CHAPTER 1

GENERAL INTRODUCTION

The general goal of this PhD research was to study the phylogeny and systematics of *Macaranga*, *Mallotus* and other genera in the subfamily Acalyphoideae s.s. of the angiosperm family Euphorbiaceae s.s. In this introduction, the Euphorbiaceae and the genera studied are briefly reviewed, and the phylogenetic positions of these genera are discussed in the light of the recent molecular phylogenetic discoveries. The methods used are discussed, and the research questions are posed. Finally, an outline of the thesis is given.

EUPHORBIACEAE

The family Euphorbiaceae, as traditionally delimited (Euphorbiaceae s.l., Webster, 1994a), has c. 330 genera and c. 8000–9000 species, and is, therefore, one of the largest angiosperm families. Some of its genera are extremely species rich, and can be truly counted among the angiosperm 'giant genera' (e.g., *Euphorbia* L. > 2000 spp.; *Croton* L. > 800 spp.). The family has a pantropical to temperate distribution and displays a mind-boggling range of morphological diversity. The growth forms vary from small annual herbs and geophytes to shrubs, climbers, lianas and tall rain forest canopy trees (Webster, 1987). Ecological adaptions in the family are also many — there is even an aquatic species, *Phyllanthus fluitans* Benth. Generally the family has adapted to fragmented, shifting, and extreme niches (e.g., disturbed forest habitats and deserts; Webster, 1994a). Also, several economically important species can be found in the Euphorbiaceae (e.g., cassava, *Manihot esculenta* Crantz; the commercial rubber tree, *Hevea brasiliensis* (Willd. ex A.Juss.) Müll.Arg; and the castor oil plant, *Ricinus communis* L.).

  The classification of Euphorbiaceae has a complex history (reviewed by Webster, 1987). The latest complete system in the family was devised by Webster (1975; 1994b) and subsequently slightly modified by Radcliffe-Smith (2001). This system was strongly influenced by pollen characters and divides the family into five subfamilies. Three of the subfamilies were characterized by one ovule per locule (uniovulates: Acalyphoideae, Crotonoideae and Euphorbioideae) and two of them by two ovules per locule (biovulates: Phyllanthoideae and Oldfieldioideae). As is no surprise for a large and diverse family like Euphorbiaceae, there have been several proposals to redefine the boundaries of the family and to recognize a varying number of segregate families (see Radcliffe-Smith, 1987; Webster, 1987). Notably, the separation of biovulate subfamilies from uniovulate Euphorbiaceae s.s. was suggested based on, among others, seed anatomical differences (Corner, 1976; Meeuse, 1990). Lately, molecular phylogenetic studies have

accumulated evidence showing the non-monophyly of Euphorbiaceae s.l. (e.g., Chase et al., 1993; Savolainen et al., 2000; Soltis et al., 2000; Chase et al., 2002; Wurdack, 2002; Soltis et al., 2003; Davis & Chase, 2004). These findings were reflected in the original APG classification of angiosperms, where three euphorbiaceous families were recognized in the order Malpighiales of the eurosid I clade (APG, 1998). Finally, in the APG-II system (APG II, 2003) the Euphorbiaceae s.l. was divided into 2 uniovulate families (Euphorbiaceae s.s. and a small segregate family Pandaceae) and 3 biovulate families (Putranjivaceae from subf. Phyllantoideae tribe Drypeteae, Phyllanthaceae from the rest of subf. Phyllanthoideae, and Picrodendraceae from subf. Oldfieldioideae). Further studies with wider taxon sampling (Wurdack et al., 2004; Kathriarachchi et al., 2005; Samuel et al., 2005; Wurdack et al., 2005) have recovered these families as monophyletic with a few aberrant taxa excluded. However, no studies so far have resulted in a strongly supported hypothesis about relationships among these families or between them and other families in the order Malpighiales (Davis et al., 2005; Tokuoka and Tobe, 2006; Davis et al., 2007).

THE PHYLOGENY OF SUBFAMILY ACALYPHOIDEAE AND TRIBE ACALYPHEAE

Wurdack (2002) and Wurdack et al. (2005; see also Tokuoka & Tobe, 2006; Tokuoka, 2007) studied the phylogeny of Euphorbiaceae s.s. using *rbcL* and *trnL-F* sequence data and a sample of c. 70% of the genera. Although their partly unsupported results prevented them from proposing a new classification for the entire family, several important conclusions were drawn (see Fig. 1.1). First, few genera previously assigned to the subfamily Acalyphoideae were found to be successive sisters to the rest of Euphorbiaceae s.s. These taxa in this 'acalyphoid grade' were subsequently excluded to make the subfamily monophyletic. Second, the subfamily Crotonoideae was also found to be non-monophyletic. Furthermore, the correspondence between the phylogenetic results and the tribal level of the classification (Webster, 1994b) was relatively low, as 14 of the 35 tribes included were either para- or polyphyletic (Wurdack et al., 2005).

The original research plan for this thesis was to study the phylogeny of the tribe Acalypheae (subfamily Acalyphoideae), a group with several ecologically, economically and evolutionary significant genera (e.g., *Acalypha* L., *Macaranga* Thouars, *Mallotus* Lour. and *Ricinus* L.). First, a skeleton phylogeny of the tribe was to be constructed, and then the most important clade found was to be selected for a lower level study. However, the results of the above-mentioned study in Euphorbiaceae s.s. phylogeny made it clear that this plan was not feasible as PhD project, because the 20 originally sampled Acalypheae genera were scattered in seven different clades (Fig 1.1; see also Fig. 3 in Wurdack et al., 2005). Studying the phylogeny of this tribe would have thus not been possible without sampling practically the whole subfamily Acalyphoideae (with c. 120 genera!) as well. Therefore, the decision was made to abandon the first part of the work plan and focus the research primarily on one of the clades consisting mainly of Acalypheae genera: the A1 clade (sensu Wurdack et al., 2005) in the 'core acalyphoid' group.

Fig. 1.1. The phylogenetic relationships in the Euphorbiaceae s.s., with an emphasis on the subfamily *Acalyphoideae* s.s. The results from a parsimony reanalysis of the Acalyphoideae s.s. clade (*rbcL* and *trnL-F* data from Wurdack et al. 2005, with few newly sequenced taxa, for details and methods see chapters 2 and 7) embedded in the main phylogenetic structure in the family (after Wurdack et al. 2005). The members of tribe Acalypheae are indicated with stars, and the genera studied in this thesis are printed in boldface. Bootstrap frequencies are given above branches.

A1 CLADE, SUBTRIBE ROTTLERINAE, *MACARANGA* AND *MALLOTUS*

Apart from the basal genera *Cleidion* Blume and *Blumeodendron* Kurz, the A1 clade has a highly supported subclade with genera assigned to subtribes Rottlerinae and Macaranginae (tribe Acalypheae; Webster, 1994b). This 'Rottlerinae clade' is the subject of phylogenetic and taxonomic studies presented in Chapters 2–5. From four Rottlerinae and Macaranginae genera sampled in the original study of Wurdack et al. (2005), three, *Macaranga*, *Mallotus* and *Trewia* L., are part of this clade, and only one, *Rockinghamia* Airy Shaw, was found not to be closely related. When the Rottlerinae clade was chosen as the subject of phylogenetic studies in this thesis, a pilot study (see Chapter 2) was conducted to investigate if the five unsampled Rottlerinae genera are members of the clade as well. Analyzing the *rbcL* and/or *trnL-F* sequences of these missing genera together with the dataset of Wurdack et al. (2005) revealed that they are indeed part of this clade (Fig. 1.1).

The bulk of the species in the Rottlerinae clade belong to two large genera, *Macaranga* (c. 260 spp.) and *Mallotus* (c. 150 spp.). The other genera, *Coccoceras* Miq., *Cordemoya* Baill., *Deuteromallotus* Pax & K. Hoffm., *Neotrewia* Pax & K. Hoffm., *Octospermum* Airy Shaw and *Trewia*, have only 1–4 species each (as circumscribed by Webster, 1994b). *Macaranga* and *Mallotus* are palaeotropical genera of shrubs, trees and occasionally woody climbers with a prominent ecological role in especially the secondary forests of South-East Asia (Keßler, 2000; Slik et al., 2003a; Eichhorn, 2006). Furthermore, *Macaranga* has attracted special research attention, because some of its species are myrmecophytic, i.e., live in obligatory symbiosis with ants. These myrmecophytic *Macaranga* species provide housing and food bodies for ants and receive protection from herbivores, competing climbers and fungal pathogens (e.g., Ridley, 1910; Fiala, 1996; Blattner et al., 2001; Davies et al., 2001). *Macaranga* and *Mallotus* are also biogeographically interesting: most species occur in an area from India through Malesia to Australia and southwest Pacific, but both genera have representatives in Africa and Madagascar as well.

Prior to the research presented here, the ecological role and value as forest disturbance indicators of both *Macaranga* and *Mallotus* have been studied at the Universiteit Leiden branch of Nationaal Herbarium Nederland (NHN-L; Slik et al., 2003a; Slik, 2005), simultaneously with taxonomic and phylogenetic studies, especially in *Mallotus* (Bollendorff et al., 2000; Slik et al., 2000; Slik & Van Welzen, 2001a, b; Van Welzen & Chayamarit, 2001; Van Welzen & Phattarahirankanok, 2001). The phylogenetic study of *Macaranga* and *Mallotus*, based on morphological characters (Slik & Van Welzen, 2001a), had an intriguing result: it suggested that the genus *Macaranga* is deeply nested in the genus *Mallotus*. Severe taxonomic rearrangements would thus have been needed to make the genera monophyletic, but the authors refrained from making formal changes because of the low support in the phylogenetic tree (Slik and Van Welzen, 2001a). Furthermore, the small related Rottlerinae genera, many of which clearly resemble *Mallotus* morphologically, were not included. Studies concentrating on the phylogenetic relationships in *Macaranga*, *Mallotus* and related genera were thus clearly needed.

The studies presented in this thesis were complemented by Soraya Sierra's PhD research (also at NHN-L). Her work concentrated on the genera *Mallotus* and *Cor-*

demoya Baill., providing taxonomic revisions in *Mallotus* (Sierra & Van Welzen, 2005; Sierra et al., 2005; Sierra & van Welzen, 2006; Sierra et al., 2007), a revision and new circumscription for *Cordemoya* (Sierra et al., 2006; based on the results presented in Chapter 2), and a morphological dataset for a joint study on the phylogeny of *Mallotus* (Chapter 5).

Apart from the studies concentrating on the clade of *Macaranga*, *Mallotus* and relatives, Malesian species of the genus *Cleidion* were revised (Chapter 6). This pan-tropical genus of c. 33 species also belongs to the A1 clade (Fig. 1.1), although only with moderate support. Furthermore, a study on the genus *Afrotrewia* was conducted (Chapter 7). This monotypic African genus was previously placed *incertae sedis* because of lack of extant specimens (Radcliffe-Smith, 2001). The collections recently discovered by Dr. F. Breteler (WAG) made it possible to determine the phylogenetic position of this genus. *Afrotrewia* was found to belong to the 'core acalyphoid' clade as well (Fig. 1.1).

METHODS

A phylogenetic approach to systematics was followed in this thesis: only monophyletic taxa — taxa encompassing all the descendants of a common ancestor — are recognized (Hennig, 1966; Judd et al., 1999). To create a phylogenetic hypothesis to evaluate the monophyly of taxa, three different methods (optimality criteria to choose the 'best' phylogenetic tree) were employed: maximum parsimony, Bayesian inference and maximum likelihood. In the maximum parsimony (cladistic) analyses, the most parsimonious solution, i.e., a tree requiring the least number of character changes, is preferred, and no specific evolutionary model is used (see e.g., Kitching et al., 1998). On the other hand, in the Bayesian inference of phylogeny and in the maximum likelihood method, an explicit model of character change is employed, and the competing evolutionary hypotheses can be ranked based on their statistical probabilities (e.g., Holder & Lewis, 2003).

For the phylogenetic analyses of this study, either DNA sequences only (Chapters 2 & 7) or a combination of sequence data and morphological and anatomical characters (Chapter 5) were used. Both plastid and nuclear markers were sequenced (with the exception of Chapter 7, which focused on higher taxonomic level). Employing genetic data from both plastid and nuclear genome, and in general from several non-linked loci, makes it possible to detect processes confounding the phylogenetic inference (e.g., hybridization, cpDNA introgression and lineage sorting; Wendel & Doyle, 1998). In addition to the nucleotide characters, the indel events were taken into account by coding them as absence/presence characters (Simmons & Ochoterena, 2000).

Compared to morphology, molecular data can easily provide a large number of characters whose homologies are often relatively easy to determine with confidence and are currently, therefore, ubiquitously used in phylogeny reconstruction. The limitations of morphological data were demonstrated in earlier phylogenetic study of *Mallotus* and *Macaranga* (Slik and van Welzen, 2001a), which suffered from low ratio of terminal taxa to characters and high number of polymorphic characters. However, the number of characters can be increased by also studying micromorphology and anatomy and by employing quantitative characters in addition to the commonly used qualitative characters (Chapter 5).

After inspecting the results of separate analyses of data from different sources for notable incongruences (see Materials & Methods of Chapter 2), combined analyses of all available data were conducted (total evidence approach; Kluge, 1989), allowing all characters to interact and contribute to the resulting phylogenetic tree. This approach has been shown to result in reduced analysis times and higher clade support when compared to analyzing datasets separately (Soltis et al., 1998).

Species concept—Taxonomists and evolutionary biologists working at species level are confronted with difficult questions: What is a species? How can one recognize it? Rivers of ink have been spilled while debating on these conceptual problems. Competing species concepts keep proliferating (Mayden, 1997; Hey, 2001). This 'species problem' stems partly from the conflict between two ideas: the idea of species as a category (species taxon) and the idea of species as an evolutionary entity. The species categories are distinct, but evolutionary entities in nature are not necessarily so. Therefore, whichever species concept is preferred, it is not always possible to equate species taxa with the evolutionary entities (Hey, 2001).

In the taxonomic revisions of this thesis, the species are recognized by morphological divergence. This 'morphological species concept' commonly used by taxonomists can be defined, for example, as follows: "Species are the smallest groups that are consistently and persistently distinct, and distinguishable by ordinary means" (Cronquist, 1988). Definitions like this (see also Van Steenis, 1957) tend to be rather loose, and, because each group of organisms represents a unique evolutionary case study and because human observers are limited and differ in their capacities for perception and judgment, a certain amount of subjectivity cannot be avoided in species recognition. Moreover, species devised and delimited in the process of taxonomic work should be seen as hypotheses of evolutionary entities. These hypotheses can always be further tested with new data (e.g., new collections) or with methods more directly examining the evolutionary processes, e.g., with population genetic studies (Hey et al., 2003).

RESEARCH QUESTIONS

The following main research questions are addressed in this thesis:

1) Are *Macaranga* and *Mallotus* monophyletic, or is *Macaranga* nested inside *Mallotus* as suggested by Slik & Van Welzen (2001a)?

2) What is the phylogenetic position of the small genera in subtribe *Rottlerinae* in relation to *Macaranga* and *Mallotus*?

3) What are the evolutionary relationships within *Macaranga* and *Mallotus*, and are the infrageneric groups in these genera monophyletic?

4) How are the *Macaranga* and *Mallotus* species occurring in Africa and Madagacar related to those in Asia and what kind of biogeographical scenario could explain the Afro-Asian distribution pattern?

5) What is the species composition of the genus *Cleidion* in Malesia?

6) What is the phylogenetic position of the *incertae sedis* genus *Afrotrewia*?

OUTLINE OF THE THESIS

In Chapter 2, the phylogeny of *Macaranga*, *Mallotus* and related genera was studied using sequence data from plastid (*trnL-F*) and nuclear (ITS, *ncpGS* and *phyC*) markers. The results show that *Macaranga* is a monophyletic genus and that the main clades inside the genus are rather homogenous geographically. On the other hand, the genus *Mallotus* proved to be paraphyletic for two reasons: firstly, *Mallotus* sections *Hancea* and *Oliganthae* form a separate clade together with the genera *Cordemoya* and *Deutero-mallotus*, and secondly, the genera *Coccoceras*, *Neotrewia*, *Octospermum* and *Trewia* are part of the main *Mallotus* clade. The phylogenetic relationships inside the main *Mallotus* clade were, however, poorly resolved and/or supported.

In Chapter 3, the main phylogenetic results from Chapter 2 are reflected in the classification by merging three small Asiatic genera, *Neotrewia*, *Octospermum* and *Trewia*, with *Mallotus* (the other necessary taxonomic rearrangements have been conducted in separate papers; Bollendorff et al., 2000; Sierra et al., 2006). A taxonomic treatment with descriptions and distribution maps is also given for the transferred species.

Chapter 4 accompanies the taxonomic changes made in the previous chapter by providing a proposal to conserve the name *Mallotus* against *Trewia*. This was done to prevent numerous name changes otherwise necessary, because *Trewia*, as an older name, has priority over *Mallotus*.

In Chapter 5 the phylogeny of newly circumscribed *Mallotus* (*Mallotus* s.s.) was studied in more detail. This was done by increasing taxon sampling significantly and creating a new dataset consisting of molecular (plastid *matK* and nuclear *gpd*), macromorphological and leaf anatomical characters. Also quantitative morphological and anatomical data were utilized in the phylogeny reconstruction. The currently used infrageneric classification was evaluated against the acquired phylogeny, and several new clades with morphological synapomorphies were identified.

Chapter 6 represents the taxonomic revision of the Malesian species in *Cleidion*, a genus rather closely related to the genera studied in the previous chapters. A full taxonomic treatment with keys, synonymies, descriptions, distribution maps and illustrations is given.

Chapter 7 concentrates on *Afrotrewia*, a previously insufficiently known genus of Euphorbiaceae from Africa. Based on newly found and collected material an amended description and neotypification were given for it. Also the pollen structure was studied and the *rbcL* and *trnL-F* sequences were analyzed to find its phylogenetic position in the family.