario outlined above. While the Indian Plate continued drifting northwards after the initial collision with the Eurasian Plate it came relatively close to Western Malesian areas (Briggs, 2003; van Welzen et al., 2005), and palaeopalynological and phylogenetic studies indicate an exchange of floral elements between India and Malesia and especially from India to Southeast Asia as early as the Middle Eocene, maybe facilitated by the establishment of a moist rainforest corridor between India and Southeast Asia (Morley, 2000, 2003; van Welzen et al., 2005). Phylogenetic trees and ancestral area reconstructions which indicate that continental Southeast Asian taxa are derived from

Malesian ancestors would support hypotheses involving initial diversification in Malesia and subsequent colonization of the Southeast Asian mainland.

2. Colonization of Wallacea and the crossing of purported biogeographical boundaries in Malesia: Several deep water channels and marine basins have purportedly acted as physical biogeographic barriers or filters in Malesia hindering range expansion of taxa with poor dispersal capabilities. The most prominent of these is the ancient deep water channel of the Makassar Straits separating Borneo and Sulawesi that coincide with Wallace's Line, the well-known demarcation of the faunistic divide which Alfred Russel Wallace observed in central Malesia (Fig. 3.1) (Simpson, 1977; Wallace, 1860, 1869; Whitten et al., 2002). The Makassar Straits remained a barrier even at times of low sea-levels caused by glacioeustatic fluctuations associated with the change of ice volume in the Northern Hemisphere during the Pleistocene, when the Western Malesian islands and the surrounding continental shelf formed a vast landmass, and a major landbridge was present between New Guinea and the Australian continent (Hall, 2009; Lomolino et al., 2006; Voris, 2000). Wallace's Line also demarcates the western border of Wallacea, the biotic interface region in central Malesia where the western Malesian Sunda Shelf flora and the Australian-New Guinean flora meet. The eastern border of Wallacea is located west of New Guinea and demarcated by another prominent biogeographical division, Lydekker's Line, which was set to mark the western boundary of a strictly Australian fauna (Lydekker, 1896; Simpson, 1977). According to van Welzen et al. (2005) and van Welzen and Slik (2009) the whole of Wallacea acted as biogeographic barrier or filter between the Sunda Shelf region and the New Guinean-Australian region. Reasons for this are that, firstly, the continental fragments which constitute parts of Wallacea migrated to their current position only during the Cenozoic, when Western Malesia was largely in place already (Hall, 2009); secondly, the fragments which formed Wallacea were submerged during most of their migration and substantial land emerged in Sulawesi only from the middle of the Miocene onwards (Hall, 2009); thirdly, there were no major land-bridges in Wallacea even at Pleistocene times of maximum glaciations causing substantial lowering of sea-levels (Voris, 2000); fourthly, large parts of Wallacea exhibit a dryer monsoonal climate, while the Sunda Shelf and New Guinea have a predominantly everwet climate (van Welzen and Slik, 2009; van Welzen et al., 2005). Crucial questions with regards to the biogeography of Wallacea are:

- When, by which paths and how frequently were purported barriers crossed and newly available land in the interface region of Wallacea colonized?
- Did rafting on tectonic fragments contribute to the composition of the Wallacean flora?

3. Origin and diversification of *Begonia* section *Petermannia*: The largest Asian *Begonia* section, section *Petermannia*, comprises with more than 270 species almost half of the *Begonia* species diversity in Southeast Asia. It exhibits an almost exclusively Malesian

distribution with only four species extending the geographic range of the section to Thailand, Vietnam and possibly Southeastern China (Fig. 2.21) (Hughes, 2008; Shui and Chen, 2004). Despite a decrease of sectional diversity from the west to the east of Malesia, eastern Malesian areas such as the Philippines and New Guinea show *Begonia* floras which together with the *Begonia* flora of Borneo are the most species-rich in Asia. This is mainly based on the large numbers of species in *Begonia* section *Petermannia* (Hughes, 2008). The temporal and spatial origin of this section, and the timing of and the factors contributing to the remarkable diversification in *Begonia* section *Petermannia* are unknown. Key questions with regards to this topic are:

- Where and when did the species-rich section *Petermannia* originate?
- Which factors contributed to the massive radiations in this section resulting in large numbers of predominantly narrow endemics?

3.2 Material and Methods

3.2.1 Molecular divergence age estimates

3.2.1.1 Methodological approach

To evaluate the impact of climatic and geological events on the evolution of Southeast Asian *Begonia*, a timeframe provided by molecular divergence age estimates is crucial. However, molecular divergence age estimation for Begoniaceae is problematic, as Begoniaceae fossils are absent from the known pre-Pleistocene fossil record and suitable fossils for direct calibration are lacking. Former studies have addressed this problem by using either island emergence dates as calibrations (Plana, 2002; Plana et al., 2004) or by putting Begoniaceae in a wider phylogenetic context using suitable fossils calibrations in related taxa (Clement et al., 2004, 2005; Goodall-Copestake, 2005; Goodall-Copestake et al., 2009). One of the primary aims of the study presented here was to check for the temporal coincidence of diversifications and geological events in Southeast Asia. Therefore, island emergence dates or other geological calibrations, which assume that a geological event caused a divergence and enforce temporal congruence between these events, are not used in this study to avoid the potential pitfall of circularity (Renner, 2005). This study applied a two-stage approach using fossil calibrations. In the first stage (Stage 1, Cucurbitales-Fagales dataset), a 92 taxon dataset including Begoniaceae and related taxa in the orders Cucurbitales and Fagales, and data of well-conserved to medium variable cpDNA regions (*mat*K gene, *rbc*L gene, *trn*L intron, *trn*L-F spacer), was analysed. This allowed the integration of multiple fossil based calibration points, and to estimate the divergence ages of the Begoniaceae and *Begonia* crown groups. In the second stage (Stage 2, Begoniaceae dataset), the divergence age estimates of the Begoniaceae and *Begonia* crown groups inferred in Stage 1 were applied as secondary calibration points. A 110 taxon Begoniaceae data matrix including data from three fast evolving non-coding chloroplast regions (*ndh*A

intron, *ndh*F-*rpl*32 spacer, *rpl*32-*trn*L spacer) was analysed to estimate the divergence ages of major Asian lineages.

3.2.1.2 Taxon sampling

The Cucurbitales-Fagales dataset analysed in Stage 1 comprised 92 taxa representing a low-density sampling of all families of the orders Fagales and Cucurbitales (sensu Angiosperm Phylogeny Group, 2009). Thirty-one species of Cucurbitaceae and 24 species of Begoniaceae, which broadly cover the known major taxonomic lineages and the geographic ranges of these two larger Cucurbitales families, were included.

The Begoniaceae dataset analysed in Stage 2 consisted of sequences of 110 Begoniaceae taxa including the monotypic sister genus of *Begonia*, *Hillebrandia*, and 6 African and 103 Asian taxa. Sampling of Asian species included samples of all major Asian sections of *Begonia* (sections *Coelocentrum*, *Diploclinium, Parvibegonia*, *Petermannia, Platycentrum, Reichenheimia, Sphenanthera, Symbegonia*) as well as the small Asian *Begonia* sections *Bracteibegonia*, *Haagea*, and *Ridleyella*.

3.2.1.3 DNA region sampling

For Stage 1, the data matrix included sequences of four conserved to medium variable protein-coding and non-coding cpDNA regions: the maturase K coding gene (*mat*K), the large subunit of the ribulose-1,5-bisphosphate carboxylase/oxygenase gene (*rbc*L), the *trn*L intron (*trn*L) and the *trn*L-F intergenic spacer region (*trn*L-F), all of which have been used to elucidate phylogenetic relationships within Cucurbitales and Cucurbitaceae (Schaefer et al., 2009; Zhang et al., 2006) and Fagales (Cook and Crisp, 2005; Li et al., 2004). These four regions were chosen to achieve a robust Fagales and Cucurbitales taxon sampling based on published data, to maximize sequence length while minimizing ambiguous alignment positions and ultimately achieve resolution of the interfamiliar and major infrafamiliar relationships. DNA sequences were downloaded from the nucleotide database of the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), 46 *Begonia* sequences by Goodall-Copestake (2005) were used, and 46 *rbc*L and *trn*L-F sequences were newly generated. Voucher information and Genbank numbers can be found in Appendices 2 and 3, respectively.

For Stage 2, the datamatrix comprised sequences of three fast evolving non-coding cpDNA regions: the *ndh*A intron, the *ndh*F-*rpl*32 spacer, and the *rpl*32-*trn*L spacer (Shaw et al., 2007). These three regions are distinctly more variable than other chloroplast regions used in molecular *Begonia* systematics (*mat*K gene, *pet*D gene and intron, *psb*B gene, $trnL_{UAA}$ intron), and have considerable utility for *Begonia* systematics at the inter- and infrasectional level (see Chapter 2). All sequences of these three non-coding regions were newly generated. Voucher information can be found in Appendix 2.

3.2.1.4 DNA extraction, amplification and sequencing

DNA extraction, PCR master mix composition, purification, amplification of the *ndh*A intron, and the *ndh*F-*rpl*32 and *rpl*32-*trn*L spacers, and sequencing protocols are the same as described in the methods sections of Chapter 2. Amplification of the *rbc*L gene was carried out using the primers z-1 and z1375R (Clement et al., 2004). The PCR temperature profile used was the same as in Clement et al. (2004): template denaturation at 95°C for 4 min followed by 35 cycles of denaturation at 95°C for 1 min, primer annealing at 55°C for 1 min, and primer extension at 72°C for 1 min; followed by a final extension step at 72°C for 4 min. The *trn*L-F spacer region was amplified using the universal primers e and f (Taberlet et al., 1991). The PCR temperature profile used was the same as in Zhang and Renner (2003): template denaturation at 94°C for 3 min followed by 30 cycles of denaturation at 94°C for 1 min, primer annealing at 53°C for 1 min, and primer extension at 72°C for 2 min; followed by a final extension step at 72°C for 7 min. All primers used in this study are listed in Table 3.1.

3.2.1.5 Alignment

Sequences were assembled and edited using *GeneiousPro* v4.8.2 (Drummond et al., 2010). Plastid DNA sequences of the Cucurbitales-Fagales dataset (*mat*K, *rbc*L, *trn*L intron, *trn*L-*trn*F spacer) were aligned using the multiple sequence alignment software *MUSCLE* (Edgar, 2004) implemented in *GeneiousPro* using default settings, and subsequently manually modified in *GeneiousPro*. Exon sequences of the *mat*K were translated into amino acid sequences using the translate function in *GeneiousPro* to check for the presence of pseudogenes. All *mat*K sequences exhibited open reading frames indicating that only functional genes were included in the dataset. Alignment posed few difficulties, but a highly variable region of the *trn*L intron (14-65 bp) was removed due to uncertain homology. The region of an inversion (24-44 bp) characterising all Cucurbitaceae taxa except *Austrobryonia argillicola* I.Telford and *Neoalsomitra sarcophylla* (Wall.) Hutch. was removed from the *trn*L-F spacer alignment of all 92 taxa (see Kocyan et al., 2007).

Alignment of the Begoniaceae dataset was performed as described in the methods section

Table 3.1. Primers used in this study.

DNA Region	Primer	Primer Sequence (5'–3')	Source
*ndh*A intron	ndhAx1	GCY CAA TCW ATT AGT TAT GAA ATA CC	Shaw et al., 2007
	ndhAx2	GGT TGA CGC CAM ARA TTC CA	Shaw et al., 2007
*ndh*F-*rpl*32	rpL32-R	CCA ATA TCC CTT YYT TTT CCA A	Shaw et al., 2007
	ndhF	GAA AGG TAT KAT CCA YGM ATA TT	Shaw et al., 2007
	Beg1F	TGG ATG TGA AAG ACA TAT TTT GCT	this study
	Beg2R	TTT GAA AAG GGT CAG TTA ATA ACA A	this study
*rbc*L	z-1	ATG TCA CCA CAA ACA GAA ACT AAA GCA AGT	Clement et al., 2004
	z1375R	AAT TTG ATC TCC TTC CAT ATT TCG CA	Clement et al., 2004
*rpl*32-*trn*L	$trnL^{(UAG)}$	CTG CTT CCT AAG AGC AGC GT	Shaw et al., 2007
	rpL32-F	CAG TTC CAA AAA AAC GTA CTT C	Shaw et al., 2007
*trn*L-F	e	GGT TCA AGT CCC TCT ATC CC	Taberlet et al., 1991
	f	ATT TGA ACT GGT GAC ACG AG	Taberlet et al., 1991

of chapter 2. The position and the length of inversions and excluded sites are shown in Table 3.2.

3.2.1.6 Bayesian divergence age estimation

Bayesian divergence time estimation was performed using the uncorrelated relaxed lognormal clock model implemented in the software package *BEAST* v1.5.3 (Drummond and Rambaut, 2007). In contrast to other methods of divergence time estimation, e.g. the Non-Parametric Rate Smoothing (NPRS) and Penalized Likelihood (PL) methods implemented in the software *r8s* (Sanderson, 2003), *BEAST* does not require a user-specified tree topology as input, but directly uses sequence data and calibration constraints or DNA substitution rates to generate estimates of topology and divergence ages simultaneously, accommodating topological uncertainty in the estimation of all parameters including divergence ages (Drummond et al., 2006; Drummond et al., 2007). Moreover, in contrast to NPRS and PL, the uncorrelated relaxed molecular clock methods implemented in *BEAST* does not work under the *a priori* assumption of rate autocorrelation, i.e. that evolutionary rates among branches are correlated in a phylogenetic tree because of inherited factors like life-history traits or biochemical mechanisms. This intuitive assumption can be problematic, e.g. when strong rate changes occur; when very closely related taxa are analysed and little of the rate variation between the taxa can be attributed to inherited factors and stochastic or environmental factors become more important; or when only distantly related lineages are analysed, especially with a sparse taxon sampling, and variation is so strong that rate autocorrelation from lineage to lineage breaks down (Drummond et al., 2006; Ho et al., 2005). Uncorrelated clock models treat evolutionary

Table 3.2. Excluded alignment positions and inversions. Alignment positions refer to the reference alignments in Appendix 4 (Cucurbitales-Fagales dataset, 92 taxa; Begoniaceae dataset, 110 taxa). n/a: not applicable.

Region	Aligned sites [#]	Excluded fragments [#]	Excluded aligned sites [#]	Excluded fragment (Position)	Excluded aligned sites [% aligned]	Inversions [#]	Inversion [bp (# taxa)]	Inversion (Position)
*ndh*A intron	1406	6	39	172-175, 310-312, 683-690, 719-729, 777-781, 1114-1121	2.8	0	n/a	n/a
*ndh*F-*rpl*32	1208	8	68	1493-1498, 1939-1950, 2059-2068, 2083-2088, 2157-2163, 2229-2234, 2349-2359, 2558-2567	5.6	1	309 (2) 345 (3)	1506-1932 1506-1932
*rpl*32-*trn*L	1509	10	114	2722-2724, 2740-2744, 3135-3138, 3180-3187, 3251-3261, 3308-3308, 3406-3412, 3629-3639, 3838-3892, 4111-4119	7.6	3	11 (8) 27 (1) 37 (1)	3629-3639, 3900-3966, 3900-3994
trnL intron	749	1	92	3126-3217	12.3	0	n/a	n/a
trnL-F	634	1	45	3849-3893	7.1	1	29-40 (29)	3849-3893

rates among branches as random variables, and permit post-analysis assessment of whether the data indicates autocorrelation or not (Drummond et al., 2006).

To assess whether the data behaves in a clock-like manner the estimates of the coefficient of variation parameter were assessed with *Tracer* v1.5 (Rambaut and Drummond, 2009) after the analyses of the Cucurbitales-Fagales and the Begoniaceae datasets. A frequency histogram abutting against zero indicates that a strict molecular clock cannot be rejected, while if the frequency histogram is not abutting to zero it indicates among branch heterogeneity within the data (Drummond et al., 2007). Partitions were defined *a priori* for both the Cucurbitales-Fagales dataset and the Begoniaceae dataset. For the Cucurbitales-Fagales dataset three partition strategies were employed: 1. no partitioning; 2. partitioning based on coding region (*mat*K, *rbc*L) and spacer and intron identity (*trn*L intron, *trn*L-F spacer); 3. partitioning based on spacer and intron identity and subdivision of the coding regions (*rbc*L, *mat*K) in two codon position partitions, the first one of which comprised 3rd codon positions, the second one comprised 1st and 2nd codon positions. For the Begoniaceae dataset two partition strategies were employed: No partitioning and partitioning based on spacer and intron identity (*ndh*A intron and the *ndh*F-*rpl*32, *rpl*32-*trn*L spacer). Models of sequence evolution for each partition were determined using *jModelTest* (Posada, 2008). Maximum likelihood topologies were used to estimate the optimal evolutionary model comparing 88 distinct models (11 substitution schemes, with equal or unequal base frequencies, a proportion of invariable sites, and rate variation among sites). Log-likelihoods of different models of substitution under ML tree topologies were compared using the corrected version of Akaike Information Criterion for small samples (AICc) as model selection criterion (Posada and Buckley, 2004). The AICc converges towards the AIC, when larger sampling sizes are used, and should therefore always be used regardless of the sample size (Burnham and Anderson, 2004). Table 3.3 gives an overview of the different partition strategies and the selected models. Overall performance of unpartitioned and partitioned datasets were assessed with comparison of the mean -ln*L* of all trees sampled from the posterior distribution at stationarity for each strategy, and with Bayes Factor comparison implemented in *Tracer*, which is based on smoothed estimates of marginal likelihoods (Newton and Raftery, 1994; Suchard et al., 2001). The criterion of 2ln Bayes Factor of ≥10 was used as a benchmark indicating strong evidence of one strategy over another (Kass and Raftery, 1995). A Birth-and-death process prior was selected modelling cladogenesis and extinction of lineages, and a single overall uncorrelated lognormal relaxed clock model was applied for all partitions. For analyses of both Cucurbitales-Fagales and the Begoniaceae datasets starting trees for the MCMC runs satisfying all calibration priors (see 3.2.1.7) were generated in *TreeEdit* v1.0a10 (Rambaut and Charleston, 2002) by editing trees retrieved in maximum parsimony analyses. Several short *BEAST* runs were performed for each analysis to assess the MCMC performance, and to adjust the operators as suggested by the output diagnostics. Finally, for each analysis of both datasets two separate MCMC analyses were run, each with 4×10^7 generations and

sampling every 1000th generation. Times-series plots of all parameters were analysed in *Tracer* to check for convergence and to confirm that stationarity and adequate effective sampling sizes were reached. Trees were combined in *LogCombiner* v1.5.3, setting the burn-in to 25% of the initial samples of each MCMC run. Trees were then summarized using the maximum clade credibility option in *TreeAnnotator* v1.5.3.

3.2.1.7 Fossil constraints and secondary calibrations

The concept of stem group and crown group is crucial for the correct placements of fossils on a phylogenetic tree (Renner, 2005). The crown group comprises all extant taxa of a clade, their most recent common ancestor (MRCA) and all extinct taxa which diverged after the origin of the MRCA of the living taxa. The crown group is preceded by the stem lineage which comprises all extinct taxa that are closer to their crown clade than to another crown clade. The divergence point of the stem group, i.e. crown group plus stem lineage, from its sister, is the stem node, which is older than the point of origin of the crown clade (Benton and Donoghue, 2007; Forest, 2009; Renner, 2005). Assignment of fossils to extant taxa is based on shared derived characters. These apomorphic characters evolved along the stem lineage leading from the stem node to the crown group. As it is uncertain when an apomorphy evolved along the stem lineage, according to some authors a fossil placement on a phylogenetic tree has to assign a fossil constraint to the stem node rather than the crown node to account for this uncertainty (Forest, 2009; Magallón, 2004; Renner, 2005). Fossils provide minimum constraints as the fossil record is incomplete

Table 3.3. Partitioning strategies and model selection using *jModelTest.* Models: TPM (''3-parameter model'' = K81) (Kimura, 1981), TIM (''transitional model'') (Posada, 2003), TVM (''transversional model'') (Posada, 2003), and GTR (Tavaré, 1986). +G: among-site rate variation modelled with a gamma distribution; +I: proportion of invariable sites; uf: unequal base frequencies. For different types of the TIM and TPM models see Posada (2008). Codon pos.: codon position.

Dataset	Partitions [#]	Partitions	Aligned characters [#]	Model selected (AIC)	Model selected (AICc)	Model applied (*BEAST*)
Cucurbitales-Fagales, 92 taxa	1	Combined	3910	GTR+I+G	GTR+I+G	GTR+I+G
Cucurbitales-Fagales, 92 taxa	4	*matK*	1259	TVM+G	TVM+G	GTR+G
		rbcL	1429	GTR+I+G	GTR+I+G	GTR+I+G
		*trn*L intron	656	TVM+G	TVM+G	GTR+G
		*trn*L-F	566	GTR+G	GTR+G	GTR+G
Cucurbitales-Fagales, 92 taxa	6	*matK* 1^{st}+2^{nd} codon pos.	840	TVM+G	TVM+G	GTR+G
		matK 3^{rd} codon pos.	419	TVM+G	TVM+G	GTR+G
		rbcL 1^{st}+2^{nd} codon pos.	953	TVM+G	TVM+G	GTR+G
		rbcL 3^{rd} codon pos.	476	TIM3+I+G	TPM3uf+I+G	GTR+G
		*trn*L intron	656	TVM+G	TVM+G	GTR+G
		*trn*L-F	566	GTR+G	GTR+G	GTR+G
Begoniaceae, 110 taxa	1	Combined	3893	TVM+G	TVM+G	GTR+G
Begoniaceae, 110 taxa	3	*ndh*A intron	1370	TVM+G	TVM+G	GTR+G
		*ndh*F-*rpl*32	1136	TVM+G	TVM+G	GTR+G
		*rpl*32-*trn*L	1387	TVM+G	TVM+G	GTR+G

and it is likely that the fossil record documents when a structure became abundant rather than its time of origin (Magallón, 2004). Fossil minimum constraints require the oldest relevant fossil, which shares apomorphic characters of a given group, a well supported phylogenetic hypothesis, and a well dated stratum from which the fossil originated (Benton and Donoghue, 2007). The assumption that a divergence event is older than the oldest relevant fossil, assumptions about the quality of a fossil constraint and uncertainty of fossil ages can be modelled using probability distributions (Drummond et al., 2007; Ho, 2007; Ho and Phillips, 2009). The software *BEAST* allows the specification of priors for the ages of internal nodes defined by statistical distributions such as lognormal, exponential and normal distributions. Lognormal prior distributions can be used to model the assumption that the actual divergence likely occurred sometime before the earliest appearance of fossil evidence by assigning the highest point probability for a nodal age to an older age than the fossil constraint. Exponential prior distributions also allow for older ages than the age of a fossil as the distribution exhibits a long tail of diminishing probability with growing difference between the estimated nodal age and the fossil age. Selection of this distribution is reasonable to model the assumption that, based on well-dated and abundant fossils, there is a high probability of a small time lapse between the fossil age and the actual divergence event. A normal distribution is not suitable for fossil calibration, except for few exceptions, as it models non-directional uncertainty, i.e. towards both younger and older ages. Normally distributed priors can be applied to model uncertainty of the ages of geological events, e.g. island emergence dates, or for secondary calibrations based on other molecular studies (Ho, 2007).

Fossil constraints in the anlysis of the Cucurbitales-Fagales dataset: For the assignment of numerical ages to stages of the geologic time scale, this study follows Gradstein et al. (2004). Five fossil calibration points (C1-C5) were used in this study (Table 3.4). For calibration point C1 a uniform prior was chosen to assign a maximum constraint of 125 Ma to the root node using the first occurrences of tricolpate pollen indicative of Eudicots as a provisory maximum age constraint (discussed below). For calibration point C2, which is based on fagalean fossil pollen and associated

Table 3.4. Fossil calibrations. Ma: Millions of years ago; n/a: not applicable; SD: standard deviation.

Constraint	Clade	Anchor fossil	Assigned age [Ma]	Uniform prior [min, max]	Lognormal prior [mean, SD]	Exponential prior [set-off, mean]	References
C1	Cucurbitales-Fagales	Tricolpate pollen grains	125	0, 125	0,1	n/a	Doyle and Hotton, 1991; Hughes and McDougall, 1990
C2	Core Fagales-Fagaceae	Normapolles taxa pollen grains, *Caryanthus*	96.55	n/a	n/a	96.55, 1.02	Friis et al., 2006
C3	Coryloideae crown	*Corylus*, *Carpinus*	50	50, 125	0,1	n/a	Pigg et al., 2003
C4	Cucurbitaceae crown	Cucurbitaceae seeds, *Trichosanthes*	65.5	65.5, 125	0,1	n/a	Collinson et al., 1993; Collinson, 1986
C5	Tetramelaceae-Datiscaceae	*Tetrameleoxylon*	68.05	68.05, 125	0,1	n/a	Lakhanpal, 1970; Lakhanpal and Verma, 1965

floral macrofossils, an exponential prior was chosen reflecting the assumption that based on the good fossil record from numerous localities and well dated strata, and the lack of similar fossils from older strata, the age of the oldest relevant fossils is relatively close to the actual divergence date. Calibration points C3-5 are based on few fossils or on a single fossilized structure, and actual divergence of given clades likely occurred earlier than the oldest relevant fossil record. A lognormal prior distribution seems to be the most appropriate to model this assumption, as it includes a fossil age as lower hard bound, but allows for older ages by providing a soft upper bound defined by the shape of the lognormal distribution and the age interval which contains 95% of the probability. However, defining the upper soft bound is highly problematic and setting of the mean is somewhat subjective, especially when there is inadequate paleontological data (Benton and Donoghue, 2007; Ho, 2007). Therefore, a mean of 0 and a standard deviation of 1 were chosen to define lognormal prior distributions, which provide a soft upper bound allowing older node ages than the fossil ages, but ensure that 95% of the probability fall within reasonable age intervals. Alternatively, in an independent analysis priors with uniform distributions were applied with the lower hard bound defined by the fossil ages of calibration points C3-5 and the upper boundary set to 125 Ma based on the first occurrences of tricolpate pollen indicative of Eudicots as provisory maximum age constraint.

Constraint C1: The root node was constrained to a maximum age of 125 Ma based on the fossil record of tricolpate pollen indicative of Eudicots. The fossil record of tricolpate pollen grains provides one of the firmest dates for fossil calibrations as it is based on numerous samples from an extensive stratigraphic range from the transition of the Barremian and Aptian (125 Ma) onwards (Doyle and Hotton, 1991; Hughes and McDougall, 1990), and this can been used as a provisory maximum age constraint for Eudicots (Bell et al., 2005; Magallón and Sanderson, 2001, 2005; Wang et al., 2009). Molecular data indicate that early divergent eudicot lineages and core eudicots originated in rapid temporal succession, and subsequent to the eudicot origin there is extremely rapid phylogenetic branching of major eudicot lineages (Magallón and Sanderson, 2005; Moore et al., 2010; Wang et al., 2009). A constraint of 125 Ma for the Cucurbitales-Fagales divergence likely biased the analyses towards older ages, but represents a reasonable choice for a *maximum* constraint on the Cucurbitales stem node. This constraint was applied by assigning a uniform prior with a lower boundary of zero and an upper boundary of 125 Ma to the Cucurbitales stem node.

Other studies have used considerably younger ages to constrain the Cucurbitales-Fagales split. Within the Fagales crown group, Nothofagaceae and Fagaceae are the subsequent sisters to the core Fagales, which comprise all other extant Fagales families (Li et al., 2004). Wikström et al. (2001) and Wang et al. (2009) used generative macrofossils from Santonian strata of the Gaillard Formation in central Georgia, U.S.A, described as *Antiquacupula* Sims, Herend. & P.R.Crane and *Protofagacea* Herend., P.R.Crane & Drinnan, which

exhibit cupules and prolate pollen indicating a relationship to Nothofagaceae and particularly Fagaceae (Friis et al., 2006; Herendeen et al., 1995; Sims et al., 1998), to constrain the Cucurbitales-Fagales split to 84 or 85 Ma. However, Wikström et al. (2001) already discussed that it is unclear whether this fossil constraint should be assigned to the Fagales crown node or on the stem node, and their assignment to the stem node was made to control the direction of incorporated errors and to be confident to underestimate the age of the calibration point. Cook and Crisp (2005) assigned a fossil constraint based on *Antiquacupula* and *Protofagacea* fossils to a subclade within the Fagales crown group constraining the split between the Nothofagaceae and the Fagaceae plus core Fagales. Cook and Crisp (2005) and Goodall-Copestake et al. (2009) used the first occurrence of fagalean pollen of the Normapolles type in the Cenomanian, to constrain the age of the Fagales-Cucurbitales split to 96 or 97 Ma. However, Magallón and Sanderson (2001) and Magallón and Castillo (2009) assigned this fossil constraint to the Fagales crown node, and fagalean Normapolles type pollen gains and associated floral macrofossils were confidently placed within a subclade of Fagales, the core Fagales, based on a suite of shared floral characters, the oblate pollen form and the lack of a cupule (Friis et al., 2005; Friis et al., 2006; Sims et al., 1999). Therefore, this fossil constraint is here assigned to the core Fagales stem node as constraint C2.

Constraint C2: Fagales are well represented in the fossil record (Kubitzki et al., 1993) and Friis et al. (2006) state that the stratigraphic range of fagalean micro- and macrofossils indicates that "all major fagalean lineages were present by the Cenomanian or earlier." The oldest Fagales fossils are fagalean pollen grains of the Normapolles type, which are characterised by their oblate form, and complex, protruding apertures consisting of an exo- and endoaperture (Friis et al., 2006). These pollen grains and associated floral macrofossils have been reported from numerous localities and an extensive stratigraphic range from the Cenomanian (99.6-93.6 Ma) onwards (Friis et al., 2006; Sims et al., 1999), and they share a suite of floral and palynological characters with extant families of the core Fagales (Friis et al., 2006; Sims et al., 1999). Therefore, the first occurrences of fagalean Normapolles type pollen grains and associated fossil floral structures in the middle Cenomanian was used as fossil constraint by assigning an exponential prior to core Fagales stem node (offset: 96.55 Ma, SD: 1.02), i.e. the divergence of Fagaceae and core Fagales. A standard deviation of 1.02 was chosen so that 95% of the probability is contained in an interval between the midpoint and the upper boundary of the Cenomanian (96.55-99.6 Ma).

Constraint C3: Generative Fagales macrofossils from the middle Eocene Republic flora of northeastern Washington State, U.S.A., have been assigned to the modern genera *Corylus* L. and *Carpinus* L. representing the oldest fossils of the subfamily Coryloideae of the Betulaceae (Pigg et al., 2003). These fossils were found in the Klondike Mountain Formation, which was dated to 50-49 Ma by radiometric Argon-Argon dating (Pigg et al.,

2003). This fossil data has been used in molecular studies to constrain the crown group age of Coryloideae to 50 Ma (Cook and Crisp, 2005). Following Cook and Crisp (2005) this fossil constraint was assigned to the MRCA of the crown group of the Coryloideae. A lognormal prior distribution was selected with an off-set based on the radiometric dates (50 Ma).

Constraint C4: Seed fossils from the Uppermost Paleocene and Lower Eocene London Clay have been assigned to Cucurbitaceae and the extant Cucurbitaceae genus *Trichosanthes* L. (Chandler, 1964; Collinson, 1986; Collinson et al., 1993; Schaefer et al., 2008). This fossil data has been used in molecular studies to constrain the crown group age of Cucurbitaceae to 65 Ma, which is at the upper boundary of the Paleocene (Schaefer et al., 2009). This study follows Schaefer et al. (2009) by assigning a lognormal prior to the Cucurbitaceae crown node with a lower bound of 65.5 Ma, which is at the Paleocene-Eocene boundary.

Constraint C5: *Tetrameles*-like fossil bark described as *Tetrameleoxylon prenudiflora* Lakhanpal & Verma has been described from the Deccan intertrappean beds at Mohgaonkalan in India (Lakhanpal, 1970; Lakhanpal and Verma, 1965). These beds have been dated to the Maastrichtian (65.5-70.6) (Khajuria et al., 1994). *Tetrameleoxylon prenudiflora* exhibits "characteristics consistent with *Tetrameles*" (E. Wheeler, personal communication cited in Zhang et al., 2006), and this fossil constraint has either been placed on the split of Tetramelaceae and Datiscaceae (Goodall-Copestake et al., 2009: 68 Ma; Schaefer et al., 2009: 68 or 65.5 Ma; Zhang et al., 2007: 68 Ma) or on the split of *Tetrameles* R.Br. and *Octomeles* Miq. (Clement et al., 2004, 2005: 55 Ma). Following Zhang et al. (2007), Goodall-Copestake et al., 2009, and Schaefer et al. (2009) this fossil constraint is here assigned to the MRCA of Datiscaceae and Tetramelaceae using the midpoint of the Maastrichtian (68.05 Ma) as lower boundary of the lognormal prior distribution.

To explore the sensitivity of the molecular age estimates of the Begoniaceae and *Begonia* crown group to priors C3-5, which are based on only few fossils or single fossilized structures, alternative MCMC runs which tested different fossil combinations were performed. These included the more reliable fossil calibration points C1-2, and all combinations of calibrations points C3-5. MCMC runs used the same setting as described 3.2.1.5, employing the four partitions scheme and used lognormally distributed priors for calibration points C3-5.

Secondary constraints in the analysis of the Begoniaceae dataset: Secondary calibration, i.e. the use of molecular age estimates from previous analyses to calibrate a molecular clock, is problematic as errors of the primary analysis are propagated and likely magnified in the secondary analysis (Graur and Martin, 2004; Renner, 2005). In the absence of

Table 3.5. Secondary calibrations. HPD: highest posterior density date range; Ma: Millions of years ago, n/a: not applicable; SD: standard deviation.

Constraint	Clade	Basis	Uniform Prior [min, max]	Normal Prior [mean, SD]
S1	*Begonia* crown	Mean = Average of mean values of primary analyses 1 and 2. SD chosen to incorporate the 95% HPDs of primary analyses 1 and 2	n/a	26.0, 4
S2	*Hillebrandia-Begonia*	Bounds = Mean age estimates of primary analyses 1 and 2	40.1, 41.8	n/a

suitable fossils for calibrations as is the case in *Begonia*, secondary calibration provides a useful source of calibration information, but has to be implemented with caution.

As outlined above, Begoniaceae and *Begonia* crown node age estimates were calculated in two primary analyses of the Cucurbitales-Fagales dataset. The first primary analysis included five fossil calibrations and used lognormally distributed priors of calibration points C3-5. The second primary analysis included five fossil calibrations and used uniform priors of calibration points C3-5. The mean estimates and 95% highest posterior density date ranges (HPDs) of the Begoniaceae and *Begonia* crown node ages were used to define secondary calibration points for the analyses of the Begoniaceae dataset. Two different priors were applied in independent analyses (Table 3.5).

Secondary calibration S1: To account for the uncertainty of the age estimates of the two primary analyses the age prior of the MRCA of the *Begonia* crown group was modelled as normal distribution with its mean set to 26.0 Ma, which is the average of the mean age estimates of the *Begonia* crown group divergence of the two primary analyses of the Cucurbitales-Fagales data set. The 95% confidence interval was specified by setting the standard deviation to 4 (95% probability interval: 18.2-33.8 Ma) including most of the ranges of the 95% HPDs of the two primary analyses (primary analysis 1: 18.3-34.0 Ma; primary analysis 2: 18.2-34.3 Ma).

Secondary calibration S2: Alternatively, the root age prior was modelled as a uniform distribution, having lower and upper boundaries equal to the mean age estimates of the Begoniaceae crown group divergence of the two primary analyses of the Cucurbitales-Fagales data set. This approach uses a secondary calibration point based on only the highest probability densities of the two primary analyses. The lower boundary was set to 40.1 Ma, and the upper boundary was set to 41.8 Ma to specify this prior.

3.2.2 Biogeographic analyses

3.2.2.1 Area delimitation

Seven areas based on the geological history of Asia and the extant distributions of *Begonia*

and *Hillebrandia* species were considered in the analysis (Fig. 3.1): 1. Hawaii, USA, to which *Hillebrandia sandwicensis* is endemic; 2. Africa including the Yemenite Socotra Archipelago; 3. Continental Asia and adjacent major islands (Sri Lanka, Taiwan, Hainan); 4. the Sunda Shelf region, which extends from just north of the Thai-Malay border, through the Malay Peninsula, Sumatra, Borneo, and Java to Bali; 5. Wallacea, which comprises Sulawesi and adjacent islands, and the Lesser Sunda Islands and the Maluku Islands between Wallace's and Lydekker's Line; 6. the Philippines including Palawan; 7. New Guinea and adjacent islands (islands east of Lydekker's Line, the Bismarck Archipelago, Solomon Islands).

Most *Begonia* species are narrow endemics and the distributions of only few species span more than one of the defined regions. The analysed dataset contains only five taxa which show wider distributions. Four of these more widespread taxa exhibit fleshy fruits and belong to difficult species complexes. *Begonia longifolia* is the most widespread species of *Begonia* in Asia. Its distribution ranges from continental Asia (Nepal, Bhutan, India, China, Laos, Vietnam, Thailand and Peninsula Malaysia) through Sumatra, Java and Bali, to Sulawesi (Hughes, 2008; Tebbitt, 2003). *Begonia aptera*, which is closely related to *Begonia longifolia*, is mainly distributed on Sulawesi and the Moluccas, but has been recently collected in Western New Guinea (George Argent, Royal Botanic Garden Edinburgh, UK, *pers. com.*). *Begonia multangula* is distributed on Java and the Lesser Sunda Islands and maybe Sulawesi, and taxa in the *B. rieckei* complex exhibit a wide distribution in eastern Malesia including Sulawesi, the Philippines and New Guinea (Hughes, 2008). Finally, *Begonia fenicis* Merr. is the only *Begonia* species which exhibits a distribution including both the Philippines (islands of the Taiwan-Luzon arc) and Taiwan. As likelihood and Bayesian reconstructions require the assignment of single discrete character states, the samples of the five more widespread taxa were assigned to the areas of their collection localities. Alternatively, five additional composite area states were defined: 8. Continental Asia plus the Philippines (*Begonia fenicis*), 9. Sunda Shelf region plus Wallacea (*B. multangula*); 10: Wallacea plus New Guinea (*B. aptera*); 11. Continental Asia plus Sunda Shelf region plus Wallacea (*B. longifolia*); 12. Wallacea plus Philippines plus New Guinea (*B. rieckei*).

3.2.2.2 Likelihood ancestral area reconstructions

Ancestral areas were reconstructed using the likelihood method implemented in *Mesquite* v2.7.2 (Maddison and Maddison, 2009). In contrast to parsimony analyses, likelihood reconstructions can account for time-proportional branch-length information, and estimate relative probability values for each nodal reconstruction. To account for phylogenetic uncertainty the "Trace over trees" option was selected, and 1000 randomly chosen trees from the stabilized part of the MCMC of the *BEAST* analysis (Begoniaceae dataset, 110 taxa, secondary calibration S1) were included as input trees. Likelihood reconstructions optimize the area states at each node which maximize the probability of arriving at the

observed extant areas of the terminals, given a model of evolution. The Mk1 model (Markov k-state 1 parameter model) (Lewis, 2001) was selected. Under this model any particular change is equally probable, and the rate of change is the only parameter. The "trace over trees" option calculates for every node the likelihoods of all area states averaged over all trees possessing the node in the 1000 input tree sample. Ancestral area reconstructions were mapped on the maximum clade credibility chronogram derived from the *BEAST* analysis of the 110 taxon Begoniaceae dataset using secondary calibration S1.

3.2.2.3 Bayesian ancestral area reconstructions

Bayesian ancestral area reconstructions were performed in *BEAST* based on an analysis of the 110 taxon Begoniaceae dataset using the continuous-time Markov chain (CTMC) model for discretized diffusion specified by Lemey et al. (Lemey et al., 2009), considering diffusion among the seven areas specified in 3.2.2.1 (K=7). The CTMC model for discretized diffusion is the equivalent of the GTR model for nucleotide substitutions, and allows for $K^*(K-1)/2$ different diffusion rates. The model incorporates two main parameters, the relative rate parameter, which describes how often diffusion between two locations occurs during evolution in relation to other location transitions, and the geosite model parameter, which rescales location transitioning into time units. A gamma prior (shape=1.0) was chosen for the rates parameter and an exponential prior (mean=1) for the geosite model parameter following recommendations by Lemey et al. (2009). The parameters were sampled from simultaneously estimated time-scaled phylogenies. Bayesian inference settings included two separate MCMC analyses, each run for 4×10^7 generations, parameter sampling every 1000th generation, using an uncorrelated relaxed lognormal clock model, and a birth-and-death process prior. The *Begonia* stem node was calibrated using the secondary calibration S1 (see Table 3.5).

3.3 Results

3.3.1 Molecular age estimates: Cucurbitales-Fagales dataset

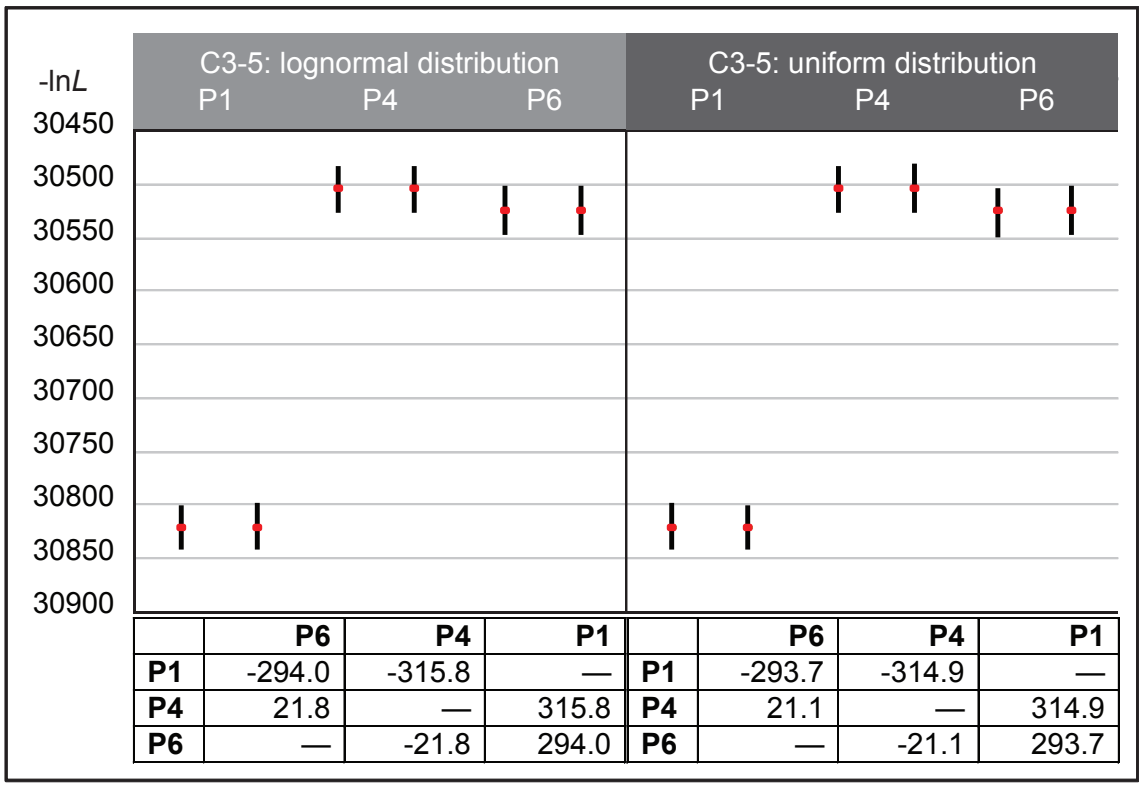
Summary statistics of the concatenated 92-taxon Cucurbitales-Fagales dataset and its partitions are shown in Table 3.6. Sequence alignment of the four combined plastid regions (*mat*K, *rbc*L, *trn*L intron, *trn*L-*trn*F spacer) yielded a dataset of 3910 aligned characters. Of the analysed regions, the *rbc*L fragment exhibited the lowest variability with 27.4 percent of polymorphic sites and 16.6 percent of potentially parsimony informative sites, and the third codon position partition of the *rbc*L fragment showed the least variability of all six data partitions with 17.2 percent of polymorphic sites and 11.1 percent of potentially parsimony informative sites. The *mat*K fragment and its codon position partitions exhibited the highest variability with 54.6 percent of polymorphic sites and 36.9 percent of potentially parsimony informative sites.

Table 3.6. Descriptive statistics of the Cucurbitales-Fagales and the Begoniaceae datasets.

Dataset	Partition	Aligned positions [#]	Fragment length [bp]	Variable sites [# (%)]	Parsimony informative sites [# (%)]
Cucurbitales-Fagales, 92 taxa	*mat*K	1259	750-1203	687 (54.6)	465 (36.9)
	matK 1st+2nd codon pos.	840	500-803	458 (54.5)	310 (36.9)
	matK 3rd codon pos.	419	250-400	229 (54.6)	155 (37.0)
	rbcL	1429	994-1425	392 (27.4)	237 (16.6)
	rbcL 1st+2nd codon pos.	953	664-950	310 (32.5)	184 (19.3)
	rbcL 3rd codon pos.	476	330-475	82 (17.2)	53 (11.1)
	*trn*L intron	656	277-533	291 (44.4)	177 (27.0)
	*trn*L-F	566	185-389	290 (51.2)	179 (31.6)
	Combined	3910	2809-3467	1660 (42.5)	1058 (27.1)
Begoniaceae, 110 taxa	*ndh*A intron	1370	1076-1188	327 (23.9)	147 (10.7)
	*ndh*F-*rpl*32	1136	772-991	378 (33.3)	203 (17.9)
	*rpl*32-*trn*L	1387	439-1099	387 (27.9)	200 (14.4)
	Combined	3893	2448-3254	1092 (28.1)	550 (14.1)

Estimates of the coefficient of variation of the analysis of the Cucurbitales-Fagales dataset were checked in *Tracer* to assess whether the data behaves in a clock-like manner. The frequency histogram did not abut against zero indicating considerable among branch rate heterogeneity, which indicated that a relaxed molecular clock approach was appropriate (Drummond et al., 2007). Nucleotide model selection under the AIC and its corrected version for small sample sizes (AICc) did not differ for most partitions with the exception of the smallest and least variable partition, the 3rd codon position partition of the *rbc*L fragment (Table 3.3). For this partition the TIM3 (transitional model; Posada, 2003), was selected under the AIC, while the slightly less parameter-rich TPM3uf model (= K81 model, Kimura, 1981) was selected under the AICc. More complex models, the TVM (transversional model; Posada 2003) or the GTR model (General time reversible model; Tavaré, 1986), were selected for all other partitions. Partitioning improved mean *–lnL* values considerably and the analyses using a partition strategy with four partitions based on gene, intron and spacer identity provided distinctly better explanations of the data than both the more complex partition strategy including partitions based on codon positions for the coding regions and unpartitioned analyses according to Bayes Factor comparison (Fig. 3.2). The subsequent presentation of the results of the analyses of the Cucurbitales-Fagales dataset will be limited to the analyses using four data partitions.

The maximum clade credibility chronogram resulting from the analyses using five fossil calibrations is presented in Fig. 3.3. Mean date estimates and 95% HPDs for the *Begonia* stem and crown nodes derived from nine analyses using different combinations of constraints and different calibration prior distributions are shown in Figure 3.4. The 95% HPD date range associated with each estimate varied depending on the combination and prior distributions of the applied constraints. In the analyses using lognormal priors for constraints C3-5 and eight alternative sets of constraints, the HPDs of the *Begonia* stem node ages estimates extended from 21-54 Ma, and the HPDs of the *Begonia* crown



	P6	P4	P1
P1	-294.0	-315.8	—
P4	21.8	—	315.8
P6	—	-21.8	294.0

	P6	P4	P1
P1	-293.7	-314.9	—
P4	21.1	—	314.9
P6	—	-21.1	293.7

Fig. 3.2. Divergence time estimation: Bayes Factor comparison of different partitioning strategies for the Cucurbitales-Fagales dataset. Charts show −ln*L* values of two independent *BEAST* runs for three different partitioning strategies (P1, P4, P6) and two different calibration prior settings (C3-5 with lognormal distribution; C3-5 with uniform distribution, see Table 3.4). The bars indicate 95% highest posterior density date ranges, the red dots indicate mean values. The table shows ln Base Factors calculated from harmonic means of likelihoods in *Tracer* v1.5. A positive value of > 5 indicates strong evidence against alternative hypotheses (partition strategies indicated in the first column are compared with partition strategies indicated in subsequent columns).

node ages estimates extended from 14-34.0 Ma. The two primary analyses using all five constraints but differing in using lognormal priors or alternatively uniform prioirs for constraints C3-5 produced similar mean values for the *Begonia* stem node (lognormal prior analysis: 40 Ma, uniform prior analysis: 41.8 Ma) and crown node (lognormal prior analysis: 26 Ma, uniform prior analysis: 26 Ma) with slightly older ages estimates and broader HPDs in the analyses using uniform priors. Highest posterior density credibility date ranges of the *Begonia* stem node extended from 28 to 54 Ma in the analysis using lognormal priors, and from 28 to 57 Ma in the analysis using uniform priors. Highest posterior density credibility sets of the crown node extended from 18 to 34 Ma in the analysis using lognormal priors and from 18 to 34 Ma in the analysis using uniform priors. The results of the analyses using nine alternative sets of calibration constraints show the sensitivity of the analysis to different calibration points (Fig. 3.4). The absence of constraints C3-5 had a strong impact on age estimates and resulted in distinctly younger age estimates than in the analyses including all five constraints. Inclusion of constraints C3-5 pushed the estimated ages back in time, and the absence or presence of constraint C3 (Coryloideae crown group) had the strongest impact, while the constraint on the Tetramelaceae-Datiscaceae split (C5), which is the phylogenetically closest constraint to the Begoniaceae crown clade, had the weakest impact on the age estimates.

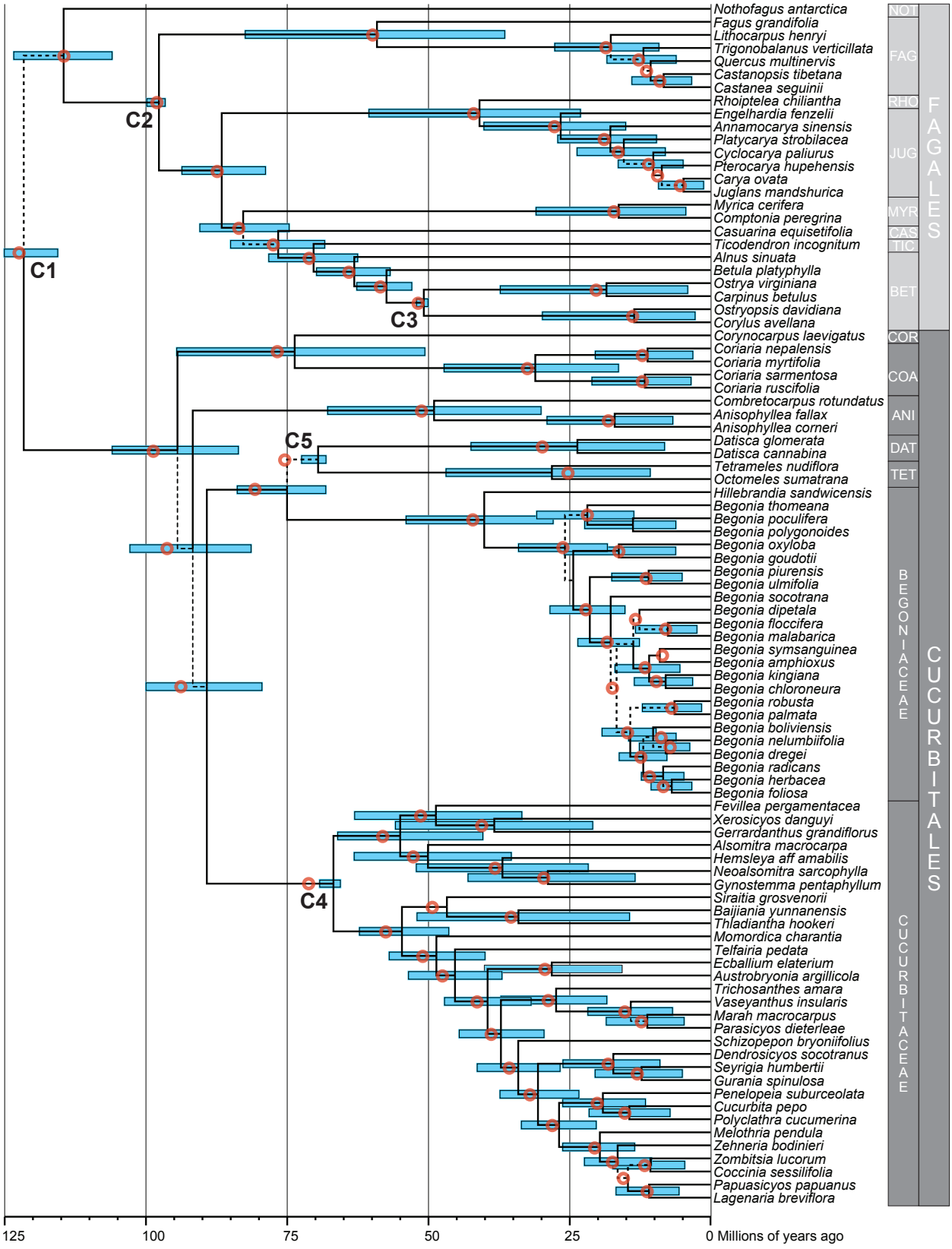


Fig. 3.3. Maximum clade credibility chronogram: Cucurbitales-Fagales dataset. Divergence time estimations based on an analysis of the Cucurbitales-Fagales dataset using *BEAST* and five fossil calibrations (C1-5, using lognormal prior distributions for C3-5, see Table 3.4.). Node heights indicate mean ages and node bars indicate 95% HPD date ranges. Mean ages derived from an alternative analysis using five fossil calibrations (C1-5, using uniform distributed priors for C3-5, see Table 3.4) are mapped on the chronogram as red circles, the centres of which indicate the mean age estimates. Broken lines indicate branches which lead to nodes with a PP < 0.95. ANI: Anisophylleaceae, BET: Betulaceae, CAS: Casuarinaceae, COA: Coriariaceae, COR: Corynocarpaceae, DAT: Datiscaceae, FAG: Fagaceae, JUG: Juglandaceae, MYR: Myricaceae, NOT: Nothofagaceae, RHO: Rhoipteleaceae, TET: Tetramelaceae, TIC: Ticodendraceae.

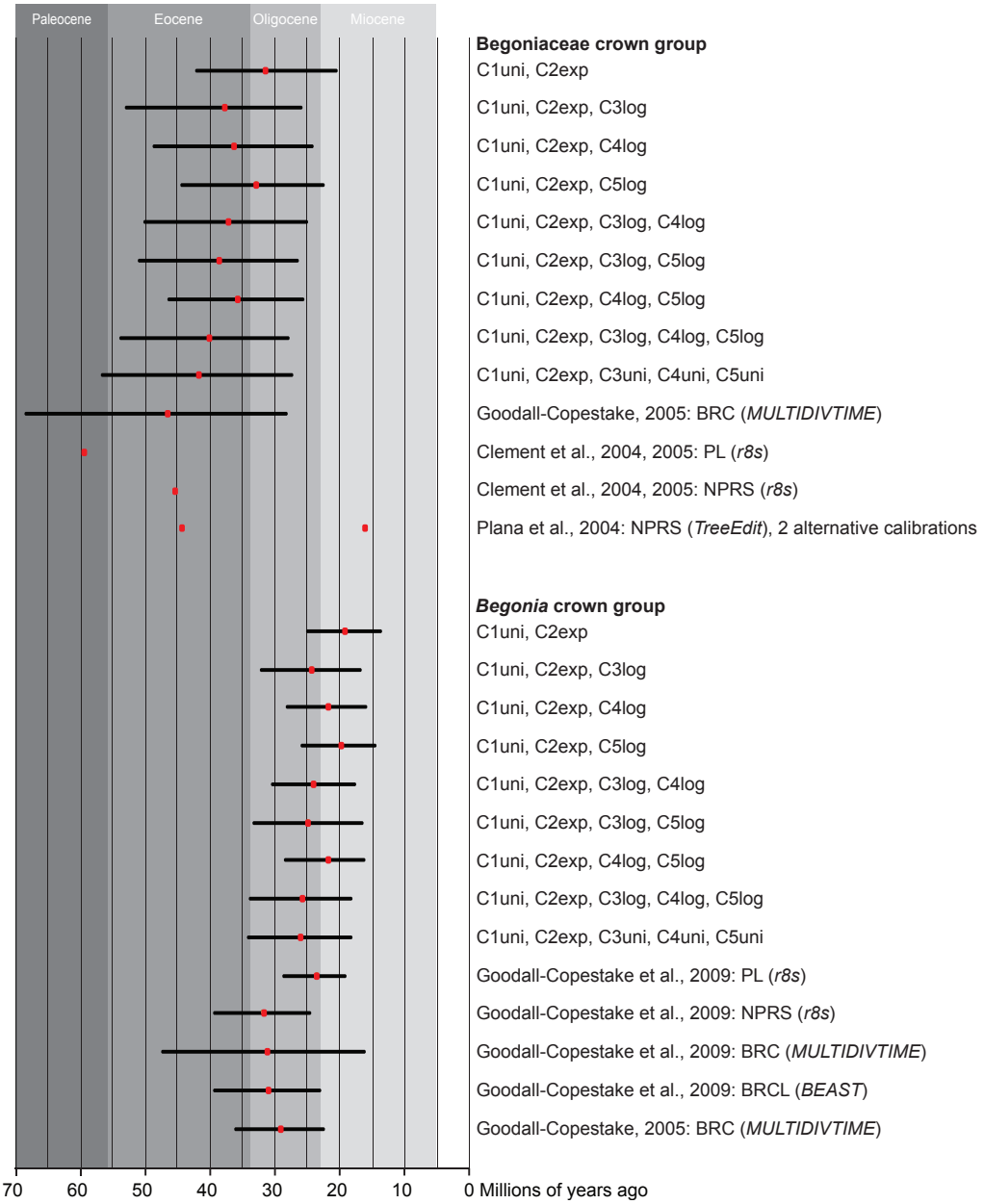


Fig. 3.4. Divergence age estimates for the MRCAs of the Begoniaceae and *Begonia*. Divergence time estimations are based on analyses of the Cucurbitales-Fagales dataset using *BEAST*. Analyses using different combinations of fossil calibrations are indicated (C1-5, see Table 3.4). Published age estimates of previous studies are indicated (Clement et al., 2004, 2005; Goodall-Copestake, 2005; Goodall-Copestake et al., 2009; Plana et al., 2004). Red dots indicate mean estimates, and bars indicate 95% HPDs and 95% confidence ranges. BRC: Bayesian relaxed clock method; BRCL: Bayesian uncorrelated relaxed lognormal clock method; exp: exponential prior distribution; log: lognormal prior distribution; NPRS: Non-Parametric Rate Smoothing method; PL: Penalized Likelihood method; uni: uniform prior distribution. Geological epochs are indicated in different shades of grey.

3.3.2 Molecular age estimates: Begoniaceae dataset

Summary statistics of the concatenated 110-taxon Begoniaceae dataset and its partitions are shown in Table 3.6. Sequence alignment of the three combined plastid regions (*ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L) yielded a dataset of 3893 aligned characters. Of the analysed regions, the *ndh*A intron exhibited the lowest variability with 23.9 percent of polymorphic sites and 10.7 percent of potentially parsimony informative sites. The *ndh*F-*rpl*32 fragment exhibited the highest variability with 33.3 percent of polymorphic sites

and 17.9 percent of potentially parsimony informative sites.

Estimates of the coefficient of variation of the analysis of the Begoniaceae dataset were checked in *Tracer* to assess whether the data behaves in a clock-like manner. The frequency histogram did not abut against zero indicating considerable among branch rate heterogeneity. Nucleotide model selection under the AIC and its corrected version for small sample sizes (AICc) did not differ for the three partitions, and a complex model, the TVM (transversional model; Posada 2003), was selected for all partitions (Table 3.3). Partitioning improved mean –ln*L* values considerably and the analyses using a partition strategy with three partitions based on intron and spacer identity provided distinctly better explanations than analyses of unpartitioned datasets according to Bayes Factor comparison (lnBayes Factors 46.8 for the analyses using calibration S1; lnBayes Factors 46.9 for the analyses using calibration S2). The subsequent presentation of the results of the analyses of the Cucurbitales-Fagales dataset will be limited to the analyses using three data partitions.

Figure 3.5 shows a maximum clade credibility chronogram resulting from the analysis using secondary calibration S1 (MRCA *Begonia*, normally distributed prior, mean: 26 Ma, SD: 4, 95% probability interval: 18.2-33.8). Mean values of the analysis using calibration S2 (MRCA Begoniaceae, uniform prior, lower bound: 40.8, upper bound: 41.8) are mapped on the chronogram, and mean values within the 95% HPD date ranges for the divergences of several major clades are shown in Figure 3.6. Age estimates for diversifications within a well supported clade comprising Asian and Socotran *Begonia* species differ slightly between the two analyses constraining either the root node (calibration S2) or the Begoniaceae crown node (calibration S1), with mean age estimates differing by no more than 1.01 Ma, and generally slightly older ages and slightly wider HPDs in the analyses constraining the root node (Fig. 3.6). Age estimates in the following presented results are given as mean values and 95 % HPDs in brackets, and estimates of the analysis using calibration S1 are followed by estimates of the analysis employing calibration S2. The age of the stem lineage of Asian plus Socotran *Begonia* was estimated as *c.* 18 (11-25) and *c.* 20 (12-28) Ma old, the crown group age estimates are *c.* 15 (9-21) and *c.* 16 (10-23) Ma. The phylogenetic relationships of two small, early divergent clades comprising species in sections *Haagea*, *Peltaugustia*, *Reichenheimia* and one species unplaced to section, are not well supported. Apart from these two early divergent lineages, two well supported main clades can be differentiated: Clades A and B. The crown group age of Clade A, which includes species in sections *Parvibegonia*, *Diploclinium*, and *Platycentrum* s.l. (containing section *Sphenanthera*), is estimated as *c.* 13 (8-19) Ma and *c.* 14 (9-21) Ma old. The crown group age of a well supported subclade containing species of section *Platycentrum* s.l. is estimated as *c.* 5 (3-7) Ma and *c.* 5 (3-8) Ma. The crown group age estimates of Clade B, which includes species of sections *Coelocentrum*, *Ridleyella*, *Reichenheimia*, *Diploclinium*, *Bracteibegonia*, *Petermannia*

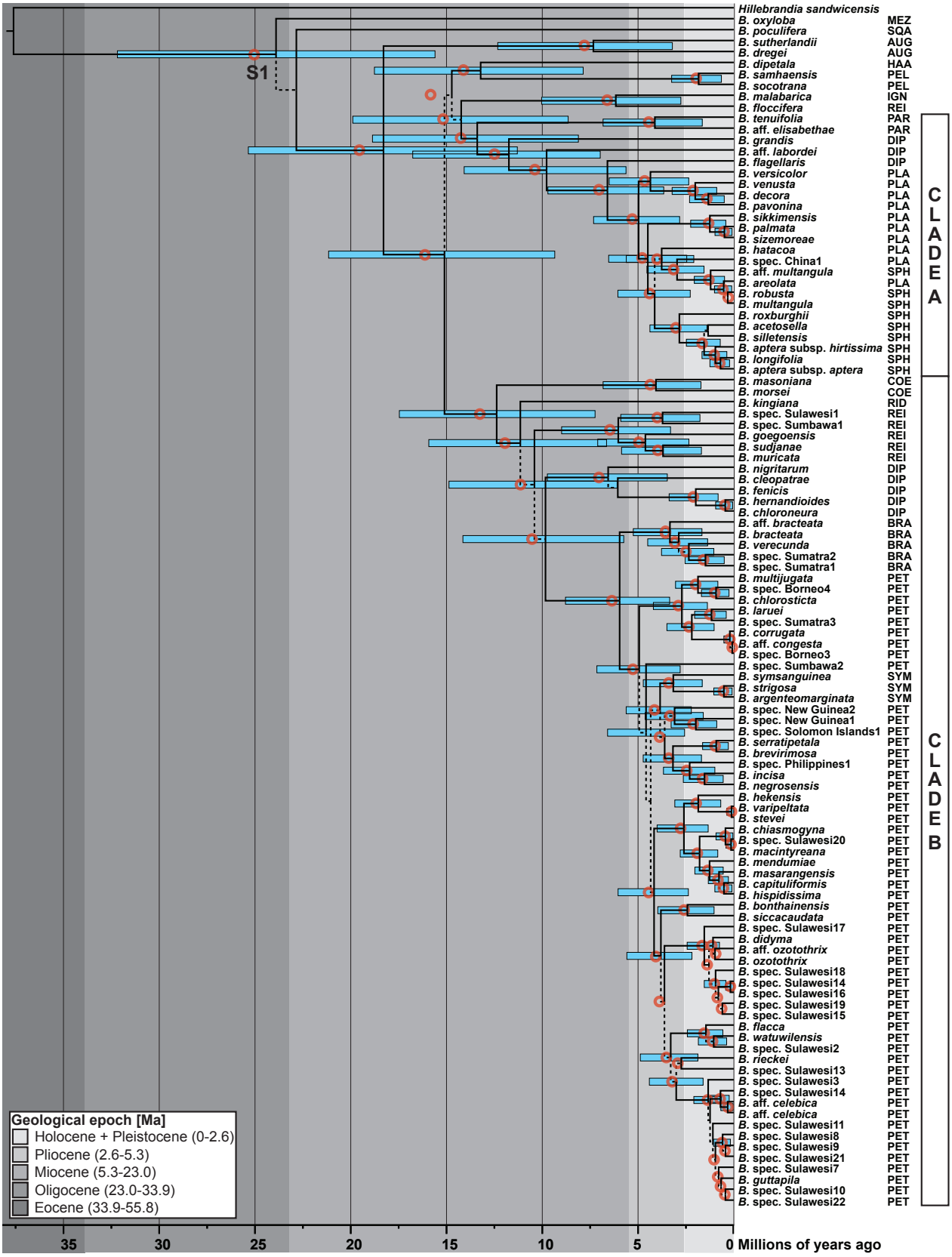


Fig. 3.5. Maximum clade credibility chronogram: Begoniaceae dataset. Divergence time estimations based on an analysis of the Begoniaceae dataset using *BEAST* and secondary calibration S1 (see Table 3.5). Node heights indicate mean ages and node bars indicate the 95% HPD date ranges. Mean ages derived from an alternative analysis using secondary calibration S2 (see Table 3.5) are mapped on the chronogram as red circles, the centre of which indicates the mean age estimates. Broken lines indicate branches which lead to nodes with a PP < 0.95. Geological epochs are indicated in different shades of grey. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, HAA: *Haagea*, IGN: unplaced to section, MEZ: *Mezierea*, PAR: *Parvibegonia*, PEL: *Peltaugustia*, PET: *Petermannia*, PLA: *Platycentrum*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQU: *Squamibegonia*, SYM: *Symbegonia.*

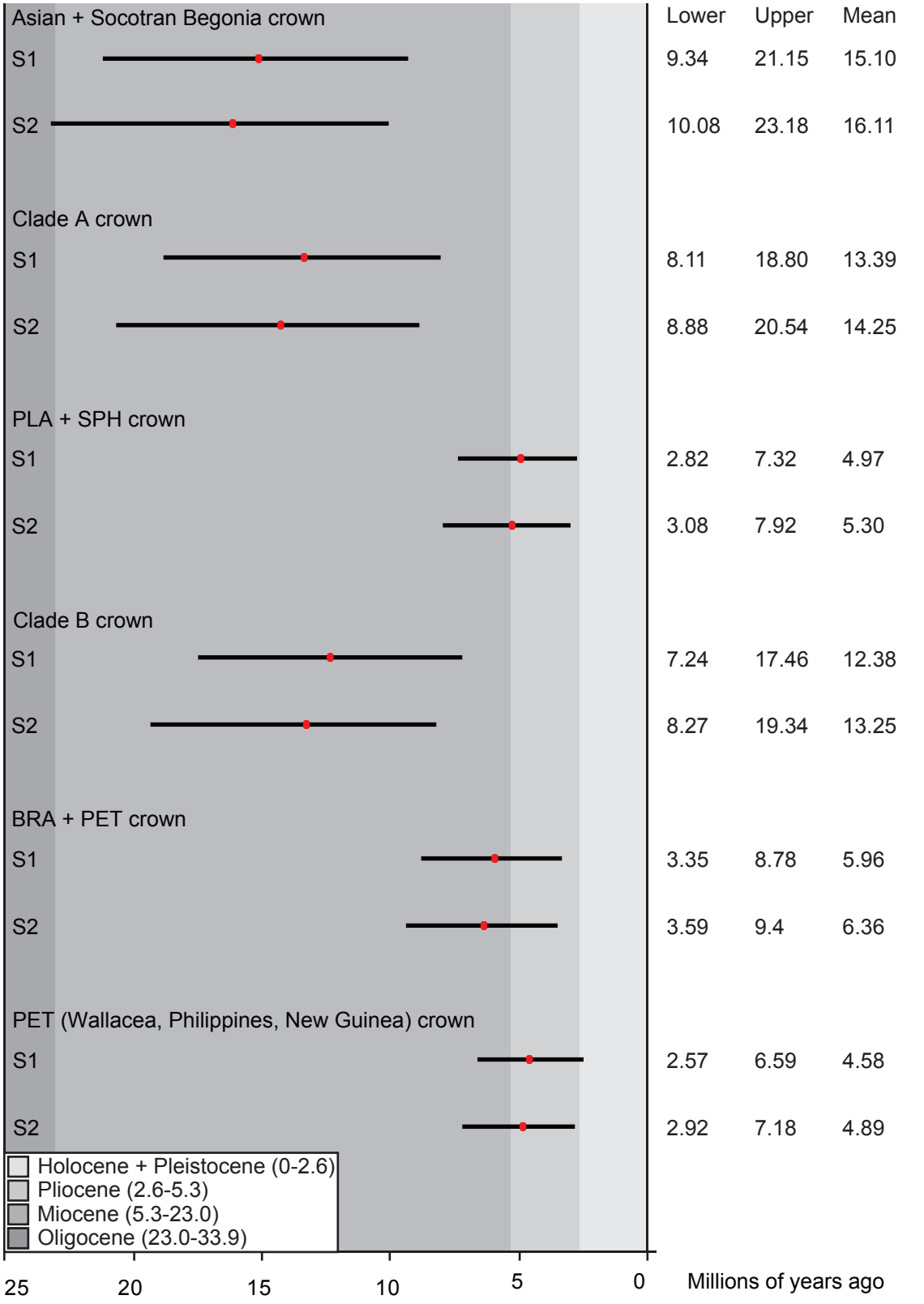


Fig. 3.6. Divergence age estimates: Begoniaceae dataset. Divergence time estimations are based on analyses of the Begoniaceae dataset using *BEAST*. Red dots indicate mean estimates, and bars indicate 95% HPD date ranges. Analyses using different secondary calibrations are indicated (S1-2, see Table 3.5). The table indicates lower and upper bounds of the 95% HPDs and mean values. Geological epochs are indicated in different shades of grey.

s.l. (inclusive section *Symbegonia*), are c. 12 (7-18) Ma and *c.* 13 (8-19) Ma. The crown group age estimates for the subclade comrising species of sections *Bracteibegonia* and *Petermannia* s.l. are *c.* 6 (3-9) Ma and *c.* 6 (4-9) Ma. Mean age estimates and associated HPDs of diversifications in the species-rich sections *Platycentrum* and *Petermannia* fall predominantly in the Plio-Pleistocene range.

3.3.3 Ancestral area reconstructions

Figure 3.7 shows a maximum clade credibility chronogram derived from the Bayesian ancestral area reconstruction in *BEAST* using the seven area states delimited in 3.2.2.1. Branches are coloured according to the most probable area state of their descendent nodes, and area state posterior probability (PP) distributions are indicated. Figure 3.8 shows the maximum clade credibility chronogram resulting from the analysis of the Begoniaceae dataset in *BEAST* (see 3.3.2) with branches coloured according to the area state which received the highest proportional likelihood at their descendent nodes according to likelihood ancestral area reconstructions in *Mesquite* using the seven main area states plus the five composite area states delimited in 3.2.2.1. Proportional likelihoods (PL) of ancestral area reconstructions are indicated.

Bayesian and likelihood reconstructions produced similar results, and the use of five additional composite area states for the five more widespread taxa had a negligible impact on nodal reconstructions except for the reconstruction at the stem and crown node of a clade comprising taxa in the *Begonia longifolia* complex (discussed below).

Mean age estimates indicate an area transition from Africa to Asia between the Early and Middle Miocene and the analyses reconstruct continental Asia as most probable ancestral area at the crown node (PP: 0.93; PL: 0.96) of a well supported clade containing Socotran and Asian *Begonia* species. Within this clade, two early divergent lineages, both of which show only poorly supported phylogenetic relationships, form two clades which comprise species with distributions in South India/Sri Lanka and the Yemenite Socotra Archipelago. Reconstructions at the crown node of the clade which includes one species from South India and the only two species known from the Socotra Archipelago, indicate continental Asia as the most probable ancestral area (PP: 0.93, PL: 0.96) for this clade.

Continental Asia is also the most probable ancestral area for both Clade A (stem node PP: 0.99, PL: 1) and Clade B (stem node PP: 0.96, PL: 0.96). Within Clade A, most nodal reconstructions indicate continental Asia as most probable ancestral area, with several lineages showing area transitions to the Sunda Shelf and Wallacea. The Wallacean species in Clade A belong to two recently diverged clades comprising some more widespread taxa placed in the polyphyletic section *Sphenanthera*. The first clade is comprised of *Begonia longifolia*, and the two subspecies of *B. aptera*, and ancestral area reconstructions at the

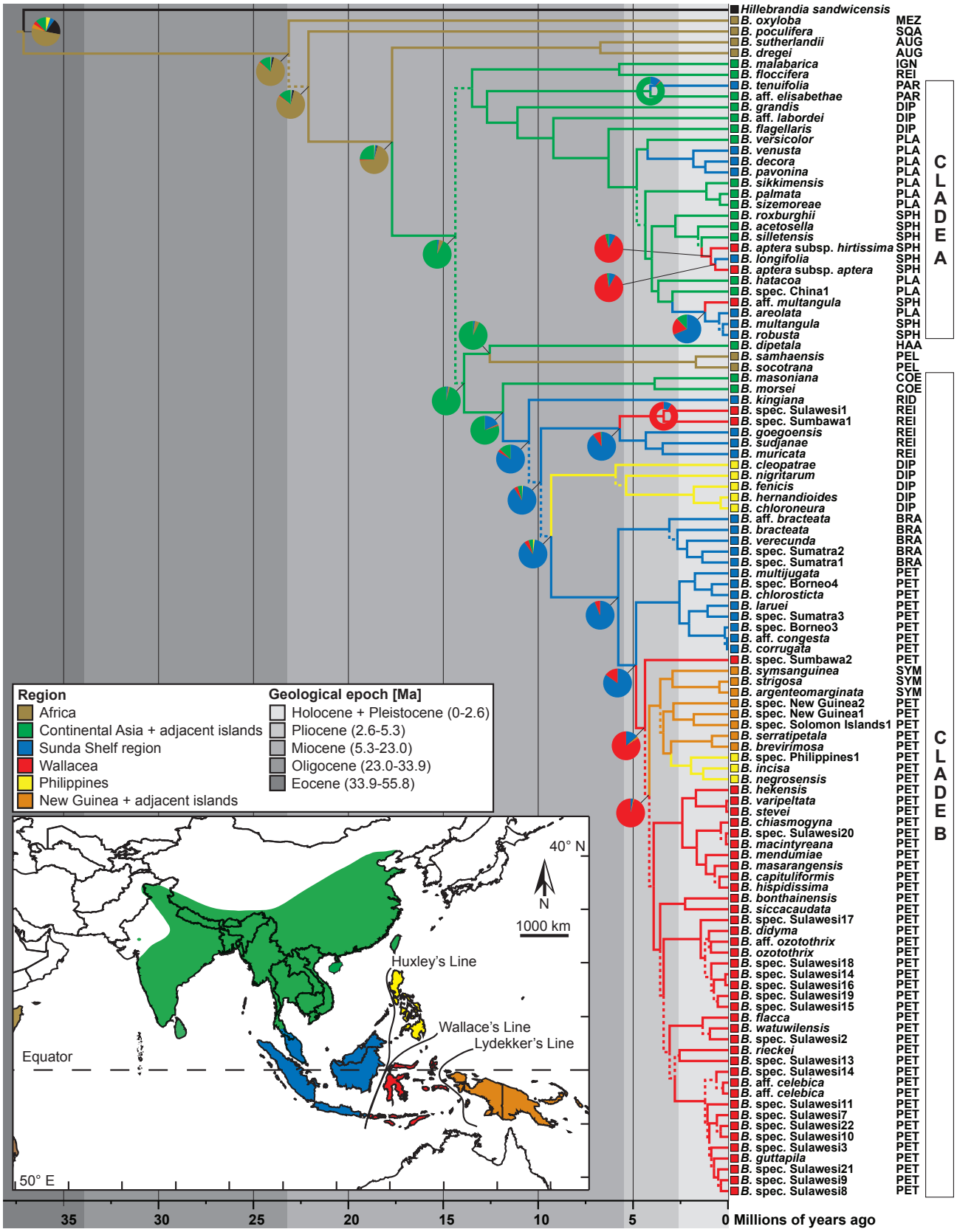


Fig. 3.7. Bayesian ancestral area reconstructions. Maximum clade credibility chronogram derived from a phylogeographic analysis in *BEAST* using secondary calibration S1 (see Table 3.5) and seven area states as defined in 3.2.2.1. Branches are coloured according to the most probable area state of their descendent nodes, and area state posterior probability (PP) distributions are indicated for nodes at which the most probable area state received < 0.98 PP. Broken lines indicate branches which lead to nodes with a clade PP < 0.95. Geological periods or epochs are indicated in different shades of grey. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, HAA: *Haagea*, IGN: unplaced to section, MEZ: *Mezierea*, PAR: *Parvibegonia*, PEL: *Peltaugustia*, PET: *Petermannia*, PLA: *Platycentrum*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQU: *Squamibegonia*, SYM: *Symbegonia.*

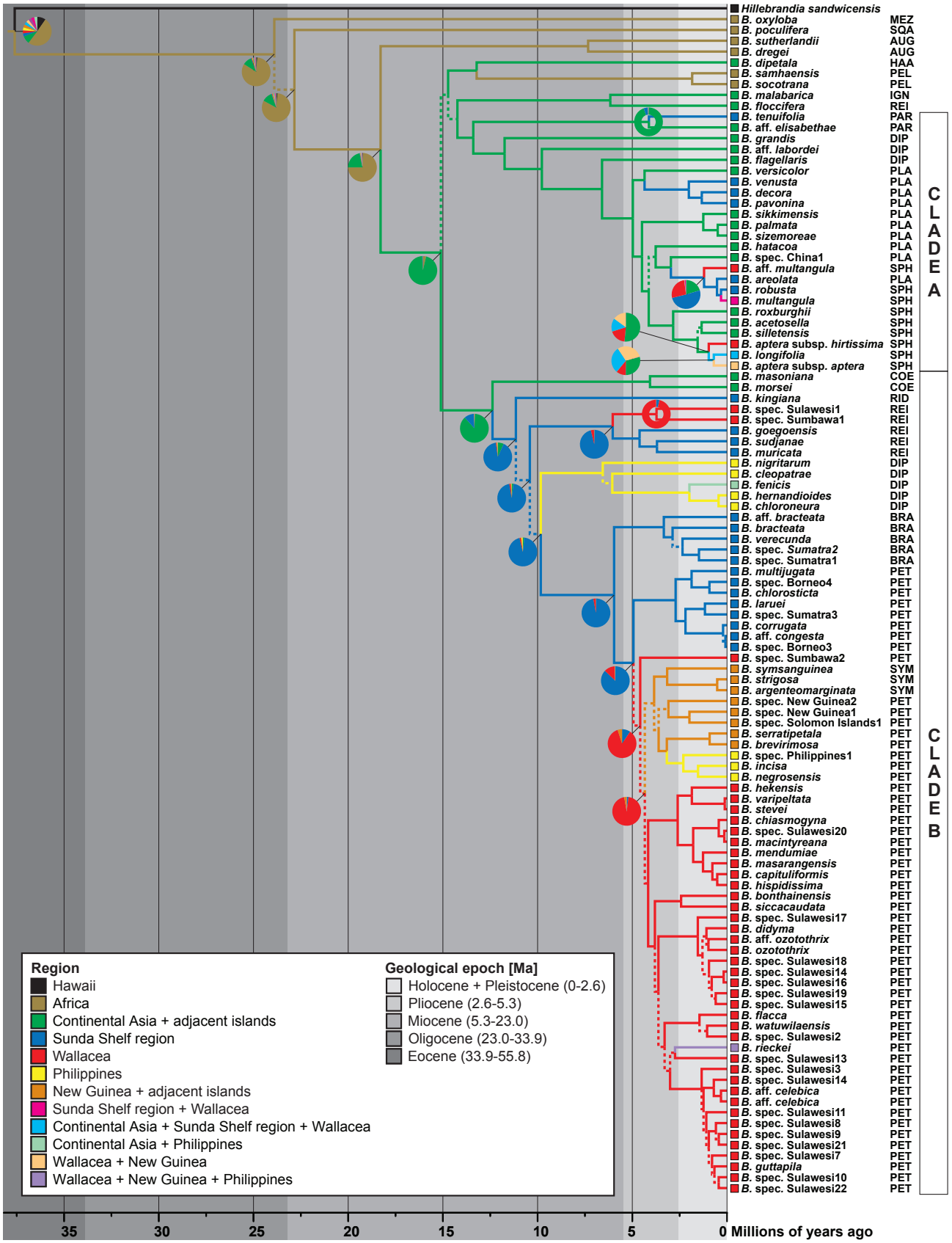


Fig. 3.8. Likelihood ancestral area reconstructions. Maximum clade credibility chronogram derived from divergence age estimates in *BEAST* using secondary calibration S1 (see Table 3.5). Branches are coloured according to the area state which received the highest proportional likelihood at their descendent nodes based on likelihood ancestral area reconstructions in *Mesquite* using the seven main area states plus the five composite area states defined in 3.2.2.1. Broken lines indicate branches which lead to nodes with a clade PP < 0.95. Geological epochs are indicated in different shades of grey. Sectional placement of taxa is indicated by the following abbreviations: AUG: *Augustia*, BRA: *Bracteibegonia*, COE: *Coelocentrum*, DIP: *Diploclinium*, HAA: *Haagea*, IGN: unplaced to section, MEZ: *Mezierea*, PAR: *Parvibegonia*, PEL: *Peltaugustia*, PET: *Petermannia*, PLA: *Platycentrum*, REI: *Reichenheimia*, RID: *Ridleyella*, SPH: *Sphenanthera*, SQU: *Squamibegonia*, SYM: *Symbegonia*.

crown node of the clade are equivocal, especially in the analyses using five additional composite area states for wider distributions. However, reconstructions at the stem nodes indicate continental Asia as most probable area state (PP: 0.99, PL: 1) for the *Begonia longifolia* complex. The second clade comprises *Begonia multangula* and *B. robusta*, and reconstruction at the crown node (PP: 1, PL: 1) and two subsequent deeper nodes support the Sunda Shelf area as most probable ancestral area. Mean divergence age estimates indicate that the divergences of the stem lineages of these two clades occurred in the Pleistocene.

Within Clade B, only species in section *Coelocentrum,* which form the sister clade to the rest of the clade, show a continental Asian distribution, while all other clades are predominantly or exclusively Malesian. The reconstructions indicate several area transitions from the Sunda Shelf region to the Philippines and to Wallacea, although statistical support is not always strong: 1. The Sunda Shelf is the most probable ancestral area both for the stem node (PP: 0.90, PL: 0.95) and for the crown node (PP: 0.84, PL: 0.96) of a clade comprising species of section *Reichenheimia* including two Sumatran and one Javanese species as well as one species from Sumbawa (Lesser Sunda Islands) and one species from Buton Island (Southeast Sulawesi). 2. The Sunda Shelf region is the most probable ancestral area reconstructed for the stem node (PP: 0.89, PL: 0.95) of a clade comprising five Philippine species of section *Diploclinium*. 3. In the clade comprising species of sections *Bracteibegonia* and *Petermannia* s.l. (including section *Symbegonia*) species in section *Bracteibegonia* with distribution on the Sunda Shelf form the sister clade to section *Petermannia*, and within section *Petermannia* Bornean and Sumatran species form the sister clade to a clade comprising Sulawesian, New Guinean and Philippine species. The Sunda Shelf region is the most probable area reconstruction at the stem node (PP: 0.89, PL: 0.95) and at the crown node (PP: 0.93, PL: 0.97) of the *Bracteibegonia-Petermannia* clade, as well as at the crown node (PP: 0.84, PL: 0.87) of the *Petermannia* clade indicating an area transition from the Sunda Shelf region to Wallacea. Within the eastern Malesian *Petermannia* clade, Wallacea is the most probable ancestral area, and Philippine and New Guinean taxa form a well supported clade, with New Guinea as most probable ancestral area reconstruction at the stem node (PP: 0.98, PL: 0.98) of the clade comprising the Philippine samples. Mean divergence estimates indicate that the two independent transitions to Wallacea in Clade B occurred in the Late Miocene (in section *Reichenheimia*) and Pliocene (in section *Petermannia*), respectively.

The predominant trend of area transitions between continental Southeast Asia, the Sunda Shelf region and Wallacea, as well as across Wallacea is from west to east (Table 3.7).

Table 3.7. Area state transitions. Based on likelihood ancestral area reconstructions in *Mesquite* (see Fig. 3.8). PL: Proportional likelihood. Taxon abbreviations: COE: *Coelocentrum*, DIP: *Diploclinium*, PEL: *Peltaugustia*, PET: *Petermannia*, REI: *Reichenheimia*.

Area State Transitions	Node	Geological time frame	Directionality	Reconstruction support [PL]
Africa ↔ Continental Asia				
Africa → Continental Asia	Asian + Socotran *Begonia* stem	Early Miocene	West → East	0.96
Continental Asia → Socotra	PEL stem	Middle Miocene	East → West	0.96
Continental Asia ↔ Sunda Shelf region				
Continental Asia → Sunda Shelf region	Sister clade of COE stem	Middle Miocene	West → East	0.88
	B. tenuifolia divergence	Pliocene	West → East	0.94
	B. venusta/*B. decora*-*B. pavonia* stem	Pliocene	West → East	0.99
	B. aff. *multangula*/*B. areolata*/ *B. robusta*-*B. multangula* stem	Pliocene	West → East	0.99
Wallace's Line				
Continental Asia → Wallacea	*B. aptera*-*B. longifolia* stem	Pleistocene	West → East	1
Sunda Shelf → Wallacea	REI (Sulawesi + Sumbawa) stem	Late Miocene	West → East	0.96
	PET (Wallacea + New Guinea + Philippines) stem	Pliocene	West → East	0.87
	B. aff. *multangula* divergence	Pleistocene	West → East	0.51
New Guinea → Philippines	PET (Philippines) stem	Pliocene	East → West	0.95
Huxley's Line				
Continental Asia → Wallacea	*B. aptera*-*B. longifolia* stem	Pleistocene	West → East	1
Sunda Shelf → Wallacea	REI (Sulawesi + Sumbawa) stem	Late Miocene	West → East	0.95
	PET (Wallacea + New Guinea + Philippines) stem	Pliocene	West → East	0.87
	B. aff. *multangula* divergence	Pleistocene	West → East	0.51
Sunda Shelf → Philippines	DIP (Philippines) stem	Late Miocene	West → East	0.95
New Guinea → Philippines	PET (Philippines) stem	Pliocene	East → West	0.98
Lydekker's Line				
Wallacea → New Guinea	PET (New Guinea + Philippines) stem	Pliocene	West → East	0.95
New Guinea → Philippines	PET (Philippines) stem	Pliocene	East → West	0.98

3.4 Discussion

3.4.1 Molecular age estimates for the Begoniaceae and *Begonia* crown group divergences

Estimates of the coefficient of variation of the analysis of the Cucurbitales-Fagales dataset indicated considerable among branch rate heterogeneity, which suggested that a relaxed molecular clock approach was appropriate (Drummond et al., 2007).

The analyses which only included fossil calibration constraints on the root node of the phylogenetic tree and the stem node of the core Fagales yielded distinctly younger age estimates than the analyses using three additional fossil constrains on divergences in Fagales and Cucubitales. The absence or presence of a constraint on the Coryloideae crown group divergence had the strongest impact and yielded ages for the Begoniaceae crown group divergence, which were 6.4 Ma older than when only the root node and core Fagales constraints were applied (Fig. 3.4). Given the good quality of the Coryloideae fossils, which are derived from well dated strata (Pigg et al., 2003), it is unlikely that an erroneous placement or dating errors are causing the strong impact of this single constraint on divergence estimates. It rather indicates the sensitivity of the relaxed molecular cock method to a relatively old age constraint (50 Ma) on the divergence of a strongly derived clade within the Fagales. This highlights the merits of including multiple fossil constraints at different

hierarchical levels, and the potential importance of the inclusion of fossil constraints in groups which are phylogenetically distant to a node of particular interest.

Several previous studies, which primarily investigated the historical biogeography of Begoniaceae (Goodall-Copestake, 2005), *Hillebrandia* (Clement et al., 2004, 2005), *Begonia* (Goodall-Copestake et al., 2009), and African *Begonia* (Plana et al., 2004) provide age estimates for the Begoniaceae crown divergence (Fig. 3.4). The mean age estimates of *c.* 40 and *c.* 42 Ma, derived here from the analyses of sequence data from four plastid regions and a low-density sampling of all families within Cucurbitales and Fagales, and five fossil constraints (Fig. 3.3-4), indicate that the Begoniaceae crown group diverged in the Eocene. These age estimates are slightly younger than an estimate of 46 Ma by Goodall-Copestake (2005), who used *rbc*L data, three fossil calibrations, and the Bayesian relaxed molecular clock method implemented in *MULTIDIVTIME* (Thorne and Kishino, 2002). The results from Clement et al. (2004, 2005) based on *rbc*L data, two fossil calibrations, and the non-parametric rate smoothing method (NPRS) implemented in *r8s* (Sanderson, 1997), indicate a similar timeframe (45 Ma), while their analyses using penalised likelihood (PL; Sanderson, 2002) resulted in distinctly older age estimates of 59 Ma. Plana et al. (2004) used ITS data, alternative single geological calibrations based on island emergence ages, and the NPRS method, and their analyses resulted in point estimates of either 16 or 43 Ma for the divergence of the Begoniaceae crown group depending on the employed geological calibration. However, Plana et al. (2004) cautioned that their age estimates were tentative as they used only island emergence dates to calibrate the divergence of single island endemics rather than island diversifications, and that large standard deviations were expected for the estimates of the Begoniaceae crown group divergence. Goodall-Copestake et al. (2009), who rigorously tested the impact of different methods, single or combined DNA region analyses and fossil calibrations on the age estimates, provide the most robust previous estimates for the divergence age of the *Begonia* crown group. The mean age estimates of their analyses of *rbc*L and 18S datasets using NPRS, PL, a Bayesian relaxed molecular clock method assuming autocorrelated rates (Thorne and Kishino, 2002) and the uncorrelated relaxed lognormal clock method implemented in *BEAST* fall within the boundaries of the Oligocene, i.e. 23-34 Ma (Fig. 3.4). The mean age estimates of *c.* 26 Ma derived from the analyses of sequence data of four plastid markers and the Fagales-Cucurbitales dataset (Figs. 3.3-4) fall within this range.

Despite the variety of data sources, taxon samplings, calibrations and methods, all of which are factors which have a substantial impact on molecular age estimates (Goodall-Copestake et al., 2009; Renner, 2005; Rutschmann, 2006), several previous studies correspondingly suggest that the divergence of the Begoniaceae crown group likely occurred in the Middle Eocene (Clement et al., 2004, 2005; Goodall-Copestake, 2005; Plana et al., 2004). The age estimates presented here (Fig. 3.3-4), which are based on a

more robust DNA region sampling and a more robust taxon sampling in Cucurbitales and Fagales in comparison to previous studies, suggest a similar timeframe, which provides some confidence that useful secondary calibrations can be derived from these estimates.

3.4.2 Temporal and spatial diversification patterns in Asian *Begonia*

3.4.2.1 Dispersal from Africa to Asia and early diversification in the Socotran-Asian clade

Intriguingly, the only two *Begonia* species known from the Yemenite Socotra Archipelago, which is more than 2000 km away from the distribution area of their closest relative in South India and Sri Lanka, fall into an early divergent subclade of a well supported Socotran-Asian clade in the cpDNA phylogenetic tree (Fig. 3.5). Goodall-Copestake (2005) hypothesised that *Begonia* dispersal from Africa to Asia may have occurred either by sweepstake dispersal, or, considering the monophyly of Socotran and Asian *Begonia*, adaptations to seasonal climates in Socotran *Begonia* and paleoclimatic reconstructions of periods of warm and moist conditions during the Paleogene and Neogene (Zachos et al., 2001), via an Arabian corridor during a time when more hospitable conditions existed than at present. The mean age estimates for the divergences of the Asian-Socotran *Begonia* stem lineage (18 Ma, 95% HPD: 11-25) and crown group (15 Ma, 95% HPD: 9-21) indicate that an area transition from Africa to Asia may have occurred in the Early to Middle Miocene (Figs. 3.5, 3.7-8), a period of predominantly moist and warm climates in Southeast Asia and with a global warm phase, the Middle Miocene climatic optimum, peaking at *c.* 17-15 Ma (Zachos et al., 2001). Reconstructions of rainforest distributions suggest that this warm phase led to the expansion of megathermal vegetation in Asia with rainforests spreading as far north as southern Japan and as far east as the northwest of the Indian subcontinent (Morley, 2007). However, evaporite and calcrete deposits suggest dry and warm conditions for extensive areas of western Asia and the Arabian Peninsula during this period (Morley, 2007; Scotese, 2003). An overland migration of *Begonia* species from Africa to Asia through this dry area seems unlikely, as most extant Asian *Begonia* species require shady, humid habitats, although some lineages in sections *Parvibegonia* and continental Asian *Diploclinium* show adaptations, such as tubers, which allow them to survive seasonally dry conditions. Kürschner (1986) and Kürschner et al. (2006) hypothesised that former exchange and migrations of the Indo-Malayan flora across Arabia have occurred in the past, and they interpreted the presence of putatively relictual mesic-African and mesic-Indo-Malayan elements in putative refuges including several fog oases, i.e. sea facing escarpments along the coastal mountains of the Southern Arabian Peninsula which capture rain and fog precipitation brought by the moist air from the sea, as indicative of former migrations. However, clear links between the Socotran flora, which shows strongest affinities to the geographically close African and Arabian floras, and the tropical Indian and tropical Southeast Asian flora are rare, and only few Indo-Malayan genoelements, like the genera *Livistona* R.Br. (Arecaceae) and *Wendlandia*

Bartl. ex DC. (Rubiaceae), have been recognized (Kilian et al., 2004; Mies, 1996; Miller and Morris, 2004). Molecular divergence age estimates for the genus *Livistona*, which predominantly exhibits an Asian-Australian distribution, suggest that the split between *Livistona carinensis* (Chiov.) J.Dransf. & N.W.Uhl, which is distributed in Somalia, Djibouti and Yemen, and an Asian-Australian clade likely occurred in the Early to Middle Miocene (Crisp et al., 2010). The Mid Miocene Climatic Optimum was followed by gradual cooling until the early Pliocene (*c.* 6 Ma) (Zachos et al., 2001) and was associated with a decline in CO_2 levels, which may have directly impacted the productivity of terrestrial vegetations (Kürschner et al., 2008). This global cooling and the associated decline in CO_2 levels resulted in the retraction of megathermal rainforests to the tropical zone and the expansion of grasslands and deserts across much of the lower and mid-latitudes (Morley, 2007). Yuan et al. (2005) investigated the historical biogeography of the genus *Exacum* L. (Gentianaceae) and hypothesised, based on phylogenetic data, molecular divergence age estimates and biogeographical analyses, that *Exacum* originated on Madagascar, dispersed to the South Indian-Sri Lankan region where the colonizer underwent a range expansion to northern India, mainland Southeast Asia and Socotra-Arabia during favourable moist and warm conditions in the Miocene. Subsequent to diversifications during these favourable conditions *Exacum* species retreated from many areas, possibly as a consequence of the expansion of drier vegetation after the Mid Miocene Optimum (Yuan et al., 2005). A similar scenario can be hypothesised for the early diversifications in Socotran-Asian *Begonia*. Within the well supported Socotran-Asian clade of the cpDNA phylogenetic tree (Fig. 3.5) five South Indian-Sri Lankan and Socotran species form two early divergent clades, whose relationships to the two main subclades A and B, which comprise all other species in the phylogenetic tree, are only poorly supported. The clade which comprises the only two *Begonia* species known from Socotra is not resolved as the sister clade to Asian *Begonia*, but falls in a well supported clade with *B. dipetala*, which is distributed in South India and Sri Lanka, and the biogeographic analyses suggest continental Asia rather than Africa as the likely ancestral area at the stem node of the Socotran lineage (Fig. 3.7-8). The conspicuous long-branches which support the crown group of Socotran *Begonia* in cpDNA phylogenetic trees (Fig. 2.3-4) suggests extinction of taxa of the stem lineage of the Socotran clade and/or a long isolation of the Socotran lineage. The geographic structure in the phylogenetic trees and the apparent paleoendemism of the Socotran lineage can be explained by long-distance dispersal from the Sri Lankan-South Indian region to the Socotra-Arabian region, or by a range expansion and diversifications of South Indian-Sri Lankan *Begonia* lineages over North India to continental Southeast Asia in the east and suitable habitats in the drier Arabian-Socotran region in the west, with subsequent extinctions in the Socotran-Arabian region, possibly as a consequence of the expansion of drier vegetation after the Mid Miocene Optimum. However, these hypothesised scenarios are tentative, as extensive extinction may have obscured the directionality of migrations. A broader taxon sampling of Sri Lankan-South Indian *Begonia* and data from additional DNA regions are needed

to resolve the relationships of the early divergent clades in Socotran-Asian *Begonia*, to mitigate potential long-branch attraction artefacts associated with the conspicuously long branches in the Socotran and South Indian-Sri Lankan lineages, and to ultimately further elucidate biogeographic patterns in the early diversifications within this group.

3.4.2.2 Colonization of the Malay Archipelago

The biogeographical analyses support the hypothesis that subsequent to an initial diversification of *Begonia* in South Asia and continental Southeast Asia, multiple migrations occurred from continental Southeast Asia to the Malay Archipelago (Figs. 3.7-8). Mean estimates of the stem node (12 Ma, 95% HPD: 7-18) and crown node ages (11 Ma, 95% HPD: 7-16) of a large, predominantly Malesian clade, which comprises sections *Ridleyella*, *Bracteibegonia*, *Petermannia* s.l. (including section *Symbegonia*) and Malesian lineages in the polyphyletic sections *Reichenheimia* and *Diploclinium*, indicate the Middle or Late Miocene as the likely timeframe for the origin of this group. Moreover, one area transition occurred in section *Parvibegonia*, and at least three area transitions are indicated for section *Platycentrum* s.l. (including section *Sphenanthera*). Stem and crown node estimates of the Malesian lineages in sections *Parvibegonia* and *Platycentrum* indicate the Plio- to Pleistocene as likely timeframe of these transitions. Palaeogeographical reconstructions suggest that substantial parts of the Sunda Shelf region, including the Malay Peninsula, were terrestrial throughout the Miocene and onwards (Fig. 1.2) (Hall, 2001, 2009). Extensive land connections existed between the Malay Peninsula and emergent parts of the Sunda Shelf during substantial intervals in the Miocene and Pliocene, and the Western Malesian islands and the surrounding continental shelf formed a vast landmass during long phases of low sea-levels caused by glacioeustatic fluctuations associated with the change of ice volume in the Northern Hemisphere during the Pleistocene (Voris, 2000; Woodruff, 2010). Thus, overland migrations in phases during which continuous suitable habitats existed sufficiently explain the pattern of multiple area transitions from continental Asia to the Sunda Shelf region (Fig. 3.7-8.).

3.4.2.3 Current distributions and diversification patterns in Southeast Asian **Begonia**

Sections *Bracteibegonia* and *Petermannia*: The large section *Petermannia* contains with *c.* 270 species almost half of the *Begonia* species diversity in Southeast Asia, and its distribution spans the biotic interface region of Wallacea as well as the wider Malay Archipelago (Fig. 2.20). Species in section *Bracteibegonia*, which is exclusively distributed in Western Malesia (Fig. 2.20), form the sister clade to a clade which comprises all 59 accessions of species placed in section *Petermannia* in the cpDNA phylogenetic trees (Figs. 3.5, 3.7-8). This *Petermannia* clade exhibits strong geographic structure. Ancestral area reconstructions and divergence age estimates for the stem lineage (6 Ma, 95% HPD: 3-9) and the crown group (5 Ma, 95 % HPD: 3-7) of the *Petermannia* clade suggest that this large section originated in the Malesian Sunda Shelf region in the Late Miocene

or Early Pliocene. Within the *Petermannia* clade, western Malesian taxa form a well supported clade, which is the sister to a clade which comprises Wallacean, New Guinean and Philippine taxa (Fig. 3.7-8). The biogeographical analyses and mean divergence age estimates for the stem lineage (5 Ma, 95% HPD: 3-7) and crown group (5 Ma, 95% HPD: 3-7) of this clade suggest that dispersal from the Sunda Shelf region across Wallace's Line to Wallacea likely occurred in the Early Pliocene. The biogeographic analyses further suggest a spatial sequence of dispersals and diversifications form the Sunda Shelf region through Wallacea to New Guinea, and from New Guinea to the Philippines. Geological data indicates that there was never one clear overland track which allowed biotic exchange between the Sunda Shelf region, Sulawesi, the Philippines and New Guinea by overland migration in the Late Miocene, Pliocene, or even during phases of past sea-levels 120 m below present levels during the Pleistocene (Hall, 2001, 2009; Voris, 2000). However, the period from the Late Miocene onwards offered opportunities for dispersal to and across Wallacea as substantial land masses emerged in Sulawesi, and the emergence of numerous volcanic islands along the Sunda Arc, the Banda Arc and the Halmahera Arc offered potential avenues for dispersal by island hopping (Hall, 2001, 2009). A migration from New Guinea to the Philippines in the Pliocene may have been facilitated by the emergence of volcanic islands of the Halmahera Arc (Hall, 2009), and putative genealogical connections of taxa between Northern New Guinea, Halmahera and Southern Philippines have been described by Michaux (2001) as the Melanesian arc track.

The strong geographic structure of the phylogenetic trees supports the hypothesis that subsequent to the dispersal to the major islands massive autochthonous radiations occurred. Mean diversification rates of 1.14 species/Ma (0.77-1.99 species/Ma)[1] for the crown group of *Begonia* section *Petermannia* fall into the range of exceptionally high diversification rates which have been recently reported from several rapid plant and animal radiations (see Scherson et al., 2008; Valente et al., 2010, and references therein). The analyses included only low density samplings of New Guinean and Philippine species placed in *Begonia* section *Petermannia*, but also 38 accessions of Sulawesian species placed in the section representing a geographically robust sampling including over 50% of the known species as well as several putatively new species from the island (Thomas et al., 2009a; Thomas et al., 2009b). Within this sample there is no indication of exchange between neighbouring islands, except for the presence of *Begonia rieckei*. *Begonia rieckei* is the oldest published name in a taxonomically difficult species complex, which includes two taxa endemic to Sulawesi, *B. koordersii* Warburg ex L.B.Sm. & Wassh. and

[1]This estimate is based on a mean age estimate of 4.9 Ma (the rate range in brackets is based on the lower and upper bounds of the 95% HPD; i.e. 2.8 and 7.2 Ma) for the crown group diversification in section *Petermannia* (Fig. 3.5), and a conservative assumption of 270 extant species in the section, and assumes a random speciation Yule model: $r = [\ln(n1) - \ln(n0)]/ t$, where r = rate of species diversification; $n0$ = initial species diversity, here taken as 1; $n1$ = extant species; t = mean age estimate or upper or lower bound of the 95% HPD of the crown group divergence.

B. strictipetiolaris Irmsch., but also *B. rieckei* (Sulawesi, Moluccas, New Guinea), *B. pseudolateralis* Warb. (Philippines), *B. brachybotrys* Merr. & L.M.Perry (New Guinea and surrounding islands), and *B. peekelii* Irmsch. (Bismarck Archipelago) (Hughes, 2008). According to Hughes (2008), these taxa may be best considered as one widespread species, as they show only minor morphological differences. The presence of *Begonia rieckei* nested within a clade of Sulawesi endemics indicates that this species complex originated on the island, and subsequently became widespread east of Huxley's Line (Fig. 3.1). *Begonia rieckei* shows conspicuous fruit and inflorescence syndromes, which are unusual in section *Petermannia*, and which may partially explain the colonization success of this species. First, their inflorescences and partial inflorescences are bisexual and do not exhibit effective dichogamy, while the inflorescences of the vast majority of species in section *Petermannia* are characterised by protogyny and a clear separation of a basal female part and a distal male part, or separate female and male inflorescences (Irmscher, 1914; Thomas et al., 2009b). It has not been investigated whether *Begonia rieckei* is autogamous, but numerous *Begonia* species were shown to be self-fertile in cultivation (East, 1940), and some *Begonia* species which lack effective dichogamy were found to be highly selfing in the wild (Ågren and Schemske, 1993). Autogamy is expected to be beneficial to colonization success, because a single individual is sufficient for colonisation, and autogamous plant lineages are well represented in current invasive floras (see Harmon-Threatt et al., 2009 and citations therein). Second, the fruits of *Begonia rieckei* exhibit a fleshy pericarp and reduced wings, while the vast majority of species in the section exhibit thin-walled capsules with well developed wings and a dry, membraneous pericarp at maturity. Tebbitt et al. (2006) hypothesised that some fleshy-fruited species in section *Platycentrum* s.l. are dispersed by bats and other animals, and that zoochory may be a factor contributing to the unusually wide distributions of some taxa in the section. It can be speculated that zoochory may also occur in the fleshy-fruited, widespread *Begonia rieckei* complex. However, despite its wide distribution, the ecology of this species is only poorly known, and more cryptic physiological characteristics such as wider tolerances towards different soil types, pHs, light intensities, humidity levels and habitat disturbance may also explain the widespread distribution of this species.

Section *Reichenheimia*: Malesian species placed in section *Reichenheimia* form a strongly supported clade in the phylogenetic trees, and are apparently only distantly related to Indian, Sri Lankan and continental Asian species placed in this section (Fig, 3.5; see discussion of phylogenetic relationships in Chapter 2). This group is mainly distributed in the Sunda Shelf region, and most species have been described from the Malay Peninsula and Sumatra (Hughes, 2008; Kiew, 2005), while it is apparently absent from Continental Asian regions north of the Thai-Malay Peninsula, which show a monsoonal seasonal climate with pronounced dry seasons. Only *Begonia muricata* Blume, and a few putatively new species extend the distributional range to islands east of Wallace's Line including the Lesser Sunda Islands (Sumba, Sumbawa), Buton and Wowoni, which lie

off the southeastern coast of the Southeastern arm of Sulawesi, and Seram (Fig. 2.19) (Hughes, 2008, for records from Southeast Sulawesi see Coode s.n. at K, L). The Sunda Shelf region is reconstructed as the likely ancestral area at stem and crown nodes of the Malesian *Reichenheimia* clade and divergence age estimates of the crown and stem nodes of a clade comprising two unidentified species from Sumbawa and Southeast Sulawesi indicate that migration from the Sunda Shelf region to Wallacea occurred sometime in the Late Miocene or Pliocene (Fig, 3.5, 3.7-8).

Section *Diploclinium*: Malesian species placed in the polyphyletic section *Diploclinium* include a large radiation on the Philippines (> 40 spp.) (Rubite, 2010), five species described from Borneo and seven species on New Guinea (Fig. 2.21) (Hughes, 2008; Hughes et al., 2010). The analyses only included samples from the Philippines (including Palawan), which limits their power to elucidate the historical biogeography of this group. Philippine species placed in section *Diploclinium* form a well supported clade in a subclade of Clade B which comprises several other major Malesian lineages (Fig. 3.5). The Sunda Shelf region is reconstructed as likely ancestral area at the stem node of this clade, but additional data from the Bornean and New Guinean species placed in section *Diploclinium* are needed to test whether Bornean and New Guinean lineages are nested within and derived from lineages of the large Philippine radiation, or whether Bornean taxa form early divergent lineages within this group, which would evoke an origin on Borneo and dispersal from Borneo to the Philippines along probable avenues for dispersal such as Palawan and the Sulu Archipelago (Atkins et al., 2001; Evans et al., 2003).

Section *Platycentrum* s.l.: Malesian species in section *Platycentrum* s.l. are predominantly distributed on the Malay Peninsula, Sumatra, and Java (Hughes, 2008). Only one species has been described from Borneo, and only few species extend the range of the section to the east of Wallace's Line (Fig. 2.19). These eastern Malesian lineages belong to two taxonomically difficult complexes of fleshy-fruited species: the *Begonia longifolia* complex (Tebbitt, 2003) and the *Begonia robusta* complex (Hughes, 2008). *Begonia longifolia* is the most widespread species in Asia and its distribution ranges from Northeastern India, Bhutan, China, Indo-China, through the Malay Peninsula, Sumatra, Java, Bali to Sulawesi (Hughes, 2008; Tebbitt, 2003). Based on morphological data and current distributions, Tebbitt (2003) hypothesised that *Begonia longifolia* originated in the mountainous region between northeastern India and northern Vietnam, and migrated from there along mountain corridors to Malesia and Taiwan. Population genetic and phylogeographic approaches using data from highly variable markers and multiple accessions are needed to elucidate the historical biogeography of this species, but the biogeographical analyses clearly corroborate the hypothesis that ancestors of the *Begonia longifolia* complex first diversified on the Southeast Asian mainland and that the Malay Archipelago was subsequently colonized during the Pleistocene. The analyses indicate similar temporal and spatial patterns for the *Begonia robusta* complex, the distribution

of which extends from Sumatra and Java to the Lesser Sunda Islands, and Sulawesi (Hughes, 2008) (Fig. 3.7-8). Species in both these groups seem to have colonized large areas in Malesia in a relatively short period of time during the Pleistocene (Fig. 2.19), and the fleshy fruits and associated putative zoochory by bats and other animals (Tebbitt et al., 2006), as well as relatively wide ecological tolerances of some species (Tebbitt, 2003) may, to a certain extent, explain their colonisation success. Moreover, the relatively recent origins of these groups in combination with putative good dispersal capabilities facilitating gene flow between populations may explain the relatively poor morphological differentiation within these widespread species complexes.

Section *Parvibegonia*: Only two of *c.* 30 species placed in section *Parvibegonia* were included in the analyses, but current distribution patterns in combination with the phylogenetic data and divergence age estimates provide the basis for some hypotheses about the historical biogeography of this group. Section *Parvibegonia* is diverse on the Asian mainland and the Malay Peninsula, but only five species extend the distributional range of the section to the island of Banka, which lies off the east coast of Sumatra, and to Java and the Lesser Sunda Islands (Fig. 2.19) (Hughes, 2008; Kiew, 2005). Many Malesian species in section *Parvibegonia* can resist drought by dying down during the dry season and by resprouting from tuberous perennating organs in the next rainy season. Although they are perennials, these species show a similar life strategy as drought avoiding annual plants. Most species in the section are small plants, i.e. a single plant produces a relatively low quantity of biomass, mature flowers are often developed relatively soon after resprouting and germination, and they profusely regenerate from the seed bank in the rainy season (see species descriptions in Kiew, 2005). These adaptations seems to be essential to survive dry seasonal extremes or more pronounced seasonal monsoonal climates with several dry months, not just on continental Southeast Asia (Phutthai et al., 2009), but also in the North of the Malay Peninsula (Kiew, 2005), eastern Java and the Lesser Sunda Islands (*pers. obs.*). In contrast to this, the majority of Malesian *Begonia* species depend on moist, shady habitats and more or less everwet conditions. The current distribution of species in section *Parvibegonia* may be partially correlated with the current climate in Malesia, which is characterised by predominantly everwet conditions on the Sunda Shelf in the west and the Sahul Shelf in the east, while parts of Wallacea and Java exhibit a drier monsoonal, seasonal climate (van Welzen et al., 2005). Section *Parvibegonia* is absent from large parts of the Sunda Shelf including Sumatra except for Bangka Island, and Borneo. However, four species have been described from Java, and *Begonia tenuifolia* extends the range of the section to the east of Wallace's Line to the islands of Nusa Tenggara Barat (Fig. 2.20) (Hughes, 2008). The record of the widespread and variable species *Begonia sinuata* from Bangka Island is interesting, as paleogeographic reconstructions indicate that Bangka Island was part of a land corridor, which connected the Malay Peninsula with Borneo and southwest Sumatra and at times with eastern Java in the Late Miocene and Pliocene. Continental Asia is reconstructed as the most probable ancestral area at the stem

node of the *Parvibegonia* clade (Fig. 3.7-8), and it can be hypothesised that ancestors of the Malesian lineages in section *Parvibegonia*, which were likely preadapted to seasonal climates, successfully colonized the Malesian Sunda Shelf region from the mainland via the Malay Peninsula and possibly via a land corridor which included Bangka Island. This migration may have been facilitated by phases of low sea-levels and more widespread seasonal climates caused by glacioeustatic fluctuations in the Pliocene and Pleistocene. It can be further speculated that phases of more widespread everwet conditions may have resulted in retractions of the distributions of Malesian species in section *Parvibegonia* because they were outcompeted by species which were better adapted to these climatic conditions. Van Welzen et al. (2005) hypothesised similar scenarios for the historical biogeography of some grass species in the genus *Arthraxon* P.Beauv. which show disjunct distributions on the Asian mainland and in the more seasonal areas of Java and Wallacea, but are largely absent from the everwet Sunda and Sahul Shelf regions apart from several locally dryer areas most of which are influenced by mountain rain shadows.

3.4.2.4 Drivers of diversification in Southeast Asian Begonia

Molecular age estimates indicate that lineages in *Begonia* section *Petermannia* have diversified rapidly since the Pliocene with diversification peaking in the Pleistocene (Figs. 3.7-8), and massive autochthonous radiations occurred on Borneo (*c.* 90 spp.), Sulawesi (*c.* 35 spp.), New Guinea (*c.* 70 spp.) and the Philippines (*c.* 65 spp.) resulting in almost entirely endemic *Begonia* floras (Girmansyah, 2009; Girmansyah et al., 2009; Hughes, 2008; Hughes and Coyle, 2009; Hughes et al., 2010; Hughes et al., 2009; Kiew and Sang, 2009; Thomas et al., 2009a; Thomas et al., 2009b; Thomas and Hughes, 2008). The timing of diversifications coincides with substantial orogenesis on Sulawesi, New Guinea and the Philippines (Hall, 2001, 2009). Many Southeast Asian species in *Begonia* are local endemics restricted to rather narrow altitudinal ranges in single mountain chains or even a single peak or valley (Hughes, 2008; Kiew, 2005; Sands, 2001). Examples for local endemism come from the mountain flora of Mt. Kinabalu (4094 m) on Borneo, from which 17 indigenous *Begonia* species have been described (Sands, 2001). Fourteen of the 17 species, most of which are confined to altitudinal ranges in upland and montane rain forests, are only known from the Kinabalu massif (Sands, 2001). Similar patterns can be observed in the Cameron Highlands of the Malay Peninsula (Kiew, 2005), and the less well-studied mountain floras of Sulawesi and New Guinea. Three *Begonia* species placed in section *Petermannia* (*B. hekensis* D.C.Thomas, *B. stevei* M.Hughes, *B. varipeltata* D.C.Thomas) were described from Gunung Hek in eastern Central Sulawesi (Hughes, 2006; Thomas et al., 2009b; Thomas and Hughes, 2008). These species are morphologically very distinct from each other with regards to both vegetative and generative characters (see Appendix 1), but all of them fall into a well supported clade in the cpDNA phylogenetic tree indicating a local radiation, in which a high degree of morphological differentiation evolved in a relatively short period of time (Fig. 3.5). Considering the preponderance of narrow endemics in Southeast Asian

Begonia, the majority of which are confined to upland or montane primary rainforests, and the strong genetic and morphological differentiation of subpopulations of some *Begonia* species at very local scales indicating very limited dispersal capabilities and limited gene flow between populations (Hughes and Hollingsworth, 2008; Hughes et al., 2003; Matolweni et al., 2000), it can be hypothesised that the presence of geologically dynamic highlands on Borneo and the formation and patchy distribution of suitable habitats by orogenesis on Sulawesi, the Philippines and New Guinea was likely a crucial factor in the diversification in *Begonia* in the Southeast Asian tropics. Moreover, the concentration of rapid diversification in Southeast Asian *Begonia* in the Plio- and Pleistocene coincides with pronounced climate and sea-level fluctuations, which have been proposed as crucial drivers of diversifications in the tropics (e.g. Haffer, 1969, 1997: Amazonian forest birds; Gorog et al., 2004: Southeast Asian murine rodents; Harris et al., 2000: *Aframomum* K.Schum., Zingiberaceae; Janssens et al., 2009: *Impatiens* L., Balsaminaceae; Quek et al., 2007: Southeast Asian ants; Richardson et al., 2001: *Inga* Mill., Fabaceae; Sosef, 1994: African *Begonia*, Begoniaceae). Fragmentation and replacement of rainforests by seasonal vegetation due to drier and cooler climates during the Pleistocene may have been less pronounced in Southeast Asia and may have followed different patterns than in other tropical areas (Cannon et al., 2009; Morley, 2000, 2007), but climate fluctuations may still have provided isolating mechanisms when montane rainforests expanded and contracted and *Begonia* populations were forced to migrate following their required habitat conditions, which may have resulted in frequent isolation and amalgamation of populations. Moreover, eustatic sea-level changes and associated land exposure and submergence may have had an impact on the diversification of lowland rainforest species, especially in the Sunda Shelf region and the Philippines. Paleogeographic reconstructions show that land bridges connected the current major islands in the Sunda Shelf region during substantial phases of the Pliocene, and that there were almost continuous land connections between continental Southeast Asia and the islands of Sumatra, Java and Borneo for most of the last 2 Ma (Hall, 2009; Woodruff, 2010), which may have facilitated migrations between the islands and expansion of the distributions of lowland species of *Begonia*. Fragmentation of more widespread populations during phases of high sea-levels may have led to disjunct distributions and allopatric speciation. Van Welzen et al. (van Welzen et al., 2005) pointed out that phylogenetic trees of various plant taxa distributed on the Sunda Shelf indicate a pattern of vicariance and exchange between the western Malesian islands. Multiple exchanges between Sumatra and Borneo are also indicated in the phylogenetic trees, in which Sumatran and Bornean species placed in *Begonia* section *Petermannia* form mixed assemblages within a Western Malesian clade.

3.5 Summary and Conclusions

The geographically structured phylogenetic trees and the divergence age estimates suggest a biogeographical scenario involving initial diversification of Southeast Asian *Begonia* on the Asian mainland in the Miocene and multiple subsequent dispersals into Malesia. Overland migration sufficiently explains dispersal from continental Asia to Malesia and migrations across the Sunda Shelf region, as substantial parts of the Sunda Shelf region were terrestrial and connected with continental Asia throughout most of the Miocene and onwards and potentially provided contiguous habitats suitable for *Begonia* (Fig. 1.2) (Hall, 2001, 2009). Biogeographic analyses further indicate that the predominant directional trend of dispersals between continental Asia and Malesia as well as within Malesia has been from west to east, including at least four independent area transitions from continental Asia and the Sunda Shelf region across Wallace's Line to Wallacea dating from the Late Miocene to the Pleistocene (see Table 3.7). The inferred large scale diversification patterns seem to be correlated with the geological processes in the region. While parts of western Malesia were terrestrial throughout the Oligocene, Miocene and onwards, substantial land east of Wallace's Line, in Wallacea and New Guinea, only emerged during the Late Miocene and Pliocene (Hall, 2001, 2009). The divergence ages suggest that once land became available in Wallacea, at least four *Begonia* lineages independently dispersed into the region east of Wallace's Line. However, the monophyly of a large eastern Malesian clade, which comprises all samples of Sulawesian, New Guinean and Philippine species placed in section *Petermannia*, as well as the monophyly of the well sampled Sulawesian species of this section, show that despite high species numbers on the major islands and island groups in eastern Malesia, there is no indication for frequent dispersal between Wallacea and the Malesian Sunda Shelf region. Dispersal across the ancient deep water channels separating intervening islands of the Sunda Shelf and Wallacea and subsequent successful colonisation of Wallacean islands seem to have been rather infrequent events, suggesting that the water bodies which have separated the Sunda Shelf region from Wallacea have been distinct, yet porous barriers to dispersal in *Begonia*. Dispersal of multiple *Begonia* lineages from the Sunda Shelf region to Wallacea occurred within the last 10 Ma, and likely within a Plio- to Pleistocene timeframe. Dated phylogenetic trees of taxa which are well represented with regard to species numbers on all major Malesian islands are still rare, but diversification patterns of several other plant taxa show similar spatial and temporal sequences with predominantly west to east migrations, and dispersal from the Sunda Shelf region to Sulawesi and New Guinea in the Middle to Late Miocene (Muellner et al., 2008: *Aglaia* Lour., Meliaceae; Su and Saunders, 2009: *Pseuduvaria* Miq., Annonaceae; Poulsen, 2009: *Etlingera* Giseke, Zingiberaceae; Twyford, 2009: *Rhododendron* L. subgenus *Vireya*, Ericaceae). The period from the Late Miocene onwards offered opportunities for dispersal to and across Wallacea to New Guinea as substantial land masses emerged in Sulawesi and New Guinea, and the emergence of volcanic islands

along the Sunda Arc, the Banda Arc and the Halmahera Arc offered potential avenues for dispersal by island hopping (Hall, 2001, 2009). Tectonic migration, i.e. rafting on continental microfragments has been proposed as another potential mechanism aiding dispersal into and within Malesia (Ladiges et al., 2003; Michaux, 1991; Morley, 2000; 1998). The Makassar Strait extended by block faulting and subsidence, separating a fragment from Borneo and forming a deep water channel in the Eocene, and the separated fragment subsequently amalgamated with elements of volcanic, ophiolitic, and Australian continental origin to form the composite island of Sulawesi (Hall, 2001, 2009; Morley, 2000; Moss and Wilson, 1998; Ridder-Numan, 1996). However, the molecular divergence age estimates indicate that the extension of the Makassar Straits in the Eocene long preceded the diversification of *Begonia* within Malesia, and palaeogeographic reconstructions suggest that the continental microfragments which amalgamated to form parts of Sulawesi were submerged during long phases of the migration to their current position (Hall, 2001, 2009). Thus, it is unlikely that tectonic migration was an important factor in the dispersal and diversification of *Begonia* in Malesia.

Lineages in the largest Asian section, section *Petermannia*, diversified rapidly since the Pliocene with diversification peaking in the Pleistocene. The timing of diversification coincides with massive orogenesis on Sulawesi and New Guinea, as well as pronounced sea-level and climate fluctuations. It can be hypothesised that a complex interplay of extrinsic and intrinsic factors including the presence and formation of suitable microhabitats by orogenesis, cyclical vicariance by frequent habitat fragmentations and amalgamations caused by sea-level and climate fluctuations, as well as only weakly developed mechanisms to maintain species cohesion in fragmented habitats in *Begonia* could have driven speciation in allopatry and could have resulted in the remarkable *Begonia* species diversity found in Asia today.

CHAPTER 4. General conclusions

Studies of the diversifications in giant genera such as *Astragalus* L., *Begonia*, *Croton* L., and *Impatiens*, have provided insights into the processes which underlie modern patterns of biodiversity (Sosef, 1994: African *Begonia*; Wojciechowski et al., 1999: *Astragalus*; Plana et al., 2004: African *Begonia*; Berry et al., 2005: *Croton*; Scherson et al., 2008: *Astragalus*; Jannsens et al., 2009: *Impatiens*). However, because of their sheer size and their morphological complexity, robust alpha-taxonomic foundations and a good understanding of the phylogenetics of major lineages within mega-diverse genera have often remained elusive, impeding biogeographical and evolutionary research. Considerable progress has recently been made on the systematics and phylogenetics of species-rich Asian *Begonia*, which has a great potential as model group to study patterns of diversifications in the Asian tropics. Important advances in the research on this group include the conspectus of sections of *Begonia* by Doorenbos et al. (1998), the well resolved framework phylogeny of Begoniaceae by Goodall-Copestake (2005), and *The Southeast Asian Begonia Database* (Hughes and Pullan, 2007) and *The Annotated Checklist of Southeast Asian Begonia* (Hughes, 2008), which have greatly facilitated alpha-taxonomic work in the region. The alpha-taxonomic foundation of Asian *Begonia* studies has been improved by the revision of *Begonia* from the Malay Peninsula (Kiew, 2005), the recent *Flora of China* treatment of the genus (Ku et al., 2007), and over 50 publications describing new species of Asian *Begonia* which have been published since Golding and Wasshausen's synopsis (2002) (see *ISI Web of Knowledge*: http://apps.isiknowledge.com). Several other flora treatments and regional revisions have recently been completed or are underway (Grierson, 1991: Bhutan; Sangeeta Rajbhandary, Tribhuvan University, Kathmandu, Nepal, *pers. com.*: Nepal; Thamarat Phutthai, Prince of Songkla University, Hat Yai, Thailand, *pers. com.*: Thailand; Mark Hughes, Royal Botanic Garden Edinburgh, UK, *pers. com.*: Sumatra; Girmansyah, 2009: Bali and Lombok; Deden Girmansyah, Cibinong Science Centre, Cibinong, Indonesia, *pers. com.*: Java and the Lesser Sunda Islands; Harry Wiriadinata, Deden Girmansyah, Cibinong Science Centre, Cibinong, Indonesia, *pers. com.*: Sulawesi; Hughes and Coyle, 2009, Hughes and Coyle, 2010: Palawan; Rubite, 2010: section *Diploclinium* on the Philippines; Gagul, 2009: section *Symbegonia* on New Guinea). Moreover, molecular phylogenetic studies have identified several polyphyletic Asian sections and the homoplasy of some characters traditionally used to define sections (Badcock, 1998; Forrest, 2001; Forrest and Hollingsworth, 2003; Forrest et al., 2005; Tebbitt et al., 2006). However, the poorly resolved deeper internal relationships within published phylogenies of Asian *Begonia*, most of which are based on fast evolving nrDNA markers, have been an impediment to a better understanding of the evolution and biogeography of this group (Badcock, 1998; Forrest, 2001; Forrest and Hollingsworth, 2003; Forrest et al., 2005; Tebbitt et al., 2006).

The results of the phylogenetic reconstructions based on ITS and non-coding chloroplast sequence data presented here (Chapter 2), demonstrate the value of moderately fast evolving non-coding cpDNA markers in phylogenetic reconstructions of Southeast Asian *Begonia*. The phylogenies derived from analyses of the non-coding plastid data presented in this thesis provide for the first time a reasonably resolved and supported phylogenetic framework for Asian *Begonia*, which clarifies aspects of the character evolution within this species-rich and morphologically complex taxon. Some taxonomic consequences can immediately be derived from the results, and the formal description on sectional level of two monophyletic, morphologically distinct predominantly Malesian groups of species currently placed in sections *Diploclinium* and *Reichenheimia* are underway (Rubite, 2010; Mark Hughes, Royal Botanic Garden Edinburgh, UK, *pers. com.*). However, before a comprehensive, stable and natural reclassification of Asian *Begonia* can be achieved further phylogenetic studies including a broader taxon sampling of several Indian, Sri Lankan and continental Asian lineages which were are currently placed in the highly polyphyletic sections *Diploclinium* and *Reichenheimia* are needed.

Whole chloroplast genomes of sixteen *Begonia* species, including the Socotran *Begonia socotrana* and the Asian *B. varipeltata* and *B. venusta*, have been sequenced and assembled in the Kidner lab at the Royal Botanic Garden Edinburgh (Catherine Kidner, Nicola Burton, Royal Botanic Garden Edinburgh, UK, *pers. com.*), and comprehensive comparisons of coding and non-coding cpDNA region variability based on this sample will greatly facilitate the identification of suitable markers for phylogenetic analyses at various taxonomic levels. This research will also pave the way for phylogenetic work in *Begonia* based on whole or large scale chloroplast genome data, which may have the power to resolve the basal relationships within the Socotran-Asian clade, as well as the relationships between section *Ridleyella*, Malesian *Diploclinium*, Malesian *Reichenheimia* and the monophyletic group which comprises sections *Bracteibegonia* and *Petermannia*, which have remained elusive. However, the observed conflicting positions of some taxa in nrDNA ITS and cpDNA gene trees indicate that reticulation may have been an important factor in the evolution of *Begonia* (Chapter 2). A single genome approach, even if based on large quantities of cpDNA sequence data, has very limited power to detect and investigate the processes which cause phylogenetic incongruence in *Begonia*. Independent sources of molecular phylogenetic data such as sequence data from mitochondrial regions and unlinked nuclear genes are needed to further investigate the impact of reticulation in the evolutionary history of *Begonia*. Once multiple independent datasets are generated, phylogenetic network methods may be more appropriate than standard MP, ML and Bayesian phylogenetic reconstruction methods to detect and quantify factors such as reticulation and homoplasy which can lead to gene tree conflict (Huson and Bryant, 2006; Vriesendorp and Bakker, 2005).

The molecular divergence age estimates in combination with biogeographical analyses based on non-coding cpDNA sequence data indicated broad scale patterns of the temporal and spatial diversification of *Begonia* in Southeast Asia, and allowed hypothesising about potential geological and climatic correlates (Chapter 3). The inferred west to east trend in the colonization of the Malay Archipelago and in dispersal across Malesia, and the likely timing of crossings of purported barriers to dispersal in Malesia from the Late Miocene onwards are largely concordant with both hypotheses about the palaeogeography of the region (Hall, 2009) and spatio-temporal diversification patterns observed in other taxa (Muellner et al., 2008: *Aglaia*, Meliaceae; Su and Saunders, 2009: *Pseuduvaria*, Annonaceae; Poulsen, 2009: *Etlingera*, Zingiberaceae; Twyford, 2009: *Rhododendron* subgenus *Vireya*). The preponderance of narrow endemics, the limitation to primary habitats, and apparently low dispersal capabilities of *Begonia* make the genus not only an excellent model group to investigate biogeographical patterns at the intercontinental (Goodall-Copestake, 2005) or continental scale (Plana et al., 2004: Africa; this study: Asia), but also at much narrower, regional levels. For example, the phylogenetic relationships of the geographically robust sample of *Begonia* section *Petermannia* from Sulawesi included in the cpDNA phylogeny exhibits distinct subclades of species endemic to the Northern arm, the Southwestern arm and the Southeastern arm of the island. Phylogenetic and phylogeographic studies of patterns of *Begonia* diversification based on geographically robust samplings may provide valuable insights into the geological and evolutionary processes underlying current *Begonia* distribution patterns on the major Malesian islands, and could be used to test hypotheses about postulated local areas of endemism and the factors shaping them (Evans et al., 2003).

The rationale of the biogeographical inference approach adopted here (Chapter 3) was to focus on the detection of pattern congruence between cladogenesis, molecular divergence ages, and geological or climatic events indicated by palaeogeographical and climatic reconstructions. This approach did not directly include geographical constraints, e.g. a higher likelihood of area transitions between neighbouring areas than between more distant areas, or constraints placed on the timing of area transitions based on the presence or absence of windows of dispersal opportunity as indicated by postulated emergence or submergence of land bridges in Malesia in palaeogeographic reconstructions, in the inferences. An alternative approach is outlined by Sanmartin et al. (2008) and Ree and Sanmartin (2009), who convincingly argue for an integration of organismal phylogenies, divergence age estimates, and information about historical landscapes in inferences of the range evolutions of taxa. The methodologies of their model approaches are in an early stage. Modelling can be a powerful tool for the inference of range evolution on volcanic island chains or archipelagos, whose geological history is relatively simple and well understood, but complex region histories as in Southeast Asia make model parameter definition more problematic (Ree and Sanmartin, 2009). However, using the existing detailed geological and palaeogeographic reconstructions of Southeast Asia (Hall, 2002,

2009) as a framework for model approaches, and rigorously testing the sensitivity of results to parameter variation, may be a way to integrate disparate sources of available data and to test hypotheses about the range evolution of *Begonia* lineages in the region (Ree and Webb, 2009).

The active and often cooperative research on Southeast Asian *Begonia* has resulted in a comprehensive checklist (Hughes, 2008), an online database which provides protologue, type and specimen information and pictures (Hughes and Pullan, 2007), as well as in a large body of alpha-taxonomic studies and flora treatments. Given this taxonomic foundation, the existing and growing taxonomic expertise, and the insights into the phylogenetics and historical biogeography presented here, the ambitious aim of a modern conspectus of Southeast Asian *Begonia* seems to be within reach.

5. References

Ågren, J., & Schemske, D.W. 1991. Pollination by deceit in a neotropical monoecious herb, *Begonia involucrata*. *Biotropica* 23:235-241.

Ågren, J., & Schemske, D.W. 1993. Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata*. *Evolution* 47:125-135.

Alfaro, M.E., Karns, D.R., Voris, H.K., Brock, C.D., & Stuart, B.L. 2008. Phylogeny, evolutionary history, and biogeography of Oriental-Australian rear-fanged water snakes (Colubroidea: Homalopsidae) inferred from mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution* 46:576-593.

Álvarez, I., & Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29:417-434.

Angiosperm Phylogeny Group. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141:399-436.

Angiosperm Phylogeny Group. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161:105-121.

Atkins, H., Preston, J., & Cronk, Q.C.B. 2001. A molecular test of Huxley‘s line: *Cyrtandra* (Gesneriaceae) in Borneo and the Philippines. *Biological Journal of the Linnean Society* 72:143-159.

Badcock, Z. 1998. A phylogenetic investigation of *Begonia* L. section *Knesebeckia* (Klotzsch) A.D.C. PhD thesis. University of Glasgow, Glasgow.

Barkley, F.A., & Baranov, A.I. 1972. The sections of the Begoniaceae. *Buxtonian* 1:1-8.

Beck, R.A., Burbank, D.W., Sercombe, W.J., Riley, G.W., Barndt, J.K., Berry, J.R., Afzal, J., Khan, A.M., Jurgen, H., Metje, J., Cheema, A., Shafique, N.A., Lawrence, R.D., & Khan, M.A. 1995. Stratigraphic evidence for an early collision between Northwest India and Asia. *Nature* 373:55-58.

Bell, C.D., Soltis, D.E., & Soltis, P.S. 2005. The age of the angiosperms: A molecular timescale without a clock. *Evolution* 59:1245-1258.

Benton, M.J. 2009. The Red Queen and the Court Jester: Species diversity and the role of biotic and abiotic factors through time. *Science* 323:728-732.

Benton, M.J., & Donoghue, P.C.J. 2007. Paleontological evidence to date the tree of life. *Molecular Biology and Evolution* 24:26-53.

Bergsten, J. 2005. A review of long-branch attraction. *Cladistics* 21:163-193.

Berry, P.E., Hipp, A.L., Wurdack, K.J., Van Ee, B., & Riina, R. 2005. Molecular phylogenetics of the giant genus *Croton* and tribe Crotoneae (Euphorbiaceae sensu stricto) using its and *trn*L-*trn*F DNA sequence data. *American Journal of Botany* 92:1520-1534.

Borsch, T., & Quandt, D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* 282:169-199.

Bouman, F., & de Lange, A. 1983. Structure, micromorphology of *Begonia* seeds. *Begonian* 50:70-78, 91.

Brandon-Jones, D. 2001. Borneo as a biogeographic barrier to Asian-Australasian migration. Pp. 365-372 in: Metcalfe, I., Smith, J.M.B., Morwood, M., & Davidson, I., (eds), *Faunal and floral migrations and evolution in SE Asia-Australasia*. A.A. Balkema Publishers, Lisse.

Briggs, J.C. 2003. The biogeographic and tectonic history of India. *Journal of Biogeography* 30:381-388.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D., & Rodrigues, A.S.L. 2006. Global biodiversity conservation priorities. *Science* 313:58-61.

Brown, J.M., Hedtke, S.M., Lemmon, A.R., & Lemmon, E.M. 2010. When trees grow too long: Investigating the causes of highly inaccurate Bayesian branch length estimates. *Systematic Biology* 59:145-161.

Burnham, K.P., & Anderson, D.R. 2004. Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research* 33:261-304.

Burt-Utley, K. 1985. A revision of Central American species of *Begonia* section *Gireoudia* (Begoniaceae). *Tulane Studies in Zoology and Botany* 25:1-131.

Cannon, C.H., & Manos, P.S. 2003. Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). *Journal of Biogeography* 30:211-226.

Cannon, C.H., Morley, R.J., & Bush, A.B.G. 2009. The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proceedings of the National Academy of Sciences of the United States of America* 106:11188-11193.

Carlquist, S. 1985. Wood anatomy of Begoniaceae, with comments on raylessness, paedomorphosis, relationships, vessel diameter, and ecology. *Bulletin of the Torrey Botanical Club* 112:59-69.

Chandler, M.E.J. 1964. *The Lower Tertiary floras of Southern England IV: A summary and survey of findings in the light of recent botanical observations*. British Museum, London.

Chiang, T.Y., Hong, K.H., & Peng, C.I. 2001. Experimental hybridization reveals biased inheritance of the internal transcribed spacer of the nuclear ribosomal DNA in *Begonia* × *taipieiensis*. *Journal of Plant Research* 114:343-351.

Clarke, C.B. 1879. Begoniaceae. Pp. 635–656 in: Hooker, J.D., (ed), *Flora of British India*. L. Reeve & Co., London.

Clement, W.L., Tebbitt, M.C., Forrest, L.L., Blair, J.E., Brouillet, L., Eriksson, T., & Swensen, S.M. 2004. Phylogenetic position and biogeography of *Hillebrandia sandwicensis* (Begoniaceae): A rare Hawaiian relict. *American Journal of Botany* 91:905-917.

Clement, W.L., Tebbitt, M.C., Forrest, L.L., Blair, J.E., Brouillet, L., Eriksson, T., & Swensen, S.M. 2005. Phylogenetic position and biogeography of *Hillebrandia sandwicensis* (Begoniaceae): A rare Hawaiian relict. *American Journal of Botany* 92:Errata.

Colinvaux, P.A., De Oliveira, P.E., & Bush, M.B. 2000. Amazonian and neotropical plant communities on glacial time-scales: The failure of the aridity and refuge hypotheses. *Quaternary Science Reviews* 19:141-169.

Collinson, M.E. 1986. The Felpham flora: A preliminary report. *Tertiary Research* 8:29-32.

Collinson, M.E., Boulter, M.C., & Holmes, P.R. 1993. Magnoliophyta ('Angiospermae'). Pp. 809-841 in: Benton, M.J., (ed), *The fossil record*. Chapman and Hall, London.

Cook, L.G., & Crisp, M.D. 2005. Not so ancient: The extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society*, Series B, Biological Sciences 272:2535-2544.

Craib, W.G. 1930. Contributions to the flora of Siam, additamentum XXVII. *Bulletin of Miscellaneous Information, Royal Botanic Gardens, Kew*:161-174.

Crisp, M.D., Isagi, Y., Kato, Y., Cook, L.G., & Bowman, D. 2010. *Livistona* palms in Australia: Ancient relics or opportunistic immigrants? *Molecular Phylogenetics and Evolution* 54:512-523.

Cronn, R., Small, R.L., Haselkorn, T., & Wendel, J.F. 2003. Cryptic repeated genomic recombination during speciation in *Gossypium gossypioides*. *Evolution* 57:2475-2489.

Daly, M.C., Cooper, M.A., Wilson, I., Smith, D.G., & Hooper, B.G.D. 1991. Cenozoic plate tectonics and basin evolution in Indonesia. *Marine and Petroleum Geology* 8:2-21.

De Candolle, A. 1864. Begoniaceae. Pp. 266–408 in: De Candolle, A., (ed), *Prodromus systematis naturalis regni vegetabilis*. Masson and Sons, Paris.

de Lange, A., & Bouman, F. 1992. Studies in Begoniaceae III. Seed micromorphology of the genus *Begonia* in Africa: Taxonomic and ecological implications. *Wageningen Agricultural University Papers*:1-82.

de Lange, A., & Bouman, F. 1999. Seed micromorphology of neotropical Begonias. Smithsonian Institution Press, Washington, D.C.

de Wilde, J.J.F.E., & Plana, V. 2003. A new section of *Begonia* (Begoniaceae) from West Central Africa. *Edinburgh Journal of Botany* 60:121-130.

Doorenbos, J., Sosef, M.S.M., & de Wilde, J.J.F.E. 1998. The sections of *Begonia*: Including descriptions, keys and species lists. Agricultural University, Wageningen.

Doyle, J.A., & Hotton, C.L. 1991. Diversification of early angiosperm pollen in a cladistic context. Pp. 169-195 in: Blackmore, S., & Barnes, S.H., (eds), *Pollen and spores: Patterns of diversification*. Clarendon Press, Oxford.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., & Wilson, A. 2010. *Geneious*. Version 4.8.4. Software.

<http://www.geneious.com/>.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4:699-710.

Drummond, A.J., Ho, S.Y.W., Rawlence, N., & Rambaut, A. 2007. *A rough guide to BEAST 1.4*. Online publication. <http://beast.bio.ed.ac.uk/Main_Page>.

Drummond, A.J., & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7:214.

East, E.M. 1940. The distribution of self-sterility in the flowering plants. *Proceedings of the American Philosophical Society* 82:449-518.

Edgar, R.C. 2004. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.

Erixon, P., Svennblad, B., Britton, T., & Oxelman, B. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology* 52:665-673.

Evans, B.J., Brown, R.M., McGuire, J.A., Supriatna, J., Andayani, N., Diesmos, A., Iskandar, D., Melnick, D.J., & Cannatella, D.C. 2003. Phylogenetics of fanged frogs: Testing biogeographical hypotheses at the interface of the Asian and Australian faunal zones. *Systematic Biology* 52:794-819.

Fazekas, A.J., Steeves, R., & Newmaster, S.G. 2010. Improving sequencing quality from PCR products containing long mononucleotide repeats. *BioTechniques* 48:277-285.

Feliner, G.N., & Rossello, J.A. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* 44:911-919.

Felsenstein, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Systematic Zoology* 27:401-410.

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

Forest, F. 2009. Calibrating the Tree of Life: Fossils, molecules and evolutionary timescales. *Annals of Botany* 104:789-794.

Forrest, L.L. 2001. Phylogeny of Begoniaceae. PhD thesis. University of Glasgow, Glasgow.

Forrest, L.L., & Hollingsworth, P.M. 2003. A recircumscription of *Begonia* based on nuclear ribosomal sequences. *Plant Systematics and Evolution* 241:193-211.

Forrest, L.L., Hughes, M., & Hollingsworth, P.M. 2005. A phylogeny of *Begonia* using nuclear ribosomal sequence data and morphological characters. *Systematic Botany* 30:671-682.

Friis, E.M., Pedersen, K.R., & Crane, P.R. 2005. When Earth started blooming: Insights from the fossil record. *Current Opinion in Plant Biology* 8:5-12.

Friis, E.M., Pedersen, K.R., & Schönenberger, J. 2006. Normapolles plants: A prominent component of the Cretaceous rosid diversification. *Plant Systematics*

and Evolution 260:107-140.

Frodin, D.G. 2004. History and concepts of big plant genera. *Taxon* 53:753-776.

Gagul, J. 2009. Systematics of *Begonia* sect. *Symbegonia*. MSc thesis. University of Port Moresby, Port Moresby.

George, W. 1981. Wallace and his line. Pp. 3-8 in: Whitmore, T.C., (ed), *Wallace's Line and plate tectonics*. Oxford University Press, Oxford.

Girmansyah, D. 2009. A taxonomic study of Bali and Lombok *Begonia* (Begoniaceae). *Reinwardtia* 12:419-434.

Girmansyah, D., Wiriadinata, H., Thomas, D.C., & Hoover, W.S. 2009. Two new species and one new subspecies of *Begonia* (Begoniaceae) from Southeast Sulawesi, Sulawesi, Indonesia. *Reinwardtia* 13:69-74.

Golding, J., & Wasshausen, D.C. 2002. *Begoniaceae*, 2nd ed. National Museum of Natural History, Washington, D.C.

Goodall-Copestake, W. 2005. Framework phylogenies for the Begoniaceae. PhD thesis. University of Glasgow, Glasgow.

Goodall-Copestake, W.P., Harris, D.J., & Hollingsworth, P.M. 2009. The origin of a mega-diverse genus: Dating *Begonia* (Begoniaceae) using alternative datasets, calibrations and relaxed clock methods. *Botanical Journal of the Linnean Society* 159:363-380.

Gorog, A.J., Sinaga, M.H., & Engstrom, M.D. 2004. Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). *Biological Journal of the Linnean Society* 81:91-109.

Gradstein, F.M., Ogg, J.G., Smith, A.G., Bleeker, W., & Lourens, L.J. 2004. A new Geologic Time Scale, with special reference to Precambrian and Neogene. *Episodes* 27:83-100.

Graham, S.W., Reeves, P.A., Burns, A.C.E., & Olmstead, R.G. 2000. Microstructural changes in noncoding chloroplast DNA: Interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. *International Journal of Plant Sciences* 161:S83-S96.

Graur, D., & Martin, W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics* 20:80-86.

Grierson, A.J.C. 1991. Begoniaceae. Pp. 237-246 in: Grierson, A.J.C., & Long, D.G., (eds), *Flora of Bhutan*. Royal Botanic Garden Edinburgh, Edinburgh.

Haffer, J. 1969. Specification in Amazonian forest birds. *Science* 165:131-137.

Haffer, J. 1997. Alternative models of vertebrate speciation in Amazonia: An overview. *Biodiversity and Conservation* 6:451-476.

Hall, R. 1998. The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. Pp. 99-131 in: Hall, R. & Holloway, J. D., (eds.), *Biogeography and geological evolution of SE Asia*. Backhuys Publishers, Leiden.

Hall, R. 2001. Cenozoic reconstructions of SE Asia and the SW Pacific: changing

patterns of land and sea. Pp. 35-56 in: Metcalfe, I., Smith, J.M.B., Morwood, M. & Davidson, I.D., (eds.), *Faunal and Floral Migrations and Evolution in SE Asia-Australasia*. A.A. Balkema Publishers, Lisse.

Hall, R. 2002. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: Computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences* 20:353-431.

Hall, R. 2009. Southeast Asia‘s changing palaeogeography. *Blumea* 54:148-161.

Harmon-Threatt, A.N., Burns, J.H., Shemyakina, L.A., & Knight, T.M. 2009. Breeding system and pollination ecology of introduced plants compared to their native relatives. *American Journal of Botany* 96:1544-1550.

Harris, D.J., Poulsen, A.D., Frimodt-Møller, C., Preston, J., & Cronk, Q.C.B. 2000. Rapid radiation in *Aframomum* (Zingiberaceae): Evidence from nuclear ribosomal DNA internal transcribed spacer (ITS) sequences. *Edinburgh Journal of Botany* 57:377-395.

Herendeen, P.S., Crane, P.R., & Drinnan, A.N. 1995. Fagaceous flowers, fruits and cupules from the Campanian (Late Cretaceous) of central Georgia, USA. *International Journal of Plant Sciences* 156:93-116.

Hershkovitz, M.A., & Zimmer, E.A. 1996. Conservation patterns in angiosperm rDNA ITS2 sequences. *Nucleic Acids Research* 24:2857-2867.

Hewitt, G.M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences* 359:183-195.

Heywood, V.H. 2007. *Flowering plant families of the world*. Firefly Books, Buffalo, N.Y.

Ho, S.Y.W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology* 38:409-414.

Ho, S.Y.W., & Phillips, M.J. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* 58:367-380.

Ho, S.Y.W., Phillips, M.J., Drummond, A.J., & Cooper, A. 2005. Accuracy of rate estimation using relaxed-clock models with a critical focus on the early metazoan radiation. *Molecular Biology and Evolution* 22:1355-1363.

Huelsenbeck, J.P., Larget, B., Miller, R.E., & Ronquist, F. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology* 51:673-688.

Huelsenbeck, J.P., & Rannala, B. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. *Systematic Biology* 53:904-913.

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

Hughes, C.E., Eastwood, R.J., & Bailey, C.D. 2006. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences* 361:211-225.

Hughes, M. 2002. Population structure and speciation in *Begonia* L. PhD thesis. University of Glasgow, Glasgow.

Hughes, M. 2006. Four new species of *Begonia* (Begoniaceae) from Sulawesi. *Edinburgh Journal of Botany* 63:191-199.

Hughes, M. 2008. An annotated checklist of Southeast Asian *Begonia*. Royal Botanic Garden Edinburgh, Edinburgh, U.K.

Hughes, M., & Coyle, C. 2009. *Begonia* section *Petermannia* on Palawan (Philippines), including two new species. *Edinburgh Journal of Botany* 66:205-211.

Hughes, M., & Hollingsworth, P.M. 2008. Population genetic divergence corresponds with species-level biodiversity patterns in the large genus *Begonia*. *Molecular Ecology* 17:2643-2651.

Hughes, M., & Miller, A.G. 2002. A new endemic species of *Begonia* (Begoniaceae) from the Socotra archipelago. *Edinburgh Journal of Botany* 59:273-281.

Hughes, M., & Pullan, M. 2007. *Southeast Asian Begonia Database*. Online database, accessed 7.6.2010, <http://elmer.rbge.org.uk/begonia/>.

Hughes, M., Coyle, C., & Rubite, R.R. 2010. A revision of *Begonia* section *Diploclinium* (Begoniceae) on the Philippine island of Palawan, including five new species. *Edinburgh Journal of Botany* 67:123-140.

Hughes, M., Girmansyah, D., Ardi, W.H., & Nurainas. 2009. Seven new species of *Begonia* from Sumatra. *Gardens' Bulletin (Singapore)* 61:29-44.

Hughes, M., Hollingsworth, P.M., & Miller, A.G. 2003. Population genetic structure in the endemic *Begonia* of the Socotra archipelago. *Biological Conservation* 113:277-284.

Hughes, N.F., & McDougall, A.B. 1990. Barremian-Aptian angiospermid pollen records from southern England. *Review of Palaeobotany and Palynology* 65:145-151.

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

Irmscher, E. 1914. Die Verteilung der Geschlechter in den Inflorescenzen der Begoniaceen unter Berücksichtigung der morphologischen Verhältnisse. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* 50:556–577.

Irmscher, E. 1925. Begoniaceae. Pp. 548–588 in: Engler, A., & Prantl, K., (eds), *Natürlichen Pflanzenfamilien*, 2nd ed. Wilhelm Engelmann, Leipzig.

Irmscher, E. 1929. Die Begoniaceen der Malaiischen Halbinsel. *Mitteilungen aus dem Institut für Allgemeine Botanik in Hamburg* 8:86-160.

Irmscher, E. 1939. Die Begonien Chinas und Ihre Bedeutung für die Frage der Formbildung in polymorphen Sippen. *Mitteilungen aus dem Institut für Allgemeine Botanik in Hamburg* 10:431–557.

Irmscher, E. 1961. Monographische Revision der Begoniaceen Afrikas: I. Sekt. *Augustia* und *Rostrobegonia* sowie einige neue Sippen aus anderen Sektionen. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* 81:106-188.

Janssens, S.B., Knox, E.B., Huysmans, S., Smets, E.F., & Merckx, V.S.F.T. 2009. Rapid radiation of *Impatiens* (Balsaminaceae) during Pliocene and Pleistocene: Result of a global climate change. *Molecular Phylogenetics and Evolution* 52:806-824.

Jayasuriya, A.H.M. 1983. Begoniaceae. Pp. 137-152 in: Dassanayake, M.D., & Fosberg, F.R., (eds), *Flora of Ceylon*. Amerind Publishing, New Delhi.

Jin, X.B., & Wang, F.H. 1994. Style and ovary anatomy of Chinese *Begonia* and its taxonomic and evolutionary implications. *Cathaya* 6:125-144.

Jobes, D.V., & Thien, L.B. 1997. A conserved motif in the 5.8S ribosomal RNA (rRNA) gene is a useful diagnostic marker for plant internal transcribed spacer (ITS) sequences. *Plant Molecular Biology Reporter* 15:326-334.

Judd, W.S., Stevens, P.F., Campbell, C.S., Kellogg, E.A., & Donoghue, M.J. 2008. *Plant systematics: A phylogenetic approach*, 3rd ed. Sinauer Associates, Sunderland.

Jukes, T.H., & Cantor, C.R. 1969. Evolution of protein molecules. Pp. 21-132 in: Munro, H.M., (ed), *Mammalian protein metabolism*. Academic Press, New York. p 21-132.

Kass, R.E., & Raftery, A.E. 1995. Bayes Factors. *Journal of the American Statistical Association* 90:773-795.

Katoh, K., Asimenos, G., & Toh, H. 2009. Multiple Alignment of DNA Sequences with MAFFT. *Bioinformatics for DNA Sequence Analysis*:39-64.

Kelchner, S.A. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* 87:482-498.

Kelchner, S.A. 2002. Group II introns as phylogenetic tools: Structure function, and evolutionary constraints. *American Journal of Botany* 89:1651-1669.

Khajuria, C.K., Prasad, G.V.R., & Manhas, B.K. 1994. Palaeontological constraints on the age of Deccan Traps, Peninsular India. *Newsletter on Stratigraphy* 31:21-32.

Kidner, C.A., & Umbreen, S. in press. Why is leaf shape so variable? *International Journal of Developmental Biology*.

Kiew, R. 1998. Niche partitioning in limestone Begonias in Sabah, Borneo, including two new species. *Gardens' Bulletin (Singapore)* 50:161-169.

Kiew, R. 2001a. The limestone Begonias of Sabah, Borneo: Flagship species for conservation. *Gardens' Bulletin (Singapore)* 53:241-286.

Kiew, R. 2001b. Towards a limestone flora of Sabah. *Malayan Nature Journal* 55:77-93.

Kiew, R. 2004. *Begonia sizemoreae* Kiew (Begoniaceae), a handsome new *Begonia* from Vietnam. *Gardens' Bulletin (Singapore)* 56:95-100.

Kiew, R. 2005. *Begonias of Peninsular Malaysia*. Natural History Publications (Borneo), Kota Kinabalu.

Kiew, R., & Sang, J. 2009. Seven new species of *Begonia* (Begoniaceae) from the Ulu Merirai and Bukit Sarang limestone areas in Sarawak, Borneo. *Gardens' Bulletin (Singapore)* 60:351-372.

Kiew, R., Teo, L.L., & Gan, Y.Y. 2003. Assessment of the hybrid status of some Malesian

plants using Amplified Fragment Length Polymorphism. *Telopea* 10:225-233.

Kilian, N., Hein, P., & Hubaishan, M.A. 2004. Further notes on the flora of the southern coastal mountains of Yemen. *Willdenowia* 34:159-182.

Kimura, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences of the United States of America* 78:454-458.

Klotzsch, J.F. 1854. Begoniaceen: Gattungen und Arten. *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin*: 121-255.

Knoop, V., & Müller, K. 2009. *Gene und Stammbäume : Ein Handbuch zur molekularen Phylogenetik*, 2nd ed. Spektrum, Heidelberg.

Kocyan, A., Zhang, L.B., Schaefer, H., & Renner, S.S. 2007. A multi-locus chloroplast phylogeny for the Cucurbitaceae and its implications for character evolution and classification. *Molecular Phylogenetics and Evolution* 44:553-577.

Ku, C.Z., Peng, C.I., & Turland, N.J. 2007. Begoniaceae. Pp. 153-207 in: Wu, Z.Y., Raven, P.H., & Hong, D.Y., (eds), *Flora of China*. Science Press, Beijing.

Ku, S.M. 2006. Systematics of *Begonia* sect. *Coelocentrum* (Begoniaceae) of China. MSc thesis. National Cheng-Kung University.

Ku, T.C. 1999. Begoniaceae. Pp. 126-269, 401-402 in: Ku, T.C., (ed), *Flora Reipublicae Popularis Sinica*. Science Press, Beijing.

Kubitzki, K., Rohwer, J.G., & Bittrich, V. 1993. *The families and genera of vascular plants*, vol. 2. Springer, Berlin, New York.

Kürschner, H. 1986. Omanisch-makranische Disjunktionen: Ein Beitrag zur pflanzengeographischen Stellung und zu den florengenetischen Beziehungen Omans. . *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* 106:541-562.

Kürschner, H., Kilian, N., Hein, P., & Mukram, A. 2006. The Adenio obesi-Sterculietum africanae, a relic Arabian mainland community vicarious to the Socotran Adenium-Sterculia woodland. *Englera*:79-96.

Kürschner, W.M., Kvacek, Z., & Dilcher, D.L. 2008. The impact of Miocene atmospheric carbon dioxide fluctuations on climate and the evolution of terrestrial ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 105:449-453.

Ladiges, P.Y., Udovicic, F., & Nelson, G. 2003. Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. *Journal of Biogeography* 30:989-998.

Lakhanpal, R.N. 1970. Tertiary floras of India and their bearing on the historical geology of the region. *Taxon* 19:675-694.

Lakhanpal, R.N., & Verma, J.K. 1965. Fossil wood of *Tetrameles* from the Deccan Intertrappean beds of Mohgaonkalan, Madhya Pradesh. *Palaeobotanist* 14:209-213.

Le Corff, J., Ågren, J., & Schemske, D.W. 1998. Floral display, pollinator discrimination,

and female reproductive success in two monoecious *Begonia* species. *Ecology* 79:1610-1619.

Legro, R.A.H., & Doorenbos, J. 1969. Chromosome numbers in *Begonia*. *Netherlands Journal of Agricultural Science* 17:189-202.

Legro, R.A.H., & Doorenbos, J. 1971. Chromosome numbers in *Begonia* 2. *Netherlands Journal of Agricultural Science* 19:176-183.

Legro, R.A.H., & Doorenbos, J. 1973. Chromosome numbers in *Begonia* 3. *Netherlands Journal of Agricultural Science* 21:167-170.

Leins, P. 2000. *Blüte und Frucht*. E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.

Lemey, P., Rambaut, A., Drummond, A.J., & Suchard, M.A. 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* 5.

Lewis, P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50:913-925.

Li, R.Q., Chen, Z.D., Lu, A.M., Soltis, D.E., Soltis, P.S., & Manos, P.S. 2004. Phylogenetic relationships in Fagales based on DNA sequences from three genomes. *International Journal of Plant Sciences* 165:311-324.

Lieberman, B.S. 2005. Geobiology and paleobiogeography: Tracking the coevolution of the Earth and its biota. *Palaeogeography, Palaeoclimatology, Palaeoecology* 219:23-33.

Linder, C.R., & Rieseberg, L.H. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* 91:1700-1708.

Liu, J.S., & Schardl, C.L. 1994. A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. *Plant Molecular Biology* 26:775-778.

Löhne, C., & Borsch, T. 2005. Molecular evolution and phylogenetic utility of the petD group II intron: A case study in basal angiosperms. *Molecular Biology and Evolution* 22:317-332.

Lomolino, M.V., Riddle, B.R., & Brown, J.H. 2006. Biogeography, 3rd ed. Sinauer Associates, Sunderland.

Lydekker, R. 1896. A geographical history of mammals. University Press, Cambridge.

Maddison, W.P., & Maddison, D.R. 2009. *Mesquite: A modular system for evolutionary analysis*. Version 2.72. Software. <http://mesquiteproject.org/>.

Magallón, S.A. 2004. Dating lineages: Molecular and paleontological approaches to the temporal framework of clades. *International Journal of Plant Sciences* 165:S7-S21.

Magallón, S.A., & Castillo, A. 2009. Angiosperm diversification through time. *American Journal of Botany* 96:349-365.

Magallón, S.A., & Sanderson, M.J. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762-1780.

Magallón, S.A., & Sanderson, M.J. 2005. Angiosperm divergence times: The effect of genes, codon positions, and time constraints. *Evolution* 59:1653-1670.

Marshall, D.C. 2010. Cryptic failure of partitioned Bayesian phylogenetic analyses: Lost

in the land of long trees. *Systematic Biology* 59:108-117.

Matolweni, L.O., Balkwill, K., & McLellan, T. 2000. Genetic diversity and gene flow in the morphologically variable, rare endemics *Begonia dregei* and *Begonia homonyma* (Begoniaceae). *American Journal of Botany* 87:431-439.

Merrill, E.D. 1912. The Philippine species of *Begonia*. *Philippine Journal of Science, Section C, Botany* 6:369-406.

Metcalfe, I. 2001. Palaeozoic and Mesozoic tectonic evolution and biogeography of SE Asia-Australasia. Pp. 15-34 in: Metcalfe, I., Smith, J.M.B., Morwood, M., & Davidson, I., (eds), *Faunal and floral migrations and evolution in SE Asia-Australasia*. A. A. Balkema Publishers, Lisse.

Metcalfe, I. 2002. Permian tectonic framework and palaeogeography of SE Asia. *Journal of Asian Earth Sciences* 20:551-566.

Michaux, B. 1991. Distributional patterns and tectonic development in Indonesia: Wallace reinterpreted. *Australian Systematic Botany* 4:25-36.

Michaux, B. 2001. Dispersal versus vicariance, artifice rather than contest. Pp. 311–318 in: Metcalfe, I., Smith, J.M.B., Morwood, M., & Davidson, I., (eds), *Faunal and floral migrations and evolution in SE Asia-Australasia*. A. A. Balkema Publishers, Lisse.

Mies, B.A. 1996. The phytogeography of Socotra: evidence for disjunctive taxa, especially with Macaronesia. Pp. 83-105 in: Dumont, H.J., (ed), *Proceedings of the First International Symposium on Socotra Island: Present and Future*. United Nations Publications, New York.

Miller, A.G., & Morris, M. 2004. *Ethnoflora of the Sqotra Archipelago*. The Royal Botanic Garden Edinburgh, Edinburgh.

Moore, M.J., Soltis, P.S., Bell, C.D., Burleigh, J.G., & Soltis, D.E. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences of the United States of America* 107:4623-4628.

Morley, R.J. 1998. Palynological evidence for Tertiary plant dispersals in the SE Asian region in relation to plate tectonics and climate. Pp. 211-234 in: Hall, R. & Holloway, J. D., (eds.), *Biogeography and geological evolution of SE Asia*. Backhuys Publishers, Leiden.

Morley, R.J. 2000. *Origin and evolution of tropical rain forests*. John Wiley & Sons, Chichester.

Morley, R.J. 2003. Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology Evolution and Systematics* 6:5-20.

Morley, R.J. 2007. Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests. Pp. 1-31 in: Flenley, J.R., & Bush, M.B., (eds), *Tropical rainforest responses to climatic changes*. Springer, Berlin.

Morrison, D.A. 2007. Increasing the efficiency of searches for the maximum likelihood tree in a phylogenetic analysis of up to 150 nucleotide sequences. *Systematic*

Biology 56:988-1010.

Moss, S.J., & Wilson, M.E.J. 1998. Biogeographic implications of the Tertiary palaeogeographic evolution of Sulawesi and Borneo. Pp. 133-163 in: Hall, R. & Holloway, J. D., (eds.), *Biogeography and geological evolution of SE Asia*. Backhuys Publishers, Leiden.

Müller, K. 2005. SeqState: Primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics* 4:65-69.

Muellner, A.N., Pannell, C.M., Coleman, A., & Chase, M.W. 2008. The origin and evolution of Indomalesian, Australasian and Pacific island biotas: Insights from Aglaieae (Meliaceae, Sapindales). *Journal of Biogeography* 35:1769-1789.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., & Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403:853-858.

Neale, S., Goodall-Copestake, W., & Kidner, C. 2006. The evolution of diversity in *Begonia*. Pp. 606-611 in: Teixeira da Silva, J.A., (ed), *Floriculture, ornamental and plant biotechnology: Advances and topical issues*. Global Science Books, Isleworth.

Newton, M.A., & Raftery, A.E. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society, Series B, Methodological* 56:3-48.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., & Swofford, D.L. 2008. AWTY (Are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581-583.

Oginuma, K., & Peng, C.I. 2002. Karyomorphology of Taiwanese *Begonia* (Begoniaceae): Taxonomic implications. *Journal of Plant Research* 115:225-235.

Peng, C.I., & Chen, Y.K. 1991. Hybridity and parentage of *Begonia buimontana* Yamamoto (Begoniaceae) from Taiwan. *Annals of the Missouri Botanical Garden* 78:995-1001.

Peng, C.I., & Chiang, T.Y. 1998. Hybridity and parentage of *Begonia taipeiensis* Peng (Begoniaceae) from Taiwan. *American Journal of Botany* 85:174.

Peng, C.I., & Chiang, T.Y. 2000. Molecular confirmation of unidirectional hybridization in *Begonia* × *taipeiensis* Peng (Begoniaceae) from Taiwan. *Annals of the Missouri Botanical Garden* 87:273-285.

Peng, C.I., & Ku, S.M. 2009. *Begonia* × *chungii* (Begoniaceae), a new natural hybrid in Taiwan. *Botanical Studies* 50:241-250.

Peng, C.I., & Sue, C.Y. 2000. *Begonia* × *taipeiensis* (Begoniaceae), a new natural hybrid in Taiwan. *Botanical Bulletin of Academia Sinica* 41:151-158.

Peng, C.I., Chen, Y.K., & Leong, W.C. 2005. Five new species of *Begonia* (Begoniaceae) from Taiwan. *Botanical Bulletin of Academia Sinica* 46:255-272.

Petit, R.J., Brewer, S., Bordács, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csaikl, U.M., van Dam, B., Deans, J.D., Espinel, S., Fineschi, S., Finkeldey, R., Glaz, I., Goicoechea, P.G., Jensen, J.S., König, A.O., Lowe, A.J., Madsen,

S.F., Mátyás, G., Munro, R.C., Popescu, F., Slade, D., Tabbener, H., de Vries, S.G.M., Ziegenhagen, B., de Beaulieu, J.L., & Kremer, A. 2002. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management* 156:49-74.

Phutthai, T., Sands, M., & Sridith, K. 2009. Field surveys of natural populations of *Begonia* L. in Thailand. *Thai Forest Bulletin (Botany)*, Special Issue:186-196.

Pickett, K.M., & Randle, C.P. 2005. Strange Bayes indeed: Uniform topological priors imply non-uniform clade priors. *Molecular Phylogenetics and Evolution* 34:203-211.

Pigg, K.B., Manchester, S.R., & Wehr, W.C. 2003. *Corylus*, *Carpinus*, and *Palaeocarpinus* (Betulaceae) from the middle Eocene Klondike Mountain and Allenby formations of northwestern North America. *International Journal of Plant Sciences* 164:807-822.

Pigram, C.J., & Davies, P.J. 1987. Terranes and the accretion history of the New Guinea orogen. *BMR Journal of Australian Geology and Geophysics* 10:193-212.

Plana, V. 2002. Systematics and biogeography of the Afro-malagasy fleshy-fruited *Begonia*. PhD thesis. University of Glasgow, Glasgow.

Plana, V. 2003. Phylogenetic relationships of the Afro-Malagasy members of the large genus *Begonia* inferred from trnL intron sequences. *Systematic Botany* 28:693-704.

Plana, V., Gascoigne, A., Forrest, L.L., Harris, D., & Pennington, R.T. 2004. Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. *Molecular Phylogenetics and Evolution* 31:449-461.

Posada, D. 2003. Using Modeltest and PAUP* to select a model of nucleotide substitution. Pp. 6.5.1–6.5.14 in: Baxevanis, A.D., Davison, D.B., Page, R.D.M., Petsko, G.A., Stein, L.D., & Stormo, G.D., (eds), *Current Protocols in Bioinformatics*. John Wiley & Sons, New York.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* 25:1253-1256.

Posada, D., & Buckley, T.R. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53:793-808.

Poulsen, A.D. 2009. The ginger genus *Etlingera* and Wallace‘s Line. Abstract and talk. *Southeast Asian Gateway Evolution (SAGE) 2009*, Royal Holloway University, Egham, UK.

Quek, S.P., Davies, S.J., Ashton, P.S., Itino, T., & Pierce, N.E. 2007. The geography of diversification in mutualistic ants: A gene‘s-eye view into the Neogene history of Sundaland rain forests. *Molecular Ecology* 16:2045-2062.

Raes, N., & van Welzen, P.C. 2009. The demarcation and internal division of Flora Malesiana: 1857-present. *Blumea* 54:6-8.

Rambaut, A., & Charleston, M. 2002. *TreeEdit*. Version 1.0a10. Software. <http://tree.bio.ed.ac.uk/software>.

Rambaut, A., & Drummond, A.J. 2009. *Tracer*. Version 1.5. Software. <http://tree.bio.ed.ac.uk/software>.

Ree, R.H., & Sanmartin, I. 2009. Prospects and challenges for parametric models in historical biogeographical inference. *Journal of Biogeography* 36:1211-1220.

Ree, R.H., & Webb, C.O. 2009. Using models of geographic range evolution to integrate phylogenetic and geological evidence in studies of historical biogeography. Abstract and talk. *Southeast Asian Gateway Evolution (SAGE) 2009*, Royal Holloway University, Egham, UK

Renner, S.S. 2005. Relaxed molecular clocks for dating historical plant dispersal events. *Trends in Plant Science* 10:550-558.

Renner, S.S. 2006. Rewardless flowers in the angiosperms and the role of insect cognition in their evolution. Pp. 123-144 in: Waser, N.M., & Ollerton, J., (eds), *Plant-pollinator interactions: From specialization to generalization*. University of Chicago Press, Chicago.

Richardson, J.E., Pennington, R.T., Pennington, T.D., & Hollingsworth, P.M. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* 293:2242-2245.

Ridder-Numan, J.W.A. 1996. Historical biogeography of the Southeast Asian genus *Spatholobus* (Legum.-Papilionoideae) and its allies. *Blumea* 0:1-144.

Ronquist, F., & Huelsenbeck, J.P. 2003. *MrBayes 3*: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.

Ronquist, F., Huelsenbeck, J.P., & van der Mark, P. 2005. *MrBayes 3.1: Manual*. University of California, USA.

Roos, M.C. 1993. State of affairs regarding Flora Malesiana: progress in revision work and publication schedule. *Flora Malesiana Bulletin* 11:133-142.

Rubite, R.R. 2010. Systematic studies on Philippine *Begonia* L. secion *Diploclinium* (Lindl.) A. D.C. (Begoniaceae). PhD thesis. De La Salle University, Manila.

Rutschmann, F. 2006. Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Diversity and Distributions* 12:35-48.

Rutschmann, F., Eriksson, T., Schönenberger, J., & Conti, E. 2004. Did crypteroniaceae really disperse out of India? Molecular dating evidence from *rbc*L, *ndh*F, and *rpl*16 intron sequences. *International Journal of Plant Sciences* 165:S69-S83.

Sanderson, M.J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14:1218-1231.

Sanderson, M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution* 19:101-109.

Sanderson, M.J. 2003. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19:301-302.

Sands, M.J.S. 2001. Begoniaceae. Pp. 147-163 in: Beaman, J.H., Anderson, C., & Beaman, R.S., (eds), *The plants of Mount Kinabalu* Royal Botanic Gardens Kew, Kew.

Sands, M.J.S. 2009. The Begonias of New Guinea - An overview. *Blumea* 54:272-277.

Sang, T., Crawford, D.J., & Stuessy, T.F. 1995. Documentation of reticulate evolution in Peonies (*Peonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. *Proceedings of the National Academy of Sciences of the United States of America* 92:6813-6817.

Sanmartin, I., Van der Mark, P., & Ronquist, F. 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography* 35:428-449.

Savile, D.B.O., & Hayhoe, H.N. 1978. The potential effect of drop size on efficiency of splash-cup and spring board dispersal devices. *Canadian Journal of Botany* 56:127-128.

Schaefer, H., Heibl, C., & Renner, S.S. 2009. Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society, Series B, Biological Sciences* 276:843-851.

Schaefer, H., Telford, I.R.H., & Renner, S.S. 2008. *Austrobryonia* (Cucurbitaceae), a new Australian endemic genus, is the closest living relative to the Eurasian and Mediterranean *Bryonia* and *Ecballium*. *Systematic Botany* 33:125-132.

Schemske, D.W., Ågren, J., Le Corff, J., Lloyd, D.G., & Barrett, S.C.H. 1996. Deceit pollination in the monoecious, Neotropical herb *Begonia oaxacana* (Begoniaceae). *Floral biology: Studies on floral evolution in animal-pollinated plants*:292-318.

Scherson, R.A., Vidal, R., & Sanderson, M.J. 2008. Phylogeny, biogeography, and rates of diversification of New World Astragalus (Leguminosae) with an emphasis on South American radiations. *American Journal of Botany* 95:1030-1039.

Scotese, C.R. 2003. *PALAEOMAP Project*. Online resource, accessed 7.6.2010, <http://www.scotese.com/>.

Shaw, J., Lickey, E.B., Schilling, E.E., & Small, R.L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* 94:275-288.

Shinde, D., Lai, Y.L., Sun, F.Z., & Arnheim, N. 2003. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)(n) and (A/T)(n) microsatellites. *Nucleic Acids Research* 31:974-980.

Shui, Y.-M., & Chen, W.-H. 2004. Revision to sect. *Petermannia* of *Begonia* (Begoniaceae) in China. *Acta Botanica Yunnanica* 26:482-486.

Shui, Y.M., Peng, C.I., & Wu, C.Y. 2002. Synopsis of the Chinese species of *Begonia* (Begoniaceae), with a reappraisal of sectional delimitation. *Botanical Bulletin of Academia Sinica* 43:313-327.

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

Simmons, M.P., Pickett, K.M., & Miya, M. 2004. How meaningful are Bayesian support values? *Molecular Biology and Evolution* 21:188-199.

Simpson, G.G. 1977. Too many lines: Limits of oriental and Australian zoogeographic regions. *Proceedings of the American Philosophical Society* 121:107-120.

Sims, H.J., Herendeen, P.S., & Crane, P.R. 1998. New genus of fossil Fagaceae from the Santonian (Late Cretaceous) of central Georgia, USA. *International Journal of Plant Sciences* 159:391-404.

Sims, H.J., Herendeen, P.S., Lupia, R., Christopher, R.A., & Crane, P.R. 1999. Fossil flowers with Normapolles pollen from the Upper Cretaceous of southeastern North America. *Review of Palaeobotany and Palynology* 106:131-151.

Small, R.L., Cronn, R.C., & Wendel, J.F. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17:145-170.

Sosef, M.S.M. 1994. Refuge Begonias: Taxonomy, phylogeny and historial biogeography of Begonia sect. Loasibegonia and sect. Scutobegonia in relation to glacial rain forest refuges in Africa. Studies in Begoniaceae V. *Wageningen Agricultural University Papers*: 94:1-306.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.

Su, Y.C.F., & Saunders, R.M.K. 2009. Evolutionary divergence times in the Annonaceae: Evidence of a late Miocene origin of *Pseuduvaria* in Sundaland with subsequent diversification in New Guinea. *BMC Evolutionary Biology* 9:(2 July 2009).

Suchard, M.A., Weiss, R.E., & Sinsheimer, J.S. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. *Molecular Biology and Evolution* 18:1001-1013.

Swofford, D.L. 2002. PAUP* 4.0b10: Phylogenetic Analysis Using Parsimony (* and other methods). Sinnauer Associates, Sunderland.

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

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Pp. 57–86 in: Miura, R.M., (ed), *Some mathematical questions in biology - DNA sequence analysis*. American Mathematical Society, Providence.

Tebbitt, M.C. 2003. Taxonomy of *Begonia longifolia* Blume (Begoniaceae) and related species. *Brittonia* 55:19-29.

Tebbitt, M.C. 2005. Begonias: Cultivation, identification, and natural history. Timber Press, Portland.

Tebbitt, M.C., & Dickson, J.H. 2000. Amended descriptions and revised sectional assignment of some Asian Begonias (Begoniaceae). *Brittonia* 52:112-117.

Tebbitt, M.C., & MacIver, C.M. 1999. The systematic significance of the endothecium

in Begoniaceae. *Botanical Journal of the Linnean Society* 131:203-221.

Tebbitt, M.C., Lowe-Forrest, L., Santoriello, A., Clement, W.L., & Swensen, S.M. 2006. Phylogenetic relationships of Asian *Begonia*, with an emphasis on the evolution of rain-ballist and animal dispersal mechanisms in sections *Platycentrum*, *Sphenanthera* and *Leprosae*. *Systematic Botany* 31:327-336.

Teo, L.-L., & Kiew, R. 1999. First record of a natural *Begonia* hybrid in Malaysia. *Gardens' Bulletin (Singapore)* 51:103-118.

Tesfaye, K., Borsch, T., Govers, K., & Bekele, E. 2007. Characterization of Coffea chloroplast microsatellites and evidence for the recent divergence of C-arabica and C-eugenioides chloroplast genomes. *Genome* 50:1112-1129.

Thomas, D.C., & Hughes, M. 2008. *Begonia varipeltata* (Begoniaceae): a new peltate species from Sulawesi, Indonesia. *Edinburgh Journal of Botany* 65:369-374.

Thomas, D.C., Ardi, W.H., Hartutiningsih, & Hughes, M. 2009a. Two new species of *Begonia* (Begoniaceae) from South Sulawesi, Indonesia. *Edinburgh Journal of Botany* 66:229-238.

Thomas, D.C., Ardi, W.H., & Hughes, M. 2009b. Two new species of *Begonia* (Begoniaceae) from Central Sulawesi, Indonesia. *Edinburgh Journal of Botany* 66:103-114.

Thorne, J.L., & Kishino, H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51:689-702.

Twyford, A., & Richardson, J.E. 2009. Uncovering the historical biogeography of *Rhododendron* subgenus Vireya (Ericaceae). Abstract and talk. *Southeast Asian Gateway Evolution (SAGE) 2009*, Royal Holloway University, Egham, UK.

Uddin, A. 2007. Distribution and status of Indian *Begonia* L. species. *Journal of Economic and Taxonomic Botany* 31:591-597.

Valente, L.M., Savolainen, V., & Vargas, P. 2010. Unparalleled rates of species diversification in Europe. *Proceedings of the Royal Society, Series B, Biological Sciences* 277:1489-1496.

van Welzen, P.C., & Slik, J.W.F. 2009. Patterns in species richness and composition of plant families in the Malay Archipelago. *Blumea* 54:166-171.

van Welzen, P.C., Slik, J.W.F., & Alahuhta, J. 2005. Plant distribution patterns and plate tectonics in Malesia. *Kongelige Danske Videnskabernes Selskab Biologiske Skrifter* 55:199-217.

Vogel, S. 1975. Mutualismus und Parasitismus in der Nutzung von Pollentragen. *Verhandlungen der Deutschen Zoologischen Gesellschaft* 68:102-110.

Vogel, S. 1993. Betrug bei Pflanzen: die Täuschblumen. Akademie der Wissenschaften und der Literatur, Mainz.

Vogel, S. 1998. Remarkable nectaries: structure, ecology, organophyletic perspectives IV. Miscellaneous cases. *Flora* 193:225-248.

Voris, H.K. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography* 27:1153-1167.

Vriesendorp, B., & Bakker, F.T. 2005. Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *Taxon* 54:593-604.

Wallace, A.R. 1860. On the Zoological Geography of the Malay Archipelago. *Journal of the Proceedings of the Linnean Society* 4:172-184.

Wallace, A.R. 1863. On the Physical Geography of the Malay Archipelago. *Journal of the Royal Geographical Society* 33:217-234.

Wallace, A.R. 1869. The Malay Archipelago. Macmillan, London.

Wallace, A.R. 1876. The Geografihical Distribution of Animals. Macmillan, London.

Wang, H.C., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., & Soltis, D.E. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences of the United States of America* 106:3853-3858.

Warburg, O. 1894. Begoniaceae. Pp. 121-150 in: Engler, A., & Prantl, K., (eds), *Natürlichen Pflanzenfamilien*. Wilhelm Engelmann, Leipzig.

Whitten, T., Mustafa, M., & Henderson, G.S. 2002. The ecology of Sulawesi. Periplus, Jakarta.

Wikström, N., Savolainen, V., & Chase, M.W. 2001. Evolution of the angiosperms: Calibrating the family tree. *Proceedings of the Royal Society, Series B, Biological Sciences* 268:2211-2220.

Willis, K.J., & Whittaker, R.J. 2000. Paleoecology - The refugial debate. *Science* 287:1406-1407.

Wojciechowski, M.F., Sanderson, M.J., & Hu, J.M. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA trnL intron data. *Systematic Botany* 24:409-437.

Woodruff, D.S. 2003. Neogene marine transgressions, palaeogeography and biogeographic transitions on the Thai-Malay Peninsula. *Journal of Biogeography* 30:551-567.

Woodruff, D.S. 2010. Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today‘s patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity and Conservation* 19:919-941.

Wu, Z., & Raven, P.H. 2008. *Flora of China*: Illustrations. Science Press, Beijing.

Yuan, Y.M., Wohlhauser, S., Möller, M., Klackenberg, J., Callmander, M.W., & Kupfer, P. 2005. Phylogeny and biogeography of *Exacum* (Gentianaceae): A disjunctive distribution in the Indian Ocean Basin resulted from long distance dispersal and extensive radiation. *Systematic Biology* 54:21-34.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292:686-693.

Zhang, L.B., & Renner, S.S. 2003. The deepest splits in Chloranthaceae as resolved by chloroplast sequences. *International Journal of Plant Sciences* 164:S383-S392.

Zhang, L.B., Simmons, M.P., Kocyan, A., & Renner, S.S. 2006. Phylogeny of the

Cucurbitales based on DNA sequences of nine loci from three genomes: Implications for morphological and sexual system evolution. *Molecular Phylogenetics and Evolution* 39:305-322.

Zhang, L.B., Simmons, M.P., & Renner, S.S. 2007. A phylogeny of Anisophylleaceae based on six nuclear and plastid loci: Ancient disjunctions and recent dispersal between South America, Africa, and Asia. *Molecular Phylogenetics and Evolution* 44:1057-1067.

Zwickl, D. 2006. *GARLI, Genetic algorithm for rapid likelihood inference*. Version 0.951. Software. <www.bio.utexas.edu/faculty/antisense/garli/Garli.html>.

Appendix 1. Additions to the *Begonia* flora of Sulawesi, Indonesia.

EDINBURGH JOURNAL OF BOTANY 65 (3): 369–374 (2008) 369

doi:10.1017/S096042860800509X

BEGONIA VARIPELTATA (*BEGONIACEAE*): A NEW PELTATE SPECIES FROM SULAWESI, INDONESIA

D. C. THOMAS[1] & M. HUGHES[2]

A new species of *Begonia* (*Begoniaceae*), *B. varipeltata* D.C.Thomas, is described from the Indonesian island of Sulawesi. It exhibits peltate leaves, which are rare in *Begonia* section *Petermannia*, to which it belongs.

Keywords. *Begonia*, new species, peltate, Sulawesi.

INTRODUCTION

The Indonesian island of Sulawesi (formerly more commonly known as Celebes) has been of prime interest to biogeographical research for almost 150 years (e.g. Wallace, 1860; Whitmore, 1981; Van Balgooy, 1987; Evans *et al.*, 2003; Mendum & Atkins, 2004; Van Welzen *et al.*, 2005). There are several reasons for this: (i) the island's complex geological history, in which fragments of different tectonic plates were amalgamated (Hall, 2002); (ii) Sulawesi's location at the western border of Wallacea, an interface region where Asian and Australian biotas meet (Van Welzen *et al.*, 2005); and (iii) Sulawesi's apparent floristic and faunistic separation from its nearest neighbouring island, Borneo (Wallace, 1860; Whitmore, 1981; Van Welzen *et al.*, 2005).

The taxonomy of many species-rich genera on Sulawesi (e.g. *Cyrtandra* and *Aeschynanthus* (*Gesneriaceae*) – Mendum & Atkins, 2004; *Begonia* (*Begoniaceae*) – Hughes, 2006; Hughes & Pullan, 2007), which would be ideal subjects for biogeographical studies, is very poorly known and revisions are needed. Thirty-one indigenous *Begonia* species have been described from Sulawesi (Doorenbos *et al.*, 1998; Doorenbos, 2000; Tebbitt, 2005; Hughes, 2006; Hughes & Pullan, 2007), but 'it is likely that a complete account of Sulawesi *Begonia* will more than double this total' (Hughes & Pullan, 2007). Recent expeditions to Sulawesi organised by the Royal Botanic Garden Edinburgh (RBGE) have brought to light several new *Begonia* species (Hughes, 2006), and another new species, which was brought into cultivation at RBGE, is described below. The majority of *Begonia* species from Sulawesi are classified in *Begonia* section *Petermannia* (27 species); one introduced

[1] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, Scotland, UK. E-mail: d.thomas@rbge.ac.uk
[2] Singapore Botanic Gardens, 1 Cluny Road, Singapore 259569. E-mail: mark_hughes@nparks.gov.sg

species, *Begonia hirtella* Link, belongs to *Begonia* section *Doratometra* (Doorenbos *et al.*, 1998; Hughes & Pullan, 2007), and the remaining species belong to *Begonia* section *Sphenanthera* (to the '*Begonia longifolia* Blume complex'; see Tebbitt, 1997, 2003). The new species described below is classified in *Begonia* section *Petermannia* because it exhibits typical characters of the section: protogynous inflorescences, bifid placentae and anthers with unilaterally positioned slits (Fig. 1).

All available *Begonia* specimens from B, BM, E, K, L and SING (plus photographic duplicates from BO) have been consulted, and hence it must be assumed, at least until more intensive collecting in Sulawesi may reveal otherwise, that this species has a very restricted range. Figure 2 shows the collecting localities in Central Sulawesi (Tengah).

Cultivated material grown from the same seed collection as the holotype was used to supplement the description.

Begonia varipeltata D.C.Thomas, **sp. nov.** Sect. *Petermannia*. **Fig. 1.**

Begoniae macintyreanae M.Hughes similis a qua caulibus pendulis (non erectis) et foliis aliquantis peltatis (non omnibus basifixis) differt. – Type: Cultivated at the Royal Botanic Garden Edinburgh, from seed collected in the wild (Indonesia, Sulawesi, Tengah, Luwuk District, Bunta Subdistrict, Sumber Agung Village, Sungai SPA, 00°09′12.6″S, 122°09′28.8″E, 200 m), 24 ix 2007, *D.C. Thomas* 07-21 (holo E; iso L).

Perennial, monoecious herb. *Stems* woody at base, to 12 mm across, erect for 30–70 cm in the basal part, but pendent in the more distal part, the pendent part up to 54 cm long, internodes 3–10.2 cm long, glabrous. *Leaves* alternate, excentrically peltate or both excentrically peltate and basifixed leaves present; stipules 14–23 × 4–10 mm, oblong or narrowly elliptic, cymbiform with abaxially prominent midrib forming a thin, c.1–2 mm long appendage at the apex, caducous; petioles 1.2–4.2 cm long, glabrous; lamina very asymmetric, ovate, narrowly elliptic or oblong, 9.7–21.5 × 3.1–6.9 cm, margin irregularly serrate to dentate, glabrous, dark green above and pale green below, venation palmate-pinnate, base cordate (basifixed leaves) with not or only slightly overlapping lobes, apex acuminate. *Inflorescences* bisexual or unisexual, protogynous, cymose, composed of (0–)1 basal, female (partial-) inflorescence with 2 female flowers and (0–)1–6 distal, male (partial-) inflorescence(s), each with 2–7 cymose branching points, dichasial or with dichasial branching in the basal part and monochasial branching in the distal part, with c.5–30 flowers. *Male flowers*: pedicels 2–14 mm; tepals 2, broadly ovate to subcircular, 6–10 × 5–11 mm, base cuneate to truncate or tepal margin convex at base, apex rounded, white or pinkish, glabrous; androecium of c.25–30 stamens, yellow, symmetric, filaments slightly fused at the base, unequal, the longer in the middle, anthers obovate, slightly longer than to c.2 times shorter than the filaments, c.0.7–1.2 mm long, dehiscing through unilaterally positioned slits > ½ as long as the anther, connective not extended. *Female flowers*: pedicels 9–26 mm; tepals 5, obovate, unequal, 9–16 mm long,

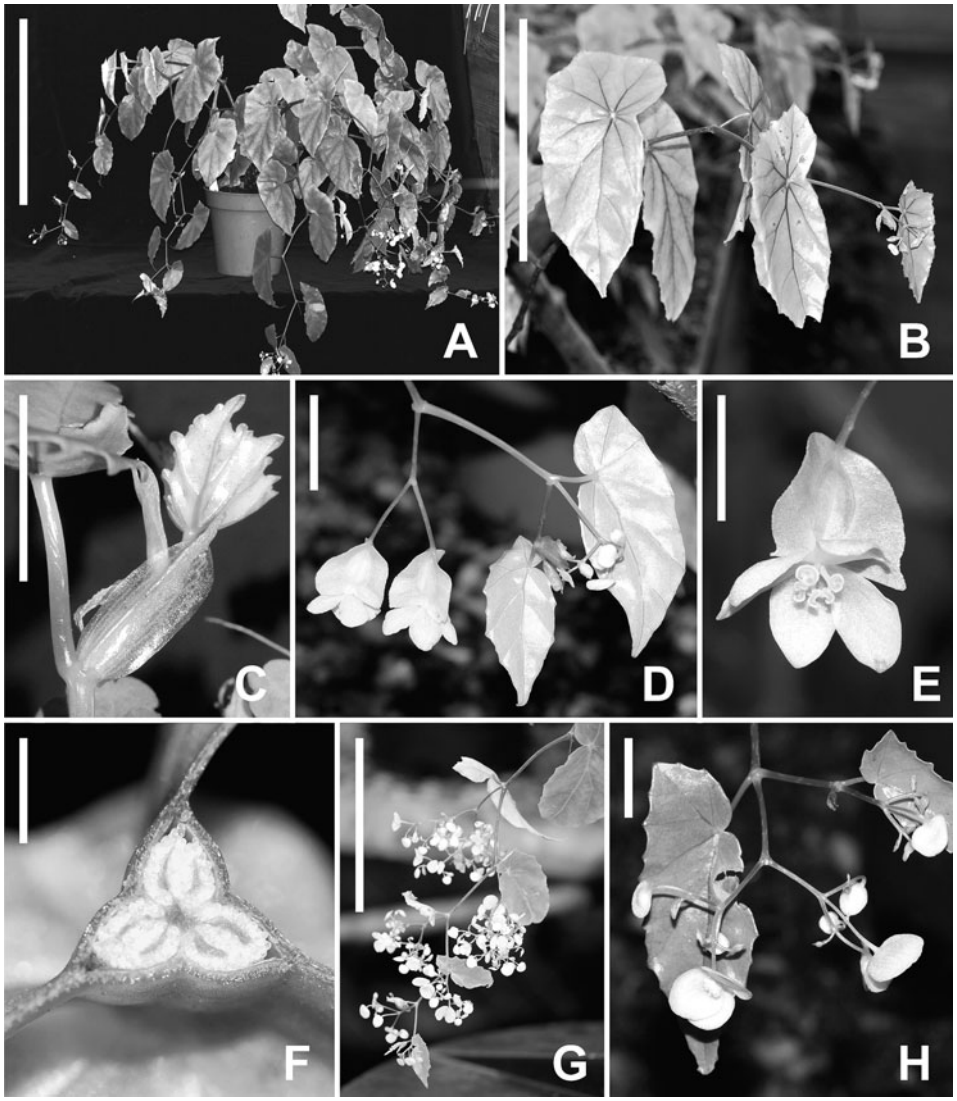


FIG. 1. *Begonia varipeltata* D.C.Thomas. A, habit (scale bar = 30 cm); B, peltate leaves (scale bar = 15 cm); C, stipule (scale bar = 2 cm); D, female partial inflorescence (scale bar = 2 cm); E, female flower (scale bar = 1.5 cm); F, ovary, cross-section, 3-locular with bifid placentae (scale bar = 2 mm); G, male inflorescence of 6 dichasial–monochasial male partial inflorescences (scale bar = 10 cm); H, male partial inflorescence with dichasial branching pattern (scale bar = 1.2 cm). A, E, G, H: *D.C. Thomas* 07-22; B, C, D, F: *D.C. Thomas* 07-21.

the four outer 5–12 mm wide, the innermost 3–7 mm wide, white, glabrous; ovary 8–12 mm long, 3-locular, placentation axile, placentae bifid, ovary 3-winged, wings subequal, 12–14 mm long and 5–11 mm wide at the widest point (in the most distal part), apex truncate or with convex margins, base with convex margins or cuneate,

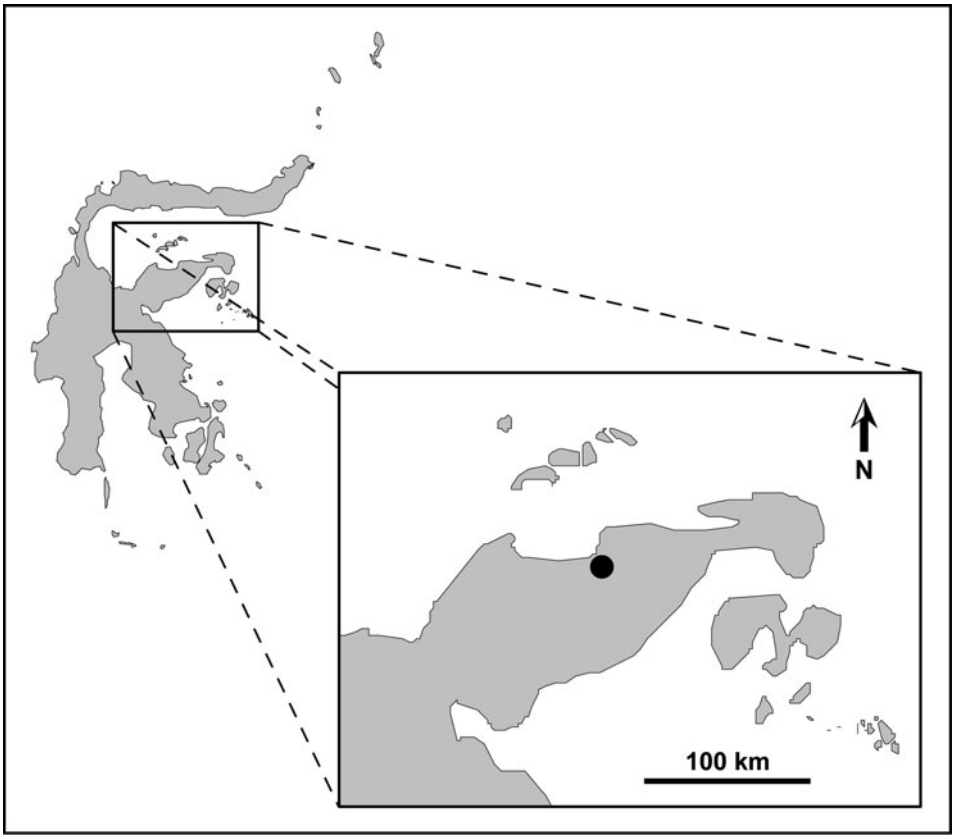


FIG. 2. Distribution of *Begonia varipeltata* in eastern Central Sulawesi (Tengah). ● = collection locality.

glabrous; style fused only in the very basal part, 3-branched, each stylodium bifurcate in the stigmatic region, deciduous, stigmatic surface a twice spirally twisted papillose band, greenish-yellow or yellow. *Fruits* pendulous, 3-winged, wing shape as for ovary, dehiscent, drying pale brown, glabrous. *Seeds* barrel-shaped, c.0.3 mm long, collar cells c.⅔ the length of the seed.

Distribution. Endemic to Sulawesi. Known only from two collections from the Luwuk District, Bunta Subdistrict, Sumber Agung, Sungai SPA in eastern Central Sulawesi (Tengah) (Fig. 2).

Habitat. Growing on rock walls along the sides of river banks and in disturbed primary forest at low altitudes (two collections at 92 m and 200 m, respectively).

Proposed IUCN conservation category. VU D2. The likelihood that this species has a very restricted range in an area which shows clear signs of anthropogenic disturbance, especially timber harvesting, and which has no legal protection as a national park, wildlife or forest reserve, makes it 'prone to the effects of human activities or

stochastic events within a very short time period in an uncertain future' (IUCN, 2001).

Additional specimens examined. SULAWESI. **Tengah**: Luwuk District, Bunta Subdistrict, Sumber Agung, Sungai SPA, 92 m, 24 ii 2004, *Hendrian, M. Newman, S. Scott, M. Nazre Saleh & D. Supriadi* 858 (E!); Cultivated RBGE, from seed collected in the wild in Luwuk District, Bunta Subdistrict, Sumber Agung Village, Sungai SPA, 00°09′12.6″S, 122°09′28.8″E, 200 m, *D.C. Thomas* 07-22 (E).

The epithet *varipeltata* refers to the variable transition of petiole and lamina in this species, which ranges from basifixed, to strongly excentrically peltate to almost centrally peltate. This great variation is similar to the condition described for *Begonia amphioxus* Sands (Sands, 1990). However, *Begonia amphioxus* clearly differs from *B. varipeltata* in both vegetative and generative morphology – for example, *B. amphioxus* has male flowers with four tepals, two-locular and usually two-winged ovaries, and distinctly narrower, spotted leaves (see Sands, 1990). Peltate species are rare in *Begonia* section *Petermannia* (only *Begonia amphioxus* and *B. baramensis* Merr.), and no peltate species of *Begonia* have been described from Sulawesi before. However, an analysis of all available *Begonia* specimens from Sulawesi in B, BM, E, K, L and SING (plus photographic duplicates from BO) showed that there are several undescribed peltate *Begonia* species from Sulawesi, which will be described in a subsequent paper.

ACKNOWLEDGEMENTS

We are grateful to Steve Scott for his excellent care of the *Begonia* living collections at the Royal Botanic Garden Edinburgh, to the curators of B, BM, BO, E, K, L and SING for allowing us access to herbarium material and for sending photographic duplicates, and to Prof. Ruth Kiew and an anonymous reviewer for their constructive reviews. The support of the first author's PhD project by the M. L. MacIntyre Trust is gratefully acknowledged.

REFERENCES

DOORENBOS, J. M. (2000). *Begonia siccacaudata* (Begoniaceae), a new species from Sulawesi. *Blumea* 45(2): 399–402.

DOORENBOS, J. M., SOSEF, S. M. & DE WILDE, J. J. F. E. (1998). *The sections of* Begonia *including descriptions, keys and species lists*. Studies in Begoniaceae VI. Wageningen Agricultural University Papers 98(2). Wageningen: Wageningen Agricultural University.

EVANS, B. J., BROWN, R. M., MCGUIRE, J. A., SUPRIATNA, J., ANDAYANI, N., DIESMOS, A., ISKANDAR, D., MELNICK, D. J. & CANNATELLA, D. C. (2003). Phylogenetics of fanged frogs: Testing biogeographical hypotheses at the interface of the Asian and Australian faunal zones. *Syst. Biol.* 52(6): 794–819.

HALL, R. (2002). Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions and animations. *J. Asian Earth Sci.* 20(4): 353–434.

HUGHES, M. (2006). Four new species of *Begonia* (*Begoniaceae*) from Sulawesi. *Edinburgh J. Bot*. 63(2&3): 191–199.

HUGHES, M. & PULLAN, M. (2007). *South East Asian* Begonia *Database*. Electronic publication accessible via www.rbge.org.uk.

IUCN (2001). *IUCN Red List Categories and Criteria, Version 3.1*. IUCN Species Survival Commission. Gland, Switzerland and Cambridge, UK: IUCN.

MENDUM, M. & ATKINS, H. J. (2004). The *Gesneriaceae* of Sulawesi I: An introduction. *Edinburgh J. Bot*. 60(3): 299–304.

SANDS, M. J. S. (1990). Six new Begonias from Sabah. *Kew Mag*. 7(2): 57–85.

TEBBITT, M. C. (1997). *A systematic investigation of* Begonia *section* Sphenanthera *(Hassk.) Benth. & Hook. f*. Unpublished PhD thesis, University of Glasgow, Scotland, UK.

TEBBITT, M. C. (2003). Taxonomy of *Begonia longifolia* Blume (Begoniaceae) and related species. *Brittonia* 55(1): 19–29.

TEBBITT, M. C. (2005). Three new species and a new subspecies of *Begonia* (*Begoniaceae*) from Asia. *Edinburgh J. Bot*. 61(2&3): 97–107.

VAN BALGOOY, M. M. J. (1987). A plant geographical analysis of Sulawesi. In: WHITMORE, T. C. (ed.) *Biogeographical Evolution of the Malay Archipelago*, pp. 94–102. Oxford: Clarendon Press.

VAN WELZEN, P. C., SLIK, J. W. F. & ALAHUHTA, J. (2005). Plant distribution patterns and plate tectonics in Malesia. *Biol. Skr*. 55: 199–217.

WALLACE, A. R. (1860). On the zoological geography of the Malay Archipelago. *Zool. J. Linn. Soc*. 4: 172–184.

WHITMORE, T. C. (1981). *Wallace's Line and Plate Tectonics*. Oxford: Clarendon Press.

Received 14 January 2008; accepted for publication 20 February 2008

EDINBURGH JOURNAL OF BOTANY 66 (1): 103–114 (2009) 103

doi:10.1017/S0960428609005320

TWO NEW SPECIES OF *BEGONIA* (*BEGONIACEAE*) FROM CENTRAL SULAWESI, INDONESIA

D. C. Thomas[1], W. H. Ardi[2] & M. Hughes[3]

Two new species of *Begonia* (*Begoniaceae*), *Begonia ozotothrix* and *Begonia hekensis*, are described from the Indonesian island of Sulawesi. Both species belong to *Begonia* section *Petermannia*. *Begonia ozotothrix* is unusual amongst Asian *Begonia* in having branched trichomes on the stems, petioles and the abaxial lamina surfaces, and it is unusual amongst species of *Begonia* section *Petermannia* in having extremely compressed cymose-subumbellate male partial inflorescences.

Keywords. *Begonia* section *Petermannia*, inflorescence, new species, Sulawesi, trichome.

Introduction

Thirty-two indigenous species of *Begonia* L. have been reported from the Indonesian island of Sulawesi (Table 1; Hughes, 2008; Thomas & Hughes, 2008). A revision of two difficult species complexes containing some widespread taxa, the *Begonia longifolia* Blume complex and the *Begonia rieckei* Warb. complex (Table 1), may result in a reduction of the number of currently accepted names in the Sulawesi *Begonia* flora through synonymy, reduction to infraspecific rank and correction of misidentifications (see discussions on synonymy and species boundaries in these two complexes in Tebbitt, 1997, 2003; Tebbitt & Dickson, 2000; Hughes, 2008). However, most species from Sulawesi are local endemics and morphologically very distinct, and a close examination of all available *Begonia* specimens from Sulawesi from A, B, BM, BO, CEB, E, K, L and SING indicates that there are numerous endemic species awaiting description. The majority of *Begonia* species from Sulawesi are classified in *Begonia* section *Petermannia* (Klotzsch) A.DC. (Doorenbos *et al.*, 1998; Hughes, 2008) (28 species, including the four closely related or conspecific taxa in the *B. rieckei* complex). Four species have been classified in *Begonia* section *Sphenanthera* (Hassk.) Warb. (including the three closely related or conspecific taxa in the *B. longifolia*

[1] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, Scotland, UK. E-mail: d.thomas@rbge.ac.uk
[2] Center for Plant Conservation, Bogor Botanic Garden, Jl. Ir. H. Juanda No. 13, Bogor 16003, Indonesia. E-mail: wisnu.handoyo.ardi@lipi.go.id
[3] Singapore Botanic Gardens, 1 Cluny Road, Singapore 259569. E-mail: mark_hughes@nparks.gov.sg

TABLE 1. Indigenous *Begonia* species of Sulawesi

Section	Species/species complex
Petermannia	*Begonia bonthainensis* Hemsl.
	Begonia capituliformis Irmsch.
	Begonia carnosa (Teijsm. & Binn.) Teijsm. & Binn.
	Begonia celebica Irmsch.
	Begonia chiasmogyna M.Hughes
	Begonia cuneatifolia Irmsch.
	Begonia flacca Irmsch.
	Begonia gemella Warb. ex L.B.Sm. & Wassh.
	Begonia grandipetala Irmsch.
	Begonia hekensis D.C.Thomas
	Begonia heteroclinis Miq. ex Koord.
	Begonia hispidissima Zipp. ex Koord.
	Begonia humilicaulis Irmsch.
	Begonia imperfecta Irmsch.
	Begonia insularum Irmsch.
	Begonia macintyreana M.Hughes
	Begonia masarangensis Irmsch.
	Begonia mendumiae M.Hughes
	Begonia ozotothrix D.C.Thomas
	Begonia rachmatii Tebbitt
	'*Begonia rieckei* Warb. complex'
	Begonia koordersii Warb. ex L.B.Sm. & Wassh.
	Begonia pseudolateralis Warb.
	Begonia rieckei Warb.
	Begonia strictipetiolaris Irmsch.
	Begonia sarasinorum Irmsch.
	Begonia siccacaudata J.Door.
	Begonia sphenocarpa Irmsch.
	Begonia stevei M.Hughes
	Begonia strachwitzii Warb. ex Irmsch.
	Begonia varipeltata D.C.Thomas
Sphenanthera	'*Begonia longifolia* Blume complex'
	Begonia aptera Blume
	Begonia longifolia Blume
	Begonia renifolia Irmsch.
	Begonia robusta Blume

complex, and *B. robusta* Blume), which was shown to be paraphyletic with respect to *Begonia* section *Platycentrum* (Klotzsch) A.DC. (Tebbitt *et al.*, 2006).

All recent expeditions to Sulawesi organised by the Royal Botanic Garden Edinburgh (RBGE) have brought to light some new species of *Begonia* (Hughes, 2006; Thomas & Hughes, 2008). This is not surprising given that, firstly, the mega-diverse genus *Begonia* has a centre of diversity in Southeast Asia; secondly, fewer botanical collections have been made on Sulawesi than on any other major island in Indonesia, and from several large regions of Sulawesi only a very small number of

specimens has been collected (Kessler *et al.*, 2002; Cannon *et al.*, 2007); and thirdly, Sulawesi *Begonia* have never been revised. Two further species collected on a joint expedition of the RBGE and Bogor Botanic Garden are described below. They are classified in *Begonia* section *Petermannia* as they exhibit typical characters of the section: protogynous, two-flowered female inflorescences, three-locular ovaries with axile placentation and bilamellate placentae, fruits with equal or subequal wings, and anthers with unilaterally positioned slits (Figs 1, 2). All available *Begonia* specimens from A, B, BM, BO, CEB, E, K, L and SING have been consulted, and hence it must be assumed, at least until more intensive collecting on Sulawesi may reveal otherwise, that these two species have restricted ranges and are endemic to Central Sulawesi (Sulawesi Tengah) (Fig. 3).

Species Descriptions

Begonia ozotothrix D.C.Thomas, **sp. nov.** Sect. *Petermannia*. **Figs 1**, **3–5.**

Ab aliis speciebus sectionis *Petermanniae* in caule, petiolis et in laminae facie abaxiali pilos ramosos habenti differt. – Type: Indonesia, Sulawesi, Sulawesi Tengah, Tojo Una-una District, close to Watusongo Village, Gunung Katopas, on wet rock wall at riverbank, 01°10′17.9″S, 121°28′40.5″E, 615 m, 7 v 2008, *D.C. Thomas & W.H. Ardi* 08-67 (holo E; iso BO, CEB).

Perennial, monoecious, erect herb, to 75 cm tall, hairy with microscopic, c.0.05–0.2 mm long, simple trichomes on all vegetative parts and heterotrichous on the stems, petioles and the veins of the abaxial lamina surface by the addition of a few interspersed multicellular, multiseriate, branched trichomes, c.0.3–1.8 mm long. *Stems* branched; internodes 2.9–7.9 cm long, hairy. *Leaves* alternate; stipules 20–32 × 6–20 mm, very asymmetric, oblong to narrowly elliptic, cymbiform with abaxially prominent midrib forming a thin, short appendage at the apex, persistent, abaxially densely hairy; petioles 4.1–18.6 cm long, hairy; lamina basifixed, 14.5–24.8 × 7.2–15.3 cm, very asymmetric, elliptic, base cordate with non- or only very slightly overlapping lobes, apex acuminate, margin dentate, the teeth bristle-pointed, adaxial and abaxial surface hairy, adaxial surface mid green and abaxial surface pale green, or adaxial surface dark green and abaxial surface reddish, venation palmate-pinnate. *Inflorescences* protogynous; *female inflorescences* basal to male inflorescences or solitary, 2-flowered, subtending leaves foliose, peduncles 1–5 mm long, bracts (subtending the pedicels of the female flowers) 8–11 × 6–7 mm, ovate to elliptic, abaxially hairy; *male inflorescences* distal to one female inflorescence, composed of 1–5 strongly compressed, cymose-subumbellate partial inflorescences, subtending leaves bracteose, c.8–11 × 7–9 mm, elliptic, abaxially hairy, peduncles 2–25 mm, hairy, bracts of only the most basal dichotomous branching developed, 3–10 × 2–6 mm, elliptic, abaxially hairy, caducous, each partial inflorescence branching once dichotomously at the base, then 1–2 times dichasially and the lateral branches of the most distal dichasia branching (0–)1–4 times monochasially. *Male flowers*: pedicels 2–15 mm, hairy; tepals 2, white or pink,

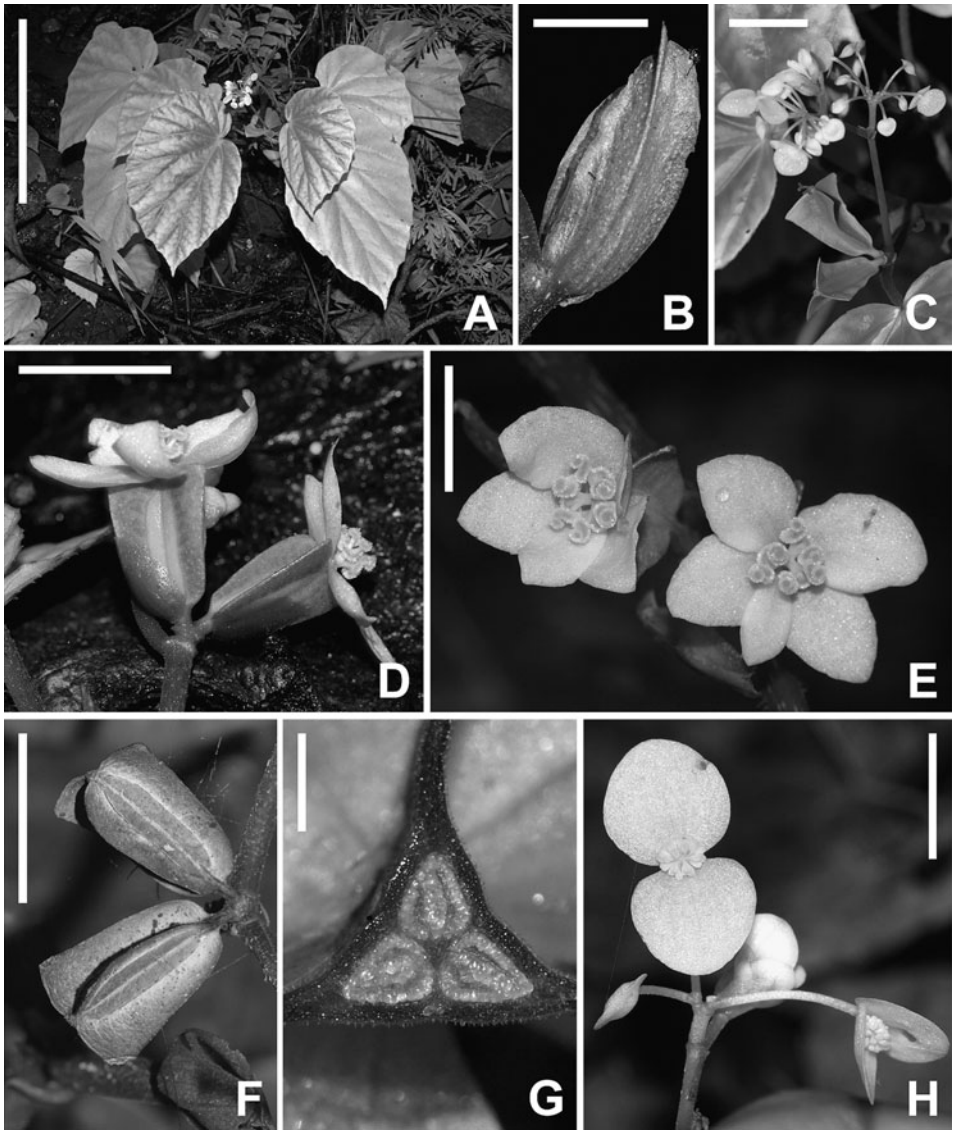


FIG. 1. *Begonia ozotothrix* D.C.Thomas. A, habit (scale bar = 20 cm); B, stipule (scale bar = 12 mm); C, inflorescence (scale bar = 2 cm); D, female inflorescence (scale bar = 2 cm); E, female flowers (scale bar = 12 mm); F, capsules (scale bar = 2 cm); G, ovary, cross-section, three-locular with axillary, bilamellate placentae (scale bar = 2 mm); H, male flowers (scale bar = 10 mm). A, B, C, F, H: *D.C. Thomas & W.H. Ardi* 08-58; D, E, G: *D.C. Thomas & W.H. Ardi* 08-53.

8–11 × 9–12 mm, broadly ovate, base slightly cordate or with convex margins, apex rounded, abaxially sparsely hairy to glabrescent; androecium of c.25–35 stamens, yellow, filaments c.0.4–1.6 mm long, slightly fused at the very base, unequal, longer in the middle of the androecium, anthers c.0.8–1.4 mm long, obovate, dehiscing through

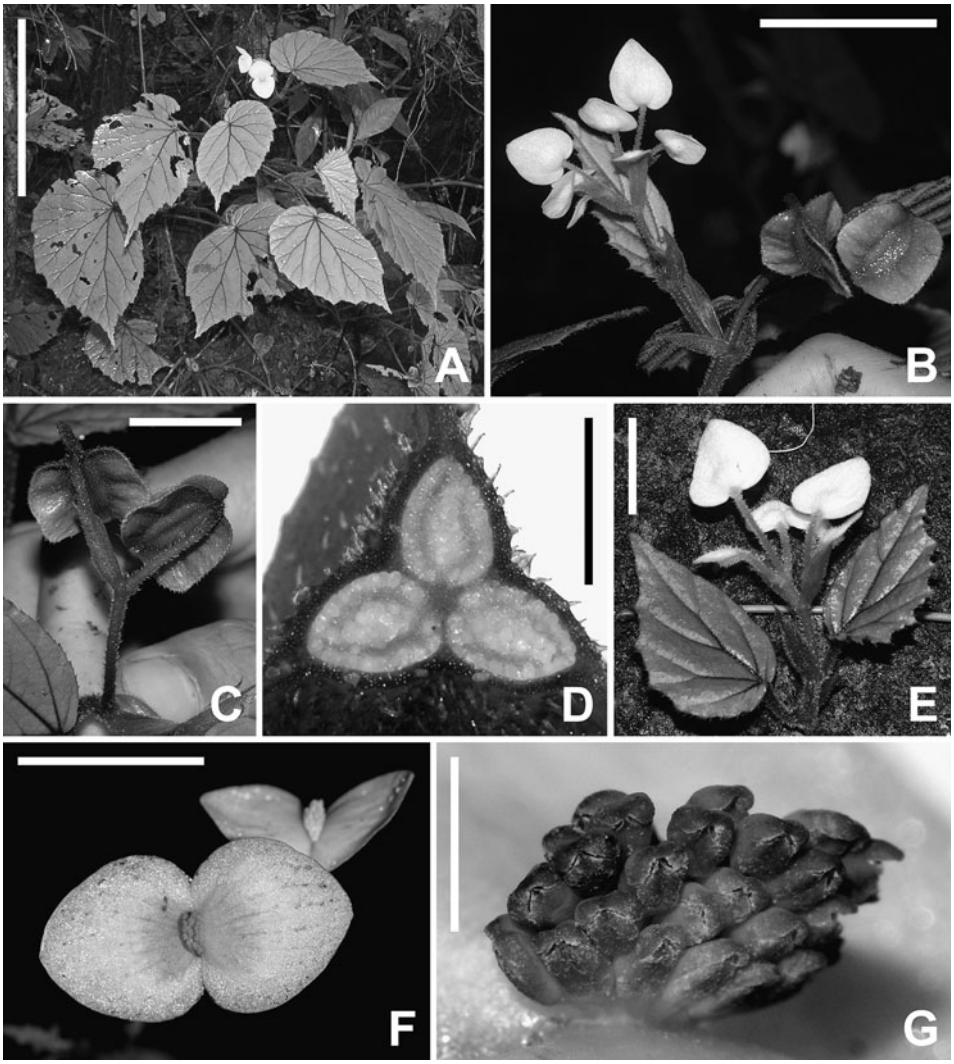


FIG. 2. *Begonia hekensis* D.C.Thomas. A, habit (scale bar = 7 cm); B, male inflorescence and infructescence (scale bar = 3 cm); C, infructescence (scale bar = 1.5 cm); D, ovary, cross-section, three-locular with axillary, bilamellate placentae (scale bar = 3 mm); E, male inflorescence with subtending leaves (scale bar = 1.5 cm); F, male flowers (scale bar = 2 cm); G, androecium, anther dehiscing through short, unilaterally positioned slits (scale bar = 4 mm). A–G: *D.C. Thomas & W.H. Ardi* 08-43.

unilaterally positioned slits c.1/2 as long as the anther, connective not projecting. *Female flowers*: pedicels 1–4 mm, hairy; tepals 5, unequal to subequal, the four larger ones ovate or elliptic, 9–14 × 7–9 mm, the innermost obovate or narrowly elliptic, 8–11 × 3–8 mm, white or pale pink, abaxially hairy; ovary 12–21 × 10–18 mm, locules 3, placentation axile, placentae bilamellate, wings 3, narrowly triangular, rounded at the

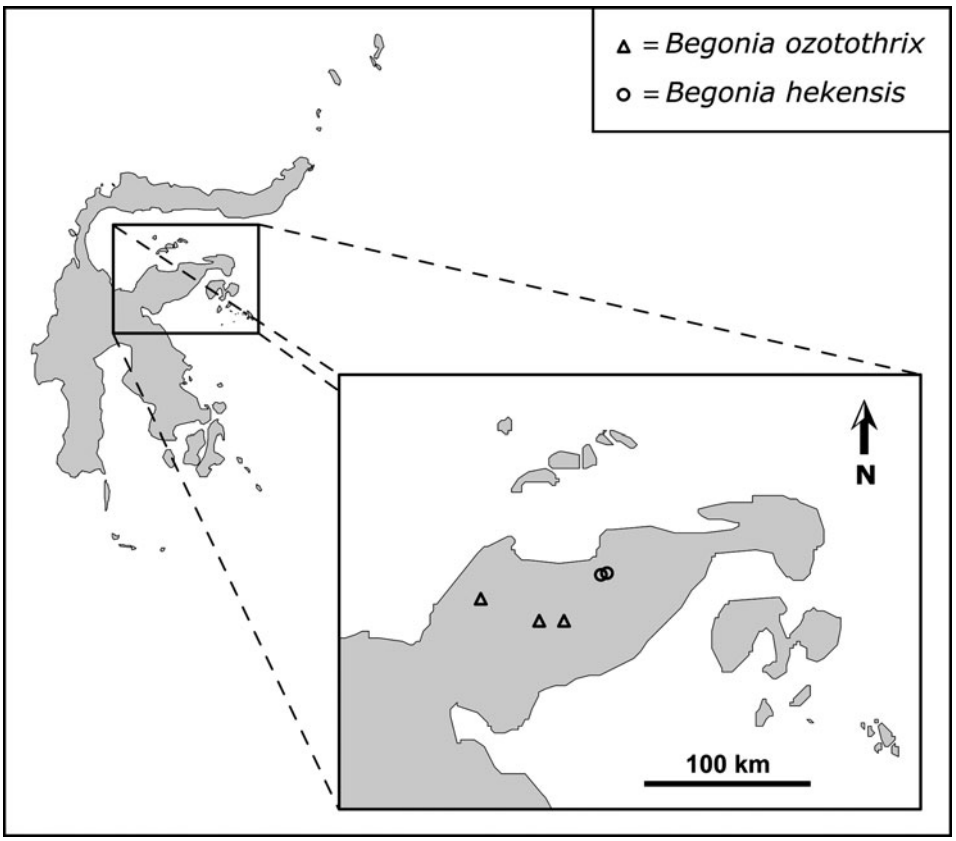


FIG. 3. Distribution of *Begonia ozotothrix* and *Begonia hekensis* in eastern Central Sulawesi (Tengah), Indonesia.

base, widest at the apex, hairy, style fused only in the most basal part, 3-branched, each stylodium bifurcate in the stigmatic region, stigmatic surface a spirally twisted papillose band, the style and stylodia pale yellow, the stigma bands orange. *Fruits*: capsules cylindric, 19–22 × 4–6 mm (without wings), on stout, erect, 1–4 mm long, hairy pedicels, dehiscent, splitting along the wing attachment, drying pale brown, hairy to glabrescent, wing shape as for ovary, wings 5–10 mm wide at the widest point (at the apex), hairy to glabrescent. *Seeds* ellipsoidal, c.0.3 mm long, collar cells c.1/2–3/4 of the length of the seed.

Distribution. Indonesia, Sulawesi, Central Sulawesi (Sulawesi Tengah), Gunung Katopas, and the lowland rainforest area between the villages of Bulan Jaya and Uwetangko, and the upland rainforest north of Uwetangko (Fig. 3).

Habitat. This species grows in the herb layer or on wet rock walls in lowland and upland primary rainforest, often at the sides of rivers or small streams, between c.300 and 800 m.

Proposed IUCN conservation category. LC. This species is locally common and more than 20 populations were observed along a c.40 km walking trail between the villages of Bulan Jaya and Uwetangko, and the upland rainforest north of Uwetangko. Additionally, several populations were observed on Gunung Katopas. While the lowland rainforest area of this species is fragmented by plantations (cacao and others), most of the upland populations were found in areas which are very difficult to access and show no or only slight signs of anthropogenic disturbance by rattan collecting.

Additional specimens examined. SULAWESI. **Tengah**: Tojo Una-una District, close to Bulan Jaya village, side of track through primary lowland rainforest, at forest margin close to river, 01°17′32.4″S, 121°57′11.6″E, 369 m, 21 iv 2008, *D.C. Thomas & W.H. Ardi* 08-52 (BO, CEB, E); close to Bulan Jaya village, disturbed lowland rainforest, margin of *Theobroma* plantation, 01°17′30.7″S, 121°57′02.6″E, 370 m, 21 iv 2008, *D.C. Thomas & W.H. Ardi* 08-53 (BO, CEB, E); between the villages of Bulan Jaya and Linkasa, on wet rock wall next to small stream, 01°17′24.4″S, 121°56′10.3″E, 364 m, 21 iv 2008, *D.C. Thomas & W.H. Ardi* 08-56 (BO, CEB, E); close to Uwetangko village, primary lowland rainforest, next to small waterfall, 01°17′09.6″S, 121°48′67.0″E, 322 m, 22 iv 2008, *D.C. Thomas & W.H. Ardi* 08-58 (BO, CEB, E); Watusongo Village, Gunung Katopas, on wet rock, primary rainforest margin at river side, 01°10′29.3″S, 121°28′36.3″E, 750 m, 11 v 2008, *D.C. Thomas & W.H. Ardi* 08-72 (BO, CEB, E); Watusongo Village, Gunung Katopas, on vertical, wet rock at the side of a small stream, primary rainforest, 01°10′14.9″S, 121°28′49.3″E, 625 m, 11 v 2008, *D.C. Thomas & W.H. Ardi* 08-74 (BO, CEB, E).

The epithet '*ozotothrix*' (from Greek *ozotos* – branched, and *thrix* – hair) refers to the very unusual branched, multicellular, multiseriate trichomes found on the stem, the petioles and the abaxial lamina surfaces of *Begonia ozotothrix* (Fig. 4). Begonias with branched trichomes are rare outside of Africa and have not previously been reported for *Begonia* section *Petermannia*. Stellate hairs have been described in only two Asian sections, *Begonia* section *Parvibegonia* A.DC. (*B. sinuata* Wall. ex Meisn.) and *Begonia* section *Diploclinium* (Lindl.) A.DC. (*B. cladotricha* M.Hughes). *Begonia picta* Sm. (*Begonia* section *Diploclinium*) has branched, flattened scale-like hairs on the capsules. *Begonia calcicola* Merr. and *B. oxysperma* A.DC. (both in *Begonia* section *Diploclinium*, although the latter is usually classified in the monotypic *Begonia* section *Baryandra* A.DC.) have hairs with a broad and flat stalk divided at the apex into few to several thinner branches on the vegetative parts (Doorenbos *et al.*, 1998; Hughes, 2007). The morphology of the male inflorescences of *Begonia ozotothrix* is noteworthy as they exhibit a strong reduction syndrome. The male inflorescence morphology predominantly found in *Begonia* section *Petermannia* is characterised by cymose-dichasial branching with only very few or no monochasial branchings in the most distal part, clearly developed axes, and small bracts usually subtending the lateral branches in the cymose inflorescence (Irmscher, 1914; Doorenbos *et al.*, 1998). However, Irmscher (1914) has already indicated that there are several variations of this typical syndrome in the huge section *Petermannia*. In *Begonia ozotothrix* the axes are strongly compressed resulting in a subumbellate appearance of the cymose inflorescences, only the bracts subtending the peduncles and the bracts subtending the lateral branches of the basal dichotomous branchings of the male partial inflorescences are developed, and

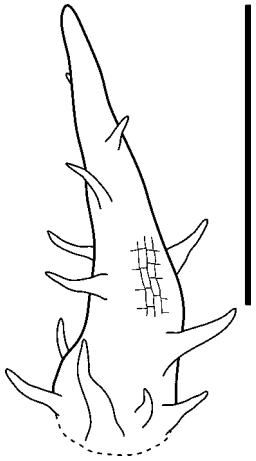


FIG. 4. Branched, multicellular trichome from the petiole of *Begonia ozotothrix* (scale bar = 500 μm) (*D.C. Thomas & W.H. Ardi* 08-58).

there are only one or sometimes two basal dichasial branchings and up to four distal monochasial branchings (Figs 1C, 5). Similar compressed subumbellate syndromes have been described for the Sulawesi endemics *Begonia siccacaudata* J.Door., which shows male partial inflorescences with only one basal dichasial branching and the end flower flanked by two monochasia (Doorenbos, 2000), and *Begonia mendumiae* M.Hughes, which shows male inflorescences with compressed monochasial partial inflorescences (Hughes, 2006). An analysis of herbarium material from Sulawesi (A, B, BM, BO, CEB, E, K, L and SING) shows that similar subumbellate male inflorescences are characteristic for several undescribed species from Sulawesi, and this reduction syndrome might represent a synapomorphy for several species derived from an endemic radiation on Sulawesi. However, phylogenetic analyses of morphological and/or molecular data are needed to test this hypothesis.

Begonia hekensis D.C.Thomas, **sp. nov.** Sect. *Petermannia*. **Figs 2, 3.**

Begoniae hispidissimae Zipp. ex Koord. similis a qua pedunculis inflorescentiarum feminearum longioribus, pedicellis capsularum valde deflexis differt. – Type: Indonesia, Sulawesi, Sulawesi Tengah, Luwuk District, Bunta Subdistrict, Sumber Agung, Gunung Hek, riverbank near small waterfall, 01°01′72.2″S, 122°11′54.7″E, 1009 m, 12 iv 2008, *D.C. Thomas & W.H. Ardi* 08-43 (holo E; iso BO, CEB).

Perennial, monoecious, erect herb, to c.100 cm tall, hairy with up to c.1.2 mm long, multicellular, multiseriate, simple trichomes and microscopic, glandular trichomes on all vegetative parts. *Stems* branched; internodes 2.2–8.3 cm long, densely hairy. *Leaves* alternate; stipules 8–28 × 2–10 mm, narrowly ovate, cymbiform with abaxially

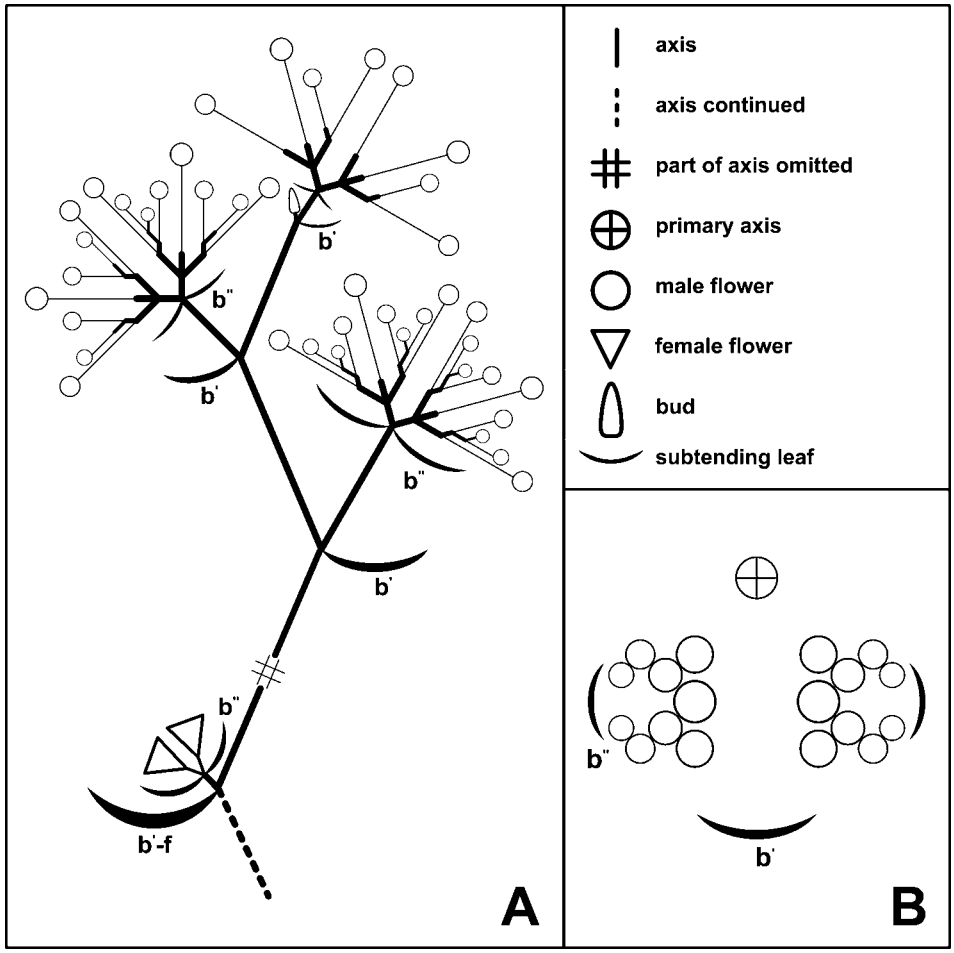


FIG. 5. Schematic inflorescence branching pattern of *Begonia ozotothrix*. A. Schematic branching of the inflorescence. The axes of the branches of the cymose-subumbellate male partial inflorescences are elongated in relation to the original in order to illustrate the branching pattern. The sequence of anthesis in the male partial inflorescences is indicated by circle size: the larger the size, the earlier the anthesis. B. Diagram of a male partial inflorescence. b′, bracteose leaf subtending the male partial inflorescence; b′-f, foliose leaf subtending the female inflorescence; b″, bracteose leaf subtending the branches of the male partial inflorescence, or subtending the pedicels of the female flowers.

prominent midrib forming a thin, short appendage at the apex, persistent, abaxially densely hairy; petioles 0.7–11.2 cm long, densely hairy; lamina basifixed, 2.5–15.2 × 1.1–8.2 cm, very asymmetric, ovate or elliptic, base cordate with non- or only very slightly overlapping lobes, apex acuminate, margin dentate to serrate, teeth bristle-pointed, abaxial surface hairy, adaxial surface sparsely hairy, adaxial surface mid green and abaxial surface pale green, venation palmate-pinnate. *Inflorescences* protogynous;

female inflorescences basal to male inflorescences or solitary, 2-flowered, subtending leaves foliose, peduncles 2.6–3 cm long (in fruit), bracts (subtending the pedicels of the female flowers) c.16 × 6 mm, narrowly elliptic, abaxially hairy; *male inflorescences* distal to one female inflorescence or solitary, subtending leaves frondose-bracteose (lamina strongly reduced in size), peduncle 9–12 mm long, bracts (subtending the lateral branches) 4–12 × 1–6 mm, oblong, the basal ones abaxially hairy, the distal ones glabrous, a once-branched dichasium or with one dichotomous branching at the base, and each of the two resulting branches branching once dichasially, sometimes the lateral branches of the dichasia branching once monochasially. *Male flowers*: pedicels 4–23 mm, hairy; tepals 2, white, 11–18 × 10–18 mm, broadly ovate to subcircular, base cordate or tepal margin convex at base, apex rounded, abaxially sparsely hairy; androecium of c.24–38 stamens, yellow, filaments c.0.4–2 mm long, slightly fused at the very base, unequal, longer in the middle of the androecium, anthers c.1–2 mm long, obovate or oblong, dehiscing through unilateral positioned slits < 1/2 as long as the anther, connective not projecting. *Female flowers*: unknown. *Fruits*: capsules ellipsoid, 14–17 × 5–8 mm (without wings), on apically strongly deflexed, 18–24 mm long, hairy pedicels, dehiscent, splitting along the wing attachment, drying brown, hairy, locules 3, placentation axile, placentae bilamellate, wings 3, sublunate, base rounded, widest in the middle to subapical part, subequal, one slightly larger than the other two, 7–8 mm wide in the widest part, the smaller two 6–7 mm in the widest part, hairy. *Seeds* ellipsoidal, c.0.3–0.4 mm long, collar cells c.1/3–1/2 of the length of the seed.

Distribution. Indonesia, Sulawesi, Central Sulawesi (Sulawesi Tengah), Gunung Hek (Fig. 3).

Habitat. This is an upland species which grows in the herb layer of primary rainforests, often along the sides of small streams, at c.850–1200 m.

Proposed IUCN conservation category. VU D2. This species is known only from Gunung Hek and has a very restricted range in an area which has no legal protection as a national park or nature reserve. Although the forest is in good condition in this area at around 1000 m, there are clear signs of anthropogenic disturbance, especially selective timber harvesting and rattan collection, at slightly lower altitudes. Therefore, the populations are 'prone to the effects of human activities or stochastic events within a very short time period in an uncertain future' (IUCN, 2001).

Additional specimens examined. SULAWESI. **Tengah**: Luwuk District, Bunta Subdistrict, Sumber Agung, Gunung Hek, Sungai Hek, between Cabang Tiga and Agathis Camp, 01°01′10″S, 122°10′30″E, 980 m, 1 iii 2004, *Hendrian, M. Newman, S. Scott, M. Nazre Saleh & D. Supriadi* 1015 (E); Sumber Agung, Gunung Hek, side of steep track, 01°01′58.2″S, 122°10′90.9″E, 870 m, 10 iv 2008, *D.C. Thomas & W.H. Ardi* 08-30 (BO, CEB, E); Sumber Agung, Gunung Hek, small isle in Sungai Hek, 01°01′81.2″S, 122°11′35.0″E, 1080 m, 11 iv 2008, *D.C. Thomas & W.H. Ardi* 08-33 (BO, CEB, E); Sumber Agung, Gunung Hek, side of small tributary of Sungai Hek, 01°01′76.0″S, 122°11′42.4″E, 993 m, 12 iv 2008, *D.C. Thomas & W.H. Ardi* 08-41 (BO, CEB, E).

The epithet '*hekensis*' is composed of 'Hek', a reference to Gunung Hek where the type material was collected, and '*-ensis*' (Latin – originating from).

Begonia hekensis is morphologically similar to *Begonia hispidissima* and *Begonia masarangensis* Irmsch. These three species exhibit a character combination which differentiates them from most other Sulawesian *Begonia* section *Petermannia* species: densely hairy stems and petioles, few-flowered male inflorescences, male flowers with abaxially hairy tepals, and short, hairy ovaries and capsules. However, *Begonia hekensis* can be easily differentiated from *Begonia masarangensis* by its ovate to elliptic leaves with a dentate to serrate margin and the compressed, purely, or at least partially, dichasially branching male inflorescences (versus oblong to narrowly elliptic leaves with double serrate margin and purely monochasially branching male inflorescences). *Begonia hekensis* differs from *B. hispidissima* by the apically strongly deflexed pedicels of the fruits (Figs 2B–C) and the peduncles of the female inflorescences which may be up to 3 cm long (versus not or only slightly deflexed pedicels and peduncles up to c.1.5 cm long in *B. hispidissima*).

Acknowledgements

We are grateful to Steve Scott for his excellent care of the *Begonia* living collections at the Royal Botanic Garden Edinburgh, to Dr Irawati (Bogor Botanic Garden) and Dr Eko Baroto Walujo (Herbarium Bogoriense) for their support of the RBGE expeditions to Sulawesi, and to the curators of A, B, BM, BO, CEB, E, K, L and SING for allowing us access to herbarium material. The support of the first author's PhD project by the M. L. MacIntyre Trust is gratefully acknowledged.

References

Cannon, C. H., Summers, M., Harting, J. R. & Kessler, P. J. A. (2007). Developing conservation priorities based on forest type, condition, and threats in a poorly known ecoregion: Sulawesi, Indonesia. *Biotropica* 39: 747–759.

Doorenbos, J. M. (2000). *Begonia siccacaudata* (Begoniaceae), a new species from Sulawesi. *Blumea* 45: 399–402.

Doorenbos, J. M., Sosef, S. M. & De Wilde, J. J. F. E. (1998). *The sections of* Begonia *including descriptions, keys and species lists*. Studies in Begoniaceae VI. Wageningen Agricultural University Papers 98(2). Wageningen: Wageningen Agricultural University.

Hughes, M. (2006). Four new species of *Begonia* (*Begoniaceae*) from Sulawesi. *Edinburgh J. Bot*. 63: 191–199.

Hughes, M. (2007). *Begonia cladotricha* (*Begoniaceae*): A new species from Laos. *Edinburgh J. Bot*. 64: 101–105.

Hughes, M. (2008). *An annotated checklist of Southeast Asian* Begonia. Edinburgh: Royal Botanic Garden Edinburgh.

Irmscher, E. (1914). Die Verteilung der Geschlechter in den Inflorescenzen der Begoniaceen unter Berücksichtigung der morphologischen Verhältnisse. *Bot. Jahrb. Syst*., Suppl. 50: 556–577.

IUCN (2001). *IUCN Red List Categories and Criteria, Version 3.1*. IUCN Species Survival Commission. Gland, Switzerland and Cambridge, UK: IUCN.

KESSLER, P. J. A., BOS, M. M., SIERRA DAZA, S. E. C., KOP, A., WILLEMSE, L. P. M., PITOPANG, R. & GRADSTEIN, S. R. (2002). *Checklist of woody plants of Sulawesi, Indonesia*. Blumea, Suppl. 14.

TEBBITT, M. C. (1997). *A systematic investigation of* Begonia *section* Sphenanthera *(Hassk.) Benth. & Hook. f*. Unpublished PhD thesis, University of Glasgow, Scotland, UK.

TEBBITT, M. C. (2003). Taxonomy of *Begonia longifolia* Blume (Begoniaceae) and related species. *Brittonia* 55(1): 19–29.

TEBBITT, M. C. & DICKSON, J. H. (2000). Amended descriptions and revised sectional assignment of some Asian Begonias (Begoniaceae). *Brittonia* 52(1): 112–117.

TEBBITT, M. C., FORREST, L. L., SANTORIELLO, A., CLEMENT, W. L. & SWENSEN, S. M. (2006). Phylogenetic relationships of Asian *Begonia*, with an emphasis on the evolution of rain-ballist and animal dispersal mechanisms in sections *Platycentrum*, *Sphenanthera* and *Leprosae*. *Syst. Bot*. 31(2): 327–336.

THOMAS, D. C. & HUGHES, M. (2008). *Begonia varipeltata* (*Begoniaceae*): A new peltate species from Sulawesi, Indonesia. *Edinburgh J. Bot*. 65: 369–374.

Received 16 July 2008; accepted for publication 21 October 2008

EDINBURGH JOURNAL OF BOTANY 66 (2): 229–238 (2009) 229

doi:10.1017/S0960428609005484

TWO NEW SPECIES OF *BEGONIA* (*BEGONIACEAE*) FROM SOUTH SULAWESI, INDONESIA

D. C. Thomas[1], W. H. Ardi[2], Hartutiningsih[3] & M. Hughes[4]

Two new species of *Begonia* (*Begoniaceae*), *Begonia didyma* D.C.Thomas & Ardi and *Begonia guttapila* D.C.Thomas & Ardi, are described from the Latimojong Mountains, South Sulawesi (Sulawesi Selatan), Indonesia. Both species belong to *Begonia* sect. *Petermannia*.

Keywords. *Begonia*, new species, Sulawesi.

Introduction

The *Begonia* L. living collections at the Botanic Garden 'Eka Kaya' Bali, which hold more than 60 indigenous Indonesian species, are an important resource for *Begonia* systematics in SE Asia. Recent expeditions organised by the Botanic Garden Bali, seed exchange with numerous institutions, as well as collaborations with the Herbarium Bogoriense, Bogor Botanic Garden, the New England Tropical Conservatory and the Royal Botanic Garden Edinburgh, have led to a rapid increase in the number of species in the collection. The new accessions include several species collected on the Indonesian island of Sulawesi (Hartutiningsih, 2005).

The *Begonia* flora of undercollected Sulawesi is poorly known. Thirty-four indigenous species of *Begonia* have been reported, 30 of which are classified in *Begonia* sect. *Petermannia* (Klotzsch) A.DC., the other four belonging to *Begonia* sect. *Sphenanthera* (Hassk.) Warb. (Hughes, 2008; Thomas *et al.*, 2009). Recent expeditions have brought to light several new species (Hughes, 2006; Thomas & Hughes, 2008; Thomas *et al.*, 2009), and a close examination of all available *Begonia* herbarium specimens from Sulawesi from A, B, BM, BO, CEB, E, K, L and SING and the living collections at the Botanic Garden Bali and the Royal Botanic Garden Edinburgh indicates that there are numerous endemic species awaiting description.

[1] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, Scotland, UK. E-mail: d.thomas@rbge.ac.uk
[2] Bogor Botanic Garden, Jl. Ir. H. Juanda No. 13, Bogor 16003, Indonesia. E-mail: wisnu.handoyo.ardi@lipi.go.id
[3] Botanic Garden Bali, Candikuning, Baturiti, Tabanan 82191, Indonesia. E-mail: hartutiningsih@yahoo.co.id
[4] Singapore Botanic Gardens, 1 Cluny Road, Singapore 259569. Current address: Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, Scotland, UK. E-mail: m.hughes@rbge.ac.uk

Two new species cultivated at the Botanic Garden Bali are described below. As with the majority of species on Sulawesi, the new species belong to *Begonia* sect. *Petermannia*. In common with most members of that section they exhibit: two-tepaled male flowers, anthers with unilaterally positioned slits, five-tepaled female flowers, two-flowered female inflorescences or solitary female flowers, three-locular ovaries with axile placentation and bilamellate placentae, and fruits with equal or subequal wings (Figs 1, 2).

Species Descriptions

Begonia didyma D.C.Thomas & Ardi, **sp. nov.** Sect. *Petermannia*. **Figs 1, 3.**

Begoniae gemellae Warb. ex L.B.Sm. & Wassh. similis a qua in caule, foliis et in tepalorum faciebus abaxialibus pilos multicellulares habenti, inflorescentia mascula semper biflora (non biflora ad quinqueflora), pedunculis inflorescentiarum feminearum brevioribus et pedicellis florum feminearum brevioribus differt. – Type: Cultivated at Bali Botanic Garden from vegetative material collected in the wild (Indonesia, Sulawesi, Sulawesi Selatan, Luwu District, Latimojong Mountains, Ranteballa village, 03°21′20″S, 120°07′36″E, 1225 m), 16 v 2008, *D.C. Thomas & W.H. Ardi* 08-77 (holo E; iso BO).

Perennial, monoecious herb with prostrate to erect stems, to c.35 cm tall, with a moderate to dense indumentum of multicellular, simple trichomes up to c.2 mm long and a sparse indumentum of microscopic, glandular trichomes on all aboveground vegetative parts. *Stems* much-branched, rooting at the lower nodes; internodes c.2–6 cm long. *Leaves* alternate; stipules persistent, 10–15 × 2–5 mm, elliptic, with an abaxially prominent midrib that projects up to c.4 mm at the apex, abaxially densely hairy along the midvein; petioles 2.5–5.6 cm long; lamina basifixed, 4.2–6.7 × 2.6–4.7 cm, very asymmetric, elliptic, base cordate, lobes not overlapping, apex acuminate, margin double serrate to double dentate, the teeth bristle-pointed, adaxial surface mid green and abaxial surface pale green, the margin reddish, venation palmate-pinnate. *Inflorescences*: *female flowers* solitary, basal to or not associated with the male inflorescences, branches bearing the female flowers c.1–2 mm long, subtending leaf foliose, 2 bracts present at the base of the pedicels of the female flowers, c.2–3 × 2–3 mm, broadly ovate; *male inflorescences* distal to or not associated with the female flowers, composed of 1–2 two-flowered partial inflorescences, each a once-branched monochasium, subtending leaves foliose, peduncles 2–7 mm long, bracts (subtending the pedicels) c.2–3 × 2–3 mm, elliptic to subcircular. *Male flowers*: pedicels 17–27 mm long, sparsely hairy; tepals 2, white or white with a tinge of pink, 12–17 × 10–14 mm, broadly ovate, base slightly cordate or with convex margins, apex rounded, abaxially sparsely hairy; androecium of c.35–43 stamens, yellow, filaments c.0.6–1.4 mm long, slightly fused at the very base, anthers c.0.9–1.2 mm long, obovate, dehiscing through unilaterally positioned slits c.1/2 as long as the anther, connective not projecting. *Female flowers*: pedicels

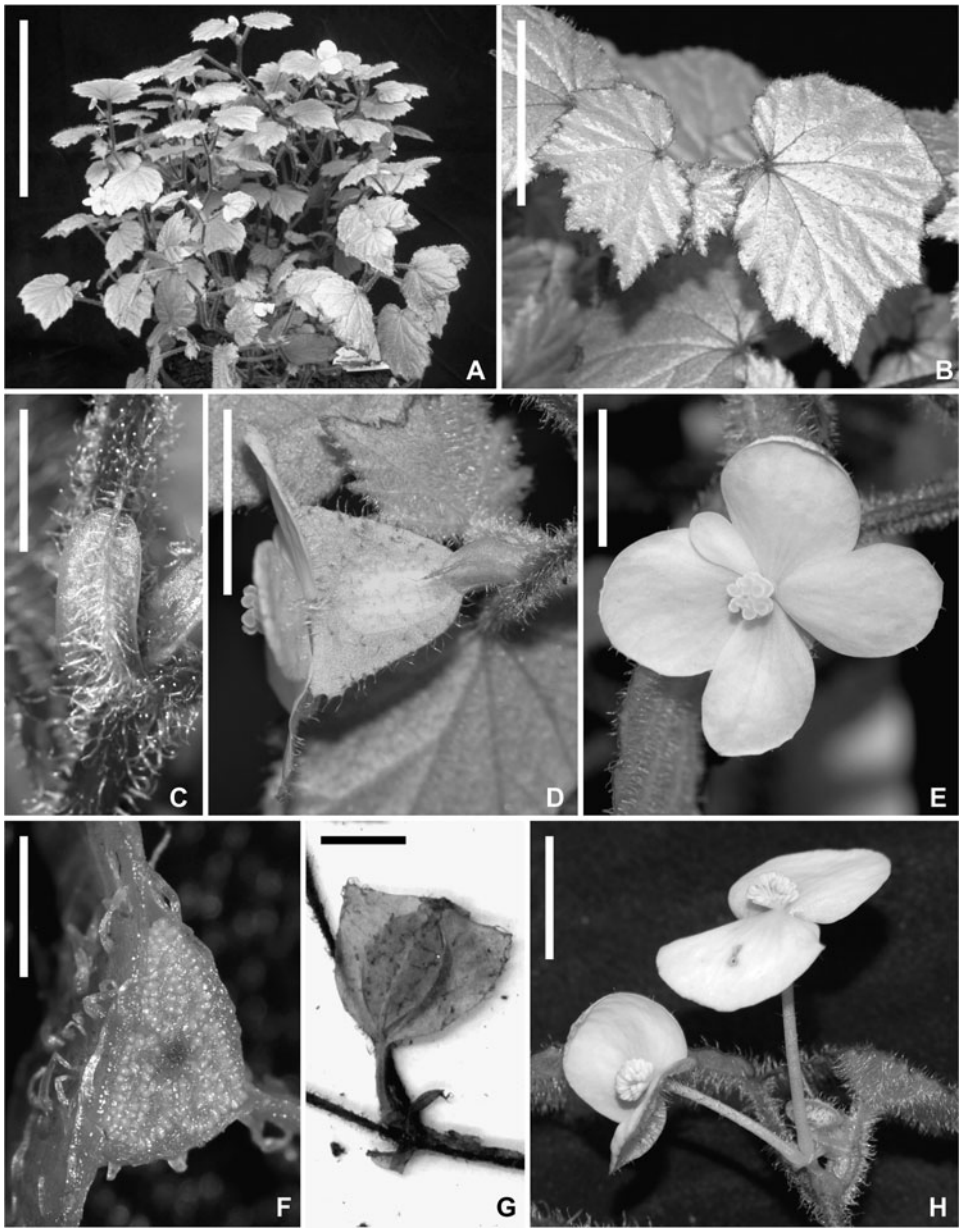


FIG. 1. *Begonia didyma* D.C.Thomas & Ardi. A, habit (scale bar = 15 cm); B, leaves (scale bar = 3 cm); C, stipule (scale bar = 6 mm); D, female flower, side view (scale bar = 12 mm); E, female flower, front view (scale bar = 10 mm); F, ovary, cross-section, three-locular with axile, bilamellate placentae (scale bar = 2 mm); G, fruit (scale bar = 5 mm); H, male inflorescence (scale bar = 10 mm). A–F, H: *D.C. Thomas & W.H. Ardi* 08-77; G: *Mogea et al.* 6596.

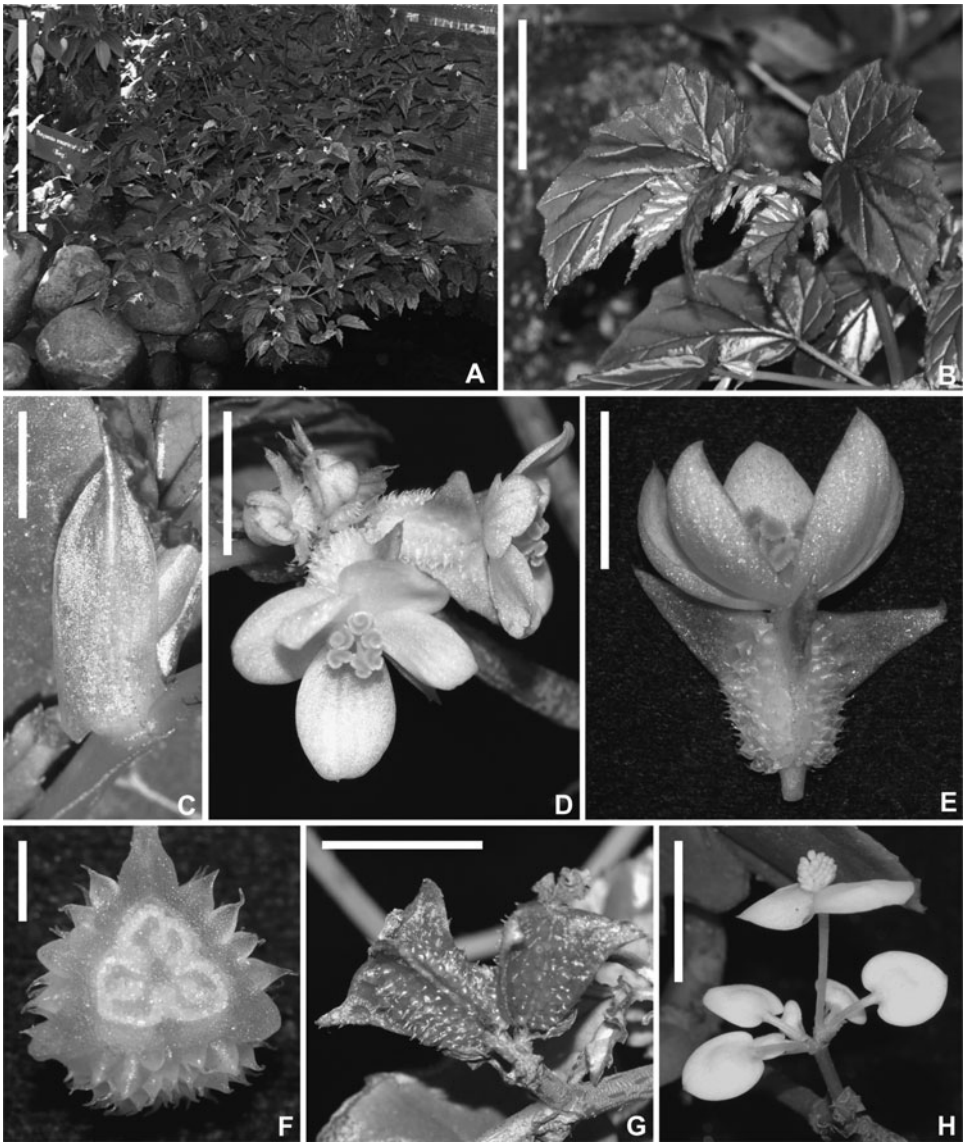


FIG. 2. *Begonia guttapila* D.C.Thomas & Ardi. A, habit (scale bar = 50 cm); B, leaves (scale bar = 5 cm); C, stipule (scale bar = 5 mm); D, female inflorescence (scale bar = 12 mm); E, female flower, side view (scale bar = 8 mm); F, ovary, cross-section, three-locular with axile, bilamellate placentae (scale bar = 2 mm); G, fruits (scale bar = 14 mm); H, male inflorescence (scale bar = 12 mm). A–H: *D.C. Thomas & W.H. Ardi* 08-81.

2–3 mm long, hairy; tepals 5, white, unequal, the two outer ones 12–13 × 10–12 mm, broadly obovate or elliptic, the two larger inner ones 13–14 × 8–9 mm, obovate, the smallest inner one 8–13 × 3–6 mm, obovate, abaxially sparsely hairy; ovary 10–13 × 11–14 mm, ellipsoid, locules 3, placentation axile, placentae bilamellate, wings 3,

narrowly triangular, rounded at base, widest at the truncate apex, hairy, style basally fused for c.1.5–2 mm, 3-branched, each stylodium bifurcate in the stigmatic region, stigmatic surface a spirally twisted papillose band, yellow. *Fruits* on thin, c.3–4 mm long, sparsely hairy pedicels; capsules ellipsoid, c.9–10 × 5–6 mm (excluding the wings), dehiscent, splitting along the wing attachment, drying pale brown, sparsely hairy, wing shape as for ovary, 4–7 mm wide at the widest point (at the apex), hairy. *Seeds* unknown.

Distribution. Indonesia, Sulawesi, South Sulawesi (Sulawesi Selatan), Luwu District, Latimojong Mountains (Fig. 3).

Habitat. Upland primary rain forest, on rocky ground, between c.1000 and 1250 m.

Proposed IUCN conservation category. VU D2. This species is only known from two collections on the eastern border of the Latimojong Forest Reserve. Despite the area's legal protection as a forest reserve, there are clear signs of anthropogenic disturbance (coffee plantations) close to the locality of this species. All available *Begonia* specimens from A, B, BM, BO, CEB, E, K, L and SING have been consulted, and hence it must be assumed, at least until more intensive collecting on

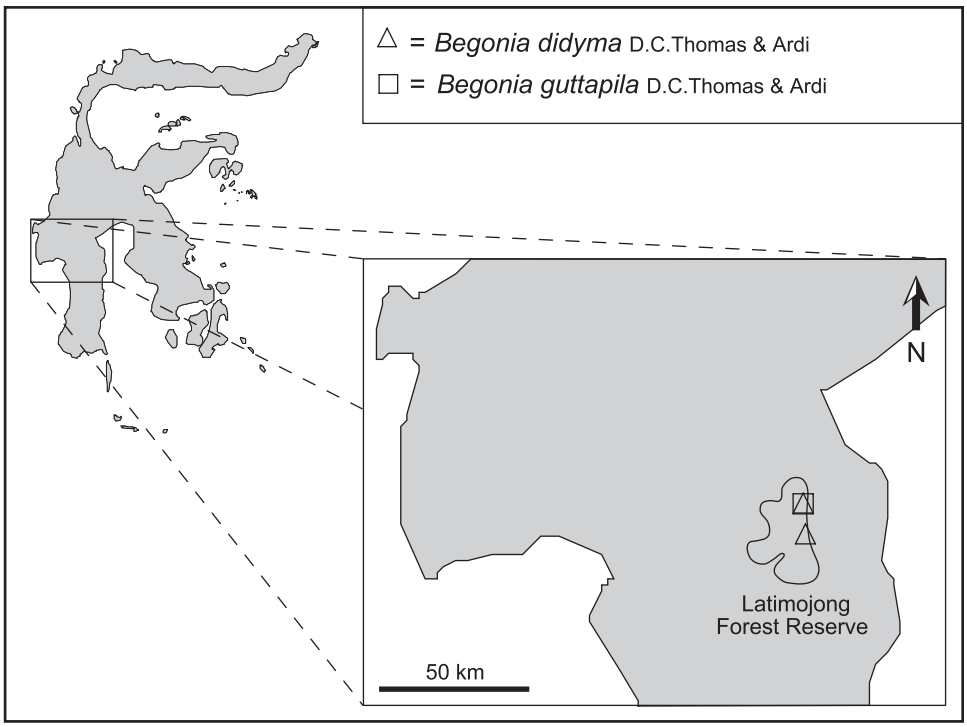


FIG. 3. Distribution of *Begonia didyma* and *Begonia guttapila* in South Sulawesi (Sulawesi Selatan), Indonesia.

Sulawesi may reveal otherwise, that this species has a very restricted range. Therefore, it is 'prone to the effects of human activities or stochastic events within a very short time period in an uncertain future' (IUCN, 2001).

Additional specimen examined. SULAWESI. **South Sulawesi**: Luwu District, Latimojong Mountain Range, Desa Lambanan Dusun Tibusan, 2 xi 1994, *J.P. Mogea, M. Amir & H. Alrasyid* 6596 (BO).

The epithet '*didyma*' (Greek *didymos* – twin) refers to the two-flowered male inflorescences of this species. *Begonia didyma* is morphologically similar to *Begonia gemella* Warb. ex L.B.Sm. & Wassh. Both species exhibit relatively thin stems rooting at the nodes, few-flowered monochasial male inflorescences and solitary female flowers. However, *Begonia didyma* can easily be distinguished from *Begonia gemella* by its moderate to dense indumentum of up to c.2 mm long, multicellular trichomes on all above-ground vegetative parts and on the tepals and the ovary of the female flowers (versus only sparsely and microscopically glandular hairy in *B. gemella*). The male inflorescences of *Begonia didyma* are strictly 2-flowered monochasial without any rudimentary, unopened flowers (Fig. 1H), while the male inflorescences of *Begonia gemella* are few-flowered monochasial (2–5-flowered). Warburg (see Koorders, 1904) probably chose the epithet '*gemella*' (Latin – twin) as reference to the male inflorescences of *Begonia gemella*, which predominantly show two fully developed flowers, but an examination of recently collected material of this species (*K. Armstrong* 364 at E), as well as the illustration in Koorders-Schumacher (1922: pl. 94) and a sketch by Irmscher on a herbarium sheet of type material of *Begonia gemella* (*S.H. Koorders* 16243β at B), show that the male inflorescences of this species comprise two to five flowers. Moreover, *Begonia didyma* exhibits very compressed branches bearing the female flowers (up to 2 mm long) and very short pedicels of the female flowers and fruits (up to 4 mm long), while in *Begonia gemella* these structures, though still short, are distinctly longer than in *B. didyma* (the branches bearing the female flowers are 5–25 mm long and the pedicels of the female flowers and fruits are 7–18 mm long).

Begonia guttapila D.C.Thomas & Ardi, **sp. nov.** Sect. *Petermannia*. **Figs 2, 3.**

Ab aliis speciebus celebicis sectionis *Petermanniae* in ovariis pilis insignibus basi bulboso habenti differt. – Type: Cultivated at Bali Botanic Garden from vegetative material collected in the wild (Indonesia, Sulawesi, Sulawesi Selatan, Luwu District, Latimojong Mountains, Ranteballa village, 03°21′33″S, 120°07′33″E, 1359 m), 16 v 2008, *D.C. Thomas & W.H. Ardi* 08-81 (holo E; iso BO).

Perennial, monoecious herb, stems first erect, but soon arching over and trailing-scrambling, to c.60 cm tall, with a sparse indumentum of microscopic, glandular hairs on all above-ground vegetative parts and a very sparse indumentum of multicellular hairs on the stems and the abaxial lamina surface or multicellular hairs absent. *Stems* branched; internodes c.2–16.5 cm long. *Leaves* alternate; stipules

14–18 × 4–6 mm, elliptic to oblong, with abaxially prominent midrib that projects shortly at the apex, caducous; petioles 1.5–5.2 cm long; lamina basifixed, 5–13 × 2.5–6 cm, very asymmetric, narrowly elliptic, elliptic, narrowly ovate or ovate, base cordate, lobes not overlapping, apex acuminate, margin double serrate, the teeth not or only slightly bristle-pointed, adaxial surface dark green and abaxial surface pale green, venation palmate-pinnate. *Inflorescences*: *female inflorescences* solitary, composed of 1–2 two-flowered partial inflorescences, subtending leaves foliose, peduncles c.2 mm long, bracts (subtending the pedicels) c.6–7 × 3–4 mm, narrowly ovate to narrowly elliptic; *male inflorescences* solitary, subtending leaves foliose, peduncles 6–9 mm, cymose-subumbellate with one dichotomous branching at the base, each of the two resulting branches branching once dichasially, the lateral branches of the dichasia branching up to three times monochasially, bracts (subtending the pedicels) c.1.5–6 × 1–2 mm, ellipsoid to oblong. *Male flowers*: pedicels 8–12 mm long, sparsely, microscopically, glandular hairy; tepals 2, white, 8–10 × 9–11 mm, broadly ovate to suborbicular, base slightly cordate or with convex margins, apex rounded, abaxially sparsely, microscopically, glandular hairy; androecium of c.38–46 stamens, yellow, filaments c.0.4–1.6 mm long, slightly fused at the very base, anthers c.0.7–1.2 mm long, obovate to oblong, dehiscing through unilaterally positioned slits > 1/2 as long as the anther, connective not projecting. *Female flowers*: pedicels 1–2 mm, sparsely, microscopically, glandular hairy; tepals 5, subequal, 10–17 × 5–9 mm, elliptic to obovate, white or pale pink, abaxially sparsely, microscopically, glandular hairy; ovary 8–12 × 12–15 mm, ellipsoid, locules 3, placentation axile, placentae bilamellate, wings 3, not developed in the basal part of the ovary, but expanding distally after c.1/3–1/2 of the ovary's length, equal, triangular, with concave margin at the base, widest at the truncate apex, microscopically, glandular hairy, the ovary surface between the wings hairy with c.0.6–2 mm long, multicellular trichomes with broad, bulbous base narrowing into a fine extended tip, style fused at the base, 3-branched, each stylodium bifurcate in the stigmatic region, stigmatic surface a spirally twisted papillose band, yellow. *Fruits* on stout, c.1–2 mm long, microscopically, glandular hairy pedicels; capsule ellipsoid, 12–15 × 5–9 mm (excluding the wings), fleshy and indehiscent, red, wings thickened and hardened, 6–9 mm wide at the widest point (at the apex), wing shape and indumentum as for ovary. *Seeds* ellipsoidal, c.0.3–0.4 mm long, collar cells c.1/2–2/3 of the length of the seed.

Distribution. Indonesia, Sulawesi, South Sulawesi (Sulawesi Selatan), Luwu District, Latimojong Mountains (Fig. 3).

Habitat. This species grows in upland primary rain forest at c.1350 m.

Proposed IUCN conservation category. VU D2. This species is only known from one collection on the eastern border of the Latimojong Forest Reserve. All available *Begonia* specimens from A, B, BM, BO, CEB, E, K, L and SING have been consulted, and hence it must be assumed, at least until more intensive collecting on

Sulawesi may reveal otherwise, that this species has a very restricted range. Therefore, it is 'prone to the effects of human activities or stochastic events within a very short time period in an uncertain future' (IUCN, 2001).

The epithet '*guttapila*' is a compound of *gutta* (Latin – a drop of fluid) and *pilus* (Latin – hair). It refers to the very unusual hairs on the ovaries and fruits of this species, which, with their bulbous base narrowing into a fine extended tip (Figs 2E–F), resemble stylised drops of water. The fruits of this species are unusual in *Begonia* sect. *Petermannia* not only because of their indumentum, but also because of their fleshy pericarp. In contrast to the dry, thin-walled capsules predominantly found in this section, *Begonia guttapila* exhibits red, fleshy and apparently indehiscent fruits, which have thickened, relatively hard wings. These characters might be adaptations to zoochory, but as for the other fleshy-fruited species of *Begonia* in SE Asia, observations of animal dispersal are lacking (Lange & Bouman, 1999; Tebbitt *et al.*, 2006), and the dispersal of the seeds of this species might be mainly by rain-wash from the decomposing fruit.

The male inflorescence morphology predominantly found in *Begonia* sect. *Petermannia* is characterised by dichasial branching, few or no distal monochasial branchings, and clearly developed axes (Irmscher, 1914; Doorenbos *et al.*, 1998). However, Irmscher (1914) also emphasised that there are several variations of this typical syndrome in the huge section *Petermannia*. The male inflorescences of *Begonia guttapila*, which are characterised by strongly compressed axes resulting in an umbel-like appearance, are similar to the subumbellate male partial inflorescences found in the Sulawesian endemic *Begonia ozotothrix* D.C.Thomas (Thomas *et al.*, 2009). An examination of herbarium material from A, B, BM, BO, CEB, E, K, L and SING shows that a similar syndrome is present in several other undescribed species from Sulawesi. Compressed subumbellate male partial inflorescences have also been described for the Sulawesian endemic *Begonia siccacaudata* J.Door., which shows male partial inflorescences with one basal dichasial branching and the end flower flanked by two monochasia (Doorenbos, 2000). Compressed cymose partial inflorescences are also present in the '*Begonia rieckei* Warb. complex', which includes two taxa endemic to Sulawesi, *B. koordersii* Warb. ex L.B.Sm. & Wassh. and *B. strictipetiolaris* Irmsch., but also *B. rieckei* (Sulawesi, Moluccas, New Guinea), *B. pseudolateralis* Warb. (Philippines), *B. brachybotrys* Merr. & L.M.Perry (New Guinea and surrounding islands), and *B. peekelii* Irmsch. (Bismarck Archipelago) (Hughes, 2008). According to Hughes (2008), these taxa may be best considered as one widespread species, as they show only minor morphological differences. Examination of material from Sulawesi and the Philippines showed that in contrast to the cymose-subumbellate male inflorescences of some Sulawesian species, and in contrast to all other species in *Begonia* sect. *Petermannia*, the cymose-subumbellate partial inflorescences of species in the *Begonia rieckei* complex comprise both male and female flowers. Despite this major difference, the subumbellate partial inflorescence architecture in the *Begonia rieckei* complex seems to indicate a close

relationship with the cymose-subumbellate taxa from Sulawesi. It is tempting to speculate that the observed variation in male inflorescence morphology of Sulawesian *Begonia* sect. *Petermannia* species may be the result of evolution from many-flowered male inflorescences with well-developed axes and predominantly dichasial branching (e.g. *B. grandipetala* Irmsch., *B. macintyreana* M.Hughes, *B. stevei* M.Hughes, *B. varipeltata* D.C.Thomas) to subumbellate, dichasial–monochasial male inflorescences with strongly compressed axes (e.g. *B. guttapila*, *B. ozotothrix*) to compressed, purely monochasial inflorescences including the two-flowered monochasial inflorescences of *B. didyma*. Other Sulawesian species, such as *Begonia chiasmogyna* M.Hughes, exhibit male inflorescences with monochasial branching and well-developed axes, which are most likely derived from dichasial–monochasial male inflorescences with well-developed axes. However, phylogenetic analyses of morphological and/or molecular data are necessary to investigate these hypotheses.

Acknowledgements

We are grateful to Dr Irawati (Bogor Botanic Garden) and Dr Eko Baroto Walujo (Herbarium Bogoriense) for their support of RBGE expeditions to Sulawesi, to I Dewa Darma (Botanic Garden Bali) for his *Begonia* collections from South Sulawesi, to the curators of A, B, BM, BO, CEB, E, K, L and SING for allowing us access to herbarium material, to Dr Robert Mill for the correction of the Latin diagnoses, and to Dr Mark Tebbitt and an anonymous reviewer for their constructive reviews of the manuscript. The support of the first author's PhD project by the M. L. MacIntyre Trust is gratefully acknowledged.

References

Doorenbos, J. M. (2000). *Begonia siccacaudata* (Begoniaceae), a new species from Sulawesi. *Blumea* 45: 399–402.

Doorenbos, J. M., Sosef, S. M. & De Wilde, J. J. F. E. (1998). *The sections of* Begonia *including descriptions, keys and species lists*. Studies in Begoniaceae VI. Wageningen Agricultural University Papers 98(2). Wageningen: Wageningen Agricultural University.

Hartutiningsih (2005). *Begonia Kebun Raya Bali*. Candikuning: UPT Balai Konservasi Tumbuhan, Kebun Raya Bali-LIPI.

Hughes, M. (2006). Four new species of *Begonia* (*Begoniaceae*) from Sulawesi. *Edinburgh J. Bot*. 63: 191–199.

Hughes, M. (2008). *An annotated checklist of Southeast Asian* Begonia. Edinburgh: Royal Botanic Garden Edinburgh.

Irmscher, E. (1914). Die Verteilung der Geschlechter in den Inflorescenzen der Begoniaceen unter Berücksichtigung der morphologischen Verhältnisse. *Bot. Jahrb. Syst., Suppl*. 50: 556–577.

IUCN (2001). *IUCN Red List Categories and Criteria, Version 3.1*. IUCN Species Survival Commission. Gland, Switzerland and Cambridge, UK: IUCN.

Koorders, S. H. (1904). Enumeratio specierum phanerogamarum minahassae. *Natuurk. Tijdschr. Ned.-Indië* 63: 90–91.

KOORDERS-SCHUMACHER, A. (1922). *Supplement op het Eerste Overzicht der Flora van N.O. Celebes ter verslag eener botanische dienstreis door de Minahassa*. Part 3. Batavia: van Visser & Co.

LANGE, A. & BOUMAN, F. (1999). *Seed micromorphology of neotropical Begonias*. Smithsonian Contr. Bot. 90. Washington, DC: Smithsonian Institution Press.

TEBBITT, M. C., FORREST, L. L., SANTORIELLO, A., CLEMENT, W. L. & SWENSEN, S. M. (2006). Phylogenetic relationships of Asian *Begonia*, with an emphasis on the evolution of rain-ballist and animal dispersal mechanisms in sections *Platycentrum*, *Sphenanthera* and *Leprosae*. *Syst. Bot*. 31: 327–336.

THOMAS, D. C. & HUGHES, M. (2008). *Begonia varipeltata* (*Begoniaceae*): a new peltate species from Sulawesi, Indonesia. *Edinburgh J. Bot*. 65: 369–374.

THOMAS, D. C., ARDI, W. H. & HUGHES, M. (2009). Two new species of *Begonia* (*Begoniaceae*) from Central Sulawesi, Indonesia. *Edinburgh J. Bot*. 66: 103–114.

Received 29 October 2008; accepted for publication 25 February 2009

Appendix 2. Accessions used for the generation of DNA sequence data. Acc. nom.: accession number of living material cultivated at botanic gardens; BaBG: Bali Botanic Garden; BoBG: Bogor Botanic Garden; BrBG: Brooklyn Botanic Garden; GBG: Glasgow Botanic Gardens; PCHW: Private collection of Harry Wiriadinata (Indonesia, Bogor); RBGE: Royal Botanic Garden Edinburgh; SBG: Singapore Botanic Gardens. Locality data of cultivated material is given when seed or vegetative material was collected in the wild. Locality data in square brackets indicates that the source locality is unknown, and is based on the known distribution of a taxon. Herbarium acronyms follow the *Index Herbariorum* at http://sciweb.nybg.org/

Taxon	Origin	Voucher (Herbarium)
Begonia acetosella Craib	Cultivated: GBG (acc. nom.: 001 073 96), Vietnam	Thomas, D. C. & Ardi, W. H. 08-105 (E)
Begonia* aff. *bracteata Jack	Indonesia, Sumatra	Wilkie, P., Hughes, M., Sumadijaya, A., Rasnovi, S., Marlan & Suhardi 621 (E)
Begonia* aff. *celebica Irmsch. 1	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-102 (E)
Begonia* aff. *celebica Irmsch. 2	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-104 (E)
Begonia* aff. *congesta Ridl.	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-05 (E)
Begonia* aff. *elisabethae Kiew	Cultivated: RBGE (acc. nom.: 20081038), Vietnam	Thomas, D. C. 08-149 (E)
Begonia* aff. *labordei H.Lév.	Cultivated: RBGE (acc. nom.: 20020477), China	Möller, M. 01-156B (E)
Begonia* aff. *multangula Blume	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-85 (E)
Begonia* aff. *ozotothrix D.C.Thomas	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-76 (E)
Begonia* aff. *propinqua Ridl.	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-06 (E)
Begonia amphioxus Sands	Cultivated: GBG (acc. nom.: 001 156 94), Malaysia, Borneo	Forrest, L. L. 141 (E)
Begonia aptera Blume **subsp. *hirtissima*** Girmansyah & D.C.Thomas	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-75 (E)
Begonia aptera Blume **susbsp. *aptera***	Indonesia, Sulawesi	Smith, P. & Galloway, L. 67 (E)
Begonia areolata Miq.	Cultivated: BoBG, Indonesia, Java	Thomas, D. C. & Ardi, W. H. 09-137 (E)
Begonia argenteomarginata Tebbitt	Cultivated: GBG (acc. nom.: 008 038 87), Papua New Guinea	Forrest, L. L. 145 (E)
Begonia boliviensis A.DC.	Cultivated: GBG (acc. nom.: 00801998) [Bolivia]	No voucher available
Begonia bonthainensis Hemsl.	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-63 (E)
Begonia bracteata Jack	Indonesia, Java	Ardi, W. H. & Thomas, D. C. 25 (E)
Begonia brevirimosa Irmsch.	Cultivated: RBGE (acc. nom.: 19821108, Papua New Guinea	Forrest, L. L. 137 (E
Begonia burbidgei Stapf	Cultivated: RBGE (acc. nom.: 2006.1666), Malaysia, Borneo	Thomas, D. C. 07-26 (E)
Begonia capituliformis Irmsch.	Indonesia, Sulawesi	Kinho, J. & Poulsen, A. 169 (E)
Begonia chiasmogyna M.Hughes	Cultivated: RBGE (acc. nom.: 20021895), Indonesia, Sulawesi	Thomas, D. C. 07-29 (E)

Begonia chloroneura P.Wilkie & Sands	Cultivated: RBGE (acc. nom.: 19972555), Philippines, Luzon Island	Forrest, L.L. 128 (E)
Begonia chlorosticta Sands	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-04 (E)
Begonia cleopatrae Coyle	Philippines, Palawan	Wilkie, P., Mendum, M., Argent, G. C. G., Cronk, Q., Middleton, D. J., Fuentes, R. & Chavez, R. V. 25373 (E)
Begonia corrugata Kiew & S.Julia	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-02 (E)
Begonia decora Stapf	Cultivated: RBGE (acc. nom.: 20021608), Malaysia, Peninsula Malaysia	Neale, S. 8C (E)
Begonia didyma D.C.Thomas & Ardi	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-77 (E)
Begonia dipetala Graham	Cultivated: GBG (acc. nom.: 001 004 86) [India, Sri Lanka]	Forrest, L. L. 239 (E)
Begonia dregei Otto & Dietr.	Cultivated: RBGE (acc. nom.: 20000902), South Africa)	McLellan, T. 415 (E)
Begonia fenicis Merr.	Cultivated: GBG (acc. nom.: 003 023 02), Philippines	Thomas, D. C. 08-119 (E)
Begonia flagellaris Hara	Nepal	Rajbhandary, S. & Bista ,S. 54 (E)
Begonia floccifera Bedd.	Cultivated: GBG (acc. nom.: 030 099 89) [India, Sri Lanka]	Forrest, L. L. 238 (E)
Begonia foliosa H.B. & K.	Cultivated: RBGE (acc. nom.: 19691804), [Columbia]	No voucher available
Begonia goegoensis N.E.Br.	Cultivated: GBG (acc. nom.: 011 125 57), Indonesia, Sumatra	Thomas, D. C. & Ardi, W. H. 08-107 (E)
Begonia grandis Dryand.	Cultivated: RBGE (acc. nom.: 19521036), China	Thomas, D. C. 08-145 (E)
Begonia goudotii A.DC.	Madagascar	Plana, V. 120 (E)
Begonia guttapila D.C.Thomas & Ardi	Cultivated: BaBG, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-81 (E)
Begonia hatacoa Buch.-Ham.	Cultivated: GBG (acc. nom.: 004 005 89), Nepal	Thomas, D. C. 08-110 (E)
Begonia hekensis D.C.Thomas	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-43 (E)
Begonia herbacea Vell.	Cultivated: RBGE (acc. nom.: 19731857), [Brazil]	Forrest, L. L. 163 (E)
Begonia hernandioides Merr.	Cultivated: GBG (acc. nom.: 006 035 89), Philippines	Forrest, L.L. 129 (E),
Begonia hispidissima Zipp. ex Koord.	Indonesia, Sulawesi	Kinho, J. & Poulsen, A. 168 (E)
Begonia incisa A.DC.	Cultivated: RBGE (acc. nom.: 006 151 95), Philippines, Luzon	Forrest, L. L. 139 (E)
Begonia kingiana Irmsch.	Cultivated: GBG (acc. nom.: 018 070 07), Malaysia, Peninsula Malaysia	Thomas, D. C. 08-102 (E)
Begonia laruei M.Hughes	Indonesia, Sumatra	Hughes, M. 1389 (E)

Begonia longifolia Blume	Cultivated: RBGE (acc. nom.: 20021614), Malaysia, Peninsula Malaysia	Neale, S. 11C (E)
Begonia macintyreana M.Hughes	Cultivated: RBGE (acc. nom.: 20021848), Indonesia, Sulawesi	Thomas, D. C. 07-28 (E)
Begonia malabarica Lam.	Cultivated: GBG (acc. nom.: 002 018 96), [India, Sri Lanka]	Forrest, L. L. 288 (E)
Begonia masarangensis Irmsch.	Cultivated: PCHW, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-131 (E)
Begonia masoniana Irmsch. ex Ziesenh.	Cultivated: RBGE (acc. nom.: 19980075), [China]	Thomas, D. C. 07-24 (E)
Begonia mendumiae M.Hughes	Cultivated: RBGE (acc. nom.: 20021912), Indonesia, Sulawesi	Thomas, D. C. 07-27 (E)
Begonia morsei Irmsch.	Cultivated: RBGE (acc. nom.: 19980076), China	No voucher available
Begonia multangula Blume	Indonesia, Bali	Thomas, D. C. & Ardi, W. H. 08-90 (E)
Begonia multijugata M.Hughes	Indonesia, Sumatra	Wilkie, P., Hughes, M., Sumadijaya, A., Rasnovi, S., Marlan & Suhardi 768 (E)
Begonia muricata Blume	Indonesia, Java	Ardi, W. H. & Thomas, D. C. 27 (E)
Begonia negrosensis Elmer	Philippines, Negros	Wilkie, P. 76 (E)
Begonia nelumbiifolia Cham. & Schlecht.	Cultivated: RBGE (acc. nom.: 19791888), Mexico	Hunt, D.R. 7516 (K)
Begonia nigritarum Steud.	Cultivated: RBGE (acc. nom.: 19991994), Philippines, Luzon Island	Thomas, D. C. 07-25 (E)
Begonia oxyloba Welw. ex Hook.f.	Cultivated: RBGE (acc. nom.: 19982761), Tanzania	Thomas, D. C. 08-141 (E)
Begonia ozotothrix D.C.Thomas	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-58 (E)
Begonia palmata D.Don	Cultivated: RBGE (acc. nom.: 20020476), China	Möller, M. 01-127 (E)
Begonia pavonina Ridl.	Cultivated: RBGE (acc. nom.: 20021611), Malaysia, Peninsula Malaysia	Neale, S. 9C (E)
Begonia pendula Ridl.	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-03 (E)
Begonia piurensis L.B.Smith & B.G.Schubert	Cultivated: GBG (acc. nom.: 00403476), Ecuador	No voucher available
Begonia poculifera Hook.f.	Cultivated: RBGE (acc. nom.: 19923143), Cameroon	Forrest, L. L. 234 (E)
Begonia polygonoides Hook.f.	Ivory Coast	van der Burg, W. J. 244 (WAG)
Begonia radicans Vell.	Cultivated: GBG (acc. nom.: 00908995) [Brazil]	No voucher available
Begonia rieckei Warb.	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-62
Begonia robusta Blume	Cultivated: BaBG, Indonesia, Java	Thomas, D. C. & Ardi, W. H. 08-133 (E)
Begonia roxburghii A.DC.	Cultivated: GBG (acc. nom.: 011 007 97), India	Thomas, D. C. 08-103 (E)
Begonia samhaensis M.Hughes & A.G.Mill.	Cultivated: RBGE (acc. nom.: 1999.0412), Yemen, Socotra	Thomas, D. C. 09-01 (E)

Begonia serratipetala Irmsch.	Cultivated: RBGE (acc. nom.: 19681637), Papua New Guinea	Forrest, L. L. 135 (E)
Begonia *siccacaudata* J.Door.	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-60 (E)
Begonia sikkimensis A.DC.	Cultivated: RBGE (acc. nom.: 20051755), India	Thomas, D. C. 08-144 (E)
Begonia silletensis (A.DC.) C.B.Clarke **subsp. *mengyangensis*** Tebbitt & K.Y.Guan	Cultivated: GBG (acc. nom.: 001 152 95), China	Thomas, D. C. 08-104 (E)
Begonia sizemoreae Kiew	Cultivated: GBG (acc. nom.: 001 014 00), Vietnam	Thomas, D. C. 08-111 (E)
Begonia socotrana Hook.f.	Yemen, Socotra	Miller, A. G. 19210/10 (E)
***Begonia* spec. Borneo 1**	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-07 (E)
***Begonia* spec. Borneo 2**	Cultivated: SBG, Malaysia, Borneo	Thomas, D. C. 09-08 (E)
***Begonia* spec. Borneo 3**	Cultivated: BoBG, Indonesia, Borneo	Thomas, D. C. & Ardi, W. H. 09-136 (E)
***Begonia* spec. Borneo 4**	Cultivated: RBGE (acc. nom.: 20030131), Malaysia, Borneo	Thomas, D. C. 07-1 (E)
***Begonia* spec. China 1**	Cultivated: RBGE (acc. nom.: 19980067), China	Forrest, L. L. 31 (E)
***Begonia* spec. New Guinea 1**	Indonesia, Papua	Armstrong, K. 351 (E)
***Begonia* spec. New Guinea 2**	Cultivated: BoBG, Indonesia, Papua	Thomas, D. C. & Ardi, W. H. 09-139 (E)
***Begonia* spec. Philippines 1**	Cultivated: RBGE (acc. nom.: 20080433), Philippines: Luzon Island	Thomas, D. C. 08-146 (E)
***Begonia* spec. Solomon Islands 1**	Solomon Islands	Pitisopa, F., Gardner, M. F., Herrington, S. 10 (E)
***Begonia* spec. Sulawesi 1**	Cultivated: PCHW, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-92 (E)
***Begonia* spec. Sulawesi 2**	Cultivated: PCHW, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-132 (E)
***Begonia* spec. Sulawesi 3**	Cultivated: BaBG, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-84 (E)
***Begonia* spec. Sulawesi 4**	Indonesia, Sulawesi	Vermeulen, J.J. 2301 (L)
***Begonia* spec. Sulawesi 7**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-108 (E)
***Begonia* spec. Sulawesi 8**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-112 (E)
***Begonia* spec. Sulawesi 9**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-120 (E)
***Begonia* spec. Sulawesi 10**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-121 (E)
***Begonia* spec. Sulawesi 11**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-125 (E)
***Begonia* spec. Sulawesi 12**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-20 (E)
***Begonia* spec. Sulawesi 13**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-110 (E)
***Begonia* spec. Sulawesi 14**	Indonesia, Sulawesi	Smith, P. & Galloway, L. 73A (E)
***Begonia* spec. Sulawesi 15**	Cultivated: BrBG, Indonesia, Sulawesi	Tebbitt, M. s.n. (E)
***Begonia* spec. Sulawesi 16**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-97 (E)

***Begonia* spec. Sulawesi 17**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-51 (E)
***Begonia* spec. Sulawesi 18**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-100 (E)
***Begonia* spec. Sulawesi 19**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-123 (E)
***Begonia* spec. Sulawesi 20**	Cultivated: BaBG, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-53 (E)
***Begonia* spec. Sulawesi 21**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-62 (E)
***Begonia* spec. Sulawesi 22**	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-78 (E)
***Begonia* spec. Sumatra 1**	Cultivated: BoBG, Indonesia, Sumatra	Thomas, D. C. & Ardi, W. H. 08-132 (E)
***Begonia* spec. Sumatra 2**	Indonesia, Sumatra	Hughes, M. 1402 (E)
***Begonia* spec. Sumatra 3**	Cultivated: BoBG, Indonesia, Sumatra	Thomas, D. C. & Ardi, W. H. 09-134 (E)
***Begonia* spec. Sumatra 4**	Cultivated: RBGE (acc. nom.: 20080584), Indonesia, Sumatra	Hughes, M. 1502 (E)
***Begonia* spec. Sumatra 5**	Cultivated: RBGE (acc. nom.: 20070752), Indonesia, Sumatra	Thomas, D. C. 07-31 (E)
***Begonia* spec. Sumbawa 1**	Cultivated: BaBG, Indonesia, Sumbawa	Thomas, D. C. & Ardi, W. H. 08-85 (E)
***Begonia* spec. Sumbawa 2**	Cultivated: BoBG, Indonesia, Sumbawa	Thomas, D. C. & Ardi, W. H. 09-138 (E)
Begonia stevei M.Hughes	Cultivated: RBGE (acc. nom.: 20040642), Indonesia, Sulawesi	Thomas, D. C. 07-30 (E)
Begonia strigosa (Warb.) L.L.Forrest & Hollingsw.	Cultivated: GBG (acc. nom.: Papua New Guinea)	Forrest, L. L. 143 (E)
Begonia sudjanae Jansson	Cultivated: GBG (acc. nom.: 026 054 99) [Indonesia, Sumatra]	Thomas, D. C. & Ardi, W. H. 08-109 (E)
Begonia sutherlandii Hook.f.	Cultivated: RBGE (acc. nom.: 2001.0167), South Africa	Thomas, D. C. 08-140 (E)
Begonia symsanguinea L.L.Forrest & Hollingsw.	Cultivated: GBG (acc. nom.: 003 127 93), Papua New Guinea	Forrest, L. L. 142 (E)
Begonia tenuifolia Dryand.	Indonesia, Bali	Thomas, D. C. & Ardi, W. H. 08-86 (E)
Begonia thomeana C.DC.	Cultivated: GBG (acc. nom.: 05407997), São Tomé	Forrest, L. L. 199 (E)
Begonia ulmifolia Willd.	Cultivated: RBGE (acc. nom.: 19682869), [Guyana]	Forrest, L. L. 169 (E)
Begonia varipeltata D.C.Thomas	Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 08-51 (E)
Begonia venusta King	Cultivated: RBGE (acc. nom.: 20021596), Malaysia, Peninsula Malaysia	Neale, Sophie 7 (E)
Begonia verecunda M.Hughes	Indonesia, Sumatra	Wilkie, P., Hughes, M., Sumadijaya, A., Rasnovi, S., Marlan & Suhardi 618 (E)
Begonia versicolor Irmsch.	Cultivated: RBGE (acc. nom.: 19980037), China	Forrest, L. L. 2 (E)

Begonia watuwilensis Girmansyah	Cultivated: BaBG, Indonesia, Sulawesi	Thomas, D. C. & Ardi, W. H. 09-55 (E)
Hillebrandia sandwicensis Oliv.	USA, Hawaii	Morris, R. 1264-2005 (E)

Appendix 3. Genbank accessions used in the analyses. Sequences were downloaded from the nucleotide database of the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/).

Taxon	ITS	*mat*K	*rbc*L	*trn*L intron	*trn*L-F spacer
Alnus sinuata (Regel) Rydb.	n/a	AY263914	AY263926	AY147067	AY147067
Alsomitra macrocarpa M.Roem.	n/a	DQ536632	DQ535780	DQ536784	DQ536784
Anisophyllea corneri Ding Hou	n/a	AY968444	AF027109	AY968559	AY968375
Anisophyllea fallax Scott-Elliot	n/a	AY935923	AY935742	AY968560	AY935779
Annamocarya sinensis (Dode) J.-F.Leroy	n/a	AY263919	AY263935	AY147080	AY147080
Austrobryonia argillicola I.Telford	n/a	EF487555	EF487549	EF487572	EF487572
Baijiania yunnanensis (A.M.Lu & Zhi Y.Zhang) A.M.Lu & J.Q.Li	n/a	DQ469138	DQ501258	DQ501270	DQ501270
Begonia alveolata Yu	AY048977	n/a	n/a	n/a	n/a
Begonia balansana Gagnep.	AF485091	n/a	n/a	n/a	n/a
Begonia boisiana Gagnep.	AF534719	n/a	n/a	n/a	n/a
Begonia cavaleriei H.Lév.	GU176060	n/a	n/a	n/a	n/a
Begonia chloroneura P.Wilkie & Sands	AF485134	n/a	n/a	n/a	n/a
Begonia cirrosa L.B.Sm. & Wassh.	AY048979	n/a	n/a	n/a	n/a
Begonia grandis Otto ex A.DC.	AF485088	n/a	n/a	n/a	n/a
Begonia henryi Hemsl.	GU176061	n/a	n/a	n/a	n/a
Begonia hernandioides Merr.	AF485135	n/a	n/a	n/a	n/a
Begonia incisa A.DC.	AF485148	n/a	n/a	n/a	n/a
Begonia labordei H.Lév.	AF485122	n/a	n/a	n/a	n/a
Begonia leprosa Hance	AY753722	n/a	n/a	n/a	n/a
Begonia longicarpa K.Y.Guan & D.K.Tian	AF485109	n/a	n/a	n/a	n/a
Begonia malabarica Lam.	AF485140	n/a	n/a	n/a	n/a
Begonia muricata Blume	AY753725	n/a	n/a	n/a	n/a
Begonia nepalensis Warb.	AY753726	n/a	n/a	n/a	n/a
Begonia nigritarum Steud.	AF534715	n/a	n/a	n/a	n/a
Begonia oxysperma A.DC.	AF485131	n/a	n/a	n/a	n/a
Begonia parvula H.Lév. & Vaniot	GU176066	n/a	n/a	n/a	n/a
Begonia rajah Ridl. & Ridl.	AF485136	n/a	n/a	n/a	n/a
Begonia serratipetala Irmsch.	AF485143	n/a	n/a	n/a	n/a
Begonia variabilis Ridl.	AY753732	n/a	n/a	n/a	n/a
Begonia versicolor Irmsch.	AF485090	n/a	n/a	n/a	n/a
Begonia wenshanensis C.M.Hu ex C.Y.Wu & T.C.Ku	AY048974	n/a	n/a	n/a	n/a
Betula platyphylla Sukaczev	n/a	AB015457	AY263927	AY147068	AY147068
Carpinus betulus L.	n/a	AY263915	AY263928	AY147070	AY147070
Carya ovata (Mill.) K.Koch	n/a	U92850	AY263931	AY147074	AY147074
Castanea seguinii Dode	n/a	AY263920	AY263937	AY147082	AY147082
Castanopsis tibetana Hance	n/a	AY263921	AY147096	AY147083	AY147083
Casuarina equisetifolia L.	n/a	AB015462	AY263930	AY147090	AY147090
Coccinia sessilifolia Cogn.	n/a	AY968446	AY968520	AY968568	AY968385
Combretocarpus rotundatus (Miq.) Danser	n/a	AY968447	AF127698	AY968561	AY968376
Comptonia peregrina (L.) J.M.Coult.	n/a	U92856	X69529	AY263905	AY263905
Coriaria myrtifolia L.	n/a	AB016459	AY968521	AY091824	AY091824
Coriaria nepalensis Wall.	n/a	AB016460	AY968522	AY091825	AY091825

Coriaria ruscifolia L.	n/a	AB016462	AB016448	AY091826	AY968380
Coriaria sarmentosa G.Forst.	n/a	AB016464	AB016450	AY091829	AY968381
Corylus avellana L.	n/a	AY263916	AY263929	AY147072	AY147072
Corynocarpus laevigata J.R.Forst. & G.Forst.	n/a	AY968448	X69731	AY968565	AY968382
Cucurbita pepo L.	n/a	DQ536666	AF206756	DQ536808	DQ536808
Cyclocarya paliurus (Batal.) Iljinsk.	n/a	AY147098	AY263942	AY147075	AY147075
Datisca cannabina L.	n/a	AB016467	AB016453	AY238600	AY968383
Datisca glomerata (C.Presl) Baill.	n/a	AY968449	L21940	AY238601	AY968384
Dendrosicyos socotranus Balf.f.	n/a	AY973018	AY973022	AY973005	AY973005
Ecballium elaterium (L.) A.Rich.	n/a	AY973019	AY973023	EU102417	EU102420
Engelhardia fenzelii Merr.	n/a	AY147099	AY147095	AY147076	AY147076
Fagus grandifolia Ehrh.	n/a	U92861	AY263936	AY147087	AY147087
Fevillea pergamentacea Cogn.	n/a	DQ536679	DQ535813	DQ536819	DQ536819
Gerrardanthus grandiflorus Gilg ex Cogn.	n/a	DQ536668	DQ535805	DQ536768	DQ536768
Gurania spinulosa Cogn.	n/a	DQ536681	DQ535815	DQ536822	DQ536822
Gynostemma pentaphyllum (Thunb.) Makino	n/a	AY968451	AY968523	EU436355	EU436355
Hemsleya amabilis Diels	n/a	EU436407	EU436384	EU436356	EU436356
Hillebrandia sandwicensis Oliv.	n/a	AY968452	U59822	AY238599	AY968379
Juglans mandshurica Maxim.	n/a	AF118033	AY263932	AY147077	AY147077
Lagenaria breviflora (Benth.) Roberty	n/a	AY935934	AY935747	AY968570	AY935788
Lithocarpus henryi Rehder & E.H.Wilson	n/a	AY263923	AY147097	AY147086	AY147086
Marah macrocarpa Greene	n/a	AY968453	AY968524	AY968571	AY968387
Melothria pendula L.	n/a	DQ536699	DQ535828	DQ536839	DQ536839
Momordica charantia L.	n/a	DQ491019	DQ535760	DQ501269	DQ501269
Myrica cerifera L.	n/a	U92857	AF119179	AY147089	AY147089
Neoalsomitra sarcophylla (Wall.) Hutch.	n/a	AY968454	AY968525	AY968572	AY973008
Nothofagus antarctica (G.Forst.) Oerst.	n/a	AY263924	AY263939	AY147091	AY147091
Octomeles sumatrana Miq.	n/a	AY968455	L21942	AY968574	AY968389
Ostrya virginiana K.Koch	n/a	AB015460	X56620	AY147069	AY147069
Ostryopsis davidiana Decne.	n/a	AY263917	AF081515	AY147071	AY147071
Papuasicyos papuanus (Cogn.) Duyfjes	n/a	EU590121	EU590122	EU590123	EU590123
Parasicyos dieterleae Lira & R.Torres	n/a	DQ536712	DQ535763	DQ536846	DQ536846
Penelopeia suburceolata Urb.	n/a	DQ536713	DQ535834	DQ536847	DQ536847
Platycarya strobilacea Siebold & Zucc.	n/a	AY147100	AY263933	AY147078	AY147078
Polyclathra cucumerina Bertol.	n/a	DQ536717	DQ535767	DQ536849	DQ536849
Pterocarya hupehensis Skan	n/a	AY263918	AY263934	AY147079	AY147079
Quercus multinervis (W.C.Cheng & T.Hong) J.Q.Li	n/a	AY263922	AY263938	AY147084	AY147084
Rhoiptelea chiliantha Diels & Hand.-Mazz.	n/a	U92852	AF017687	AY147081	AY147081
Schizopepon bryoniifolius Maxim.	n/a	AY968456	AY973025	AY973009	AY973009
Seyrigia humbertii Keraudren	n/a	AY968457	AY968526	AY973010	AY973010
Siraitia grosvenorii (Swingle) C.Jeffrey ex A.M.Lu & Zhi Y.Zhang	n/a	DQ536736	DQ535850	DQ536869	DQ536869
Telfairia pedata Hook.	n/a	DQ491021	DQ535853	DQ501271	DQ501271

Tetrameles nudiflora R.Br.	n/a	AY968458	AF206828	AY091831	AY091831
Thladiantha hookeri C.B.Clarke	n/a	DQ491022	DQ535734	DQ536780	DQ536780
Ticodendron incognitum Gómez-Laur. & L.D.Gómez	n/a	U92855	AF061197	AY147073	AY147073
Trichosanthes amara L.	n/a	EU037001	EU037000	DQ536873	EU037004
Trigonobalanus verticillata Forman	n/a	U92866	AJ235812	AY147085	AY147085
Vaseyanthus insularis Rose	n/a	DQ536748	DQ535776	DQ536880	DQ536880
Xerosicyos danguyi Humbert	n/a	AY968459	AY973026	AY968573	AY968388
Zehneria bodinieri (H.Lév.) W.J.de Wilde & Duyfjes	n/a	DQ536754	DQ535865	DQ536885	DQ536885
Zombitsia lucorum Keraudren	n/a	DQ491024	DQ501260	DQ501273	DQ501273

Appendix 4. DNA sequence data and reference alignments.

Files on the accompanying DVD comprise DNA sequence data and sequence alignments in FASTA format.

File	Contents
1 Chapter2 Reference alignment 115taxa cpDNA.fas	Alignment of concatenated *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L sequences for 115 taxa
2 Chapter2 Reference alignment 89 taxa ITS.fas	Alignment of ITS sequences for 89 taxa
3 Chapter3 Reference alignment 92 taxa cpDNA.fas	Alignment of concatenated *mat*K, *rbc*L, *trn*L intron and *trn*L-F spacer sequences for 92 taxa (Cucurbitales-Fagales dataset)
4 Chapter3 Reference alignment 110 taxa cpDNA.fas	Alignment of concatenated *ndh*A intron, *ndh*F-*rpl*32, *rpl*32-*trn*L sequences for 110 taxa (Begoniaceae dataset)
5 ITS data.fas	Unaligned ITS sequences for 65 taxa
6 *ndh*A intron data.fas	Unaligned *ndh*A sequences for 116 taxa
7 *ndh*F-*rpl*32 spacer data.fas	Unaligned *ndh*F-*rpl*32 spacer sequences for 116 taxa
8 *rbc*L data.fas	Unaligned *rbc*L sequences for 23 taxa
9 *rpl*32-*trn*L spacer data.fas	Unaligned *rpl*32-*trn*L spacer sequences for 116 taxa
10 *trn*L-F data.fas	Unaligned *trn*L-F sequences for 23 taxa

The positions of highly variable regions and inversions which were excluded from the reference alignments prior to the analysis are indicated in Tables 2.4 and 3.2.