The Phylogeny and Evolution of Two Ancient Lineages of Aquatic Plants

by

William James Donaldson Iles

B.Sc., The University of Toronto, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2013

Abstract

In my thesis I aim to improve our phylogenetic and evolutionary knowledge of two ancient and distantly related groups of aquatic flowering plants, Hydatellaceae and Alismatales. While the phylogeny of monocots has received fairly intense scrutiny for two decades, some parts of its diversification have been less frequently investigated. One such lineage is the order Alismatales, which defines one of the deepest splits in monocot evolution. Many families of Alismatales are aquatic or semi-aquatic, and they have been implicated in historical discussions of monocot origins. I evaluate inter-familial relationships in the order, considering a suite of 17 plastid genes for 31 Alismatales taxa for all 13 recognized families. This study improves on our understanding of, and confidence in, higher-order Alismatales relationships. I also uncovered convergent gene loss of plastid-encoded subunits for the NADH dehydrogenase complex. I then expand monocot coverage outside Alismatales by including unpublished and newly sequenced data for other orders. This large-scale sample facilitated a re-evaluation of monocot phylogeny and molecular dating, the latter using 25 fossil constraints. Previously included in the monocot order Poales, Hydatellaceae are a small family of ephemeral aquatics relatively recently found to be the sister group of water lilies (Cabombaceae and Nymphaeaceae). I present the first molecular phylogeny of the family and evaluate aspects of the family's morphological evolution. I show how sexual system shifts are associated with shifts in other reproductive traits. I also infer a temporal scale for Hydatellaceae diversification using a two-step Bayesian approach. I use the resulting dated tree to address biogeographic patterns and aspects of niche evolution. I show that its "Gondwanan" distribution is the result of long-distance dispersal and not continental rifting, and demonstrate strong phylogenetic niche conser-

vatism in the family. These studies expand our understanding of evolution in Hydatellaceae, and provide a substantial update to our understanding of Alismatales (and more generally monocot) phylogeny and divergence times.

Preface

All chapters benefited from input from coauthors. A version of Chapter 2 was previously published as:

> Iles W.J.D., Rudall P.J., Sokoloff D.D., Remizowa M.V., Macfarlane T.D., Logacheva M.D., and Graham S.W. 2012. Hydatellaceae (Nymphaeales): sexual-system homoplasy and a new sectional classification. *American Journal of Botany* 99: 663–676.

I produced most of the sequence data, carried out analysis, and wrote the manuscript. P.J. Rudall, D.D. Sokoloff, M.V. Remizowa and T.D. Macfarlane collected and provided plant material. M.D. Logacheva produced a small amount of sequence data. D.D. Sokoloff and M.V. Remizowa did character scoring and measurements. D.D. Sokoloff produced the sectional treatment. S.W. Graham conceived of the initial project and provided guidance. All authors helped with writing and editing.

Chapter 3 was coauthored. I conceived of the initial project, carried out most of the analyses, and wrote most of the manuscript. D.D. Sokoloff, M.V. Remizowa, S.R. Yadav, M.D. Barrett, R.L. Barrett, and T.D. Macfarlane provided accession location data for georeferencing. C. Lee carried out georeferencing and Maxent analysis, and wrote an initial draft of the corresponding material and methods section. S.W. Graham helped with writing and provided guidance. All authors helped with editing.

A version of Chapter 4 is in press as:

> Iles W.J.D., Smith S.Y., and Graham S.W. *In press.* A well-supported phylogenetic framework for the monocot order Alismatales reveals multiple losses of the plastid NADH dehydrogenase complex and a strong long-branch effect. P. Wilkin and S.J. Mayo [eds]. In *Early Events in Monocot Evolution.* Cambridge:

Cambridge University Press. pp. 1-28.

I produced most of the new sequence data (19 taxa), carried out analysis, and wrote the manuscript. S.Y. Smith produced new sequence data for six taxa. S.W. Graham conceived of the project and provided guidance. All authors helped with editing.

Chapter 5 was coauthored. I conceived of the initial project, generated sequence data for 12 new taxa, carried out analysis, and wrote the manuscript. J. Zgurski generated most of the new sequence data in the order Liliales. J.M. Saarela generated most of the new sequence data in the order Poales. G. Ross and Q. Lin each generated data for one new taxon. S.Y. Smith provided advice on fossil selection. S.W. Graham helped with writing, editing, and guidance.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor Dr Sean Graham. Working with Sean has been very formative in my approach to scientific research and writing. I would also like to thank my committee members Drs Jeannette Whitton and Mary Berbee. They have proved invaluable in helping me mature as a scientist. I would like to single out special thanks to Mary for the extensive help she provided with my thesis in the months leading to my defence.

My fellow graduate students in the Graham lab have been crucial to my success. Hardeep Rai and Jeffrey Saarela shepherded me through lab training, and provided a wealth of information and help during my initial years in the lab. I have had many fruitful conversations with Ying Chang and Vivienne Lam.

I would like to thank my collaborators for data, material, and many insightful ideas: Matthew Barrett, Russell Barrett, Maria Logacheva, Terry Macfarlane, Margarita Remizowa, Paula Rudall, Selena Smith, Dmitry Sokoloff, and Shrirang Yadav. I would also like to highlight Chris Lee who in addition to being a good friend collaborated with me on Chapter 3.

Lynn Norman braved the Deep South with me one spring to collect *Aletris* samples from such dubious places as military lands. Many other researchers kindly contributed samples.

I have made several outstanding friends at UBC. I would like to highlight Heather Esson and Kyle Demes. I would also like to thank the following people in Botany and the Biodiversity Centre for making my years at UBC worthwhile: Greg Baute, Gina Choe, Rich FitzJohn, Jonathan Griffiths, Sarah McKim, Michael Scott, Chris Sears, Kathryn Turner, and Ross Whippo.

I thank my friend Marc Jopson for help with programming.

I appreciate the funding I have received over the years from my supervisor in the form of research assistantships and from UBC in the form of scholarships and teaching assistantships.

Finally my thesis would not have been possible without support from my family, including my sister Sarah Iles, and parents Diane Donaldson and Mike Iles.

Dedication

To my parents
Michael H. Iles & Diane E. Donaldson

Chapter 1

Introduction

> "'Monocotyledons are on the whole a decadent race'... *This view is obviously bound up with the assumption that the adoption of an aquatic life is a device by which a poorly equipped species may escape from the competition of its more favoured compeers, saving itself from extinction by retirement into a quiet back-water of existence... The present writer, having begun the study of aquatics ten years ago with a full conviction of the truth of this picturesque theory, has gradually and reluctantly been forced to the conclusion that there is no sound evidence in its favour."*
>
> —Agnes Arber,1920. *Water Plants: A Study of Aquatic Angiosperms: Chapter 26, The theory of the aquatic origin of monocotyledons.*

Aquatic angiosperms are distinctive and often ubiquitous inhabitants of freshwater lakes and streams, with some specialized groups found in inter- and subtidal regions (the seagrasses). Plants characterized as aquatic display a diversity of growth forms, from emergents found along water body margins whose morphology and anatomy are relatively unmodified compared to purely terrestrial relatives, to ephemeral aquatic or completely submerged plants that display a larger degree of specialization (Sculthorpe, 1967). Aquatic plants are phylogenetically well dispersed across the angiosperms, with at least 50 and perhaps more than 100 independent origins, although the fraction of species probably comprises <2% of all angiosperms (Cook, 1990; Les et al., 1997). There are strong morphological similarities seen across a broad array of aquatic lineages, and botanists have long recognized that these often represent strong convergences (Arber, 1920; Les et al., 1997). Perhaps the two major lineages most widely associated with

the aquatic habit are the water lilies (Cabombaceae and Nymphaeaceae) and relatives, and the monocotyledons (monocots hereafter), the main subjects of this thesis.

1.1 Water lilies and Hydatellaceae

The water lilies are nearly all either submergent or floating-leaved aquatics (Cook, 1990; Sculthorpe, 1967). Their low species diversity (∼64 species) belies their ubiquity in natural environments, charismatic appearance, and horticultural importance (Slocum and Robinson, 1996). The recognition of their divergence near the base of flowering plant phylogeny (e.g., Graham and Olmstead, 2000; Jansen et al., 2007) and morphological diversity has also promoted continued scientific interest in the group's anatomy and morphology (e.g., Schneider and Carlquist, 2009; Taylor et al., 2012; Vialette-Guiraud et al., 2011; Warner et al., 2008; Zhou and Fu, 2008), phylogeny (Borsch et al., 2007, 2011, 2008; Ito, 1987; Les et al., 1999; Löhne et al., 2007; Taylor, 2008), and the age of extant members (Löhne et al., 2008; Nixon, 2008; Yoo et al., 2005). The water lilies also have one of the oldest and most extensive fossil records of any modern group of angiosperms, with old and new world specimens known from the Lower to mid-Cretaceous onward (Friis et al., 2009, 2001; Gandolfo et al., 2004; Mohr et al., 2008; Taylor et al., 2008).

One of the biggest recent realignments in angiosperm classification was the removal of Hydatellaceae from Poales and its placement in Nymphaeales as the sister taxon to Cabombaceae and Nymphaeaceae (Figure 1.1; Saarela et al. 2007). The family had traditionally been placed close to Centrolepidaceae, a poalean family with which it shares a minute basal rosette of linear leaves, what appear to be pseudanthia of reduced flowers (but see Rudall et al., 2007; Sokoloff et al., 2009b), an annual life cycle, an ephemeral aquatic habitat, and a centre of diversity in Australia (e.g., Rudall et al., 2007; Sokoloff et al., 2008a). The seedling morphology of Hydatellaceae (embryos with a single cotyledon; Sokoloff et al. 2008b; Tillich et al. 2007), and the presence of P2c-type sieve element plastids (Behnke, 2000), which

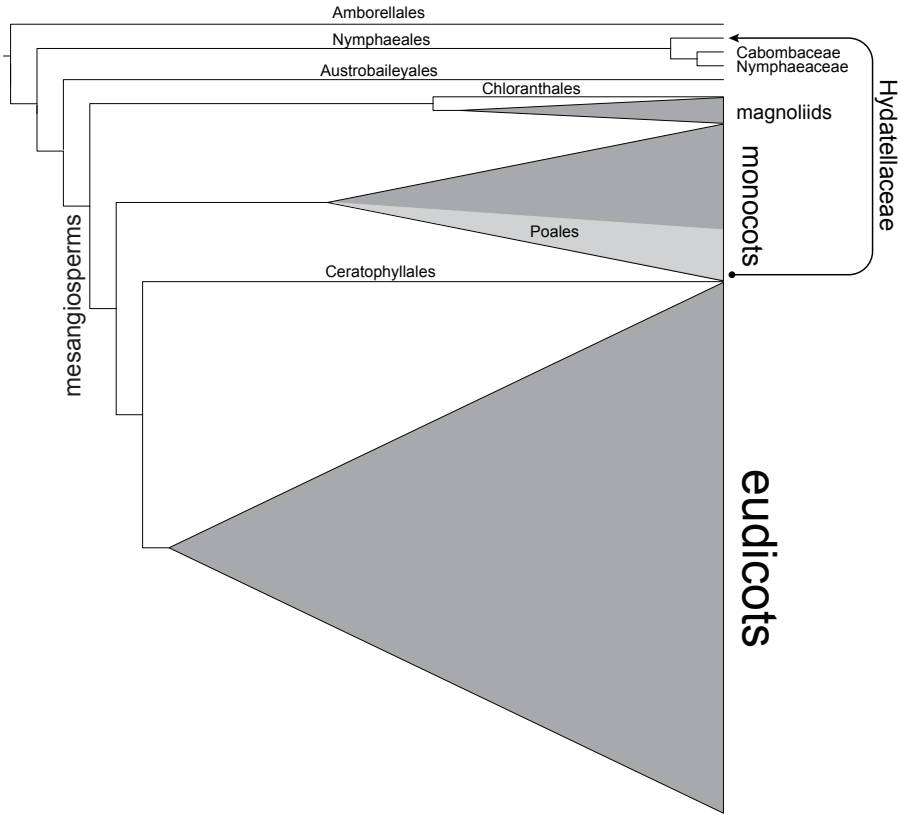


Figure 1.1: Summary of angiosperm relationships. The tree is based on molecular studies from the last decade or so (e.g., Graham and Olmstead, 2000; Jansen et al., 2007; Moore and Donoghue, 2007; Soltis et al., 2011). The placement of Hydatellaceae as the sister taxon to Cabombaceae and Nymphaeaceae in the Nymphaeales follows Saarela et al. (2007). Clade height is proportional to species richness, total diversity for angiosperms is estimated at 260 000 species (Stevens, 2001).

are found almost exclusively in monocots, were thought by some to confirm its placement in monocots (Tillich et al., 2007). However, a recent re-evaluation demonstrated that the sieve element plastids of Hydatellaceae are instead of the S-type, which is common to members of Nymphaeales (Tratt et al., 2009). The single cotyledon likely represents a convergent origin of monocotyledony with monocots (Sokoloff et al., 2008b). Despite their strong superficial resemblance to Centrolepidaceae, the reproductive anatomy of Hydatellaceae was recognized early on as being highly divergent from this family. For example, Dahlgren et al. (1985a) were uncertain about what their closest relations might be in their treatment of the monocots, placing Hydatellaceae in their own order within the Commeliniflorae. Several molecular studies that apparently confirmed its placement in Poales (see Figure 1.1) (e.g., Davis et al., 2004; Michelangeli et al., 2003) are instead inferred to be the result of a PCR artefact (Saarela et al., 2007). The new placement near the base of angiosperm phylogeny, as the sister group of water lilies (Saarela et al., 2007), has led to substantial recent research on the anatomy, morphology, and ecology of the family (e.g., Prychid et al., 2011; Remizowa et al., 2008; Rudall et al., 2007, 2008; Sokoloff et al., 2009b; Tratt et al., 2009), and informed a recent taxonomic monograph (Sokoloff et al., 2008a). However, there has been no attempt to derive a molecular phylogeny for Hydatellaceae, or to re-evaluate its taxonomy, biogeography, and evolutionary history in the light of phylogeny. This is the subject of two chapters of this thesis.

1.2 Alismatales and monocots

Unlike Nymphaeales, which are predominantly aquatic, monocots inhabit a wide variety of terrestrial and aquatic habitats. However, botanists working in Europe at the end of the nineteenth century first noticed that a higher percentage of monocots were aquatic compared with dicots, and that this pattern proved to be global rather than restricted to Europe (Henslow 1893; see Cook 1990; Les and Haynes 1995). They took this to indicate an aquatic origin for monocots as a whole (Henslow, 1893; Les and Schneider, 1995).

Other evidence interpreted as favouring the hypothesis of aquatic origin includes the absence in monocots of a vascular cambium. There were, however, many alternative theories on monocot origins and what might constitute ancestral characters, "primitive" taxa, and the closest dicot relatives. Closely connected with the aquatic-origin hypothesis for monocots was the notion that Nymphaeales and monocots share a common origin, although Les and Schneider (1995) reviewed extensively and rejected this idea. A variety of families has been proposed as defining the earliest branches of monocot phylogeny (for reviews see Dahlgren et al., 1985b; Kubitzki et al., 1998; Les and Schneider, 1995). Of particular relevance to current molecular findings (Chase et al., 1995, 2006, 2000; Fuse and Tamura, 2000; Graham et al., 2006; Tamura et al., 2004b) is that multiple taxa previously considered under Melanthiaceae or Nartheciaceae are now dispersed among many of the taxa that define the earliest splits in monocot phylogeny. For example, Thorne (1992) considered Melanthiaceae to be the "*... most archaic and least specialized monocots*," and Tamura (1998) considered Nartheciaceae to represent the "*... core stock of the monocots... several taxa including Alismatales and the Asparagales-Liliales-Commelinales complex successively branched off...* [but] *Nartheciaceae has little modified its characteristics.*" These two families have largely overlapping taxon memberships (as construed by these authors), comprising what are now referred to as Nartheciaceae (Dioscoreales), Petrosaviaceae (Petrosaviales), and Tofieldiaceae (Alismatales) in the APG system (note that Tamura 1998, recognized Melanthiaceae sensu stricto [Liliales] as distinct from the rest of this group). These taxa were highlighted for the insights they may give into the ancestral monocot form, with their equitant leaves, simple floral construction (trimerous and pentacyclic), and often apocarpous flowers considered by some to be indicative of their relative "primitiveness" (Remizowa et al., 2010; Tamura, 1998), although these speculations were not confirmed using formal ancestral-state reconstructions.

Although the sister group of the monocots and the placement of monocots among other mesangiosperms (the core clade of angiosperms see Figure 1.1; Cantino et al. 2007) remain contentious (e.g., Moore and Donoghue,

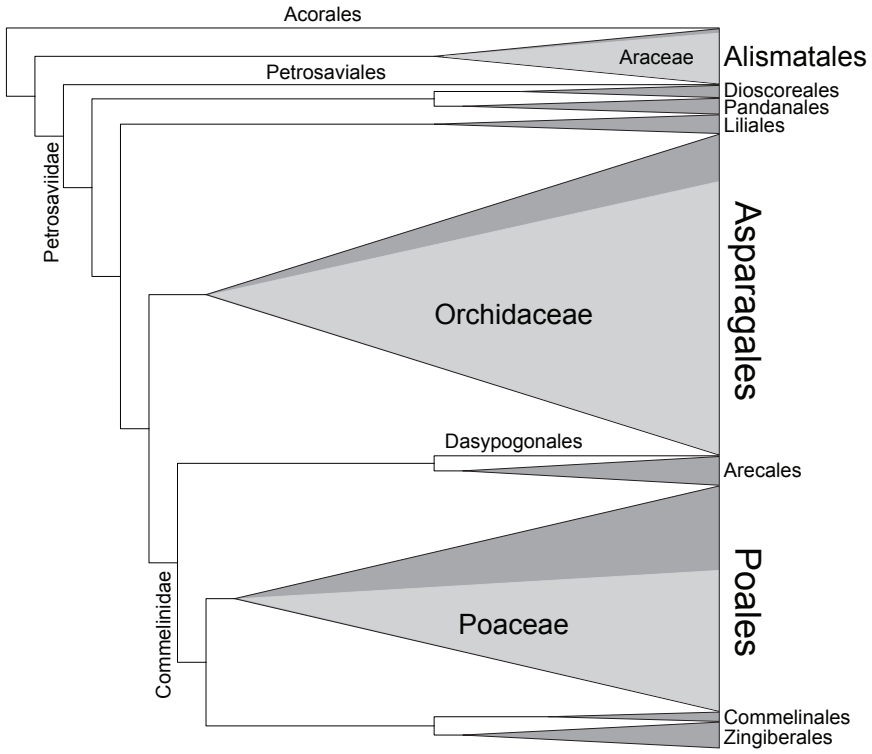


Figure 1.2: Summary of higher-order monocot relationships. The tree is based on two decades of molecular studies, including Chapter 5 and others (e.g., Chase et al., 1995, 2006, 2000; Duvall et al., 1993; Givnish et al., 2010; Graham et al., 2006). Relationships in the Commelinidae follow Givnish et al. (2010). Clade height is proportional species richness, total diversity for monocots is estimated at 58 000 species (Stevens, 2001). Note that several orders are monofamilial: Acorales = Acoraceae; Arecales = Arecaceae; Dasypogonales = Dasypogonaceae; Petrosaviales = Petrosaviaceae.

2007; Soltis et al., 2011), our current understanding of the first few splits in monocot phylogeny is potentially consistent with an aquatic or semi-aquatic origin for monocots, as groups diverging from basal nodes (*Acorus* and Alismatales) contain many aquatic and semi-aquatic members (Figure 1.2). The earliest splits in monocots also gave rise to lineages that were classified formerly as members of Melanthiaceae sensu Thorne (1992) and Nartheciaceae sensu Tamura (1998). These lineages include Tofieldiaceae in Alismatales, the order Petrosaviales, Nartheciaceae in Dioscoreales, and Campynemataceae and Melanthiaceae in Liliales (Figure 1.2). Members of these clades, especially Tofieldiaceae and Nartheciaceae, are often moisture loving and semi-aquatic. Comparatively little recent work has been done on higher-order relationships in Alismatales and the segregate families of Melanthiaceae sensu Thorne. The most thorough molecular study of the alismatids, for example, examined the plastid gene *rbc*L for 69 alismatid species, which covered all recognized 'core' families (Alismatales excluding Araceae and Tofieldiaceae), and 83% of alismatid genera (Les et al., 1997). Les et al.'s study constitutes a substantial advance in our understanding of alismatid relationships, although many of the inter-family relationships that were uncovered were relatively poorly supported. A few subsequent studies have examined mitochondrial evolution within the alismatids and relatives (Cuenca et al., 2010; Petersen et al., 2006b). These studies recovered similar relationships and similar levels of support as Les et al. (1997) and did not fundamentally change our understanding of the underlying phylogenetic relationships.

Key members of Melanthiaceae sensu Thorne have been more consistently sampled in large-scale phylogenetic studies (e.g., Chase et al., 1995, 2006, 2000; Fuse and Tamura, 2000; Graham et al., 2006; Tamura et al., 2004b) as well as in targeted family-level analyses: Melanthiaceae s.s. (e.g., Zomlefer et al., 2006); Nartheciaceae (e.g., Fuse et al., 2012; Merckx et al., 2008b); and Tofieldiaceae (e.g., Tamura et al., 2010, 2004a). Yet, despite these advances, multiple questions of family placement within their respective orders and intra-family relationships remain unresolved. The relative placement of Tofieldiaceae within Alismatales is unclear, for example; Gra-

ham et al. (2006) found Tofieldiaceae to be moderately supported as the sister group of Araceae and the alismatids, whereas Chase et al. (2006) found the family to be strongly supported as the sister group of the core alismatids only. In Nartheciaceae, the position of *Metanarthecium* is also unclear (Fuse et al., 2012; Merckx et al., 2008b).

In addition to understanding the sequence of phylogenetic diversification in monocots, it is also important to attempt to place monocot diversification in a temporal context. A time-scale of diversification would allow us to address macroevolutionary questions, including: (1) Are biogeographic hypotheses consistent with the timing of geological events such as tectonic rifting and fusion (Swenson et al., 2012)? (2) How were the patterns of diversification of different lineages related to one other (Schneider et al., 2004)? (3) How did lineages respond to past climate change (Crisp and Cook, 2007)? (4) What are (time-calibrated) rates of diversification and morphological evolution in the monocots (Ackerly, 2009; Magallón and Sanderson, 2001; Nee et al., 1994)? Monocots are substantially less well represented in the fossil record compared to other plant groups (Herendeen and Crane, 1995; Smith, 2013; Stockey, 2006); this is probably due to the low preservation potential of most monocot lineages, which are predominately insect-pollinated herbs (Herendeen and Crane, 1995). In the monocots (and many other groups with incomplete fossil records) most lineage and clade ages must be estimated by interpolating the lineages that left a fossil record. The idea of using branch lengths derived from protein or nucleotide data as proxies of relative age is an old one (Zuckerkandl and Pauling, 1962). Global clocks are often a poor fit when there is ubiquitous rate variation (Douzery et al., 2004). As rate variation is common in many groups, multiple methods to accommodate this have been developed (e.g., Drummond et al., 2006; Drummond and Suchard, 2010; Sanderson, 1997, 2002; Yang, 2006). Multiple fossil constraints are essential for calibrating local regions of the phylogeny and often result in improved fit (Sauquet et al., 2012). Despite these advances in molecular dating, there have been only a few serious attempts to date all the major lineages of monocots (Bremer, 2000; Givnish et al., 2011, 2005; Janssen and Bremer, 2004; Merckx et al., 2008a), and only in the two earli-

est papers was this the main goal. Bremer (2000) used a suite of eight fossils to date the major splits in monocot phylogeny, but subsequently one and possibly more of these were found to be erroneously placed (Crepet et al., 2004; Janssen and Bremer, 2004; Stockey, 2006). Nevertheless, the crown age for monocots inferred in this study were used either alone or in combination with other fossil constraints in subsequent analyses (Givnish et al., 2011, 2004; Janssen and Bremer, 2004; Merckx et al., 2008a). In my thesis, I revisit these questions with a more comprehensive set of molecular data and fossil calibration points.

1.3 Overview of the thesis

I address species-level relationships within Hydatellaceae in my second chapter. I use plastid and nuclear markers to sample all currently described species, in most cases sampling multiple populations per species, and use these two gene trees to construct a species-level phylogeny of the family. I then trace the evolution of several morphological characters, and use these to define new sections. I also evaluate the correlation between sexual system (dioecy vs cosexuality) and morphological characteristics of the reproductive structures using a phylogenetic ANOVA method (Garland et al., 1993).

In my third chapter I estimate the crown age of Hydatellaceae by using a recently developed method of Bayesian molecular dating (Drummond and Suchard, 2010), considering seven fossil derived constraints from across the seed plants. I use the resulting age posteriors for the two deepest splits in the family as priors in a multispecies-coalescent-based estimate of the species tree, to produce a dated family phylogeny. I use this time-calibrated phylogeny to examine questions of the family's "classical" Gondwanan distribution in India, Australia, and New Zealand (Gaikwad and Yadav, 2003; Yadav and Janarthanam, 1994, 1995), such as whether its current distribution across three land masses can be explained by ancient vicariance events, or whether it must represent relatively recent long-distance dispersals. I also consider how ecological preferences in the family may have evolved using climate preference profiles derived from Maxent (Phillips et al., 2006).

In my fourth chapter I sample representatives for each of the alismatid families for 17 plastid encoded genes, and add these to an existing monocot matrix (Graham et al., 2006; Saarela and Graham, 2010; Saarela et al., 2008, 2007). This permits inference of a robustly supported frame-work of alismatid relationships. I sample *Maundia triglochinoides* (Juncaginaceae) to evaluate its placement in alismatid phylogeny, following up on a recent report by von Mering and Kadereit (2010) who found that this genus was not part of a clade with other members of Juncaginaceae. I also characterize the convergent loss of a subset of genes in the plastid genome, which code for one of its major protein complexes (NADH dehydrogenase), and explore the potential for long branches in alismatids to mislead phylogenetic inference using plastid data.

In my fifth chapter I expand monocot sampling outside Alismatales to include Dioscoreales, Pandanales, and Liliales, by using newly sequenced and previously unpublished data from prior studies of a large plastid data set (Graham et al., 2006; Saarela, 2006; Saarela and Graham, 2010; Saarela et al., 2008, 2007; Zgurski, 2004). This resulted in a large matrix representing nearly all recognized autotrophic families of monocots (Angiosperm Phylogeny Group, 2009). I use this molecular data set to produce a comprehensive estimate of higher-order monocot relationships. I also review recent paleobotanical literature to justify the use of 17 monocot fossils (and eight additional angiosperm fossils) for dating the major splits in monocot phylogeny using the Bayesian random local clocks method (Drummond and Suchard, 2010). This provides a potentially improved framework for dating divergence times in monocots over earlier studies that used phylogenies based on a single gene and fewer fossil constraints (e.g., Bremer, 2000; Givnish et al., 2011, 2005; Janssen and Bremer, 2004).

Chapter 2

Phylogenetics and sexual-system evolution of Hydatellaceae

2.1 Summary

Species relationships are unknown in Hydatellaceae, a small family of dwarf aquatics related to water lilies that arose near the base of angiosperm phylogeny. Here I use molecular evidence to infer a species tree for the family, and apply this to reconstructing major transitions in morphology and sexual system in this early branch of angiosperms. I assembled plastid (*atp*B, *mat*K, *ndh*F, *rbc*L) and nuclear (ribosomal ITS) data for 50 samples (including outgroups), and estimated a species tree for Hydatellaceae using a Bayesian multispecies coalescent approach. I reconstructed the evolution of several morphological characters and tested for associations between sexual system and reproductive morphology using phylogenetic ANOVA. Dioecious species of Hydatellaceae have significantly greater stamen number and anther length than cosexual species, suggesting changes in male function. The perennial habit defining one subclade likely represents a reversion from annuality. Species relationships do not fall along traditional morphological divisions, but new sections proposed here are supported by fruit and seed synapomorphies. The earliest split in the family is reflected in geography and climate (i.e., tropical vs. subtropical/temperate clades). I found limited evidence of incongruence between plastid and nuclear trees, with one exception involving gene-tree non-monophyly for two close relatives (*Trithuria*

submersa, *T. bibracteata*). While the direction of sexual-system evolution is ambiguous, transitions are significantly associated with changes in involucral phyllome length and proxies of pollen production. I propose a new sectional circumscription based on fruit, seed and DNA evidence.

2.2 Introduction

Hydatellaceae are a small family of 12 aquatic species restricted to Australia, New Zealand and India, with a majority of species diversity in Australia (Sokoloff et al., 2008b; Yadav and Janarthanam, 1994). The family was traditionally placed in the monocots near or within another family of highly reduced plants, Centrolepidaceae (e.g., Gilg-Benedict, 1930; Hieronymus, 1888), and was most recently considered to be in the commelinid order Poales (Angiosperm Phylogeny Group, 1998, 2003) despite multiple morphological and anatomical features discordant with this placement (Dahlgren and Clifford, 1982; Dahlgren et al., 1985a; Hamann, 1976, 1998). Recent phylogenetic analyses of plastid, nuclear and morphological data instead support a placement of Hydatellaceae as the sister group of the water lilies, Nymphaeaceae and Cabombaceae (Saarela et al., 2007). Subsequent analyses confirmed this placement with plastid and mitochondrial data (Graham and Iles, 2009; Moore et al., 2011; Qiu et al., 2010; Soltis et al., 2011), and Hydatellaceae are now considered to be part of an expanded order Nymphaeales (Angiosperm Phylogeny Group, 2009; Rudall et al., 2007). Our new understanding of the phylogenetic placement of Hydatellaceae may have important consequences for understanding the early evolution of angiosperms (e.g., Rudall and Bateman, 2010; Sokoloff et al., 2008a). Following the discovery that Hydatellaceae belong in the basal grade of angiosperms, multiple investigations expanded our basic understanding of the anatomy, morphology and ecology of this family (Carlquist and Schneider, 2009; Friedman, 2008; Prychid et al., 2011; Remizowa et al., 2008; Rudall et al., 2007, 2009a, 2008, 2009b; Sokoloff et al., 2009a, 2008a,b, 2011, 2010; Taylor et al., 2010; Tillich et al., 2007; Tratt et al., 2009; Tuckett et al., 2010a,b). Perhaps one of the most striking cases of variation in Hydatellaceae concerns the

breadth of reproductive diversity found in it, including instances of dioecy, autogamy and apomixis (Sokoloff et al., 2008a; Taylor et al., 2010). A phylogenetic analysis of the family has not yet been performed, and as a result intrafamilial variation in anatomy, morphology, ecology and reproductive biology lacks proper evolutionary context.

The floral structures in the family, termed reproductive units (RUs) by Rudall et al. (2007), defy classical morphological interpretation. Traditionally each RU has been considered to represent a condensed inflorescence of simple unicarpellate and unistaminate perianthless flowers (Cooke, 1987; Doyle, 2008; Endress, 2008, 2010; Endress and Doyle, 2009; Hamann, 1976, 1998; Hieronymus, 1888; Hooker, 1860; Yadav and Janarthanam, 1994) as this solves a morphological problem in bisexual RUs where the anther is surrounded by pistils, an arrangement found almost nowhere else in angiosperms (but see Ambrose et al., 2006; Martínez and Ramos, 1989). However, recent developmental analysis has shown that a straightforward interpretation of RUs as either flowers or inflorescences may be problematic (Rudall et al., 2007, 2009b; Sokoloff et al., 2010). Following Sokoloff et al. (2008a), I use the term "cosexual" for species in which individual plants possess both staminate and pistillate organs (see Lloyd, 1980, 1982). Cosexual species of Hydatellaceae are composed either of hermaphrodite plants with the two kinds of unisexual RUs present on the same plant (i.e., individual plants with both pistillate and staminate RUs, a form of monoecy), or of hermaphrodite plants with only bisexual RUs. Dioecious species are composed of staminate and pistillate plants (i.e., individual plants have only one kind of unisexual RU). Two species that I provisionally code as cosexual, *Trithuria filamentosa* Rodway and especially *T. inconspicua* Cheesem., appear to be predominantly apomictic: stamens (usually in unisexual RUs on hermaphroditic individuals) are not always present or are very rare (Edgar, 1966; Pledge, 1974; Remizowa et al., 2008; Sokoloff et al., 2008a) and a majority of pollen is sterile in these two species (Remizowa et al., 2008), flowering and seed set frequently occurs while the RUs are submerged in *T. inconspicua* (Edgar, 1966; Pledge, 1974; Sokoloff et al., 2008b), and embryo development can take place without pollen tubes being present in *T.*

filamentosa (Rudall et al., 2008).

Early classifications grouped members of Hydatellaceae into two genera using two apparently co-varying characters. *Trithuria* Hook.f. was characterized as having bisexual RUs and three pericarp ribs (the latter possibly associated with dehiscence; Rudall et al., 2007), with *Hydatella* Diels defined by unisexual RUs and a lack of pericarp ribs (Cooke, 1987; Hamann, 1998). A recent revision by Sokoloff et al. (2008a) covered all the Australian and New Zealand species, only omitting *Trithuria konkanensis* Yadav & Janarthanam from India. They merged two pairs of species and described four new species. *Hydatella leptogyne* Diels was synonymized with *H. australis* Diels, as suggested by Cooke (1987). They also provided evidence that the staminate *H. dioica* D.A. Cooke belongs to the same species as the pistillate *T. occidentalis* Benth. Due to nomenclatural priority, this dioecious species should be called *T. occidentalis* (Sokoloff et al., 2008a). They also described four new species: the dioecious *T. austinensis* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall from south-western Australia, two additional dioecious species from northern Australia (*T. cookeana* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall and *T. polybracteata* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall), and one cosexual species also from northern Australia (*T. cowieana* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall). The new circumscription of *T. occidentalis* and the new species *T. cowieana* and *T. polybracteata* break the apparently co-varying characters of RU sexuality and pericarp rib presence that were used to distinguish *Trithuria* and *Hydatella* (Cooke, 1987; Hamann, 1998). Sokoloff et al. (2008a) therefore transferred all species in the family to *Trithuria*, which has nomenclatural priority.

Dioecy in angiosperms is phylogenetically widespread, with probably more than 100 distinct origins, but is also relatively rare, accounting for only ∼6% of all species (Charlesworth, 2002; Renner and Ricklefs, 1995). The correlates of dioecy and causes of the evolution of dioecy from cosexuality have garnered much theoretical attention, with the change usually considered to occur via gynodioecy or monoecy (e.g., Barrett, 2002, 2010; Charlesworth, 1999; Charnov, 1982; Renner and Ricklefs, 1995; Sakai and

Weller, 1999; Thomson and Brunet, 1990; Vamosi et al., 2003). The reverse process, the evolution of cosexuality from dioecy, has received far less theoretical attention (e.g., Delph, 2009; Wolf and Takebayashi, 2004) and there are fewer purported transitions (Case et al., 2008; Dorken et al., 2002; Obbard et al., 2006; Schaefer and Renner, 2010; Sytsma et al., 1991; Volz and Renner, 2008). The direction(s) and frequency of sexual-system change in Hydatellaceae are unclear. For example, potential non-homology of RUs with bisexual flowers prevents us from polarizing character-state changes in these structures (Rudall et al., 2007, 2009b). However, although they did not perform a formal phylogenetic analysis, Sokoloff et al. (2008a) suggested that the four dioecious species are not closely related, based on their morphological distinctiveness from each other. If multiple potentially independent switches in sexual system occurred, this would allow us to test hypotheses related to these evolutionary transitions.

Sexual systems in Hydatellaceae are currently inferred primarily from morphology alone, the evidence for which is reviewed here. Hydatellaceae have generally been considered to be abiotically pollinated by either wind or water (Gaikwad and Yadav, 2003; Hamann, 1998; Remizowa et al., 2008; Rudall et al., 2007), and recent pollination biology experiments on *Trithuria submersa* Hook.f. confirm wind- but not water-mediated pollination (Taylor et al., 2010). The four dioecious species appear to be closely related to separate cosexual species based on fruit and seed characters (Sokoloff et al., 2008a), and three of the four are only known from single populations (with the fourth known from at least six populations; Sokoloff et al. 2008a). These populations might therefore be considered to merely represent dioecious populations of species with locally varying sexual systems. However, Sokoloff et al. (2008a) listed quantitative and qualitative differences between dioecious species and their putative cosexual relatives, and so I consider these dioecious populations to represent four species based on current evidence. The degree of selfing has only been studied in two cosexual species: *T. submersa* (hermaphrodite plants with bisexual RUs) and *T. australis* (Diels) D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall (hermaphrodite plants with unisexual RUs). There is evidence of obligate

autogamy in *T. submersa* and in *T. australis* (Taylor, 2011; Taylor et al., 2010). Other cosexual species are often described as predominately selfing (*T. konkanensis*; Yadav et al., 2011) and some species or populations of species have sessile RUs, which is suggestive of obligate autogamy (*T. australis*, *T. lanterna* D.A. Cooke, *T. konkanensis* and *T. submersa*; Sokoloff et al. 2008a, 2010).

I sampled four plastid regions and the nuclear ribosomal ITS region for all 12 Hydatellaceae species for phylogenetic analysis, with multiple samples included per species where possible. I use standard methods to infer gene trees for plastid and ITS data, and apply a recently developed Bayesian method (Heled and Drummond, 2010) to reconcile the gene trees and thus estimate the species tree. I use the resulting phylogenetic evidence to assess species circumscriptions and to reconstruct the evolution of sexual systems among species. I infer the evolution of discrete morphological characters (mostly from Sokoloff et al., 2008a), using maximum likelihood (ML) and maximum parsimony (MP) ancestral-state reconstruction, targeting lifespan (perennial or annual), sexual system, RU sexuality, and a suite of fruit and seed characters thought by Sokoloff et al. (2008a) to be phylogenetically informative. I relate these transitions to a new sectional treatment, proposed here. Finally, I evaluate whether sexual system and RU sexuality are associated with various RU traits. Molecular dating estimates and a biogeographic analysis are the subject of Chapter 3.

2.3 Materials and methods

2.3.1 Taxonomic and genomic sampling

The complete plastid matrix consists of 50 samples (Appendix A) including nine outgroup taxa from Amborellales, Austrobaileyales, Nymphaeales and Mesangiospermae (Angiosperm Phylogeny Group, 2009; Cantino et al., 2007). All 12 species of Hydatellaceae were sampled, in most cases using multiple samples per species, for a total of 41 ingroup terminals. Approximately half of the Hydatellaceae samples are from herbarium specimens,

the remainder coming from recent field-collected material preserved in silica gel. Because of the general scarcity of dioecious species (Sokoloff et al., 2008a) and the poorly preserved nature of the staminate individuals in them, only pistillate individuals from *Trithuria cookeana*, *T. polybracteata*, and *T. occidentalis* were successfully sampled.

I sampled four partial plastid regions (*atp*B, *mat*K, *ndh*F and *rbc*L). From the nuclear genome I sampled the nuclear ribosomal internal transcribed spacers 1 and 2, including the intervening 5.8S ribosomal subunit (collectively referred to as 'ITS'), generally as a single amplicon. In some cases amplification either failed or I recovered fungal contaminants, which I confirmed by BLAST searches. As a result the total number of Hydatellaceae sampled in the complete nuclear matrix was reduced to 35 samples. For *Trithuria cookeana* it was necessary to substitute a different individual from the same collection that I used to produce the plastid sequences. One species, *T. occidentalis*, is missing from the nuclear matrix, despite many attempts to retrieve sequence data using different samples and primer pairs.

2.3.2 Extraction, amplification and sequencing

Genomic DNAs were extracted with the DNeasy Plant Mini Kit (Qiagen Inc., Mississauga, ON, Canada). Regions were amplified using the polymerase chain reaction (PCR) under conditions described in Graham and Olmstead (2000). In some cases Paq5000$^{\text{TM}}$(Agilent Technologies, Santa Clara, CA, USA) was used instead of *Taq* polymerase, in which case the manufacturer's suggested conditions were used. Cycle sequencing was carried out using BigDye 3.1 chemistry (Applied Biosystems, Foster City, CA, USA) using ∼1/26 reactions. Primers for amplification and sequencing were used from a wide variety of published and unpublished sources: *atp*B (Hoot et al., 1995), *mat*K (Hilu et al. 2003; Löhne et al. 2007; `http://www.kew.org/barcoding/update.html`), *ndh*F (Kim and Jansen, 1995; Olmstead and Sweere, 1994), *rbc*L (Terachi et al. 1987; Yamashita and Tamura 2000; G. Zurawski – DNAX, Inc.), ITS (Hsiao et al., 1994; White et al., 1990). For some samples complete coverage of all gene regions was not

possible. For example, in the case of *Trithuria occidentalis* only a short fragment of *mat*K was recovered. Source details are provided in Appendix A. Base calling and contig assembly were done using Sequencher version 4.2.2 (Gene Codes Corp., Ann Arbor, MI).

2.3.3 Alignment and gene-tree analyses

Finalized contigs were aligned manually using Se-Al v2.0a11 (Rambaut, 2002) according to previously published protocols (Graham and Olmstead, 2000; Graham et al., 2000). As the plastid regions represent one linkage group they were concatenated; their total aligned length is 4122 bp. The nuclear ITS region, which represents a separate linkage group, was not concatenated with the plastid regions as I found some evidence of conflict with the plastid data (see below). The total aligned length of ITS is 717 bp. Gaps were treated as missing data and indels were not coded. It was not possible to align outgroup ITS sequences to those of Hydatellaceae, and so this region was analyzed for ingroup taxa only.

I used ML (Felsenstein, 1973), MP (Fitch, 1971) and Bayesian Markov Chain Monte Carlo inference (BI; Yang and Rannala, 1997) methods to infer phylogenetic relationships. Two full sets of analyses were undertaken for the plastid dataset because the placement of the *Trithuria occidentalis* sample was variable. The recovered data for this taxon (a portion of *mat*K) appears to be sufficient to place it locally (if somewhat imprecisely) in the tree (see Results), behaviour that is not unexpected when a large fraction of data are missing in this way (Wiens and Morrill, 2011). Heuristic MP searches were carried out using PAUP* 4b10-ppc-macosx (Swofford, 2003). Characters and character-state changes were weighted equally, tree bisection-reconnection branch swapping was employed, and 100 random addition replicates were performed for each search. It was necessary to set a maximum number of saved trees to 10 000. For model-based inference methods I used the Akaike information criterion corrected for finite sample sizes (AICc; Akaike, 1973; Hurvich and Tsai, 1989; Sugiura, 1978), as implemented in jModelTest version 0.1.1 (Guindon and Gascuel, 2003; Posada,

2008), to evaluate alternative DNA substitution schemes. I did not consider a separate parameter accounting for invariable sites because low substitution rates may be accommodated by the Γ parameter (Yang, 2006). Maximum likelihood inference was carried out using Garli 1.0 Univ OSX (Zwickl, 2006), considering the models preferred by AICc: GTR+Γ_4 for the plastid matrix and TIM3+Γ_4 (a nested model of GTR+Γ) for the nuclear matrix. Model parameters were estimated from the data. At least three random ML searches were initiated per data set. Branch support for MP and ML analyses was evaluated by using the non-parametric bootstrap (BS; Felsenstein, 1985a). For MP I ran 1000 BS replicates, with the same search parameters as above, except that I limited random addition replicates to 10 and the maximum number of saved trees to 100. For ML, 500 BS replicates were run with the same search parameters as above, except that only one random ML search was initiated per BS replicate. Following Graham et al. (1998) I consider BS values <70% to be "weak," 70–89% to be "moderate," and ≥90% to be "strong."

I used MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003) to carry out BI. Because Bayesian methods appear robust to slight over-parameterization (Lemmon and Moriarty, 2004), I used the GTR+Γ_4 model for both the plastid and nuclear datasets. All other model settings were set to default. Four independent runs of four chains each were run in two pairs for 1.0×10^6 generations, sampling every 100 generations for each locus; convergence within each pair of runs was evaluated by considering the average standard deviation of split frequencies, with values below 1.0×10^{-2} taken to indicate convergence. Before combining all four runs with LogCombiner (Drummond and Rambaut, 2007), 10% of each run was discarded as burn-in, and *ln*L values were compared for consistent plateau values. For the combined runs estimated sample sizes (ESS) was calculated using Tracer 1.5 (Rambaut and Drummond, 2009) for the likelihood, posterior and prior; values over 200 were taken to indicate convergence. I consider PP values <95% to be "weak," 95–98% to be "moderate," and ≥99% to be "strong" (see Alfaro and Holder 2006, for a review on the interpretation of PP values).

2.3.4 Species-tree inference

In order to estimate the species tree I used a Bayesian multispecies coalescent approach implemented in BEAST called *BEAST (Drummond and Rambaut, 2007; Heled and Drummond, 2010). This method requires that species correspond to discretely evolving populations, and assumes that the applied species delimitations are accurate. I provisionally treat *Trithuria submersa* as comprising eastern and western "species" because the wide geographic separation (>1500 km; estimated from Sokoloff et al. 2008a) between these two groups of populations may make contemporary gene flow unlikely. I trimmed the samples present in the plastid dataset to mirror those available in the nuclear dataset (with the exception of *T. occidentalis*, which was retained in the analysis despite its absence in the nuclear dataset), for a total of 36 samples across 13 effective species. A clade comprising *T. cookeana*, *T. cowieana*, *T. konkanensis*, *T. lanterna* and *T. polybracteata* was set as an outgroup for the plastid dataset here, matching the ingroup rooting observed in the plastid gene tree (Figure 2.1), which includes outgroup taxa. Both loci were assigned GTR+Γ_4 substitution models and uncorrelated lognormal clocks. Because clock rates need set priors in the absence of fossil calibrations, the plastid locus was assigned a mean clock rate of 1.0 and the nuclear locus was estimated relative to it. A Yule prior was used for the speciation process and the population size model was set to piecewise linear and constant root. The plastid "ploidy type" was set to "mitochondrial" (as both organelles are haploid and generally uniparentally inherited) while the nuclear "ploidy type" was set to "autosomal nuclear." Priors were left to default cases or changed to the recommended distribution if it was not automatically implemented (Drummond and Rambaut 2007; `http://code.google.com/p/beast-mcmc/wiki/ParameterPriors`). Five independent trials were run for 1.0×10^7 generations, sampling every 1.0×10^3 generations. The first 10% of each trial was burned in and convergence was then checked by examining trace files for consistent plateau values for parameters and ESS above 200 in Tracer 1.5 (Rambaut and Drummond, 2009). Post burn-in trees were combined with LogCombiner 1.6.1 (Drum-

mond and Rambaut, 2007) and summarized as a maximum clade credibility tree, with node heights set to mean values using TreeAnnotator 1.6.1 (Drummond and Rambaut, 2007).

2.3.5 Ancestral character states and character associations

Sokoloff et al. (2008a) compiled a list of characters of potential taxonomic and evolutionary interest for the family. I completed this dataset by filling in character states for *Trithuria konkanensis* from the literature (Gaikwad and Yadav, 2003; Sokoloff et al., 2010; Yadav and Janarthanam, 1994), and by filling in missing character states of other species for: seed cuticle (Rudall et al., 2008), pericarp epicuticular wax (Tillich et al. 2007; D.D. Sokoloff et al., unpublished data), pericarp anatomy (D.D. Sokoloff et al., unpublished data), and leaf auricles (Rudall et al., 2007). The character matrix consists of 11 discrete and five quantitative characters. Two of the quantitative characters, maximum involucral phyllome number and maximum involucral phyllome length, have differing character states for unisexual RUs (in species that have these), and so in each case the character was also tested separately for staminate and pistillate RUs, comparing each to bisexual RUs. The eastern and western populations of *T. submersa* were also scored identically, based on direct investigation of both eastern and western members of this species.

Maximum parsimony and ML ancestral character-state reconstructions (Lewis, 2001; Schluter et al., 1997) were performed for discrete characters using Mesquite 2.6 (Maddison and Maddison, 2009). The Bayesian species tree was used as the reference phylogeny for both sets of analyses. For the MP analyses, the default settings in Mesquite were used (i.e., unordered states). For the ML analyses single parameter models (Mk1) were compared to two-parameter models (AsymmMk) using the likelihood-ratio test; the single parameter models were preferred in all cases. I also evaluated associations of several quantitative characters to sexual system (dioecy vs. cosexuality) and RU sexuality (unisexual vs. bisexual RUs) using phylogenetic ANOVA (Garland et al., 1993) as implemented in the package 'geiger' (Harmon et al.,

2008) for R (R Development Core Team, 2011). The phylogenetic ANOVA corrects for species non-independence by comparing the ANOVA F-statistic to an empirical F-statistic distribution drawn by simulating the evolution of the dependent character many times on the reference tree (only the quantitative character is simulated and resampled). The simulation uses a Brownian motion model of evolution with parameter values taken from the estimate based on the original distribution. The original quantitative data were *ln*-transformed and tested for normality and homogeneity of variance using the Shapiro-Wilk test and Bartlett's test, respectively. In one case (maximum involucral phyllome number in male or bisexual RUs) the *ln*-transformed data did not meet assumptions of normality or homogeneity of variance, and so I used the rank transformation to meet these requirements. The default number of simulations (10^3) was not always adequate to estimate very small phylogenetic P-values, and for those tests the number of simulations was increased to 10^4 or 10^5.

2.4 Results

2.4.1 Plastid and nuclear phylogenies of Hydatellaceae

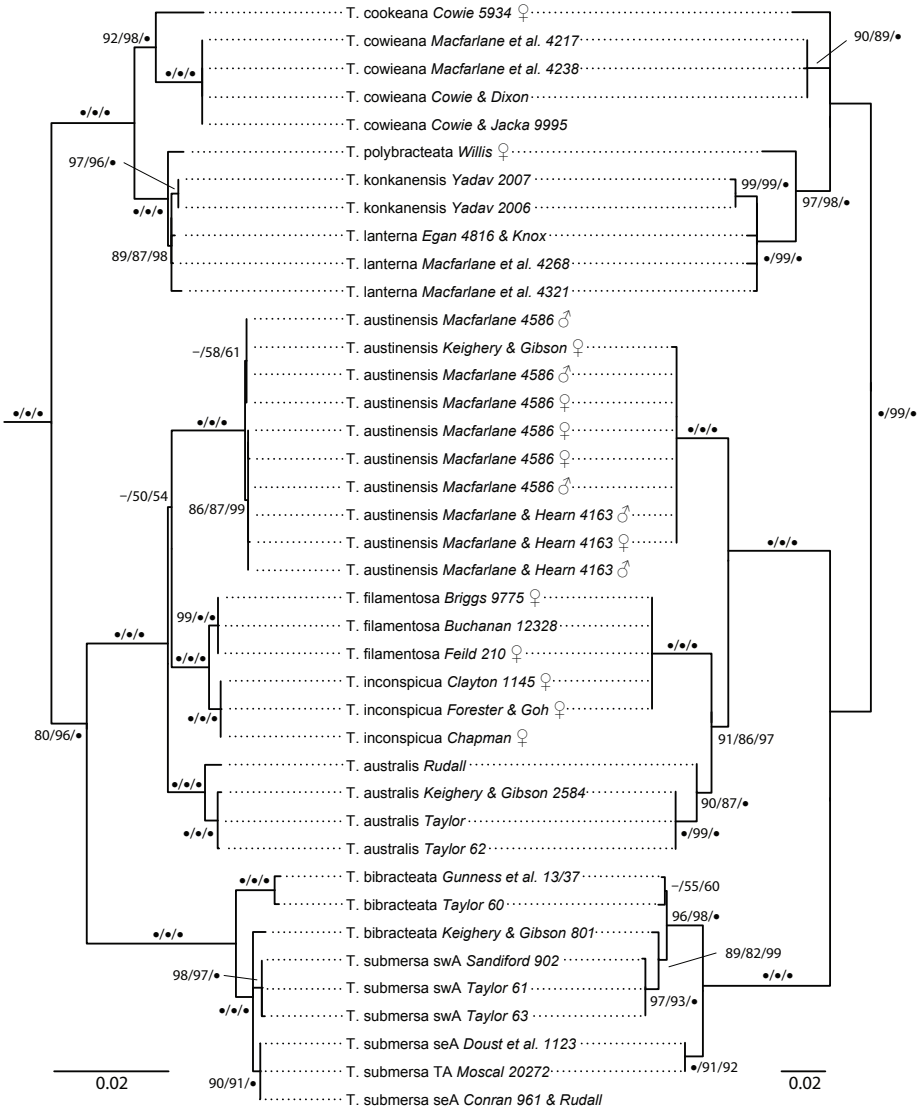
The plastid and nuclear phylogenies (Figure 2.1) are generally congruent with each other, with most branches resolved, generally with moderate to high branch support. In the plastid tree Hydatellaceae are the sister group to Cabombaceae and Nymphaeaceae (not shown), and Hydatellaceae are divided into a tropical clade (*Trithuria cookeana*, *T. cowieana*, *T. konkanensis*, *T. lanterna*, and *T. polybracteata*) and a temperate/subtropical clade containing the remaining species, with moderate to strong support for both clades (Figure 2.1). The most uncertain parts of the trees include a polytomy comprising the *T. lanterna* samples plus a small *T. konkanensis* clade, and a weakly supported conflict between the plastid and nuclear phylogenies concerning the relative positions of *T. austinensis*, *T. australis* and (*T. filamentosa* + *T. inconspicua*) (Figure 2.1). A more serious conflict concerns the reciprocal non-monophyly shown by *T. bibracteata* Stapf ex D.A. Cooke

and *T. submersa*, with the alternative non-monophyletic arrangements well supported (Figure 2.1). The placement of the short *T. occidentalis* sequence in the plastid phylogeny is somewhat ambiguous (Figure 2.2), with the second best ML placement being nearly identical in log-likelihood to the optimal placement ($\Delta = 1.688 \times 10^{-5}$). *Trithuria occidentalis* appears to depress local branch support values when it is included in the analysis (cf. Figures 2.1 and 2.2).

2.4.2 Species phylogeny of Hydatellaceae

The Bayesian multispecies coalescent tree (Figure 2.3A) is in most respects very similar to the individual plastid and nuclear trees (Figures 2.1 and 2.2), with all species falling into four well-supported clades. *Trithuria cookeana* is the sister taxon of *T. cowieana*, despite this arrangement only having strong support in the plastid tree. Within the temperate/subtropical clades *T. austinensis* appears to be weakly supported as a sister group to *T. filamentosa*, *T. inconspicua* and *T. australis*. *Trithuria occidentalis* was recovered as the sister group to *T. bibracteata* and the two geographic *T. submersa* segregates; however there is poor support for any set of relationships among the latter four taxa.

Figure 2.1 *(following page)*: Maximum likelihood phylograms of Hydatellaceae. Taxon names are followed by collection information and the gender of the sample (when the taxon is dioecious). Numbers adjacent to branches represent maximum parsimony bootstrap support values, maximum likelihood bootstrap values, and posterior probabilities (expressed as percentages), in that order (filled circle = 100, dash = <50). Scale bars represent expected substitutions per site. Left: concatenated plastid (*atp*B, *mat*K, *ndh*F, and *rbc*L) tree, excluding *Trithuria occidentalis* (Note: outgroups were included in this analysis, but are not shown here). Right: nuclear (ITS) tree.


T. cookeana Cowie 5934 ♀
T. cowieana Macfarlane et al. 4217
T. cowieana Macfarlane et al. 4238
T. cowieana Cowie & Dixon
T. cowieana Cowie & Jacka 9995
T. polybracteata Willis ♀
T. konkanensis Yadav 2007
T. konkanensis Yadav 2006
T. lanterna Egan 4816 & Knox
T. lanterna Macfarlane et al. 4268
T. lanterna Macfarlane et al. 4321
T. austinensis Macfarlane 4586 ♂
T. austinensis Keighery & Gibson ♀
T. austinensis Macfarlane 4586 ♂
T. austinensis Macfarlane 4586 ♀
T. austinensis Macfarlane 4586 ♀
T. austinensis Macfarlane 4586 ♀
T. austinensis Macfarlane 4586 ♂
T. austinensis Macfarlane & Hearn 4163 ♂
T. austinensis Macfarlane & Hearn 4163 ♀
T. austinensis Macfarlane & Hearn 4163 ♂
T. filamentosa Briggs 9775 ♀
T. filamentosa Buchanan 12328
T. filamentosa Feild 210 ♀
T. inconspicua Clayton 1145 ♀
T. inconspicua Forester & Goh ♀
T. inconspicua Chapman ♀
T. australis Rudall
T. australis Keighery & Gibson 2584
T. australis Taylor
T. australis Taylor 62
T. bibracteata Gunness et al. 13/37
T. bibracteata Taylor 60
T. bibracteata Keighery & Gibson 801
T. submersa swA Sandiford 902
T. submersa swA Taylor 61
T. submersa swA Taylor 63
T. submersa seA Doust et al. 1123
T. submersa TA Moscal 20272
T. submersa seA Conran 961 & Rudall
92/98/•
90/89/•
•/•/•
97/96/•
99/99/•
97/98/•
89/87/98
•/99/•
–/58/61
•/99/•
–/50/54
86/87/99
99/•/•
80/96/•
91/86/97
90/87/•
•/99/•
–/55/60
96/98/•
89/82/99
98/97/•
97/93/•
•/91/92
90/91/•
0.02
0.02

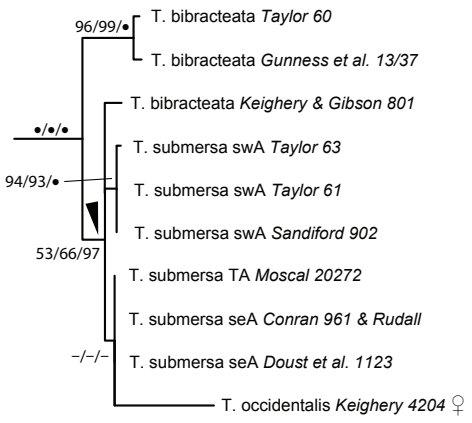


Figure 2.2: Portion of the plastid tree including *Trithuria occidentalis*. Arrowhead indicates possible placement of *T. occidentalis* with a nearly identical likelihood score to the shown optimal placement. For *T. submersa* samples: seA = south-east Australia. swA = south-west Australia.

2.4.3 Ancestral character-state reconstructions

Considering the ML reconstructions (Figure 2.4), five out of six of the fruit or seed characters (sculptured seed surface, thin seed cuticle, paired pericarp papillae, presence of pericarp epicuticular wax, and specialized endocarp cells) are potential non-homoplasious synapomorphies for various clades. These characters have unambiguous reconstructions for the most recent common ancestor of species possessing the relevant state, but some deeper nodes are equivocal, making it possible that some character states evolved slightly earlier on the tree (Figures 2.3 and 2.4). Maximum parsimony reconstructions of the same characters support them as unequivocal and non-homoplasious synapomorphies (Figures 2.3 and 2.4). It is also uncertain whether the presence of leaf auricles is a synapomorphy for the clade comprising *Trithuria filamentosa*, *T. inconspicua* and *T. australis*, or whether it originated at a deeper node, comparing ML and MP reconstructions (Figures 2.3 and 2.4). Also, preliminary observations of this character

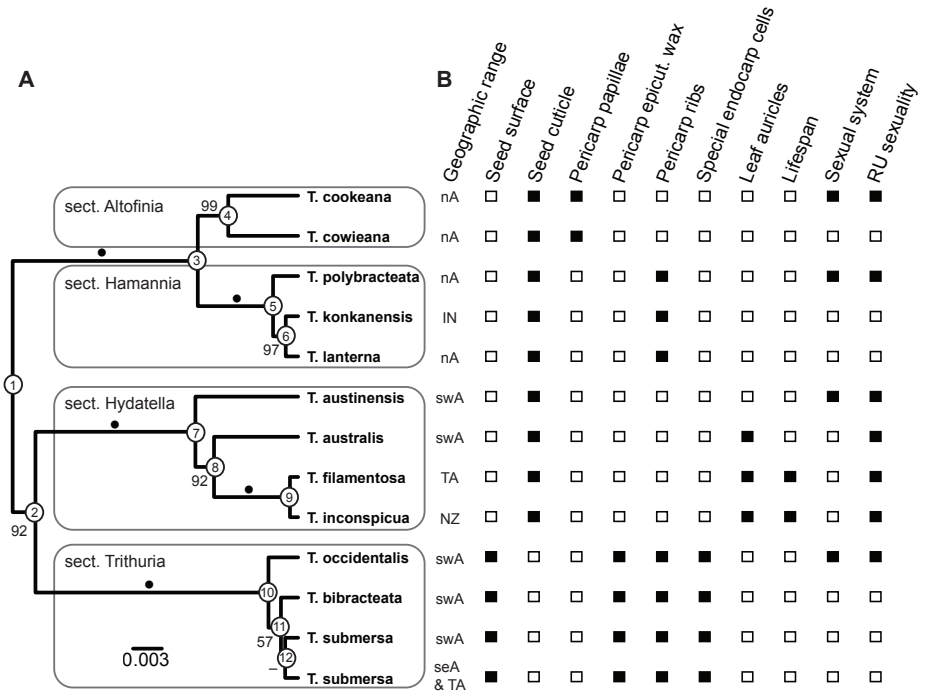


Figure 2.3: (A) Bayesian multispecies coalescent estimate of species phylogeny based on two linkage groups (plastid and nuclear). Numbers adjacent to branches represent posterior probabilities, expressed as percentages (filled circle = 100%). Numbered nodes correspond to numbers in Figure 2.4. Scale bar represents expected substitutions per site. I provisionally divided *Trithuria submersa* into two "species" based on geography (see text). Sectional names and composition are indicated. (B) Geographic ranges of species and selected character states. Geographic range: IN = India, nA = northern Australia, NZ = New Zealand, seA = south-east Australia, swA = south-west Australia, TA = Tasmania. Seed surface: black = strongly sculptured, white = smooth. Seed cuticle: black = thick, white = not thick. Paired pericarp papillae: black = present, white = absent. Pericarp epicuticular wax: black = present, white = absent. Pericarp ribs: black = present, white = absent. Specialized endocarp cells facilitating fruit dehiscence: black = present, white = absent. Leaf auricles: black = present, white = absent. Lifespan: black = perennial, white = annual. Species sexual system: black = dioecious, white = cosexual. RU sexuality: black = unisexual, white = bisexual.

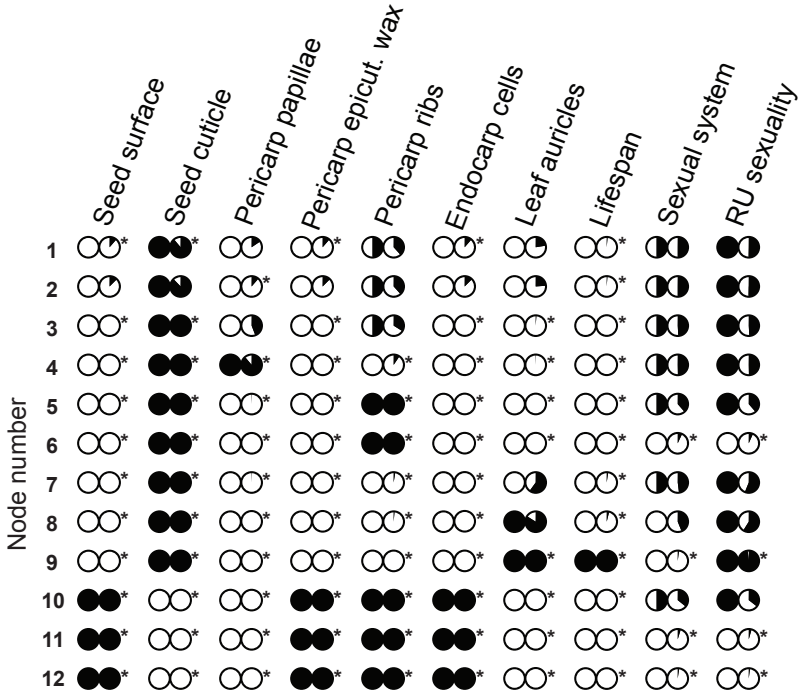


Figure 2.4: Reconstructions of ancestral character states. Characters correspond to those in Figure 2.3B, node numbers correspond to those in Figure 2.3A; character-state coding as for Figure 2.3. Left pie corresponds to MP reconstruction and right pie corresponds to ML reconstruction. For ML pies = proportional likelihoods, and an asterisk indicates a significant log likelihood difference (2.0 or more).

(not considered here) indicate that auricles are present in the very first one or two leaves of seedlings for in least one accession of *T. submersa*, and so this character needs to be investigated further. Perennial lifespan provides the only unequivocal and non-homoplasious synapomorphy in both MP and ML reconstructions (Figures 2.3 and 2.4). Three characters, pericarp ribs, sexual system and RU sexuality, are homoplasious and/or have ambiguous ML reconstructions over much of the tree (Figures 2.3 and 2.4).

2.4.4 Phylogenetic ANOVAs

Phylogenetic P-values linking morphology to sexual transitions were overall lower considering sexual system as a factor, than was the case for RU sexuality, and they were also significant in more cases. Four of the seven quantitative characters showed significant differences for in at least sexual system or RU sexuality (Table 2.1, Figure 2.5, note that data are shown untransformed). Stamen number and anther length were both larger in dioecious species and in species with unisexual RUs, compared to cosexual species or species with bisexual RUs. The longer maximum involucral phyllome length for male RUs in dioecious species compared to cosexual species was also statistically significant, as was the higher involucral phyllome number in female RUs in dioecious species compared to cosexual species. To assess whether significant results were due to the treatment of two *Trithuria submersa* "species" or the coding of the apomicts *T. inconspicua* and *T. filamentosa* as cosexuals, I simultaneously deleted one of the *T. submersa* segregates and *T. inconspicua* and *T. filamentosa*. The resulting phylogenetic P-values changed but remained significant (not shown). Finally, the phylogenetic P-values generally tended to be lower (more significant) than uncorrected ANOVA P-values for sexual system, but not for RU sexuality (Table 2.1).

Table 2.1: Phylogenetic ANOVAs of sexuality and quantitative characters. The F statistic and phylogenetic P-value are given along with the uncorrected P-value in parentheses. Phylogenetic P-values were estimated using the method of Garland et al. (1993) as implemented by 'geiger' (Harmon et al., 2008). Notes: All quantitative characters were *ln*-transformed, except for maximum bract number in male vs. bisexual RUs, which was rank-transformed. Bisex. = bisexual; i.p. = involucral phyllome; i.p.l. = involucral phyllome length; max. = maximum; no. = number.

	Max. stamen no. per RU	Max. anther l. (mm)	Max. stigmatic hair l. (mm)	Max. i.p. no. for male vs. bisex. RUs	Max. i.p. no. for female vs. bisex. RUs	Max. i.p.l. for male vs. bisex. RUs (mm)	Max. i.p.l. for female vs. bisex. RUs (mm)
Sexual system	$F_{1,11} = 14.65$ 0.00052 (0.0028)	$F_{1,11} = 19.60$ 0.00013 (0.001)	$F_{1,11} = 0.37$ 0.4 (0.56)	$F_{1,11} = 0.20$ 0.53 (0.66)	$F_{1,11} = 6.49$ 0.0045 (0.027)	$F_{1,11} = 6.55$ 0.0066 (0.027)	$F_{1,11} = 1.87$ 0.066 (0.2)
RU sexuality	$F_{1,11} = 30.31$ 0.0006 (0.0002)	$F_{1,11} = 20.35$ 0.0016 (0.0009)	$F_{1,11} = 0.085$ 0.87 (0.78)	$F_{1,11} = 0.72$ 0.59 (0.42)	$F_{1,11} = 0.65$ 0.62 (0.44)	$F_{1,11} = 6.96$ 0.054 (0.023)	$F_{1,11} = 0.17$ 0.81 (0.69)

2.5 Discussion

Individual gene trees support a majority of species circumscriptions and reconstructions based on the species tree and validate several fruit and seed characters as taxonomically useful in distinguishing new sections (Figures 2.1, 2.3, 2.4; Appendix B). Perennial habit is strongly supported as a synapomorphy for a clade of two species (Figures 2.3 and 2.4). In contrast, ancestral-state reconstructions of sexual system and RU sexuality are either ambiguous for many nodes, including the root, or are unambiguous but homoplasious (Figures 2.3 and 2.4). I find significant associations of sexual system and RU sexuality with several androecial characters and with the length and number of involucral phyllomes.

2.5.1 Gene and species phylogenies

There is a generally close correspondence between plastid and nuclear phylogenies concerning the well-supported relationships (Figure 2.1), and so the high support for most branches in the species tree is unsurprising (Figure 2.3). While there are some ambiguities within individual gene trees (such as polytomies involving *Trithuria lanterna* and *T. konkanensis*), there is only one case in which individual gene trees show non-monophyly in the case of individual species, with both arrangements strongly supported: in the plastid tree, *T. bibracteata* has two *T. submersa* clades (and likely also *T. occidentalis*) embedded, while the reverse is true for the nuclear tree, and in both cases *T. bibracteata* and *T. submersa* are non-monophyletic (*T. occidentalis* was not sampled for ITS) (Figure 2.1). Apparent non-monophyly of a species at a particular locus could represent a taxonomic (circumscription) error, but these sorts of errors cannot account for the strongly conflicting relationship between the plastid and nuclear gene trees concerning the south-east versus south-west Australian *T. submersa* samples (Figure 2.1), unless the whole group of three species represents a single species or a species complex. If all the samples in this clade do represent a single species, it would have unusually high levels of genetic and morphological variation (e.g., *T. occidentalis* is dioecious while *T. bibracteata* and

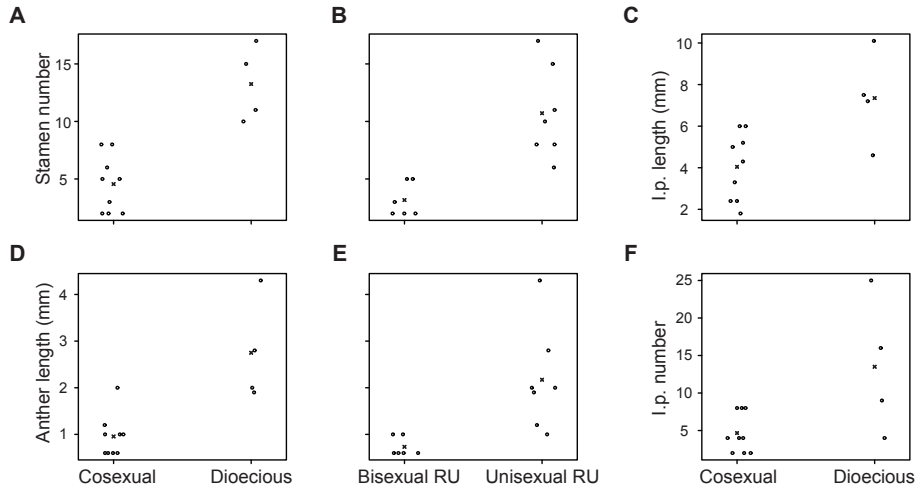


Figure 2.5: Plots of characters associated with changes in sexuality. Untransformed plots of quantitative species characters for sexual system (cosexual vs dioecious) and for RU sexuality (bisexual vs unisexual RU) (see Table 2.1). Crosses are mean values for the corresponding species grouping.

T. submersa are cosexual) compared with other species in the family (Figure 2.1; Sokoloff et al. 2008a). Non-monophyly of individual loci may also be the outcome of species-level paraphyly (e.g., reflecting recent speciation for the plesiospecies of Olmstead, 1995), or of short branches between speciation events (i.e., incomplete lineage sorting; Maddison 1997; Wakeley 2008); the latter may also lead to gene-tree incongruence. Gene-tree non-monophyly and incongruence can also result from hybridization/introgression. However, the *BEAST program (Heled and Drummond, 2010) accommodates non-monophyly and incongruence by assuming it is the result of incomplete lineage sorting (i.e., it is not able to accommodate reticulate evolution). Given the strong incongruence in the placement of the south-east Australian *T. submersa* samples between the two linkage groups, it is not possible to strongly support any set of relationships within this clade (Figure 2.3). Unpublished cytogenetic data (R.G. Kynast, P.J. Rudall et al., unpublished data) show that *T. submersa* is polyploid, as it has a somatic chromosome number of 2n = 56, which might support reticulate evolution. Additional

nuclear loci could help as tie-breakers in this region of the phylogeny, and may allow the source of incongruence to be determined.

Our study revealed relatively high levels of genetic divergence among investigated samples of *Trithuria australis* in the individual plastid and nuclear gene trees. One of these samples (collection *Rudall s.n.*; see Appendix A) differs significantly from the other three samples, although I was unable to find any clear morphological differences in this accession. It is tempting to suggest the occurrence of a cryptic species, but additional sampling of populations and genes and additional morphological work are needed to address this issue. It is perhaps not surprising that species boundaries might be difficult to define morphologically in some cases in Hydatellaceae, given its generally "reduced" morphology.

There is otherwise a close correspondence of species circumscription to phylogeny in most cases, and there is virtually no evidence of mixed sexual systems within species (D.D. Sokoloff et al., unpublished data). The only case I encountered here is a cosexual individual of *Trithuria filamentosa* (collection *Buchanan 12328*; see Appendix A) that groups with pistillate-only samples of *T. filamentosa* (Figure 2.1). The lack of any clear case of variation in sexual systems within Hydatellaceae contrasts with other groups, where among-population variation in sexual system may be common (e.g., Case et al., 2008; Obbard et al., 2006; Sarkissian et al., 2001). It is, however, possible that greater sampling will uncover unisexual individuals associated with cosexual species, or hermaphroditic individuals associated with dioecious species.

2.5.2 New classification

Our phylogenetic data support recognition of four new sections in the genus *Trithuria*, two in a tropical clade (sections *Altofinia* and *Hamannia*) and two in a subtropical/temperate clade (sections *Hydatella* and *Trithuria*), see Appendix B. It is tempting to recognize the tropical and subtropical/temperate clades as subgenera, but currently I lack a clear non-molecular synapomorphy to characterize them. Seedling morphology might act as a diagnostic

character for this in the future (see below). In contrast, each of the four sections recognized here can be readily characterized morphologically (Table 2.2; Appendix B). The two subtropical/temperate sections are *Trithuria* (comprising the type species of the genus *Trithuria submersa*, in addition to *T. bibracteata* and *T. occidentalis*) and *Hydatella* (comprising *T. filamentosa*, *T. inconspicua*, *T. australis* and *T. austinensis*). The sectional name *Hydatella* is based on the generic name *Hydatella* Diels, although it is not identical to the traditional taxonomic circumscription of that genus (Cooke, 1987). The two tropical sections, both newly described, are *Hamannia* (*T. polybracteata*, *T. lanterna* and *T. konkanensis*) and *Altofinia* (*T. cookeana* and *T. cowieana*). Each of the four sections includes a dioecious species and one to several cosexual species.

Table 2.2: Summary of the new sectional treatment of *Trithuria*. Node numbers correspond to those in Figure 2.3. Notes: [a]Possible apomorphy; [b]Collectively define a unique combination of character states.

Section	Node	Species	Defining Characters
Altofinia	4	*T. cookeana*, *T. cowieana*	Presence of paired pericarp papillae[a]
Hamannia	5	*T. konkanensis*, *T. lanterna*, *T. polybracteata*	Presence of pericarp ribs and a thick seed cuticle[b]
Hydatella	7	*T. austinensis*, *T. australis*, *T. filamentosa*, *T. inconspicua*	Absence of pericarp papillae, Absence of pericarp ribs, and a thick seed cuticle[b]
Trithuria	10	*T. bibracteata*, *T. occidentalis*, *T. submersa*	Sculptured seed[a], thin seed cuticle[a], presence of pericarp epicuticular wax[a], specialized endocarp cells facilitating dehiscence[a]

2.5.3 Morphological evolution

Prior to the revision of the family by Sokoloff et al. (2008a), Hydatellaceae were divided into two genera based on the apparent distribution of RU sexuality and pericarp ribs. This division was subsequently shown to be spurious (Sokoloff et al., 2008a). While it was clear that these characters in combination do not support any natural division of the family, it remained open whether they might act separately as non-homoplasious synapomorphies. The MP reconstruction of RU sexuality suggests that unisexual RUs are the ancestral state and there were three origins of bisexual RUs (Figure 2.4). However, ambiguity in the MP reconstruction of pericarp ribs and in ML reconstructions of both characters, supports the view that neither character is useful for defining taxa above the species level in Hydatellaceae (Figures 2.3 and 2.4). Sokoloff and colleagues also made several predictions concerning seed or fruit characters that might better reflect phylogenetic relationships. The utility of these characters is generally affirmed here (Figures 2.3 and 2.4; Sokoloff et al. 2008a): paired pericarp papillae may be synapomorphic for sect. *Altofinia* (Table 2.1), and strongly sculptured seeds for sect. *Trithuria* (Table 2.1). The latter group is also supported by a lack of thick cuticular layer on the seed, the presence of epicuticular wax on the pericarp, and by complete release of seed from the fruit caused by the occurrence of specialized endocarp cells facilitating fruit dehiscence. Maximum likelihood character-state reconstructions of these characters are slightly ambiguous at one to several nodes deeper (Figure 2.4), so it remains somewhat unclear whether they are actually synapomorphic for these clades. However, in MP reconstructions they are unambiguous synapomorphies.

Sections *Hamannia* and *Hydatella* can be defined by combinations of characters that have unambiguous MP and ML reconstructions for the relevant nodes (Figures 2.3 and 2.4; Table 2.1). Detailed study of fruit anatomy tentatively suggests that members of section *Hamannia* differ from the rest of the genus in the structure of the fruit apex, which could act as an additional synapomorphy for this section (D.D. Sokoloff et al., unpublished data). The fruit apex has not so far been investigated in two species

(*Trithuria filamentosa* and *T. austinensis*), and so phylogenetic mapping of the character is premature. Another morphological character of potentially strong phylogenetic utility is cotyledon morphology. Members of sect. *Hamannia* have a strongly reduced cotyledon bearing no sheath (Sokoloff et al., 2008b), while members of sections *Trithuria* and *Hydatella* possess a cotyledonary sheath (Sokoloff et al., 2008b; Tillich et al., 2007). Unfortunately, I currently lack data on seedlings in section *Altofinia*.

All but two characters analyzed here are directly related to reproductive characters. Leaf auricles represent a potential synapomorphy for the clade comprising *Trithuria australis*, *T. filamentosa* and *T. inconspicua*; this is unequivocal in the MP reconstruction but the ML reconstruction is equivocal for several deeper nodes (Figures 2.3 and 2.4). Perennial habit is an unequivocal synapomorphy of *T. filamentosa* and *T. inconspicua* in both ML and MP reconstructions (Figures 2.3 and 2.4). This suggests that perenniality is a derived lifespan strategy in Hydatellaceae, as the less likely scenario would require multiple origins of annuality. With the possible exception of island radiations (Carlquist, 1974) annual habit is usually considered the derived state (e.g., Stebbins, 1957) with only a few studies finding the opposite (e.g., Barrett and Graham, 1997; Tank and Olmstead, 2008). In Hydatellaceae the evolution of perennial habit is associated with a shift to more permanent aquatic environments (Cooke, 1987; Pledge, 1974; Sokoloff et al., 2008a; Wells et al., 1998) that mirrors a similar evolutionary transition seen in Pontederiaceae (Barrett and Graham, 1997), and may support a more general case of annual growth being associated with more ephemeral aquatic habitats.

2.5.4 Sexual-system evolution

The ambiguous reconstruction of sexual system and RU sexuality across most nodes, including the root node of the family (Figures 2.3 and 2.4), makes inference about the direction(s) of functional gender evolution and morphological evolution of the RU unclear. Uncertainty regarding homology between bisexual RUs in Hydatellaceae and bisexual flowers in other

angiosperms (Rudall et al., 2007, 2009b) obscures the issue of the evolution of RU sexuality. Dioecy is normally considered to be a derived state within angiosperms (Endress, 2004), but it is also present in other "early-divergent" lineages of angiosperms (i.e., Amborellaceae, Schisandraceae and some Trimeniaceae; Thien et al. 2000, 2003; Wagner and Lorence 1999), although some of these are cryptically dioecious, with complementary sterility in reproductive organs between genders (Endress, 2004). This may support ancestral hermaphrodite flowers in angiosperms (Endress, 2004), although the inferred ancestral state was equivocal in the study of Endress and Doyle (2009). Nymphaeaceae and Cabombaceae, the closest relatives of Hydatellaceae, are hermaphroditic and self-compatible (Endress, 2004; Saarela et al., 2007).

Although less common than the transition from hermaphroditism to dioecy, cosexuality has been inferred to be a potentially derived state in several angiosperm lineages (Case et al., 2008; Dorken et al., 2002; Obbard et al., 2006; Schaefer and Renner, 2010; Sytsma et al., 1991; Volz and Renner, 2008). While our data now make clear that scenarios involving a single origin of dioecy or cosexuality (Sokoloff et al., 2009a) are unlikely (Figures 2.3 and 2.4), additional data are still required to resolve the evolutionary transitions of sexual system in Hydatellaceae. These might include: (1) the morphological identity of the Hydatellaceae RU; (2) correlating sexual system with irreversible character changes; (3) discovery of additional species that may clarify the character state at the root node of the tree. To date, work on the Hydatellaceae RU has failed to conclusively demonstrate its morphological nature (Rudall et al., 2007, 2009b), and species discovery will only result from additional field, morphological and phylogenetic work. One type of character change that is thought to be difficult to reverse is change in ploidy level (Bull and Charnov 1985, but see Wolfe 2001). Indeed, this character is known to correlate with changes in sexual system in other groups, often with polyploidy resulting in a transition to dioecy from cosexuality (Miller and Venable, 2000). Alternatively, selfing may be selected for after polyploidization (Lande and Schemske, 1985; Levin, 1975; Stone, 2002). Ploidy levels in Hydatellaceae remain largely unknown, but published accounts ten-

tatively indicate some polyploid species, including the New Zealand apomict *Trithuria inconspicua* with 2n = ∼24 (de Lange et al., 2004) and the Indian cosexual *T. konkanensis* with 2n = 40 (Gaikwad and Yadav, 2003). Unpublished results (R.G. Kynast, P.J. Rudall et al., unpublished data) also show that the widespread species *T. submersa* is polyploid, in contrast to *T. australis*, which is a probable diploid with a somatic chromosome number of 2n = 14. Some common proxies of ploidy level such as pollen size (Knight et al., 2010) or guard/pavement cell size (Beaulieu et al., 2008) do not show an obvious relationship with ploidy level in Hydatellaceae (Remizowa et al. 2008; R.G. Kynast, P.J. Rudall et al., unpublished data).

The phylogenetic ANOVA tests (Garland et al., 1993) of association between sexual system and other reproductive traits indicate that there are significantly higher values in two proxies for pollen production (anther length and stamen number) in dioecious species (Table 2.1, Figure 2.5). This is congruent with a study examining pollen/ovule ratios, which were significantly higher for the dioecious *Trithuria austinensis* compared to the cosexual *T. submersa* (Taylor et al., 2010). High pollen production is common in wind-pollinated species, where it could compensate for the supposed inefficiency of wind as a pollination vector, or may alternatively reflect increased local pollen competition (Friedman and Barrett, 2009). The morphological differences between unisexual RUs (Table 2.1, Figure 2.5; Sokoloff et al. 2008a) is analogous to differences seen between flower types in some dioecious species (Eckhart, 1999), and is consistent with sexual specialization in the functions of male vs. female RUs. Reduction in pollen production in some species may also be a simple function of a high selfing rate (e.g., Ornduff, 1969; Sicard and Lenhard, 2011). Involucral phyllomes of staminate plants are longer and generally held together more tightly than those of pistillate plants (see Figures 2, 4, 5 and 6 in Sokoloff et al. 2008a); this difference could reflect aerodynamic considerations, by positioning the stamen further into the airstream (Niklas, 1985). Presumably cosexual species (with unisexual or bisexual RUs), which are probably mostly selfing, do not display this feature because gravity suffices to effect pollination (Taylor et al., 2010). Although I can speculate about the possible adaptive significance of morphological

traits of male RUs in dioecious species, it is difficult to offer any adaptive explanation for the consistent increase observed in the number of involucral phyllomes of female RUs in dioecious species compared with other species.

2.5.5 Conclusions

I proposed four new sections of Hydatellaceae and presented a generally well-resolved and well-supported phylogeny of the family that verifies previously postulated morphological synapomorphies for subclades, and identifies several new ones. I showed that transitions between sexual systems are particularly labile in the family, and that while the direction(s) of change remains unresolved, sexual-system change is clearly associated with changes in male reproductive function and in quantitative characters of the RU. Questions remain concerning species circumscription, which will require further population and genomic sampling. Hydatellaceae may also provide a very useful "early angiosperm" model system for examining sexual-system evolution and concomitant changes in RU morphology and reproductive output.

Chapter 3

Biogeography and niche conservatism in Hydatellaceae

3.1 Summary

Despite a very early divergence from most other flowering plants, the dwarf aquatic family Hydatellaceae is species-poor and displays relatively limited morphological and ecological diversity among its extant species. Its isolated phylogenetic position and Gondwanan distribution may suggest a very early history of conservatism. However, a Bayesian dating analyses indicates a relatively recent (early Miocene) origin for the crown clade, contrasting with a Lower Cretaceous split between Hydatellaceae and the water lilies. Nonetheless, the recent radiation of extant species in Hydatellaceae does not fully explain the family's apparent ecological homogeneity, as I also uncovered substantial phylogenetic niche conservatism in the family. This may be a response to extreme environmental conditions experienced by these ephemeral wetland plants. I also infer recent long-distance dispersal events from Australia to India and New Zealand, and so the highly disjunct distribution of species is not relictual. However, a deep split between tropical and subtropical/temperate species in Australia likely reflects a vicariance event resulting from the increase in aridity in Australia since the early Miocene.

3.2 Introduction

Widespread rainforest in Australia has been replaced gradually by deserts, monsoon-inundated savannah and sclerophyll (Mediterranean-type) biomes since the Eocene, spurred by increasing aridity and local responses to global cooling (e.g., Martin, 2006). This has led to considerable floristic diversification, particularly in lineages that have adapted to drier biomes (Crisp et al., 2004; Crisp and Cook, 2007). Despite this, Australian aquatic plant lineages have also persisted and diversified. The most remarkable of these may be Hydatellaceae, a small family of minute aquatic herbs found predominantly in temperate sclerophyll biomes and monsoon-inundated savannah (Crisp et al., 2004; Sokoloff et al., 2011). The family was recently recognized as an ancient fragment of flowering-plant phylogeny (Saarela et al., 2007), and has attracted considerable attention because of the insights its study may provide into the early angiosperm radiation (Friedman et al., 2012; Rudall et al., 2007, 2009b; Sokoloff et al., 2008b). Recent studies have addressed the ecology and phylogenetic diversification of Hydatellaceae (Chapter 2; Sokoloff et al., 2011; Taylor et al., 2010; Tuckett et al., 2010a,b). Most species of Hydatellaceae are annuals that live in temporary water bodies during the temperate wet winter or tropical summer wet season. Most species reproduce rapidly as standing water in ephemeral wetlands disappears at the end of the season, persisting in the seed bank between inundations (Deil, 2005; Sokoloff et al., 2011; Taylor et al., 2010). Two species that grow as permanently submerged apomictic perennials in regions with temperate climate likely represent a derived aquatic life form in the family (Chapter 2; Sokoloff et al., 2011).

The ancient origin of Hydatellaceae implied by phylogenetic studies of early angiosperms may indicate substantial opportunity for niche diversification across the family. However, the timing of its split from other angiosperms has not yet been established, and the age of the crown clade (the most recent common ancestor of extant species) is also unknown, and could be substantially younger. A recent crown-group age would substantially limit opportunities for diversification, and may be consistent with the

family's globally limited range and highly specialized aquatic life form. However, limited diversification in habitat preferences may also reflect phylogenetic niche conservatism (PNC), which is the tendency of lineages to appear more similar along environmental axes than expected by drift or a random adaptive walk in the available evolutionary time (Brownian motion, BM; Losos 2008a). I determined the posterior distribution of divergence times between Hydatellaceae and its sister group, the water lilies (Cabombaceae and Nymphaeaceae), and used this as the prior for estimating divergence times among species of Hydatellaceae (Chapter 2; Sokoloff et al., 2008a). I used the resulting time-calibrated species tree to examine the roles played by vicariance, extinction and long-distance dispersal in shaping the family's current distribution, based on multiple methods of biogeographic inference. I used estimated niches of extant species to test patterns of niche evolution, comparing Brownian motion and Ornstein-Uhlenbeck models.

3.3 Material and methods

3.3.1 Molecular dating

I dated the stem and crown ages of Hydatellaceae in the context of a seed-plant phylogenetic estimate based on plastid data, and used information from this analysis to inform divergence times in a multi-species coalescent tree of the family derived from plastid and nuclear data. For the seed-plant analysis I added *Trithuria cowieana* (*Macfarlane & al. 4217* [MW]; GenBank numbers-numbers) to an existing data set that included *Trithuria filamentosa* and *T. submersa* (Graham and Iles, 2009; Rai et al., 2008). The complete matrix consists of 13 plastid genes and 17 exemplar taxa from all extant lineages of seed plants except Gnetales, a highly divergent lineage, and includes taxa that define the earliest splits in angiosperm phylogeny (Graham and Iles, 2009). The three included species of Hydatellaceae define its two deepest phylogenetic splits (Figure 3.1; Appendix C; Chapter 2). I used BEAST version 1.6.1 (Drummond and Rambaut, 2007) to simultaneously infer divergence times and phylogeny using the Bayesian random

local clocks model (Drummond and Suchard, 2010), employing a suite of eight fossil constraints to calibrate the seed-plant analysis (Appendix C, Appendix Table C.1).

To date the species tree of Hydatellaceae I re-analyzed the data set of Chapter 2, which was previously estimated in a time-independent manner. The analysis includes all species in the family (all members of single genus, *Trithuria*; Sokoloff et al. 2008a). I employed mean and 95% credibility interval (CI) of posterior estimates of the two deepest splits in Hydatellaceae from the seed-plant analysis as normal priors for the same splits in this family-level analysis, but only for the plastid loci (the nuclear data lacked outgroup taxa due to unreliable alignment; Chapter 2). This analysis used a multi-species coalescent approach (Heled and Drummond, 2010).

3.3.2 Biogeographic reconstruction

Because different methods of biogeographic analysis make different assumptions about the causes of biotic distribution (Ronquist and Sanmartín, 2011), I reconstructed ancestral areas using three approaches: a maximum-parsimony based dispersal-vicariance method (DIVA; Ronquist 1997), a maximum likelihood (ML) based dispersal-extinction-cladogenesis method (DEC; Ree and Smith 2008), and standard ML ancestral-state reconstruction (ASR; Lewis 2001; Sanmartín et al. 2008). In all cases I used the species tree as the reference phylogeny (see Appendix Figure C.1), and considered the five major biogeographic areas that define the range of Hydatellaceae (Figure 3.2A; Appendix C). I used RASP version 2.0 Beta (Ronquist, 1997; Yu et al., 2010) to perform the DIVA analysis, and Lagrange version 20110117 (`http://www.reelab.net/lagrange/configurator/index`; Ree and Smith 2008) for the DEC analysis. For both methods, I constrained connectivity of India to northern Australia (NA), and New Zealand to southeast Australia (SEA). I also explored the effect of differing levels of area extinction (extirpation) in the DEC analysis. This was motivated by the fact that area extinction rates in Lagrange are biased towards zero, even when actual extinction rates are high (Ree and Smith, 2008), and because a speciation-extinction

analysis based on tree shape was consistent with a broad range of possible species extinction rates (0.003–0.955 Ma^{-1}, 95% credibility interval, CI; see Appendix C). While the area extinction rate in the DEC analysis is not directly comparable to the species extinction rate based on tree shape, the two should be related (for example, when a species experiences extinction in all areas encompassed by its range, the species itself becomes extinct). When vicariance events predicted in these analyses could not be explained by contemporaneous geographic division, I propose that these represent long-distance dispersal events.

The ML ASR analyses were performed using BayesTraits version 1.0 (Pagel et al., 2004). For this analysis, I assumed separate symmetrical transition rates between Australia and India or New Zealand, between southwest Australia (SWA) and SEA, and between NA and SWA/SEA and set all other transition rates to zero, reducing parameter dimensions in the ML ASR model. I tested this three-parameter model (hereafter termed the 'full ML ASR') against simpler nested models, using the corrected Akaike information criterion (AICc; Sugiura 1978). The simpler nested models were a one-rate transition model in which all allowed transitions rates of the full ML ASR model were set equal to each other ('simple ML ASR'), and a two-rate transition model with rates from Australia to India and New Zealand set equal to each other, and rates within Australia set equal to each other ('continental ML ASR').

3.3.3 Climate niche evolution

Species occurrence data for *Trithuria* were collected from online herbaria resources and augmented by recent collections (Table C.4). Specimens lacking explicit latitude and longitude coordinates were georeferenced using collection information when the locality could be determined within 1 km. Three of the four dioecious species (*T. cookeana*, *T. occidentalis*, and *T. polybracteata*) are known from only one to three populations (Sokoloff et al., 2008a) and were therefore not included in this analysis. Occurrence data ranged from 8 to 43 per species, for a total of 224 occurrences (Appendix Ta-

ble C.4). Populations of *T. submersa* from SWA and SEA were provisionally considered separately in Chapter 2, which I continue to follow here. I used the bioclim dataset of 19 derived temperature and precipitation climate variables with ∼1 km resolution (`http://www.worldclim.org`; Hijmans et al. 2005) in conjunction with Maxent version 3.3.3e (Phillips and Dudík, 2008) to model species distributions. The analysis was run using Maxent defaults, with exceptions noted in Appendix C.

To assess how climate preferences evolved, I separated out Maxent occupancy distributions for each of the 19 climate variables, for each species, using the method of Evans et al. (2009). Brownian motion (BM) and several Ornstein-Uhlenbeck (OU) models were used to fit mean climate preferences to the species tree. For the latter I considered one to three possible selection regimes, reflecting decreasing amounts of conservatism (referred to here, respectively, as 'ou1,' 'ou2,' and 'ou3'). Internal nodes on the species tree were assigned to alternative selective regimes for the latter two OU models (see Appendix Figure C.4). For the ou1 model I assumed that a single selective regime operated across the entire tree; the ou2 model assumed two selective regimes defined by the root node (one affecting the tropical north Australian clade, the other the temperate south Australian clade and the root node); the ou3 model further divided the south Australian clade into southwest vs. southeast components. The biogeographic splits used in these models correspond to major shifts inferred in the full ML ASR model (Figure 3.2B, see Appendix C). I used the R package 'ouch' to perform calculations (Butler and King, 2004). For each climate variable I used AICc to select between models. To understand how the shape of the occupancy distribution (the probability surface) affects regime estimates and model choice, I sampled 1000 replicates from the surfaces for each climate variable ('resampled climate preferences') and evaluated the model fit as above.

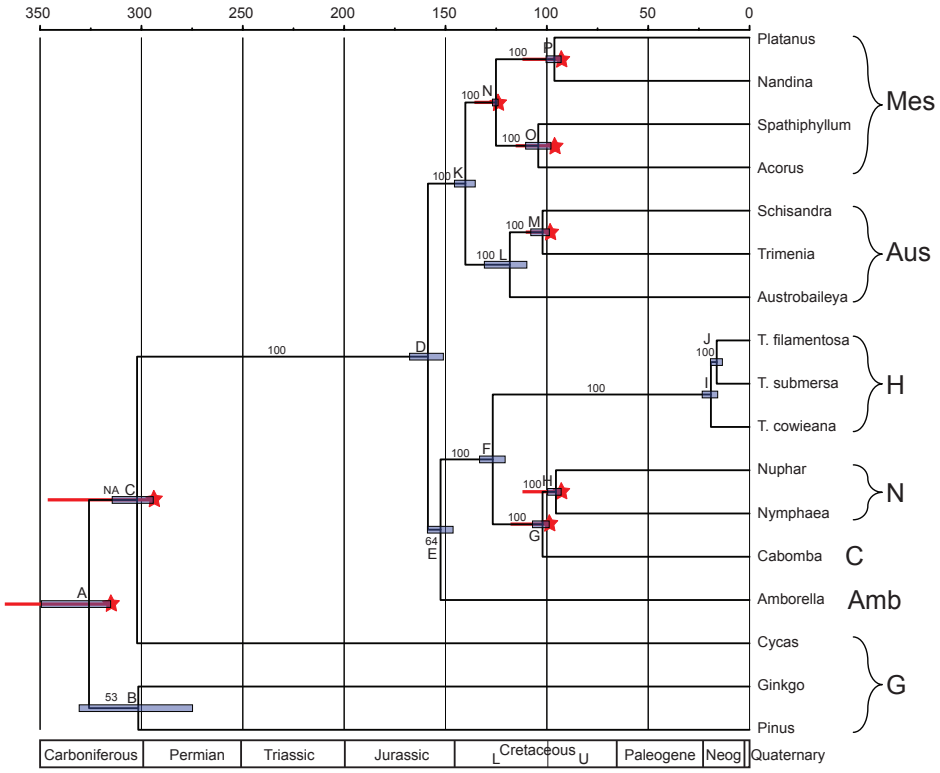


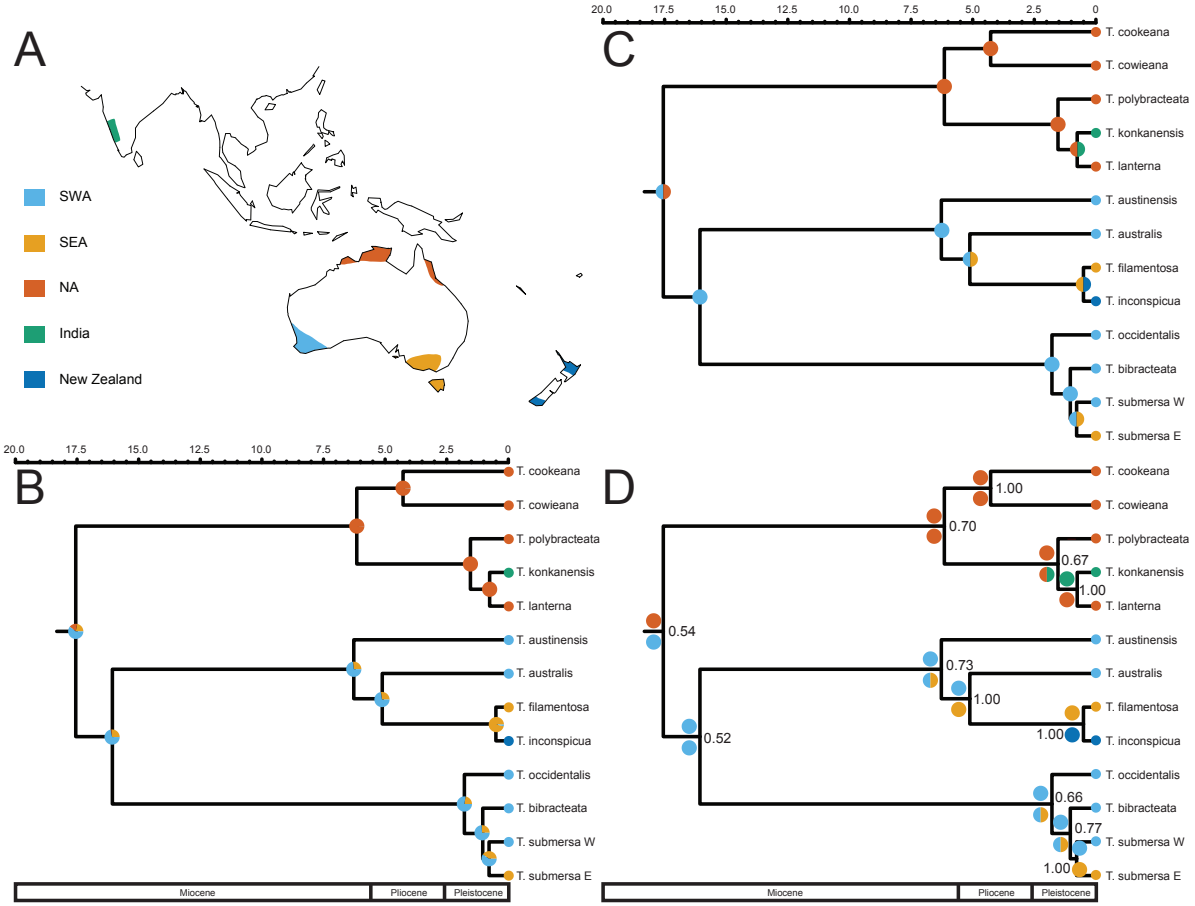
Figure 3.1: Bayesian random-local-clock dating of seed-plant phylogeny based on 13 plastid genes. Amb, Amborellaceae; Aus, Austrobaileyales; C, Cabombaceae; G, gymnosperms; H, Hydatellaceae; L, lower; Mes, mesangiosperms; Neog, Neogene; N, Nymphaeaceae; U, upper; T., *Trithuria*. Labelled nodes are referred to elsewhere (Appendix Tables C.1 and C.2). The time scale is in Ma (stratigraphic scale according International Commission on Stratigraphy 2012). Blue bars on nodes represent 95% posterior interval of ages; red stars are minimum ages of calibration fossils (Appendix Table C.1) and red bars are their assigned 95% prior age distributions. Numbers adjacent to branches are posterior probability support values (expressed as percentages).

3.4 Results and discussion

3.4.1 The age of Hydatellaceae

I infer a young crown age for Hydatellaceae around the early Miocene, and predict that most subsequent diversification in the family occurred only after the late Miocene (Figures 3.1, 3.2, C.1, Tables C.2 and C.3). The inferred phylogenetic relationships and patterns of clade support in the multi-species coalescent tree are very similar to our previous results based on the same data (Chapter 2). The phylogenetic placement of Hydatellaceae as the sister group of water lilies, Cabombaceae and Nymphaeaceae (Saarela et al., 2007), means that the stem lineage of Hydatellaceae is at least as old as the earliest Cretaceous fossils from the crown water lilies, used as fossil constraints here (see Appendix C). Several fossils have been linked to Hydatellaceae. These include the Lower Cretaceous aquatic, *Archaefructus* (Endress and Doyle, 2009; Saarela et al., 2007; Sun et al., 2002) and the Upper Cretaceous pollen, *Monosulcites riparius* (Hofmann and Zetter, 2010). However, while the ages of purported Hydatellaceae relatives are potentially consistent with them being stem-lineage fossils, their phylogenetic association with the family is uncertain (Rudall and Bateman, 2010; Sokoloff et al., 2011). The age of the crown clade of Hydatellaceae (∼19 Ma) is inferred to be much younger than the stem age (∼121 Ma), which is fairly close to the estimated age of the most recent common ancestor of flowering plants (∼151 Ma; Figure 3.1, see Table C.2).

Figure 3.2 *(following page)*: Biogeography of Hydatellaceae. A) Range map of extant species, exaggerated to aid visualization (see Appendix Table C.4). Ancestral ranges inferred using (B) the full maximum likelihood ancestral-state reconstruction, (C) maximum-parsimony based dispersal-vicariance (DIVA) analysis, and (D), a maximum-likelihood based dispersal-extinction-cladogenesis (DEC) analysis. Pie fractions in (B) represent relative likelihoods; in (C) and (D) they represent joint areas where the species is inferred to have existed in multiple areas. The relative likelihood of the best geographic range split-pair is shown in (D).


A
SWA
SEA
NA
India
New Zealand
B
C
D
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
T. cookeana
T. cowieana
T. polybracteata
T. konkanensis
T. lanterna
T. austinensis
T. australis
T. filamentosa
T. inconspicua
T. occidentalis
T. bibracteata
T. submersa W
T. submersa E
1.00
0.70
0.67
1.00
0.54
0.73
1.00
1.00
0.52
0.66
0.77
1.00
Miocene
Pliocene
Pleistocene

3.4.2 Biogeography of Hydatellaceae

There were four predicted instances of long-distance dispersal in Hydatellaceae (Australia to India, Australia to New Zealand, and two instances from southwest to southeast Australia; Figure 3.2) The inferred dispersal events likely involved selfers or apomicts, consistent with Baker's Law (Baker, 1955). Indeed, the sole New Zealand species, *T. inconspicua*, which represents one of only two perennial species of Hydatellaceae, is probably an apomict (Remizowa et al., 2008; Rudall et al., 2008), and selfing is thought to typify the Indian *T. konkanensis* (Sokoloff et al., 2008a).

The full ML ASR model (the best model as chosen by AICc, Figure 3.2B, see Appendix C) generally gave biogeographical results that are similar to the DIVA and DEC analysis (Figure 3.2C, D). The latter two analyses allow extant species and reconstruction of internal nodes (ancestral species) to encompass multiple areas. Examples of these at internal nodes facilitate identification of dispersal, extinction and vicariance events within these frameworks (in contrast, ML ASR only implicitly considers dispersal, and restricts each species range and internal node to a single area, and so divided pies indicate relative likelihood of states). The DIVA and DEC analyses indicate a continent-scale vicariance event at the root of extant Hydatellaceae (Figure 3.2C and D). The interior of Australia was still well watered in the early Miocene (up to the mid-Miocene), and although there were permanent lakes, there was also a marked dry season, indicating the potential for ephemeral aquatic habitats (Martin, 2006). The continued aridification of central Australia presumably led to this vicariance event (Figure 3.2), aided by extinction events in the interior of the continent. This resulted in the present-day biogeographic division in the family between the tropical and subtropical/temperate clades inferred in Chapter 2. The eastern edge of Australia has pockets of subtropical to tropical rainforest (Crisp et al., 2004). These may be unlikely to host Hydatellaceae, which are also unknown from surrounding areas, perhaps reflecting the weak seasonality there compared to most of the family's range.

The DIVA and DEC analyses (Figure 3.2C, D) reconstruct two vicari-

ance events between SWA and SEA (vicariance in DIVA and DEC analyses can be hypothesized when neither of the descendant lineages of an ancestrally joint range occur in the entire joint range; dispersal is inferred when the ancestral range is not a joint one between the multiple areas found in the descendants). However, the inferred ages of these events are younger than the last marine inundation and the formation of the Nullarbor Plain with eventual extreme aridification starting ∼15 Ma (Figure 3.2; Crisp and Cook 2007). Long-distance dispersals are therefore more likely explanations of the current biogeographic distribution of these species than vicariance. Similarly, the Indian and New Zealand taxa (*T. konkanensis* and *T. inconspicua*) represent relatively recent long-distance dispersal events from Australia (Figure 3.2). Both lineages are much too young to represent Gondwanan relicts, as proposed by Yadav and colleagues (Gaikwad and Yadav, 2003; Yadav and Janarthanam, 1994, 1995). This finding is consistent with studies of other groups, some of which show that a Gondwanan distribution may in part be a function of recent dispersal (e.g., Sharma and Giribet 2012; Swenson et al. 2012). These biogeographic inferences are quite sensitive to the predicted extinction rate (see Appendix C). I found that the relative likelihoods of the best range splits decrease rapidly across most nodes with higher amounts of area extinction in the DEC analysis (see Appendix Figures C.2, C.3). For higher extinction rates, it may be difficult to estimate ancestral ranges using model-based methods of biogeographic evolution. However, the cumulative relative likelihoods of inferred ranges of the immediate descendants of each node were more robust to low to moderate levels of area extinction (see Appendix Figure C.3).

3.4.3 Phylogenetic niche conservatism in Hydatellaceae

The phylogenetic niche analysis (Table 3.1) supports a constrained evolutionary trajectory in Hydatellaceae. Focusing on the mean climate preferences, models consistent with phylogenetic niche conservatism (i.e., ou1–3) were always chosen as the best model of change (i.e., Brownian motion was never chosen). There was no consistency in the degree of niche conservatism

within either temperature- or precipitation-related variables (i.e., variables 1–11 and 12–19 respectively). However, there are likely to be strong correlations between at least some climate variables here. I did not take these correlations into account (see Appendix C), but at least some related variables show consistent patterns of niche evolution (e.g., variables 5 with 10, 6 with 11, and 14 with 17; Table 3.1). The resampled climate preferences give an indication of how sensitive model choice is to underlying variability in the niche occupancy profile. Where the most commonly selected model in these replicates was different from the mean climate preference model choice, I show both in Table 3.1; when there was a resulting difference in model choice, the models chosen were usually neighbouring in complexity (e.g., ou2 and ou3).

Table 3.1: Ancestral climate variables for Hydatellaceae. The AICc score of the best model is indicated in bold in the case of mean climate preference. I considered Brownian motion (BM) and Ornstein-Uhlenbeck (OU) models, the latter with one to three selection regimes (ou1–3; ou1 = most conservative). Where the most commonly chosen model in the resampled climate preferences differed from the model of the mean climate preference, both are noted (values in parentheses are 95% intervals based on resampling). *ln*L, log likelihood; s.d., standard deviation; T, temperature; P, precipitation; Q, quarter; NA, northern Australia; SA, southern Australia; SEA, southeastern Australia; SWA, southwestern Australia. See text for sources of primary climate data and phylogeny.

Maxent climate variables	Parameter estimates	Model	*ln*L	AICc	Resampled climate preferences
1. Annual mean T (°C)	NA = 26.8 (25.3, 27.6), SEA = 11.2 (6.0, 16.3), SWA = 16.4 (14.6, 18.2)	ou3	-16.34 (-28.37, -10.55)	**57.68** (46.11, 81.75)	39%
	NA = 26.8 (25.3, 27.7), SA = 14.5 (12.8, 16.3)	ou2	-22.56 (-29.77, -18.97)	61.12 (53.95, 75.54)	48%
2. Mean diurnal range (°C)	10.6 (9.5, 11.4)	ou1	-16.11 (-23.50, -15.80)	**42.21** (41.60, 57.01)	88%
3. Isothermality	NA = 58.8 (55.7, 62.3), SEA = 48.5 (45.0, 54.4), SWA = 50.6 (48.8, 52.8)	ou3	-1.53 (-29.97, -11.41)	**28.05** (47.82, 84.93)	10%
	NA = 58.8 (55.7, 62.3), SA = 50.0 (47.8, 51.9)	ou2	-11.72 (-30.02, -14.83)	39.44 (45.67, 76.05)	65%
4. T seasonality (s.d.) (°C)	NA = 1.9 (1.4, 2.5), SA = 3.6 (3.1, 4.2)	ou2	-3.49 (-14.29, -2.41)	**22.98** (20.82, 44.58)	47%
5. Max T of warmest month (°C)	NA = 34.9 (33.6, 36.4), SEA = 20.9 (16.0, 26.4), SWA = 28.9 (26.1, 32.0)	ou3	-18.92 (-28.22, -16.46)	**62.85** (57.91, 81.45)	22%
	29.9 (28.3, 31.6)	bm	-29.57 (-39.45, -26.97)	64.86 (59.65, 84.62)	43.80%
6. Min T of coldest month (°C)	NA = 17.1 (15.1, 19.0), SEA = 3.0 (-2.0, 8.6), SWA = 6.8 (5.0, 8.5)	ou3	-13.88 (-29.94, -13.01)	**52.76** (51.02, 84.87)	31%

Maxent climate variables	Parameter estimates	Model	*ln*L	AICc	Resampled climate preferences
	NA = 17.1 (15.2, 19.0), SA = 5.3 (3.5, 7.7)	ou2	-20.32 (-30.18, -17.55)	56.64 (51.10, 76.36)	58%
7. T annual range (°C)	20.0 (17.9, 21.9)	ou1	-22.68 (-31.00, -22.18	**55.37** (54.37, 72.00)	82%
8. Mean T of wettest Q (°C)	NA = 27.4 (26.3, 28.6), SA = 10.5 (8.6, 13.1)	ou2	-21.91 (-32.30, -17.81)	**59.82** (51.62, 80.60)	72%
9. Mean T of driest Q (°C)	NA = 24.4 (22.8, 25.6), SEA = 13.9 (7.5, 18.7), SWA = 20.8 (19.0, 22.8)	ou3	-15.18 (-28.01, -13.09)	**55.35** (51.18, 81.02)	26%
	20.7 (19.2, 22.0)	ou1	-25.53 (-32.35, -23.27)	61.08 (56.53, 74.70)	44%
10. Mean T of warmest Q (°C)	NA = 29.0 (27.9, 29.9), SEA = 15.2 (10.3, 20.5), SWA = 21.4 (19.25, 23.6)	ou3	-17.6 (-28.27, -13.15)	**60.2** (51.30, 81.54)	30%
	NA = 29.0 (27.9, 30.1), SA = 19.2 (17.5, 21.4)	ou2	-23.8 (-29.20, -19.77)	63.61 (55.55, 74.40)	37%
11. Mean T of coldest Q (°C)	NA = 24.1 (22.3, 25.5), SEA = 7.0 (1.8, 12.4), SWA = 11.7 (10.1, 13.3)	ou3	-14.8 (-30.55, -11.29)	**54.6** (47.59, 86.10)	47%
	NA = 24.1 (22.4, 25.5), SA = 9.9 (8.3, 12.2)	ou2	-21.74 (-30.67, -18.17)	59.48 (52.34, 77.33)	48%
12. Annual P (mm)	1269 (997, 1562)	ou1	-81.37 (-87.22, -75.96)	**172.74** (161.92, 184.44)	88%
13. P of wettest month (mm)	288 (211, 365)	ou1	-72.56 (-78.96, -65.19)	**155.12** (140.39, 167.92)	68%
14. P of driest month (mm)	NA = 1.07 (0, 4), SEA = 96 (31, 143), SWA = 14 (7, 21)	ou3	-23.42 (-47.55, -22.10)	**71.83** (69.20, 120.10)	48%

Maxent climate variables	Parameter estimates	Model	*ln*L	AICc	Resampled climate preferences
15. P seasonality (coeff. var.)	NA = 116.9 (103.7, 128.0), SEA = 23.1 (12.6, 37.3), SWA = 65.5 (49.0, 81.7)	ou3	-36.16 (-45.64, -34.39)	**97.31** (93.78, 116.27)	50%
16. P of wettest Q (mm)	NA = 1525 (1002, 2023), SA = 378 (265, 493)	ou2	-76.74 (-83.28, -67.96)	**169.47** (151.92, 182.55)	64%
17. P of driest Q (mm)	NA = 6.39 (0, 13), SEA = 316 (109, 479), SWA = 51 (32, 72)	ou3	-41.36 (-62.12, -37.28)	**107.72** (99.57, 149.23)	45%
18. P of warmest Q (mm)	NA = 410 (239, 624), SEA = 317 (127, 478), SWA = 56 (35, 75)	ou3	-47.61 (-68.91, -51.47)	**120.22** (127.94, 162.82)	5%
	217 (154, 306)	ou1	-60.46 (-71.19, -59.81)	130.93 (129.62, 152.37)	43%
19. P of coldest Q (mm)	334 (160, 588)	ou1	-68.13 (-83.12, -62.21)	**146.28** (134.42, 176.23)	40%
	NA = 268 (-27, 1081), SA = 362 (262, 466)	ou2	-67.96 (-82.16, -56.90)	151.93 (129.80, 180.33)	45%

Phylogenetic niche conservatism (PNC) may be caused by a variety of phenomena, including stabilizing selection, niche-filling and pleiotropy (Cooper et al., 2010; Harvey and Pagel, 1991; Losos, 2008a). It has been suggested that PNC is ubiquitous in many organisms (Losos, 2008a), although in plants it has only been studied in detail in terrestrial groups (e.g., Evans et al. 2009). The relative conservatism of Hydatellaceae may contrast strongly with water lilies (its sister group). Although both include species that grow in permanent to ephemeral aquatic habitats, the latter are a cosmopolitan group with substantially greater diversity in form and ecology (Löhne et al., 2008). This apparent disparity in niche diversity may be due to the much older crown age of water lilies (approximately five times older; Figure 3.1, see Appendix Table C.2), although their degree of conservatism remains to be quantified in a phylogenetic framework. Ephemeral wetlands constrain growth to a limited, unpredictable time-frame each year, favouring rapid reproduction, annual life-span, dwarf stature and the ability to persist in a seed bank (Deil, 2005; Tuckett et al., 2010b). Convergence in life history likely occurred in other plants that occupy the ephemeral wetlands of Australia and elsewhere (Deil, 2005; Diels, 1906). Dwarf components of these ecosystems include members of Centrolepidaceae (*Centrolepis* and *Aphelia*), *Drosera* (Droseraceae), Eriocaulaceae, *Hydrocotyle* (Araliaceae), *Isoetes* (Isoetaceae), *Juncus* (Juncaceae), *Lasthenia* (Asteraceae) and *Utricularia* (Lentibulariaceae). It would be worth investigating whether the patterns shown here for Hydatellaceae are also found in these distantly related plants. The ephemeral aquatic habitat occupied by Hydatellaceae contrasts sharply with the relatively stable and mesic forests occupied by other early angiosperm lineages (e.g., Feild et al. 2009). However, while extensive extinction may have characterized the family before and after its recent radiation, the family is also a successful long-term survivor of marginal and unpredictable aquatic environments.

Chapter 4

Phylogeny and gene loss in Alismatales

4.1 Summary

The earliest phylogenetic divisions in monocot phylogeny are now fairly well understood. The phylogenetic structure of the largest of the early diverging clades, Alismatales, is also becoming clear, but the relative branching order of much its phylogenetic backbone is still poorly resolved, including inter- and intra-family relationships among the constituent families (Araceae, Tofieldiaceae, and a core alismatid group comprising ∼12 families). Here I address these problems using a subset of the plastid genome (∼14 kb per taxon) from 30+ representative taxa in the order, surveying one or more exemplar taxon per family. I recovered a strongly supported phylogenetic backbone for the order Alismatales. In general these are congruent with the backbone recovered by previous authors using a single gene (*rbc*L), but with substantially improved support in most cases. I confirm that the eastern Australian endemic *Maundia* (Juncaginaceae) is the sister group of a clade containing 4–5 families. I also resolve several sets of relationships that were unclear previously, with strong support here, including the placement of Aponogetonaceae as the sister group of the remaining tepaloid families, and a sister-group relationship between Butomaceae and Hydrocharitaceae. Rate variation in plastid genes in Alismatales is shown to have a strong impact on the relative phylogenetic arrangement of three petaloid alismatid families (Alismataceae, Butomaceae and Hydrocharitaceae) when taxon density is sparse; increased taxon sampling ameliorates this spurious

effect. Finally, I note that several taxa have internal stop codons in sampled *ndh* genes, and their phylogenetic distribution suggests several independent losses of NADH dehydrogenase function.

4.2 Introduction

The order Alismatales is a cosmopolitan and enormously diverse clade of monocotyledons, comprising ∼4500 extant species in 13 families, as currently defined (Angiosperm Phylogeny Group, 2009; Janssen and Bremer, 2004; Stevens, 2001). Some of the oldest monocot fossils (late Barremian and early Albian; 125–112 Ma) have been assigned to this lineage (Friis et al., 2004, 2010), and most phylogenetic studies (e.g., Chase et al., 2006; Givnish et al., 2010, 2006; Graham et al., 2006) resolve Alismatales as the sister group of all monocots except *Acorus* (Acorales: Acoraceae). Refining our understanding of the phylogenetic backbone of Alismatales will therefore be important for understanding the early evolutionary history of the monocots.

The overall composition of Alismatales remained relatively constant until a recent expansion to include Araceae and Tofieldiaceae (e.g., Angiosperm Phylogeny Group, 1998, 2003, 2009; Chase, 2004; Dahlgren and Clifford, 1982; Les et al., 1997; Tomlinson, 1982). This shift reflects substantial molecular systematic evidence (e.g., Chase et al., 1995, 2006, 2000; Duvall et al., 1993; Givnish et al., 2006; Graham et al., 2006; Tamura et al., 2004b) for a close relationship between subclades corresponding to these two additional families and a clade of "core" alismatid families that corresponds approximately to the order Helobiae (Engler, 1892) and subclass Alismatidae (Cronquist, 1988). Les and Tippery (2013) favour a narrower definition of the clade (as Alismatidae, with two orders, and excluding Araceae and Tofieldiaceae), but I find the broader circumscription of the order more appealing, because it underlines the evolutionary links among these diverse lineages. *Acorus* has also sometimes been recovered within Alismatales (e.g., Davis et al., 2006, 2004), but this placement may reflect substantial rate elevation in several mitochondrial genes (Cuenca et al., 2010; Mower et al., 2007; Petersen et al., 2006a,b). There have been multiple morphological

and molecular phylogenetic studies of individual families and major genera of Alismatales (Azuma and Tobe, 2011; Cabrera et al., 2008; Cusimano et al., 2011; Iida et al., 2004; Ito et al., 2010; Jacobson and Hedrén, 2007; Kato et al., 2003; Keener, 2005; Lehtonen, 2006, 2009; Lehtonen and Myllys, 2008; Les et al., 1997, 2002, 1993; Lindqvist et al., 2006; Rothwell et al., 2004; Tamura et al., 2010, 2004a; Tanaka et al., 1997, 2003; von Mering and Kadereit, 2010; Wang et al., 2007; Waycott et al., 2002, 2006; Zhang et al., 2008). However, only a few studies (e.g., Les et al., 1997) have surveyed taxa broadly enough to representatively sample the phylogenetic backbone of the order.

Les et al. (1997) provided the most comprehensive study of higher-order relationships in Alismatales. They sampled the plastid gene *rbc*L for exemplar taxa representing all families except Tofieldiaceae, and most of the genera except in Araceae. In addition to improving our knowledge of phylogenetic relationships in the order, and refining family-level circumscriptions, they were interested in reconstructing the evolution of characters that may be associated with hydrophilous (water-mediated) pollination. The core alismatid families are mostly fully aquatic (Les et al., 1997), and semi- to fully aquatic plants are also found in Araceae and Tofieldiaceae, possibly consistent with an aquatic or semi-aquatic habit for the most recent common ancestor of the monocots (e.g., Chase 2004; note that *Acorus* is also semi-aquatic). If so, terrestrial species in the order (i.e., most Araceae, some Tofieldiaceae) may therefore represent subsequent reversions in habit. The order encompasses all major aquatic life forms (i.e., emergent, floating-leaved, free-floating and submersed; Sculthorpe 1967), and includes the only fully marine angiosperms, the seagrasses, a life form that evolved several times in the order (Les et al., 1997). Morphological features linked to hydrophily and an aquatic habit are expected to have an unusually high level of homoplasy, which may have contributed to the fluidity of earlier family-level classification schemes based on morphology (see Les et al., 1997; Les and Haynes, 1995). Les et al. (1997) reconstructed the overall phylogenetic backbone of the order using a single plastid gene, and recovered multiple poorly to moderately supported branches underpinning the higher-order re-

lationships. The monophyly and extent of several families were also unclear (this latter uncertainty was partly accommodated in the APG classification systems by the expanded circumscription of several families).

A few studies have revisited their *rbc*L dataset, either alone or in combination with morphology (Chen et al., 2004a,b; Li and Zhou, 2009), but no subsequent studies have sampled the order broadly using additional genes, with the exception of a suite of papers focused primarily on mitochondrial gene evolution (Cuenca et al., 2010; Petersen et al., 2006a). Here I substantially expand the number of plastid genes sampled from exemplar species that represent the broad phylogenetic backbone of the order. My major goal is to re-examine and further refine the overall backbone of Alismatales phylogeny recovered by Les et al. (1997) by considering more plastid data per taxon. This general approach has proved to be effective for the inference of broad-scale monocot phylogeny (e.g., Givnish et al., 2010; Graham et al., 2006; Saarela and Graham, 2010; Saarela et al., 2008). I confirm much of the broad phylogenetic backbone recovered by Les et al. (1997), with some notable exceptions. I also obtain substantially improved branch support in many cases. However, I demonstrate that too limited taxon sampling can lead to spurious inference of some local relationships when using plastid genes, which may be a consequence of elevated rates of evolution in a subset of regions examined. Finally, I document and characterize multiple independent losses of plastid genes that code for two subunits of the plastid NADH dehydrogenase chlororespiratory complex.

4.3 Material and methods

4.3.1 Taxon sampling

Our main analyses focus on a set of 92 exemplar (representative) species comprising 31 species from Alismatales, 49 other monocots and 12 other angiosperms. I expanded taxon sampling in Alismatales by 26 species compared to our most recent broad study of monocot phylogeny (Saarela & Graham 2010), and included all currently recognized families in the order

(Appendix D). Our overall taxon sampling for Alismatales is generally less dense than Les et al. (1997), but the included lineages constitute a highly representative subsample of the broad backbone of Alismatales phylogeny. As far as possible I included multiple representatives per family and targeted species within families that span their deepest phylogenetic splits, at least as defined in Les et al. (1997). I included the south-eastern Australian endemic *Maundia triglochinoides* because of a recent report that it lies outside Juncaginaceae (von Mering and Kadereit, 2010), rendering that family paraphyletic as currently circumscribed (Les et al. 1997). Our most complete generic sampling in the order is in Tofieldiaceae, with four of its five genera included (only *Isidrogalvia* is not sampled).

Outside Alismatales I excluded some taxa that were included previously (Graham and Olmstead, 2000; Graham et al., 2006; Saarela and Graham, 2010; Saarela et al., 2008, 2007) to facilitate maximum likelihood analysis, but our taxon sampling is broadly representative of Petrosaviidae (Graham et al., 2013); this name was coined for the large clade that encompasses all monocots except *Acorus* and Alismatales). I also included new sequences for exemplar species from each of the following families: Acoraceae (Acorales), Bromeliaceae (Poales), Nartheciaceae (Dioscoreales), Nymphaeaceae (Nymphaeales), Orchidaceae (Asparagales), and Philesiaceae and Rhipogonaceae (Liliales); see Appendix D for details.

4.3.2 Gene sampling

I extracted total genomic DNAs from silica-gel dried leaf material (Appendix D) using standard protocols (Doyle and Doyle, 1987; Graham and Olmstead, 2000), or by using a DNeasy Plant Mini Kit (Qiagen Inc, Valencia, California, USA) for recalcitrant material. Some DNAs were provided by the Royal Botanic Gardens, Kew. In several cases I included sequences from GenBank (*Nuphar*, *Phalaenopsis*; several eudicots) or from other workers (*Vallisneria*; Appendix D). In total I sampled 17 plastid genes and associated noncoding regions (omitting several noncoding regions from analysis, see below). These genes are involved in several dif-

ferent plastid functions: photosynthesis (*atp*B, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*J, *psb*L, *psb*N, *psb*T, *rbc*L), chlororespiration (*ndh*B, *ndh*F) and protein translation (*rpl*2, *rps*7, 3'–*rps*12). Our sample includes the following multigene clusters: *psb*B-*psb*T-*psb*N-*psb*H (which I refer to here as *psb*BTNH), *psb*E-*psb*F-*psb*L-*psb*J (= *psb*EFLJ), and 3'-*rps*12-*rps*7-*ndh*B-*trn*L(CAA). I surveyed these regions using amplification and sequencing protocols noted in Graham & Olmstead (2000) and Saarela et al. (2008), and designed several modified primers for the *psb*BTNH region: modB60F (5'-CATACAGCTTTAGTTGCTGGTTGG), modB64R (5'-GGGATCAGGGATATTTCCAGCAAG), mod65R (5'-GGAAATGTTTCAAAAAAAGTAGGCA) and modB71R (5'- CCCGGCGCCACTTTACCATATTC).

4.3.3 Data assembly

I carried out base-calling and contig assembly using Sequencher 4.2.2 (Gene Codes Corp., Ann Arbor, Michigan, USA), determining gene boundaries using tobacco and *Ginkgo* reference sequences (Saarela et al. 2008). I added the new sequences to an existing alignment (Saarela et al. 2008), which I adjusted manually using Se-Al 1.0 (Rambaut 1998) following criteria described in Graham et al. (2000). I coded gaps as missing data. The total aligned length is 23 903 bp, a large portion of which consists of "offset" noncoding regions that are unique to individual taxa (for a justification of this approach see Saarela and Graham 2010). For comparison, the unaligned sequence lengths for the newly determined sequences range from 11 009 bp for *Najas* to 15 560 bp for *Stratiotes*. I recovered all 17 gene regions from most species (Appendix D). However, for a subset of taxa the *ndh* genes appear to be pseudogenes (i.e., their reading frames are interrupted by stop codons, out-of-phase indels, or both; see below). I recovered a probable *ndh*F pseudogene from *Amphibolis*, and *ndh*B pseudogenes from *Amphibolis*, *Najas*, *Posidonia* and *Thalassia* (partial sequences in several cases, see below). I could not retrieve *ndh*F for *Najas*, *Posidonia* and *Thalassia*, despite extensive attempts at amplification. The apparently pseudogenized *ndh* genes were generally straightforward to align, and so I included them in

the analyses. However, a possible *ndh*F pseudogene sequence for *Vallisneria* was so divergent that it could not be aligned reliably, and other *ndh* genes were not recovered for this taxon (M. Moore, Oberlin College, Ohio; pers. comm.).

4.3.4 Phylogenetic analysis

I focussed on coding regions and several conservative noncoding regions from the plastid IR region for the main analysis, following Saarela et al. (2007) and Graham and Iles (2009); the included noncoding regions are intergenic spacers in the contiguous region spanning 3'-*rps*12, *rps*7, *ndh*B and *trn*L, and single introns in each of *rpl*2, 3'-*rps*12 and *ndh*B. I performed heuristic maximum parsimony (MP) searches using PAUP* (Swofford, 2003) with 100 random addition replicates and tree-bisection-reconnection branch swapping, and otherwise using default settings. I used RAxML version 7.2.6 (Stamatakis, 2006b; Stamatakis et al., 2008) at the Bioportal website (`www.bioportal.uio.no`) to perform maximum likelihood (ML) analyses. jModelTest (Posada, 2008) was employed to infer the optimal DNA substitution model using the AICc (the Akaike Information Criterion, correcting for sample size) considering the full matrix or subpartitions (see below) for Alismatales only. The GTR+Γ or GTR+Γ+I models were selected in all cases (GTR is the general-time-reversible model, the gamma distribution [Γ] accounts for among-site rate heterogeneity, and the "I" parameter accommodates invariable sites). Previous analyses of the same gene set across monocots as a whole (e.g., Saarela and Graham, 2010; Saarela et al., 2008) favoured the GTR+Γ+I model. I omitted the I parameter here, as it may be adequately accounted for using the gamma distribution alone (Yang, 2006). I initiated the ML search from 104 random MP starting trees (multiples of eight are required on the Bioportal website), retaining the tree with the highest likelihood score across all searches. I also performed a partitioned ML analysis by distinguishing four partitions, one for each codon position and a separate one for the set of noncoding regions included here, but otherwise using the same settings and general DNA substitution model. I evaluated

branch support using the non-parametric bootstrap (Felsenstein, 1985b). I considered 500 (MP) or 104 (ML) bootstrap replicates using the search settings described above, but with 10 random addition replicates (MP) or a single random starting tree (ML) per bootstrap replicate. I use the terms "weak," "moderate," and "strong" to refer to bootstrap support values recovered in the ranges <70%, 70-89%, and >90% respectively (Graham et al., 1998).

In earlier unpublished analyses using fewer taxa for Alismatales it was noticed (Sean W. Graham, UBC, pers. comm.) that inferred phylogenetic relationships among three families (Alismataceae, Butomaceae and Hydrocharitaceae) depended strongly on the regions and phylogenetic criteria used, and sometimes conflicted with the main results reported here. To explore the possibility that this effect was related to taxon sampling, phylogenetic method or rate heterogeneity in plastid genes, I performed multiple ML and MP analyses for different gene and taxon subsamplings. Specifically, I ran ML and MP analyses on various subsets of the plastid genes, in addition to the full set of regions, using the search settings described above (although in some MP bootstrap analyses I set a MaxTrees limit of 1000 trees). I repeated these analyses for two different taxon densities in Alismatales, a "reduced" taxon set of 11 exemplar taxa, vs. a "dense" taxon set comprising all 31 exemplar species. I used two outgroups for these analyses: *Acorus calamus* (Acoraceae) and *Japonolirion osense* (Petrosaviaceae).

4.4 Results

4.4.1 The phylogenetic backbone of Alismatales

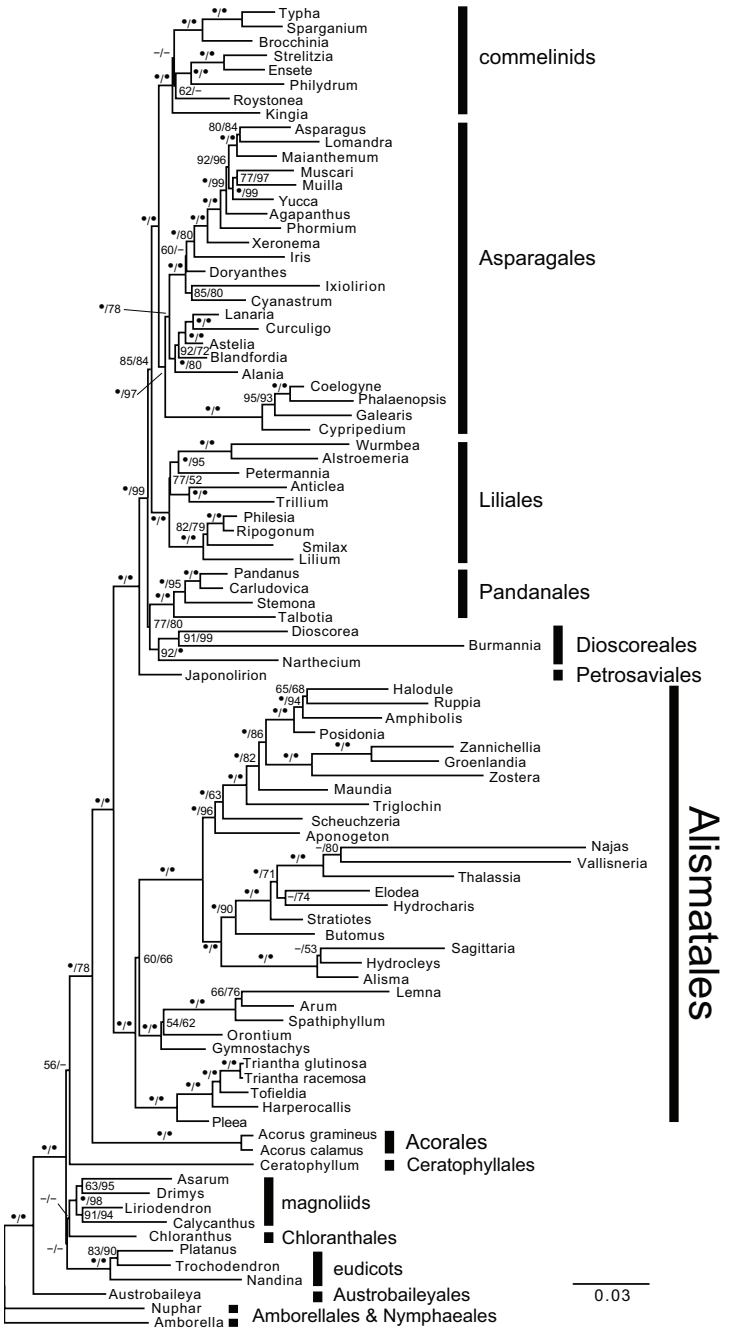
Outside Alismatales, the backbone relationships inferred from the full combined data set for 92 taxa are broadly similar to other estimates using these genes (Figure 4.1, cf. Chapter 5 and Graham et al., 2006), and so I do not discuss them further here. A portion of the (unpartitioned) ML tree representing Alismatales is presented in Figure 4.2; considering four data partitions in ML analysis did not result in a substantially different topol-

ogy (not shown; it differed in one poorly supported branch inside Araceae). The MP analysis yielded a single most parsimonious tree (tree length = 26 194 steps) that is also highly congruent with the unpartitioned ML tree for Alismatales (not shown). Unpartitioned ML and MP bootstrap values are noted beside individual branches in Figure 4.2; partitioned ML values are indicated in Table 4.1. To facilitate comparisons across analyses and to other studies I have also tabulated support values from the various ML and MP analyses for a subset of branches (Table 4.1; labelled with letters in Figure 4.2); these correspond to interfamilial relationships in the order, in addition to two branches that contradict the monophyly of Juncaginaceae and Cymodoceaceae, respectively. These major backbone relationships in Alismatales are generally strongly supported (ML) or strongly to moderately supported (MP) by the 17-gene data (summarized in the second major column in Table 4.1).

Well supported clades include Alismatales as a whole (branch a), the core alismatid clade (branch c), a "petaloid" clade (branch d, comprising three core alismatid families; Les and Tippery 2013, refer to this subclade as Alismatales), a "tepaloid" clade (branch f, comprising the remaining eight core alismatid families; Les and Tippery 2013 refer to this subclade as Potamogetonales) and most other branches (branches e and h-m). Core alismatids were distinguished as having either petaloid or tepaloid perianths by Posluszny and Charlton (1993); note that taxa lacking obvious perianths (e.g. *Halodule* and *Najas*) belong to both clades (Figure 4.2). In some cases MP bootstrap support values are marginally (10–20%) weaker than the corresponding ML values (i.e., branch g, which defines the first split in

---

Figure 4.1 *(following page)*: Placement of Alismatales in monocot phylogeny. The ML tree (-*ln*L = 158 904.167) for all 92 taxa. The tree is based on 17 plastid genes and several associated non-coding regions (see text for details). Support values based on bootstrap analysis are noted beside branches (left-hand value = unpartitioned ML, right-hand value = MP); filled circles indicate 100% bootstrap support, dashes <50% bootstrap support. Scale: substitutions per site.


Typha
Sparganium
Brocchinia
Strelitzia
Ensete
Philydrum
Roystonea
Kingia
Asparagus
Lomandra
Maianthemum
Muscari
Muilla
Yucca
Agapanthus
Phormium
Xeronema
Iris
Doryanthes
Ixiolirion
Cyanastrum
Lanaria
Curculigo
Astelia
Blandfordia
Alania
Coelogyne
Phalaenopsis
Galearis
Cypripedium
Wurmbea
Alstroemeria
Petermannia
Anticlea
Trillium
Philesia
Ripogonum
Smilax
Lilium
Pandanus
Carludovica
Stemona
Talbotia
Dioscorea
Burmannia
Narthecium
Japonolirion
Halodule
Ruppia
Amphibolis
Posidonia
Zannichellia
Groenlandia
Zostera
Maundia
Triglochin
Scheuchzeria
Aponogeton
Najas
Vallisneria
Thalassia
Elodea
Hydrocharis
Stratiotes
Butomus
Sagittaria
Hydrocleys
Alisma
Lemna
Arum
Spathiphyllum
Orontium
Gymnostachys
Triantha glutinosa
Triantha racemosa
Tofieldia
Harperocallis
Pleea
Acorus gramineus
Acorus calamus
Ceratophyllum
Asarum
Drimys
Liriodendron
Calycanthus
Chloranthus
Platanus
Trochodendron
Nandina
Austrobaileya
Nuphar
Amborella
commelinids
Asparagales
Liliales
Pandanales
Dioscoreales
Petrosaviales
Alismatales
Acorales
Ceratophyllales
magnoliids
Chloranthales
eudicots
Austrobaileyales
Amborellales & Nymphaeales
0.03

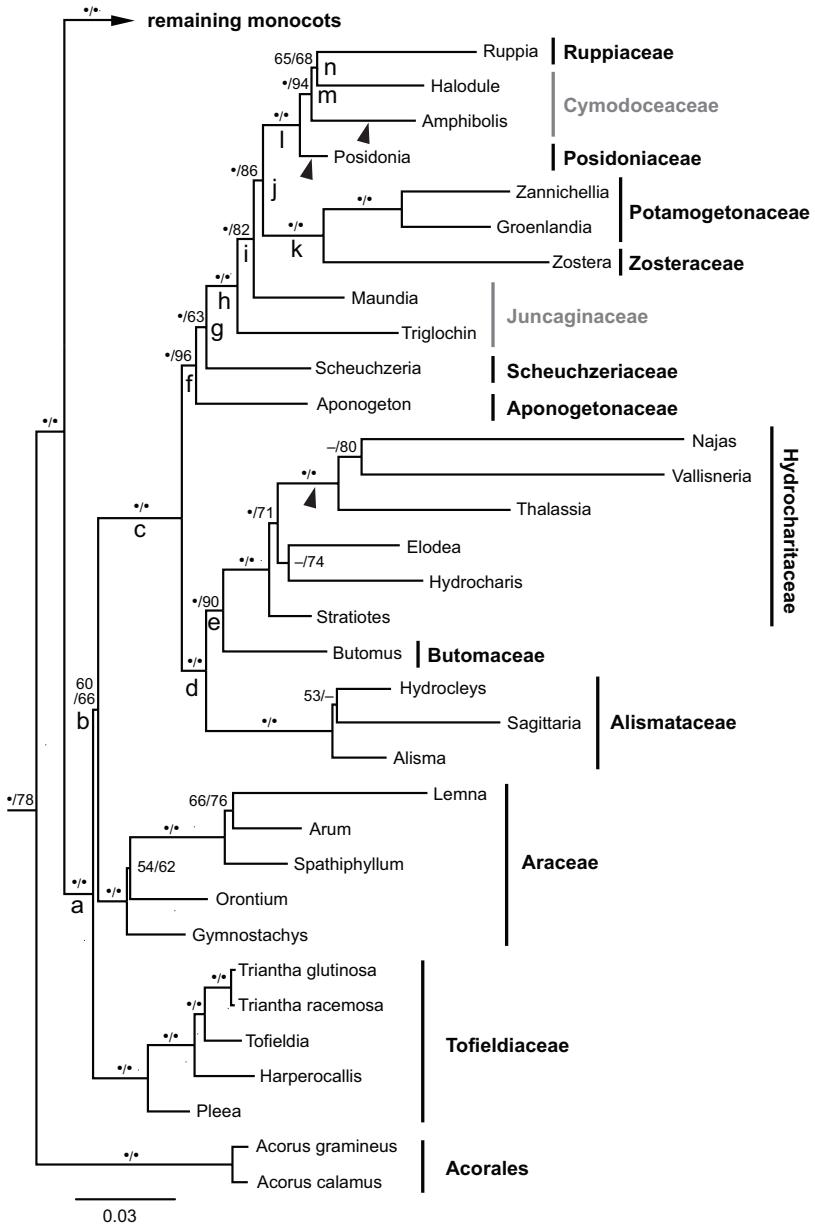


Figure 4.2: A section of the ML tree (see Figure 4.1), focused on Alismatales. Two families that are not resolved as monophyletic are noted in grey. Letter labels refer to clades noted in the text and Tables 4.1 and 4.2. Arrowheads indicate putatively independent losses of one or more *ndh* (NADH dehydrogenase subunit) loci. Scale and support values as for Figure 4.1

the tepaloid clade above its root node; branch i, which rejects monophyly of Juncaginaceae by placing *Maundia* as the sister group of five families in the tepaloid clade; branch j, for the clade comprising these five families). Two of these three are moderately supported by MP (branches i, j), but all three have strong support (94–100%) from unpartitioned and partitioned ML analyses.

A branch that contradicts the monophyly of Cymodoceaceae here (branch n, which links *Ruppia*, Ruppiaceae, a monogeneric family, with one of the two sampled genera of Cymodoceaceae, *Halodule*; Figure 4.2) is only weakly to moderately supported by all three methods. All other families with multiple exemplar species are strongly supported as monophyletic at the taxon sampling here, and several families with denser sampling also have well supported internal phylogenetic structure. Specifically, all three internal branches in Tofieldiaceae have strong support, including a placement of *Pleea* as the sister group of the remaining genera, and of *Harperocallis* as the sister group of *Tofieldia-Triantha*; two of four internal branches in Hydrocharitaceae are strongly supported, including a placement of *Stratiotes* as the sister group of other Hydrocharitaceae (Figure 4.2) at the current sparse taxon sampling for this family.

The only major relationship that is not well supported in Alismatales concerns the relative arrangement of its three major subclades: Araceae, Tofieldiaceae and the core clade of alismatid families. Branch b, recovered in the best ML trees here, depicts Araceae as the sister group of the core alismatid families (hence, Tofieldiaceae are the sister group to these two clades, as the order as a whole is also strongly supported). However, this arrangement receives relatively weak support from all three phylogenetic criteria (i.e., 62–66% support, Table 4.1). The two other possible arrangements for these three clades have been recovered elsewhere with weak to strong support (compare column 2 with columns 4–7 in Table 4.1, which summarize relevant support values in Alismatales across several other studies). One of these alternative possibilities (Araceae sister to Tofieldiaceae; clade b3) has negligible support here, but the other (Tofieldiaceae sister to core alismatids; clade b2) has poor but non-negligible support (i.e., 31–40%

of bootstrap replicates from the 17-gene data).

4.4.2 Taxon density and branch support

Our reduced taxon set (for the full 17-gene set) has only 11 exemplar taxa from Alismatales, and so several major clades are no longer applicable when compared to the full taxon set (i.e., clades i-n; Figure 4.2 and Table 4.1). I did not run a partitioned ML analysis for this taxon set. The unpartitioned ML and MP bootstrap values are lower for several clades compared to the full taxon sampling (cf. columns 2 and 3 in Table 4.2). These reductions in support are more acute for MP than ML for two clades (i.e., for e and f). However, two clades saw marginal increases with fewer taxa sampled: clade b for ML and MP (this corresponds to Araceae + core alismatids), and clade g for MP only (this is a seven-family clade of core alismatids that corresponds to all of the tepaloid families except Aponogetonaceae; Figure 4.2).

The most surprising shift with the reduced taxon sampling was clade e, (Hydrocharitaceae + Butomaceae) which is weakly (ML) to negligibly (MP) supported (column 3 in Table 4.1); this lineage of petaloid alismatids was a strongly supported clade for the full taxon sampling (Figure 4.2; column 2 in Table 4.1). In contrast, a conflicting clade (clade e2; Hydrocharitaceae + Alismataceae) that previously had negligible (<5%) support (Table 4.1), receives non-negligible bootstrap support from ML (31%) and strong support from MP (99% support) with the reduced taxon sampling. I examined the conflicting signal for these two contrasting branches by performing bootstrap analyses on subsets of the full plastid data set that correspond to individual genes or sets of genes (Table 4.2). I repeated these analyses for two different taxon densities (i.e., 11 vs. 31 ingroup taxa, the "reduced" and "dense" taxon samplings in Table 4.2), and by considering two optimality criteria (i.e. unpartitioned ML and MP). At the reduced taxon sampling, branch e, corresponding to Hydrocharitaceae + Butomaceae, is strongly supported by some data partitions for ML (e.g., coding regions only; single copy plastid genes; *ndh*B) and weakly supported by others (*psb*BTNH); no data parti-

tion supported this arrangement even moderately well for MP. In contrast, the conflicting relationship, branch e2 (corresponding to Hydrocharitaceae + Alismataceae) is moderately to strongly supported by multiple data partitions for MP at this weak taxon density, including some partitions that have strong support for the contrasting relationship for ML (e.g., the combined coding plastid regions).

Likelihood and parsimony analyses appear to converge to clade e at the dense taxon sampling (Table 4.2). This taxon sampling substantially breaks up the long terminal branches subtending *Najas* (Hydrocharitaceae) and *Sagittaria* (Alismataceae) in Figure 4.3A (note how both are divided for at least part of their length in Figure 4.2). One exception concerns the analysis of *rpl*2 alone, which recovers branch e2 for the reduced taxon sampling with strong support from ML and MP, but which in contrast has no well supported relationship at the dense taxon sampling, for either phylogenetic method (Table 4.2). This gene has the greatest disparity of branch lengths for any data partition considered here (e.g., compare the relative lengths of the terminal branches subtending *Najas* and *Sagittaria* for the single-copy coding regions combined vs. *rpl*2 alone in Figure 4.3).

4.4.3 Parallel loss of *ndh* genes in the core alismatid clade

I predict that one or more plastid *ndh* loci have been lost independently in multiple lineages of core alismatids, based on the phylogenetic distribution of taxa that have accumulated stop codons in *ndh*B, *ndh*F or both (see arrowheads in Figure 4.3). These include a loss in the petaloid alismatids (perhaps in the common ancestor of *Najas*, *Thalassia* and *Vallisneria*, which comprise a well supported subclade of Hydrocharitaceae), and two parallel losses in the tepaloid alismatid clade (one in Cymodoceaceae, one in Posidoniaceae). Although *Amphibolis* and *Posidonia* are close relatives (Figure 4.2), the former species is closely related to other sampled taxa (*Halodule* and *Ruppia*) that have retained open reading frames for *ndh*B and *ndh*F, supporting a convergent loss of function.

The internal stop codons in the *ndh*B locus are due to DNA substitutions

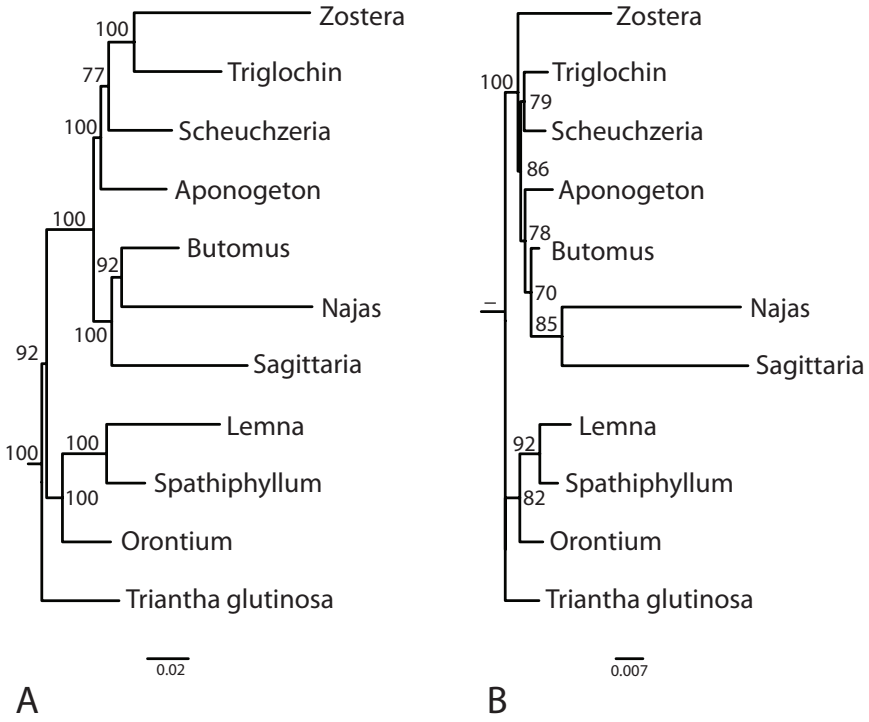


Figure 4.3: ML trees inferred for the reduced taxon set for two different plastid gene subsets, demonstrating contrasting patterns of rate heterogeneity. The taxon set includes 11 ingroup taxa (and two outgroups, *Acorus calamus* and *Japonolirion osense*, not shown here). A) Concatenated coding regions from the plastid single copy regions only; B) The *rpl*2 locus (two exons and an intron). Numbers adjacent to branches are ML bootstrap support values (dashes indicates <50% support; note the basal trichotomy in B). Scale: substitutions per site.

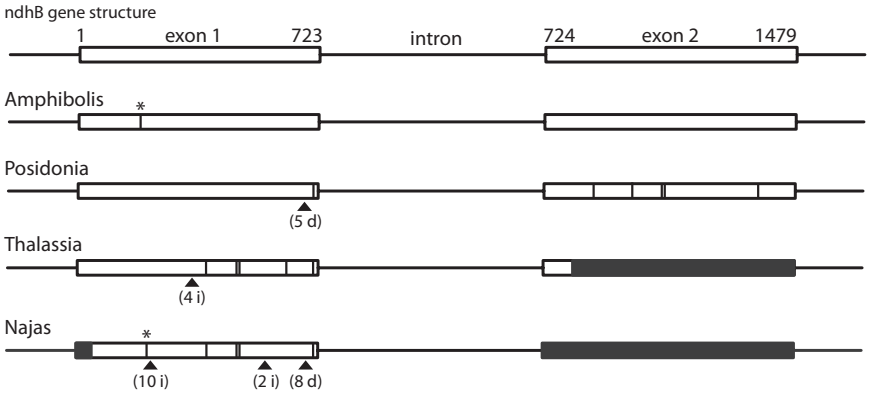


Figure 4.4: Putative pseudogenes of the plastid gene *ndh*B from *Amphibolis*, *Najas*, *Posidonia* and *Thalassia* (top panel; nucleotide positions for exon boundaries are noted relative to tobacco; the intron is not completely to scale across taxa). Vertical lines indicate predicted internal stop codons; those annotated with asterisks are inferred to be due to substitution events, whereas the remainder are a consequence of frameshifts following one or more non-triplet indel events noted below the first exon (size in bp noted; i = inferred insertion, d = inferred deletion). Shaded regions were not recovered.

in two cases (*Amphibolis*, *Najas*) and to indels in *Najas*, *Posidonia* and *Thalassia* (summarized in Figure 4.4). Reading-frame shifts resulting from these indels reveal otherwise out-of-frame stop codons that were present in their common ancestor. However, the relevant indels are not shared among them (Figure 4.4). I amplified only partial *ndh*B genes for two taxa (*Najas* and *Thalassia*) and was unable to retrieve another plastid-encoded NADH dehydrogenase subunit gene, *ndh*F, from three species (*Najas*, *Posidonia* and *Thalassia*). The *ndh*F locus recovered for *Amphibolis* also has multiple indels and stop codons. The putative *ndh*F sequence from *Vallisneria* recovered by Moore and colleagues (not included in analysis here) appears to be highly degraded.

4.5 Discussion

4.5.1 A refined phylogenetic backbone of Alismatales

Our phylogenetic inferences based on 17 slowly evolving plastid genes and several associated noncoding regions are generally highly congruent with the backbone that was inferred by Les et al. (1997) who used more ingroup taxa but only a single gene (see also Les and Tippery 2013). Only five of the internal branches that correspond to interfamilial relationships (summarized in Table 4.1) were moderately to strongly supported by Les et al., with a similar or smaller number of branches supported in other recent studies that have a moderately representative sampling of the backbone (12–37 exemplar species from the order sampled in Chase et al. 2006, Davis et al. 2006, and von Mering and Kadereit 2010; see Table 4.1). In contrast, I recover strong support for 11 of 12 of these branches using ML bootstrap analysis (for partitioned and unpartitioned ML; branches a, c–m, ignoring several poorly supported alternative relationships noted in Table 4.1). MP bootstrap support for these branches was also generally comparable for the full taxon set, although there was only moderate MP bootstrap support for branches g, i and j.

Table 4.1: Comparison of the support values for interfamilial relationships. In the case of clades i and n, support for the relationships disrupting family monophyly. Author abbreviations: Les et al. (1997; their Fig. 2) = L97; Chase et al. (2006; their Fig. 2) = C06; Davis et al. (2006; their Fig. 2) = D06; von Mering & Kadereit (2010; their Fig. 3) = vM&K10. ML-p = partitioned ML; pt = plastid. A dash ("–") means support was not noted or assessed in the corresponding study; a "<" means the branch had <50% support (<70% in von Mering & Kadereit, 2010); "na" = not applicable due to disrupted monophyly. Clade labels are depicted in Figure 4.2, except for those with a number (b2, b3, etc). Notes as follows: [a] According to APG (2009); [b] 92 exemplar species here in the full taxon set (including 61 outgroups), vs. 13 exemplars in the reduced taxon set (including two outgroups); [c] *rbc*L only; [d] Four plastid genes, two nuclear genes, one mitochondrial gene; [e] Two plastid genes, two mitochondrial genes.

		Full taxon set[b]	Reduced taxon set[b]	L97	C06	D06	vM&K10
	No. of exemplar species (Alismatales):	31 species	11 species	78 species	13 species	12 species	37 species
	No. of genes:	17 pt genes	17 pt genes	1 gene[c]	7 genes[d]	4 genes[e]	1 gene[c]
	Branch support determined using:	ML, ML-p (MP)	ML (MP)	MP	MP	MP	ML (MP)
Clade label	and description[a]						
a	Alismatales	100, 100 (100)	100 (99)	–	100	na	– (–)
b	Araceae + core alismatids	60, 62 (66)	85 (82)	–	–	na	– (–)
b2	Tofieldiaceae + core alismatids	40, 34 (31)	15 (16)	–	99	na	– (–)
b3	Tofieldiaceae + Araceae	<5, <5 (<5)	<5, <5	–	–	na	72 (<)
c	Core alismatid clade	100, 100 (100)	100 (100)	96	100	100	93 (95)
d	Petaloid clade	100, 100 (100)	100 (100)	88	87	<	92 (96)
e	Hydrocharitaceae + Butomaceae	100, 100 (90)	69 (<5)	31	–	<	<(<)
e2	Hydrocharitaceae + Alismataceae	<5, <5 (<5)	31 (99)	<	<	<	– (–)
f	Tepaloid clade	100, 100 (96)	77 (41)	77	–	–	77 (<)
g	Tepaloid clade excl. Aponogetonaceae	100, 94 (63)	92 (85)	–	–	–	– (–)

		Full taxon set[b]	Reduced taxon set[b]	L97	C06	D06	vM&K10
h	Tepaloid clade excl. Aponogetonaceae + Scheuchzeriaceae	100, 100 (100)	100 (100)	78	–	100	88 (87)
i	*Maundia* + Ruppiaceae/Cymodoceaceae + Posidoniaceae + Potamogetonaceae + Zosteraceae	100, 100 (82)	– (–)	–	–	–	<(71)
j	Ruppiaceae/Cymodoceaceae + Posidoniaceae + Potamogetonaceae + Zosteraceae	100, 100 (86)	– (–)	71	–	98	– (–)
k	Potamogetonaceae + Zosteraceae	100, 100 (100)	– (–)	100	100	–	99 (99)
l	Ruppiaceae/Cymodoceaceae + Posidoniaceae	100, 100 (100)	– (–)	40	–	–	77 (<)
m	Ruppiaceae/Cymodoceaceae	100, 91 (94)	– (–)	29	–	–	<(<)
n	*Ruppia* + *Halodule*	65, 73 (68)	– (–)	–	–	–	– (–)

Our new data provide consistently strong support for the division of the core alismatid families into "petaloid" vs. "tepaloid" clades, each comprising multiple families, and they also resolve several major branches in each case that until now have been neither well supported nor consistently resolved across studies. Butomaceae and Hydrocharitaceae are strongly supported as sister groups in the petaloid alismatids for the dense taxon sampling. In the tepaloid alismatids I find strong support for several major branches that have been recovered before but with only moderate to weak support (e.g., the clade comprising Ruppiaceae, Cymodoceaceae and Posidoniaceae). For the first time I resolve the deepest splits in the tepaloid alismatids with strong support: Aponogetonaceae and Scheuchzeriaceae are respectively the successive sister groups of the remaining families in this subclade.

The monophyly of most families of Alismatales as currently circumscribed (APG 2009) is also confirmed with strong bootstrap support here (Figure 4.2), within the limits of the current taxon sampling (an advance on previous studies in several cases). Cymodoceaceae may well be paraphyletic, although this result is not strongly supported here (Figure 4.2). I also confirm the finding of von Mering and Kadereit (2010) that the family Juncaginaceae is not monophyletic as currently construed, as the two genera that I sampled from this family, *Triglochin* and *Maundia*, are successive sister groups of five other families in the tepaloid clade (Figure 4.2). This paraphyletic arrangement has strong ML bootstrap support, and moderately strong MP bootstrap support (Figure 4.2, Table 4.1). *Maundia* may therefore deserve to be recognized as its own family Maundiaceae Nakai (see Les and Tippery, 2013; Takhtajan, 1997; von Mering and Kadereit, 2010). It may be premature to do so until the remaining genera of Juncaginaceae are included in studies using gene samples that are comparable to or larger than ours, in case further paraphyly is uncovered. However, the remainder of Juncaginaceae are monophyletic in von Mering and Kadereit (2010), although the corresponding clade was not well supported in their analyses. Les and Tippery (2013) recovered a clade that included all genera of Juncaginaceae (excluding *Maundia*), with poor support, in a study that had an expanded taxon sampling compared to the original *rbc*L-based study of Les

et al. (1997). I am currently sampling additional taxa of Juncaginaceae s.s., and in preliminary analyses, I recover its monophyly with strong support (see Chapter 5).

APG III (2009) note that an alternative to recognising an additional family (Maundiaceae) in rank-based classifications would be to lump multiple families from all or most of the tepaloid clade into one family; Aponogetonaceae and Scheuchzeriaceae might be included in the resulting family as both are small, monogeneric families (see Backlund and Bremer, 1998). However defined, the resulting family would be referred to as Potamogetonaceae Bercht & J. Presl, as this name has priority (it corresponds to Potamogetonales as circumscribed by Les and Tippery 2013). It would be an extremely heterogeneous family, if circumscribed as such.

4.5.2 Rate heterogeneity and phylogenetic inference

I observed substantial rate heterogeneity in the plastid data among different lineages of Alismatales (Figure 4.2), and the order as a whole clearly includes some of the longest branches of angiosperms considered here (Figure 4.1). Although not as extreme as the rate elevation observed in some mitochondrial genes (Mower et al., 2007), I worried that this may have a misleading impact on phylogenetic inference in at least some cases, due to long-branch artefacts (Felsenstein, 1978; Hendy and Penny, 1989). Long-branch problems can be minimized by dense taxon sampling (e.g., Heath et al., 2008; Hedtke et al., 2006; Hillis, 1998; Hillis et al., 2003; Zwickl and Hillis, 2002), but it is probably the case that additional taxon sampling here would be unlikely to break up the broad Alismatales backbone much further, as I sampled all families and used multiple exemplar species per family where this was feasible. Nonetheless, our examination of a less dense taxon sampling for these genes demonstrated that strongly supported but likely spurious findings are still possible with plastid data (Table 4.2): we know that at least one of the strongly conflicting arrangements of Alismataceae, Butomaceae and Hydrocharitaceae found using fewer taxa for different plastid data partitions or phylogenetic criteria must be incorrect. In this particular case I favour

the arrangement that places Butomaceae as the sister group of Hydrocharitaceae, as this is what I see with the dense taxon sampling (Figure 4.2) and for most data partitions (Table 4.2) for parsimony and likelihood. Nonetheless, even model-based methods may be led astray by imperfectly modelled DNA substitution events (e.g., Matsen and Steel, 2007). This is a particular concern in Alismatales as there is fairly extensive heterogeneity in rate variation apparent among different plastid regions (e.g., Figure 4.3), although not as extreme as that observed for mitochondrial genes (see Cuenca et al., 2010; Petersen et al., 2006a,b). Improved maximum likelihood models (e.g., ones that take better account of heterotachy) may help in these situations. I am also encouraged by how the denser taxon sampling led to convergence of ML and MP across most of the regions considered here, and apparently removed the strong conflict for all data partitions examined (Table 4.2). Les and Haynes (1995) suggest several morphological synapomorphies for a clade comprising only Butomaceae and Hydrocharitaceae, lending further support to this arrangement.

Table 4.2: MP and ML support values for relationships among the petaloid families of Alismatales considering different taxon densities and data partitions; A = Alismataceae; B = Butomaceae; H = Hydrocharitaceae; the "<" symbol corresponds to less than 50% bootstrap support. All listed plastid partitions lack noncoding regions unless noted. Notes as follows: [a] The reduced ingroup sampling comprises 11 members of Alismatales vs. 31 species for the dense ingroup sampling; note that both samples include only two outgroups (*Acorus calamus* and *Japonolirion osense*); [b] Inverted repeat (IR) region = *rpl*2, 3'-*rps*12, *rps*7, *ndh*B, several intergenic spacer regions (between 3'-*rps*12 and *rps*7, *rps*7 and *ndh*B, and *ndh*B to *trn*L), and three introns (one each in *rpl*2, 3'-*rps*12 and *ndh*B); [c] 12 single copy loci (i.e. 10 *psb* genes, *atp*B and *rbc*L; *ndh*F was not retrieved here for *Najas*).

	Reduced ingroup sampling[a]		Dense ingroup sampling[a]	
Clade recovered:	[H + B]	[H + A]	[H + B]	[H + A]
Clade label (Table 4.1)	e	e2	e	e2
Phylogenetic criterion:	ML (MP)	ML (MP)	ML (MP)	ML (MP)
Data partition				
All plastid genes (+IR noncoding[b])	69 (<)	<(99)	100 (91)	<(<)
Coding plastid regions only	100 (<)	<(92)	100 (87)	<(<)
Single copy plastid genes[c]	92 (<)	<(88)	100 (68)	<(<)
*atp*B	<(<)	<(<)	77 (61)	<(<)
*psb*BTNH	55 (<)	<(<)	<(62)	<(<)
*psb*DC	<(<)	69 (90)	54 (62)	<(<)
*psb*EFLJ	<(<)	<(<)	<(<)	<(<)
*rbc*L	<(<)	<(<)	54 (<)	<(<)
Entire IR region (incl. noncoding)	<(<)	73 (100)	100 (98)	<(<)
*ndh*B (IR gene, incl. intron)	92 (65)	<(<)	100 (97)	<(<)
*rpl*2 (IR gene, incl. intron)	<(<)	85 (93)	<(<)	<(<)

4.5.3 Loss of NADH dehydrogenase subunit genes

The plastid *ndh* genes encode protein subunits of the thylakoid NADH dehydrogenase complex, which is homologous at a very deep level of phylogeny with the mitochondrial NADH dehydrogenase/complex I (Shinozaki et al. 1986). The plastid complex is encoded by eleven plastid genes and additional nuclear-encoded subunits. Plastid *ndh* genes have been retained in most embryophytes and charophytes (Martín and Sabater, 2010), but appear to have been lost frequently in heterotrophic plants (e.g., for several parasitic plants, DePamphilis and Palmer 1990; Stefanović and Olmstead 2005) along with other plastid genes, apparently associated with a loss or reduction in photosynthetic capability. In monocots they have been inferred to be lost in some orchids (e.g., Chang et al., 2006; Neyland and Urbatsch, 1996), although the full extent of their loss in this mycoheterotrophic family is unclear. There are only a few putative losses in fully autotrophic plants (the suite of eleven *ndh* plastid-encoded genes is absent or pseudogenized in extant Pinaceae, Gnetales and *Erodium* of Geraniaceae; Blazier et al. 2011; Braukmann et al. 2009), and so the possibility of multiple apparently independent losses of *ndh*B and *ndh*F within a single order is surprising (note that mycorrhizae are thought to be lacking in the order as currently circumscribed; see Wang and Qiu 2006).

While a complete understanding of the function of this complex is lacking, it has been implicated in chlororespiration, programmed cell death, and protection against photooxidative stress during photosynthesis (see Martín and Sabater, 2010). The latter may be important for understanding its evolutionary loss, as tobacco plants with experimentally induced non-functional *ndh* genes grow normally under optimal conditions, but not when environmentally stressed (Martín and Sabater, 2010). In Alismatales the losses could therefore be related to altered physiological constraints in an aquatic environment, such as reduced light stress in the subtidal zone (Martín and Sabater, 2010).

The phylogenetic distribution of *ndh* pseudogenes observed here suggests at least three independent losses in the order (Figure 4.2), but the

various ways in which *ndh*B has become pseudogenized (Figure 4.4) is potentially consistent with different scenarios of loss (however, note that the lack of shared stop codons is not in itself indicative of separate losses, as a common initial loss of function may have occurred in some unsampled *ndh* subunit, with subsequent independent pseudogenization in the other subunits). At this point I can only predict loss of function based on presence of stop codons (and to some extent the general difficulty of retrieving *ndh* loci in these taxa), and make predictions of convergent loss based on phylogenetic distribution. It would therefore be useful to characterize whether other *ndh* subunit genes from the plastid genome are also pseudogenized, and to confirm loss of function of the whole complex using physiological or transcriptome-based methods. A denser taxonomic sampling would also be helpful for characterising the extent and nature of independent *ndh*B and *ndh*F losses in the order.

4.5.4 Future work on the Alismatales backbone

Although I have refined our understanding of the broad phylogenetic backbone of Alismatales, additional work remains to be done. Several major branches within the order are very short according to the plastid data (e.g., branch b, subtending Araceae + core alismatids; major intra-familial splits within Araceae, Alismataceae, Hydrocharitaceae), and the putative paraphyly of Cymodoceaceae needs further confirmation. Some of these questions may profit from consideration of more genes per taxon (e.g., of the order of the whole plastid genome). Others may benefit from expanded taxon sampling within families for the current gene sampling. Within Tofieldiaceae, for example, the sub-tropical *Pleea* and *Harperocallis* form successive sister lineages to the predominately temperate and arctic *Tofieldia* and *Triantha*, which agrees closely with Tamura et al. (2010, 2004a) and Azuma and Tobe (2011). However, one South American montane genus, *Isidrogalvia*, was not included here. This genus may be the sister group of *Harperocallis* (Azuma and Tobe, 2011). Suggestions of a placement of *Isidrogalvia* within Nartheciaceae (Tamura et al., 2004a) appear to be the

result of misidentification or contamination (Azuma and Tobe, 2011).

4.6 Conclusions

I recovered a strongly supported phylogenetic backbone for the order Alismatales here using a large subsampling of the plastid genome from exemplar taxa that span all major clades in the order. Different phylogenetic criteria yielded the same underlying set of inferred higher-order relationships. In general these are congruent with the backbone recovered by Les et al. (1997) using a single gene (*rbc*L)—but with substantially improved support, in most cases. I confirm the non-monophyly of one family with strong bootstrap support (the placement of *Maundia* renders Juncaginaceae paraphyletic; see also von Mering and Kadereit 2010). I recover strong support for intergeneric relationships in several of the more densely sampled families (e.g., Tofieldiaceae). I also resolve several sets of relationship that were unclear previously, with strong support here, including the placement of Aponogetonaceae as the sister group of the remaining tepaloid families, and a sister-group relationship between Butomaceae and Hydrocharitaceae. Rate variation in plastid genes in Alismatales is shown to have a strong impact on the inference of the relative phylogenetic arrangement of three petaloid alismatid families (Alismataceae, Butomaceae and Hydrocharitaceae) when a sparse taxon density is considered, but a denser taxon sampling apparently ameliorates this effect. This suggests that future comparative genomic studies of the order (e.g., whole-plastome studies) should aim to include a density of taxa that approaches or exceeds that used here. Several taxa have internal stop codons in sampled *ndh* genes, and their phylogenetic distribution suggests several independent losses of NADH dehydrogenase function. The phylogenetic framework recovered here should be useful for other workers interested in exploring the evolution of morphological or other molecular characters in the order, facilitating our understanding of the origin and early evolution of the major clades of monocotyledons.

Chapter 5

Dating early events in monocot phylogeny

5.1 Summary

Despite fairly intensive investigation, there are still multiple gaps in our understanding of monocot higher-order relationships, and monocots have only rarely been the focus of clade-wide molecular dating analyses. A suite of recent monocot fossil discoveries and comprehensive reviews of the monocot and angiosperm fossil record make re-evaluating monocot divergence times timely. To address these issues I added unpublished and newly sequenced taxa to an existing monocot matrix based on 17 plastid genes, representing 172 species that cover 71 out of 78 monocot families. I re-evaluated overall monocot phylogeny using maximum parsimony, partitioned maximum likelihood and Bayesian methods. To date the monocot tree I used a Bayesian random local clocks model of molecular evolution, partly constraining some node dates using age distributions based on 17 monocot fossils (and eight other angiosperm fossils). The overall higher-order monocot phylogeny inferred here was congruent with other studies, and typically at least as well supported. I inferred that most higher-order monocot diversity arose in the Lower and Upper Cretaceous (stem ages for 55 of 68 families), but only eight monocot family crown clades originate in the Cretaceous; and only one crown age for a family dated back to the Lower Cretaceous (Araceae, Alismatales).

5.2 Introduction

The earliest molecular systematic studies dramatically reorganized our understanding of phylogenetic relationships in monocots (e.g., Duvall et al., 1993), the most species-rich lineage of angiosperms besides the eudicots (see Figure 1.1). Subsequent large-scale studies have helped to solidify our understanding of monocot phylogeny (Chase et al., 1995, 2006, 2000; Davis et al., 2006, 1998, 2004; Fuse and Tamura, 2000; Givnish et al., 2006; Graham et al., 2006; Nadot et al., 1995; Savolainen et al., 2000; Soltis et al., 2007, 2011, 2000; Tamura et al., 2004b). However, these developments were foreshadowed by detailed morphological and anatomical monographs (Dahlgren and Clifford, 1982; Dahlgren et al., 1985b; Dahlgren and Rasmussen, 1983) that suggested major short-comings in earlier monocot classifications (e.g., the heterogeneous family Liliaceae sensu Cronquist 1968). The most striking recent shift in our understanding was the excision of Hydatellaceae from Poales, with the revelation that it is the sister group of Nymphaeales (Saarela et al., 2007). Since that time, additional phylogenetic studies based on whole plastid genomes (plastomes) have begun to appear, although to date the focus of these studies has been narrow or has used relatively sparse taxon samplings (Barrett et al., 2013; Givnish et al., 2010; Liu et al., 2012; Steele et al., 2012). Despite these and other recent advances, major uncertainties still remain concerning a subset of higher-order monocot relationships.

Contemporary reviews have also brought together and synthesized our current understanding of the monocot fossil record (Doyle et al., 2008; Gandolfo et al., 2000; Greenwood and Conran, 2000; Herendeen and Crane, 1995; Smith, 2013; Smith et al., 2010b; Stockey, 2006). Until recently, the oldest and best characterized fossils of monocots were palms (Arecaceae), whose age potentially extended into the Lower Cretaceous (Dransfield et al., 2008; Herendeen and Crane, 1995). Our understanding of early monocot fossils has expanded, with, for example with the discovery of fossils assignable to Araceae from the Lower Cretaceous (Friis et al., 2004, 2010) and to Triuridaceae (a mycoheterotrophic family) from the early Upper Cretaceous (Gandolfo et al., 2000, 2002, 1998). Additional examples of significant fossil

finds include the two largest monocot families: these include the first unequivocal orchid fossils (∼20 Ma; Conran et al. 2009a; Ramírez et al. 2007), and grass fossils that extend the age of this family back to the Cretaceous-Paleogene boundary (66 Ma; Prasad et al. 2005, 2011).

The first focussed attempts to date patterns of monocot diversification, considering our improved understanding of monocot phylogeny and the fossil record, were those of Bremer (2000) and Janssen and Bremer (2004). Bremer (2000) used 91 monocot terminals and eight fossils to calibrate the diversification of the earliest splits in the monocots (several fossil calibrations in this study have since been criticized; Crepet et al. 2004; Stockey 2006). A follow-up study (Janssen and Bremer, 2004) greatly expanded the taxon sampling (800+ terminals), and used the inferred date (134 Ma) from the earlier study (Bremer, 2000) to constrain the crown age of monocots. These fossils and inferred monocot ages have been reused in several other studies, sometimes (but not always) with additional fossil constraints (e.g., Bremer, 2002; Givnish et al., 2005; Merckx et al., 2008a; Vinnersten and Bremer, 2001)

In this chapter I re-evaluate higher-order monocot relationships using new data that expand on several previous studies (Graham et al. 2006; Saarela 2006; Saarela and Graham 2010; Saarela et al. 2008, 2007; Zgurski 2004; Chapter 4). I then re-examine the ages of the major monocot clades, after using a suite of 25 fossils, including 17 monocot fossils, as age constraints. I use a newly developed dating method, the random local clocks model (Drummond and Suchard, 2010), to address lineage specific rate heterogeneity seen in broad-scale phylogenetic studies. This method has an advantage over traditional local clock models in that locations of rate change do not need to be specified *a priori* (Drummond and Suchard, 2010; Rutschmann, 2006). More accurately dated phylogenies would, for example, facilitate refined characterization of lineage diversification and character evolution in taxa of interest, and permit the development of new insights into how individual plant groups interact with other biological lineages and shifting abiotic environments, such as the Cenozoic development of the grasslands that now cover ∼40% of the earth's land surface (Strömberg, 2011).

5.3 Materials and methods

5.3.1 Taxonomic and genomic assembly

I sampled 172 monocot species representing 71 of 78 monocot families. These sequences came from GenBank, previously published results (Graham and Olmstead, 2000; Graham et al., 2006; Saarela and Graham, 2010; Saarela et al., 2008, 2007), unpublished results (Saarela, 2006; Zgurski, 2004), and 14 newly sequenced samples, the latter predominantly from Dioscoreales and Pandanales (see Appendix E and references there for all source details). I attempted to sample two or more taxa for families that have more than one widely recognized genus, and where possible I included taxa that span the deepest splits in the phylogenetic history of each family (the root of each crown clade) according to recent literature. All non-commelinid families were sampled (except for several heterotrophic ones), as were most commelinid families (only excluding five of eight families in the order Zingiberales). Family and order concepts follow Angiosperm Phylogeny Group (2009), with the exception that I recognize the family Maundiaceae Nakai (see Chapter 4), and the order Dasypogonales Doweld (Givnish et al., 1999). Outgroup sampling focused on major angiosperm lineages and several gymnosperms, for a total of 58 taxa. The complete matrix consists of 230 species used in the main phylogenetic analyses. The genomic sampling consisted of the coding portions of 17 plastid genes: *atp*B, *ndh*B, *ndh*F, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*J, *psb*L, *psb*N, *psb*T, *rbc*L, *rpl*2, *rps*7, 3'*rps*12 (Graham and Olmstead, 2000). Voucher and GenBank accession details for newly sequenced samples are presented in the Appendix E. DNA extraction, amplification and sequencing methods are detailed elsewhere (Graham and Olmstead 2000; Saarela et al. 2008; Chapter 3). Contig assembly and base calling were carried out using Sequencher 4.2.2 (Gene Codes Corp., Ann Arbor, Michigan, USA), and gene boundaries were determined with reference to tobacco and *Ginkgo* sequences (Saarela et al., 2008). Coding sequences were added to an existing land-plant matrix (Saarela and Graham, 2010; Saarela et al., 2007) and aligned manually using Se-Al version 2.0a11 Carbon (Rambaut, 2002) according to previously published protocols (Graham

and Olmstead, 2000; Graham et al., 2000). The aligned matrix is 12 506 bp in length.

5.3.2 Phylogeny estimation

I performed heuristic maximum parsimony (MP) searches using PAUP* version 4.0a113 (Swofford, 2003), with 10 random addition replicates, and with other settings set to default. Parsimony branch support was estimated with 200 bootstrap replicates, using one random addition replicate per bootstrap replicate.

For model-based methods of phylogenetic inference I used PartitionFinder version 1.0.1 (Guindon et al., 2010; Lanfear et al., 2012) to define an optimal partitioning strategy. PartitionFinder considers pre-defined data subsets (in this case individual codon positions per gene, for a total of 51 subsets), and uses a greedy algorithm to find the optimal partitioning scheme and associated models of sequence evolution, here using the Bayesian information criterion (BIC; Schwarz, 1978) as a model selection criterion, as recommended by the authors (Lanfear et al., 2012). The optimal partitioning strategy here consisted of 16 subsets each with a GTR+Γ or GTR+Γ+I model of evolution (referred to below as the PF16 scheme; Table 5.1). Because the invariant sites parameter (I) may adequately accommodate the Γ parameter (Yang, 2006), I used the GTR+Γ model for all partitions. I performed phylogenetic analysis under this PF16 scheme and under several simpler schemes: (a) a non-partitioned analysis, and several partitioning schemes based on: (b) the two major plastid gene regions (single copy [SC] vs inverted repeat [IR]); (c) the three codon positions; (d) and the codon positions considered separately for the two major plastid gene regions (six subsets total). For ML phylogenetic inference I used RAxML version 7.3.1 (Stamatakis, 2006b) implemented on the CIPRES Science Gateway version 3.3 web server (Miller et al., 2010). Branch lengths were linked among partitions in all schemes considered here. I evaluated branch support for the ML analyses using 200 bootstrap replicates. The ML optimization algorithm implemented for bootstrapping with RAxML is faster than the regu-

lar ML algorithm but not as thorough (the rapid bootstrapping algorithm; Stamatakis et al. 2008) and uses the CAT approximation for site-rate heterogeneity instead of the Γ parameter (Stamatakis, 2006a).

Table 5.1: The PF16 partitioning scheme. The best partitioning strategy for 17 sampled plastid genes and their three codon positions according to PartitionFinder (Lanfear et al., 2012), the PF16 scheme. Subscripts attached to gene names refer to codon positions 1, 2, or 3.

Subset	Contains
1	*atp*B$_1$, *ndh*B$_1$, *psb*F$_1$, *rpl*2$_1$, *rpl*2$_2$, 3'*rps*12$_1$
2	*atp*B$_2$, *psb*B$_2$, *psb*T$_2$
3	*atp*B$_3$, *psb*H$_3$, *psb*L$_3$, *psb*T$_3$, *rbc*L$_3$
4	*ndh*B$_2$, *psb*D$_1$
5	*ndh*B$_3$
6	*ndh*F$_1$,*psb*F$_3$, *rpl*2$_3$
7	*ndh*F$_2$
8	*ndh*F$_3$
9	*psb*B$_1$, *psb*C$_1$, *psb*E$_1$, *psb*N$_1$, *psb*T$_1$, 3'*rps*12$_2$
10	*psb*B$_3$, *psb*C$_3$, *psb*D$_3$, *psb*E$_3$, *psb*J$_3$, *psb*N$_3$
11	*psb*C$_2$, *psb*D$_2$
12	*psb*E$_2$, *psb*F$_2$, *psb*J$_2$, *psb*L$_1$, *psb*L$_2$, *psb*N$_2$
13	*psb*H$_1$, *psb*H$_2$, *psb*J$_1$
14	*rbc*L$_1$
15	*rbc*L$_2$
16	*rps*7$_1$, *rps*7$_2$, *rps*7$_3$, 3'*rps*12$_3$

I performed Bayesian (BI) phylogenetic inference using MrBayes version 3.2.1 (Ronquist et al., 2012b). I perform partitioned inference using the PF16 scheme, with the GTR+Γ model for each subset and treated all parameters as unlinked except for branch length and topology. Four chains of 1.5×10^6 generations were run. Run statistics were evaluated in Tracer version 1.3 (Rambaut and Drummond, 2009). A 10% burnin was discarded from each chain and the resulting trimmed chains were combined. Estimated sample sizes (ESSs) exceeding 200 were taken to indicate strong chain convergence. I consider nodes with bootstrap support values <70% to be poorly supported, 70–89% to be moderately supported, and ≥90% to be well

or highly supported (Graham et al., 1998). I consider nodes with posterior probabilities (expressed as percentages) <95% to be poorly supported, 95–98% to be moderately supported, and ≥99% to be well or highly supported (Alfaro and Holder, 2006).

5.3.3 Fossil constraints

I consulted recent reviews of monocot and general seed-plant fossils (Clarke et al., 2011; Doyle and Endress, 2010; Doyle et al., 2008; Friis et al., 2011; Smith, 2013; Smith et al., 2010b) to find fossils to serve as constraints for the molecular dating analysis; the stratigraphy follows the International Commission on Stratigraphy (2012). I chose 25 fossil constraints, 17 of which are monocots. To be conservative I generally assigned fossil ages to correspond to the top of individual stratigraphic units (i.e., the youngest age in the unit) to which each was assigned. Exceptions were made in cases where absolute ages were known or when the stratigraphy of a region was poorly calibrated according to Clarke et al. (2011), in which case I used their (younger) age assignment. Details of all fossil constraints, their taxonomic placement and age are presented below (summarized in Table 5.2):

Table 5.2: Monocot fossils used as molecular dating constraints. For monocot fossils, the node heading is cross referenced with Figure 5.4–5.5. Fossil name and taxonomic placement are given, as well as the clade constrained here (most recent common ancestor; MRCA). The age of the fossil is given; in most cases this corresponds to the top of the oldest stratigraphic unit in which the fossil is found (stratigraphy follows International Commission on Stratigraphy 2012). Prior probabilities for constrained clades are given using the age assigned to the corresponding fossil and an exponential distribution with a mean value and a corresponding 95% prior distribution. Note: node numbers are only given for monocots.

Node	Fossil, higher taxon	MRCA	Age of fossil(s) (Ma)	Prior calibrations Mean (95% interval)
	Pollen, monosulcate, columellar, with a reticulate or continuous tectum, Angiosperms	Angiosperm crown	132.9	Max. 140
	Pluricarpellatia peltata, Cabombaceae	*Cabomba* & *Nymphaea*	100.5	5 (100.6-118.9)
	Unnamed seed, Trimeniaceae	*Schisandra* & *Trimenia*	98.7	5 (98.83-117.1)
	Archaeanthus linnenbergi, Magnoliaceae	*Liriodendron* & *Magnolia*	100.5	5 (100.6-118.9)
	Walkeripollis, Winteraceae	*Asarum* & *Drimys*	125	6 (125.2-147.1)
	West Brothers platanoid and *Sapindopsis*, Platanaceae	*Nelumbo* & *Platanus*	93.9	5 (94.03-112.3)

Node	Fossil, higher taxon	MRCA	Age of fossil(s) (Ma)	Prior calibrations Mean (95% interval)
	Spanomera spp., Buxaceae	*Buxus & Trochodendron*	93.9	5 (94.03-112.3)
	Tricolpate pollen, eudicots	Stem eudicots	125	6 (125.2-147.1)
1	Araceae spp. A & B	*Arum & Spathiphyllum*	93.9	5 (94.03-112.3)
2	Fossil seeds, Alismataceae	*Alisma & Sagittaria*	23.03	2 (23.08-30.41)
3	*Stratiotes* spp. Hydrocharitaceae	*Elodea & Stratiotes*	56	4 (56.1-70.76)
4	*Thalassites parkavonensis*, *Thalassia testudinum*, Hydrocharitaceae	*Thalassia*, stem lineage	33.9	3 (33.98-44.87)
5	*Cymodocea floridana*, *Thalassodendron auricula-leporis*, Cymodoceaceae	*Amphibolis, Halodule & Ruppia*	33.9	3 (33.98-44.97)
6	*Gruenbachia pandanoides*, Pandanaceae	*Pandanus & Cyclanthus*	77.8	5 (77.93-96.24)
7	*Lateropora glabra*, assignable to *Freycinetia*, Pandanaceae	*Freycinetia & Pandanus*	33.9	3 (33.98-44.97)
8	*Cyclanthus lakensis*, Cyclanthaceae	*Carludovica & Cyclanthus*	47.8	3 (47.88-58.87)

Node	Fossil, higher taxon	MRCA	Age of fossil(s) (Ma)	Prior calibrations Mean (95% interval)
9	*Dioscorea wilkinsii*, Dioscoreaceae	*Dioscorea bulbifera & D. elephantipes*	27	2 (27.05-34.38)
10	*Ripogonum tasmanicum*, Ripogonaceae	*Philesia & Ripogonum*	51.9	4 (52.0-66.66)
11	*Dendrobium winikaphyllum* and *Earina fouldenensis*, Orchidaceae	*Coelogyne & Phalaenopsis*	20	2 (20.05-27.38)
12	*Dianellophyllum eocenicum*, Xanthorrhoeaceae	*Hemerocallis & Phormium*	33.9	3 (33.98-44.97)
13	*Paracordyline kerguelensis*, Asparagaceae	*Cordyline & Trichopetalum*	22	2 (22.05-29.38)
14	*Spirematospermum chandlerae*, Zingiberales	*Ensete, Renealmia & Strelitzia*	83.6	5 (83.73-102.0)
15	*Typha*, Typhaceae	*Sparganium & Typha*	66	4 (66.1-80.76)
16	*Volkeria messelensis*, Cyperaceae	*Gahnia & Mapania*	47	3 (47.08-58.07)
17	*Matleyites indicum & Vonhueneites papillosum*, Poaceae	*Oryza & Pseudosasa*	66	4 (66.1-80.76)

Angiosperms

The earliest unequivocal angiosperm fossils are pollen grains from the Valanginian and Hauterivian in the Lower Cretaceous (reviewed in Doyle, 2005, 2009, 2012). These pollen grains have a monosulcate aperture, columellar exine, and either a continuous or reticulate tectum. Columellar exine is only found in angiosperms and is predicted to be the ancestral state for angiosperms (Doyle, 2005, 2009, 2012). Since columellar exine is plesiomorphic in extant angiosperms its presence in fossil pollen can not be used to place the fossils within the crown clade of angiosperms. However, a reticulate tectum is predicted to be the ancestral state for a clade comprising Austrobaileyales and mesangiosperms and is absent in *Amborella* and Nymphaeales (Doyle, 2005, 2012). Based on phylogenetic relationships in which *Amborella* and Nymphaeales are consecutive sister groups to the remaining angiosperms (e.g., Moore et al., 2007) it is parsimonious to consider it a derived character state within angiosperms. This combination of characters, which is found in the earliest angiospermous fossil pollen, suggests that the crown clade of angiosperms had arisen by the end of the Valanginian, 132.9 Ma (Doyle, 2012).

A rapid sequence in the further elaboration and diversification of angiospermous fossils is apparent in the stratigraphic sequences, and most of the major lineages of mesangiosperms are present by the Barremian and Aptian (e.g., Chloranthales, eudicots, magnoliids, monocots; Doyle 2012; Doyle and Endress 2010; Doyle et al. 2008; Friis et al. 2011). A difference in age of only ∼10–20 Ma is seen between the earliest crown-clade fossils and the evolution of lineages that make up the bulk of modern angiosperm diversity. Doyle (2012) has argued that the close fit between the phylogenetic sequence of diversification (e.g., Graham and Olmstead, 2000; Graham and Iles, 2009; Jansen et al., 2007; Moore et al., 2007; Qiu et al., 2010) and the morphological evolution of multiple angiospermous fossils supports a short time lag between the evolution and subsequent fossilization of extant angiosperm lineages. Conversely, if the time lag between the origin of angiosperm lineages and their subsequent fossilization were large, the order of

appearance of fossil lineages would not necessarily match their phylogenetic sequence of diversification (Doyle, 2012). If this reasoning is correct, the origin of crown angiosperms should closely predate the earliest angiosperm fossils from the Valanginian (132.9 Ma). I therefore assign a maximum age for the crown angiosperm clade of 140 Ma.

Nymphaeales

A large number of Cretaceous nymphaealean fossils have been discovered over the last dozen years (Friis et al., 2001; Gandolfo et al., 2004; Mohr et al., 2008; Taylor et al., 2008; von Balthazar et al., 2008; Wang and Qiu, 2006); however, their taxonomic placement has sometimes been controversial (Endress, 2008; Nixon, 2008; Yoo et al., 2005). Two fossils, *Monethianthus mirus* (Friis et al., 2009, 2001) and *Pluricarpellatia peltata* (Mohr et al., 2008; Taylor et al., 2008), are probably the least contested and belong in or close to crown Nymphaeaceae or Cabombaceae (respectively). *Monethianthus* was described from the Vale de Agua locality (Figuira de Foz Formation) of Portugal, and is considered to be of Late Aptian and Early Albian age (113 Ma; Dinis et al. 2008; Friis et al. 2009). Because of the complicated geological history of the region, Clarke et al. (2011) consider the youngest reliable date to be from the top of the Cenomanian (93.9 Ma). *Pluricarpellatia* was described from the Crato Formation of Brazil, which is also considered to be Late Aptian or Early Albian age (113 Ma); Clarke et al. (2011) considered this region's youngest reliable date to be from the top of the Albian (100.5 Ma). As *Pluricarpellatia* is reasoned by Clarke et al. (2011) to be older than *Monethianthus*, I consider it to constrain the minimum age of the split between Cabombaceae and Nymphaeaceae, corresponding to the top of the Albian (100.5 Ma).

Austrobaileyales

Several fossils are assignable to Schisandraceae from the middle Cretaceous (Doyle and Endress, 2010; Friis et al., 1997; Frumin and Friis, 1999). *Anacostia* spp. are the best characterized and are placed close to Schisandraceae

in morphological analyses (Doyle and Endress, 2010). They are known from several localities in eastern North America (Kenilworth, Maryland, and Puddledock, Virginia, both assignable to the Potomac Formation Zone IIB) and Portugal (Buarcos, Famalicão, and Vale de Agua, all assignable to the Figueira de Foz Formation; Clarke et al. 2011; Friis et al. 1997; Hochuli et al. 2006). Both the Potomac Formation Zone IIB and Figuira de Foz Formation are considered to be of Albian age (113–100.5 Ma; Hochuli et al. 2006). However, Clarke et al. (2011) consider the boundaries of the Potomac Formation Zone IIB to be unclear and suggest a minimum age corresponding to the top of the Cenomanian (93.9 Ma), which is the same age they assigned for the Figueira de Foz Formation (see above). Recently, a fossil seed assignable to the stem lineage of Trimeniaceae was described from the Hikagenosawa Formation of Japan (Yamada et al., 2008). This formation ranges from the middle Albian to the middle Cenomanian (∼106.8–98.7 Ma; Takashima et al. 2004), which is older than the age considered for the *Anacostia* spp. I therefore consider this unnamed trimeniaceous seed to constrain the minimum age of the split between Trimeniaceae and Schisandraceae to the middle Cenomanian (98.7 Ma).

Magnoliales

Archaeanthus linnenbergi from the Dakota Formation of Kansas (Dilcher and Crane, 1984), dated to the top of the Albian (100.5 Ma; Doyle and Endress 2010), and *Endressinia brasiliana* from the Crato Formation of Brazil (Mohr and Bernardes-de Oliveira, 2004), are the oldest fossils reliably assigned close to Magnoliales (Doyle and Endress, 2010). *Archaeanthus* may belong within Magnoliaceae, as it shares bilobed leaves with *Liriodendron* (Doyle and Endress, 2010; Friis et al., 2011), while *Endressinia* is more likely to be a crown-clade member of Magnoliales (Doyle and Endress, 2010). Despite the potentially greater age of *Endressinia*, *Archaeanthus* is likely nested more securely within Magnoliales phylogeny. I therefore consider *Archaeanthus* to define a minimum crown age for Magnoliaceae corresponding to the top of the Albian (100.5 Ma).

Canellales

Pollen tetrads (*Walkeripollis gabonensis*) from the late Barremian (∼125 Ma) of Gabon, have been consistently linked to Winteraceae (Doyle and Endress, 2010; Doyle et al., 1990). *Walkeripollis* therefore provides a minimum stem age for Canellales corresponding to the top of the Barremian (125 Ma).

Eudicots

The earliest known tricolpate pollen, a synapomorphy of eudicots, is from the late Barremian (∼125 Ma) of Africa and England (Doyle and Endress, 2010; Hughes and McDougall, 1990). It is not assignable to any extant eudicot lineage, and as a consequence, I use it to define a minimum stem age for eudicots corresponding to the top of the Barremian (125 Ma).

Proteales: Plantanaceae

A suite of male and female reproductive structures with similarity to extant Platanaceae was described from the West Brothers clay mine, Maryland (Potomac Formation Zone IIB; Crane et al. 1986); additional reproductive structures in association with leaves assignable to *Sapindopsis* were found in Bank, Virginia (Potomac Formation Zone IIB; Crane et al. 1993. These two assemblages were recovered as successive sister groups to extant Platanaceae in a morphology-based analysis (Doyle and Endress, 2010). These fossil assemblages provide a minimum age for the split between Platanaceae and Proteaceae corresponding to the top of the Cenomanian (93.9 Ma).

Buxales: Buxaceae

Two species (*Spanomera mauldinensis* and *S. marylandensis*) described from Cretaceous sediments of eastern USA are highly similar to extant Buxaceae (Drinnan et al., 1991). This is supported by a MP analysis that placed the composite taxon (*Spanomera*) as the sister group to extant Buxaceae (Doyle and Endress, 2010). *Spanomera mauldinensis* is from the Mauldin Mountain locality of Maryland, which is considered to be part of the Po-

tomac Zone III Formation, dated to the top of the Cenomanian (Hochuli et al., 2006); *S. marylandensis* is from the West Brothers locality, which is also considered to have an age corresponding to top of the Cenomanian (see above; Drinnan et al. 1991). I therefore consider these two fossil species (*S. mauldinensis* and *S. marylandensis*) to provide a minimum age for the split between Buxaceae and Trochodendraceae corresponding to the top of the Cenomanian (93.9 Ma).

Alismatales: Araceae

Several unique finds from the mid-Cretaceous of Portugal attest to the presence and diversity of Araceae at that time (Friis et al., 2004, 2010). The fossil pollen *Mayoa portugallica* is from the Torres Vedras locality of Portugal, and was taken from sediment that is part of the Lower Almargem Formation, which is usually considered to be Upper Barremian to mid-Aptian in age (Dinis et al., 2008; Friis et al., 2004). *Mayoa* is considered to be very close to extant pollen of *Spathiphyllum* and *Holochlamys* in Araceae subfamily Monsteroideae (Doyle et al., 2008; Friis et al., 2004, 2010), although Hofmann and Zetter (2010) suggest a strong likeness to *Lagenella martinii* (a presumptive gymnosperm) which has a record that spans from the Triassic to Cretaceous. More complete and unambiguous araceous material is represented by two fossil inflorescences 'Araceae fossil sp. A' and 'Araceae fossil sp. B' from the Vila Verde 2 locality of Portugal, Figueira da Foz Formation of Albian age (Dinis et al., 2008; Friis et al., 2010). These fossils are assigned to Araceae subfamilies Aroideae and Pothoideae (Friis et al., 2010). Because of some ambiguity in the taxonomic placement of *Mayoa*, I use 'Araceae fossil spp. A and B' to define the minimum age for the split among core Araceae (represented here by *Arum* and *Spathiphyllum*), corresponding to the top of the Cenomanian (93.9 Ma); see node 1 on Figure 5.4.

Alismatales: Alismataceae

Fossils with similarities to Alismataceae (including Limnocharitaceae) occur from the Upper Cretaceous onward (Smith, 2013; Stockey, 2006). The best

characterized of these are: *Cardstonia tolmanii*, *Haemonthophyllum* spp., and *Heleophyton helobiaeoides* (Erwin and Stockey, 1989; Golovneva, 1997; Riley and Stockey, 2004). These fossils show clear similarities with extant Alismataceae, but also share characteristics of other alismatalean families, and for this reason, they cannot reliably be placed within the crown clade of Alismataceae. The first fossils that can be reliably placed in the crown are fossil fruits from the late Oligocene onwards of Eurasia and North America (reviewed in Haggard and Tiffney 1997). Some of these fossil fruits belong to extant genera (Haggard and Tiffney, 1997), while others represent extinct lineages (Collinson, 1983). I therefore consider these fossil fruit taxa to constrain the minimum age of the crown clade of Alismataceae to the top of the Oligocene (23.03 Ma); see node 2 on Figure 5.4.

Alismatales: Hydrocharitaceae

Seeds recognizable as belonging to extant genera of Hydrocharitaceae first appear in the late Paleocene and increase in diversity and frequency onwards. The oldest of these belong to *Stratiotes* (Smith, 2013; Stockey, 2006), which extends into the Paleocene of England (Sille et al., 2006). As *Stratiotes* appears to be the sister taxon of the rest of the family (Figure 5.1), I consider the oldest known *Stratiotes* fossil to constrain the minimum age of the crown clade of Hydrocharitaceae, corresponding to the top of the Paleocene (56.0 Ma; Sille et al. 2006); see node 3 on Figure 5.4.

Alismatales: seagrasses

Seagrasses have a relatively extensive fossil record extending into the Upper Cretaceous, but establishing taxonomic relationships to extant lineages is often difficult, due to poor preservation and extreme morphological and anatomical convergence among different lineages (Benzecry and Brack-Hanes, 2008; Smith, 2013; Stockey, 2006; van der Ham et al., 2007). I therefore considered the well preserved and taxonomically well characterized samples from the Avon Park Formation (Claiborne Stage, Middle Eocene) of Florida. These consist of roughly six or seven seagrass species that are closely related

to extant seagrasses in Hydrocharitaceae and Cymodoceaceae (Benzecry and Brack-Hanes, 2008; Ivany et al., 1990; Lumbert et al., 1984). I use some of these fossil taxa to constrain the minimum age of the split between *Thalassia* and *Vallisneria* in Hydrocharitaceae (see node 4 on Figure 5.4), and others the minimum age of the crown group of *Amphibolis*, *Halodule* and *Ruppia* (consisting of members of both Cymodoceaceae and Ruppiaceae; see node 5 on Figure 5.4), both to the top of the Eocene (33.9 Ma).

Pandanales: Pandanaceae

Fruiting structures very similar to *Pandanus* are known from the early Campanian Grünbach Formation of Austria (Herman and Kvaček, 2010; Smith, 2013). Fossil leaves are also well known from the Grünbach Formation and other Cretaceous deposits elsewhere. However, these may be confused with certain members of Cyperaceae (Herman and Kvaček, 2010; Kvaček and Herman, 2004; Smith, 2013). Therefore I use the fossil infructescence, *Gruenbachia pandanoides*, to constrain the minimum age of the split between Cyclanthaceae and Pandanaceae to the middle Campanian (77.8 Ma; see node 6 on Figure 5.4; Smith in press). Within Pandanaceae, the fossil pollen *Lateropora glabra* from the late Eocene of New Zealand and Oligocene of Australia is assignable to *Freycinetia* (Greenwood and Conran, 2000; Macphail et al., 1994; Raine et al., 2011). I therefore use *Lateropora* to constrain the minimum age of the crown clade of Pandanaceae to the top of the Eocene (33.9 Ma); see node 7 on Figure 5.4.

Pandanales: Cyclanthaceae

Recent evaluation of infructescences from the Messel Formation of Germany demonstrates that they belong to *Cyclanthus*; seeds very similar to these are found in the early Eocene (Ypressian) of England (Smith et al., 2008). I therefore consider the English specimens (*Cyclanthus lakensis*) to constrain the minimum crown age of Cyclanthaceae to the top of the Ypressian (47.8 Ma); see node 8 on Figure 5.4.

Dioscoreales: Dioscoreaceae

Reliable fossil records of this family are known from the Eocene onwards; earlier records from the Cretaceous are less securely placed (Smith, 2013). A fruit assignable to *Dioscorea* belongs to the Late Eocene of Colorado (Manchester, 2001), but its placement with respect to extant species is unclear, and it may belong along the stem lineage of *Dioscorea*. The fossil leaf *D. wilkinii* from the Guang River flora (Oligocene) of Ethiopia is closely related to extant species of *Dioscorea* sect. *Lasiophyton* (Pan, 2007). Molecular phylogenies suggest that *D.* sect. *Lasiophyton* is closely related to *D. bulbifera*, one of the species sampled here (Wilkin et al., 2005). I therefore use *D. wilkinii* to constrain the minimum age of the split between *D. bulbifera* and *D. elephantipes* to the Guang River flora age (27 Ma; Pan 2007); see node 9 on Figure 5.4.

Liliales: Ripogonaceae

Fossil leaves from the Macquarie Harbour Formation (early Eocene) of Tasmania are assignable to *Ripogonum* based on leaf venation and stomatal patterns (Conran et al., 2009b). I therefore consider *R. tasmanicum* to constrain the minimum age of the split between *Ripogonum* (Ripogonaceae)and *Philesia* (Philesiaceae) to the middle Ypressian (51.9 Ma) (Conran et al., 2009b); see node 10 on Figure 5.4.

Asparagales: Orchidaceae

There have been several recent fossil finds of Orchidaceae from the subfamilies Orchidoideae and Epidendroideae (Conran et al., 2009a; Ramírez et al., 2007). Conran et al. (2009a) described fossil leaves with unambiguous affinities to the genera *Dendrobium* and *Earina* that come from the early Miocene of New Zealand (Foulden Hills Diatomite, 23–20 Ma). Securely placing these genera within deeper Epidendroideae phylogeny is difficult because of the extremely rapid radiation of the subfamily and a lack of phylogenetic resolution among the deeper lineages (e.g., Cameron, 2004; Freudenstein et al., 2004; Neubig et al., 2008; van den Berg et al., 2005) however a recent study

using the nuclear gene *Xhd* more confidently places *Dendrobium* and *Earina* in a clade which is the sister group of clade containing *Phalaenopsis*; *Coelogyne* is in turn part of clade which is the sister group of this larger clade (Górniak et al., 2010). I therefore use fossils *Dendrobium* and *Earina* to constrain the minimum age of the split between *Coelogyne* and *Phalaenopsis* to 20 Ma; see node 11 on Figure 5.4.

Asparagales: Xanthorrhoeaceae

A single fossil leaf very similar to those of *Dianella* is known from the middle Eocene (Nelly Creek) of Australia (Conran et al., 2003). *Dianella* belongs to subfamily Hemerocallidoideae, close to *Phormium* (Seberg et al., 2012; Steele et al., 2012). Here I consider *Dianellophyllon eocenicum* to constrain the minimum age of the split between *Phormium* and *Hemerocallis* to the top of the Eocene (33.9 Ma); see node 12 on Figure 5.4.

Asparagales: Asparagaceae

Several fossil leaves similar to those of *Cordyline* have been recovered from the Oligocene and Eocene of the Kerguélen Islands in the south Indian Ocean and Australia (Conran, 1997; Conran and Christophel, 1998). The younger of these, *Paracordyline kerguelensis*, is from ash sediments between basalt flows (22 Ma) on Kerguélen Island (Conran, 1997). The older of these, *P. aureonemoralis*, is from the lower Eocene (Golden Grove, 56.0–41.3 Ma) of Australia (Conran and Christophel, 1998). While these fossils are in general very similar to each other and to *Cordyline*, other assignments are also possible (Conran, 1997; Conran and Christophel, 1998). However, the younger fossil possesses unique cuticular sculpturing very similar to a subclade of *Cordyline*, making its relationship to this group more likely (Conran, 1997; Conran and Christophel, 1998). *Cordyline* is in subfamily Lomandroideae along with *Trichopetalum* and *Lomandra* sampled here. I therefore consider *P. kerguelensis* to constrain the minimum age of the split between *Cordyline* and *Trichopetalum* to 22 Ma; see node 13 on Figure 5.4.

Zingiberales: Zingiberaceae

Spirematospermum, which has been variously assigned to Musaceae and Zingiberaceae, has an extensive fossil record covering Eurasia and North America from the Upper Cretaceous nearly to the present day (Fischer et al., 2009; Friis et al., 2011). The oldest occurrence of the genus, *S. chandlerae*, is from the Santonian-Campanian boundary of North America (Neuse River, North Carolina) (Friis, 1988; Friis et al., 2011). Family placement is currently unclear, but evidence tends to point to an association with Zingiberaceae (S. Y. Smith, University of Michigan, pers. comm.). However, considering the uncertainty in its systematic placement, I use this fossil to constrain the minimum crown age of Zingiberales to the top of the Santonian (83.6 Ma); see node 14 on Figure 5.5.

Poales: Typhaceae

As currently circumscribed Typhaceae contain two genera, *Sparganium* and *Typha* (APG 2009). The seeds and fruits of *Typha* are distinct and the earliest fossil records are from the Upper Cretaceous and extend to the present (Smith et al., 2010b). Therefore I use these earliest records to constrain the minimum crown age of Typhaceae to the top of the Maastrichtian (66 Ma, Knobloch and Mai 1986); see node 15 on Figure 5.5.

Poales: Cyperaceae

Isolated fossil fruits assigned to Cyperaceae are common from the Eocene onwards, but potentially extend back to the Paleocene (Smith et al., 2010b). However, it is unclear if all these records are reliable, as a review of an English fossil fruit assigned to *Scirpus* revealed it to be the fossil seeds of Cyclanthaceae (Smith et al., 2008). Complete infructescences, *Volkeria messelensis*, that are unambiguously assigned to subfamily Mapanioideae were recently described from the Messel Formation (47 Ma) of Austria (Smith et al., 2009). Subfamily Mapanioideae is the sister taxon of the rest of the family (Simpson et al., 2006). Therefore I consider *V. messelensis* to constrain the minimum crown age of Cyperaceae to 47 Ma; see node 16 on

Figure 5.5.

Poales: Poaceae

The oldest Poaceae macrofossils, from the Paleocene-Eocene boundary (56 Ma), are spikelets with characters suggesting an affinity to the core Poaceae subfamilies (Crepet and Feldman, 1991; Grass Phylogeny Working Group, 2001; Grass Phylogeny Working Group II, 2012; Saarela and Graham, 2010). However, unequivocal microfossils in the form of phytoliths, sometimes still embedded within the epidermis, are known from the Cretaceous of India (Prasad et al., 2005). These Cretaceous fossils are from the Intertrappean beds of the Lameta Formation and date to the Maastrichtian-Danian boundary (66 Ma); multiple phytoliths are described with varying affinity to the core subfamilies (Prasad et al., 2005). The species *Matleyites indicum* and *Vonhueneites papillosum* are assigned to the Bambusoideae-Erharthoideae-Pooideae (BEP) clade (Prasad et al., 2005). Here I use *M. indicum* and *V. papillosum* to constrain the minimum age of the crown of the BEP clade to the top of the Maastrichtian (66 Ma); see node 17 on Figure 5.5.

5.3.4 Molecular dating

To estimate divergence times I used the Bayesian random local clocks model (Drummond et al., 2006; Drummond and Rambaut, 2007; Drummond and Suchard, 2010). This model allows branches to have different strict molecular clocks, the location and number of distinct clocks over the tree being determined using an internal model selection procedure (Drummond and Suchard, 2010). Because of the complexity of the molecular dating model and the need for tractable computation, I simplified the data set in two ways. First, I considered only genes sampled from the SC region of plastid genome (*atp*B, *ndh*F, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*J, *psb*L, *psb*N, *psb*T, *rbc*L) to remove the influence of strong heterotachy present in the IR region of the plastid genome (see Chapter 4), as this may complicate molecular dating (Wertheim et al., 2012). Second, I reduced the number of sampled taxa to 153 (with 140 monocots) from 230 (with 172 monocots). The final

taxon sampling includes only angiosperms (the three gymnosperms, *Cycas*, *Ginkgo*, and *Pinus*, are excluded), and includes 68 of 71 monocot families included in the main phylogenetic analysis.

I used the fossil constraints described in the previous section (and see Table 5.2). I used an exponential prior for age constraints, with the 95% probability density corresponding to approximately 10–20% of the fossil age. I used exponential priors because of the inherent uncertainty in minimum clade ages; in general, exponential priors are more diffuse than log-normal priors and tend to produce wider posterior distributions (Sauquet et al., 2012). I also supplied BEAST with a starting tree to initiate runs (and also used default settings in BEAST). This starting tree was created by first trimming the main ML tree (Figure 5.1–5.3) to the reduced taxon set, and then producing a fossil calibrated ultrametric tree using penalized likelihood in r8s version 1.71 (Sanderson, 2002, 2003). Clades acting as fossil constraints were forced to be monophyletic. In addition to these topological constraints I also constrained the clade comprising all angiosperms except *Amborella* to be monophyletic. Speciation was set to a pure-birth Yule process and a GTR+Γ model of molecular evolution was used, with no data partitions. Three runs of 5.0×10^7 generations were combined after discarding a 25% burnin (the total computational time was approximately six weeks). As with the MrBayes analysis, values with ESS above 200 are generally considered to indicate chain convergence, however values with ESS of 100 can be considered adequate, and are used for this analysis. To compare my ages with those depicted by other studies I used Plot Digitizer version 2.6.1 (Huwaldt, 2012) to convert figures into numerical data points when tables of node ages were not provided in those studies.

5.4 Results

5.4.1 Phylogenetic inference

The ML tree inferred using the PF16 scheme is shown in Figures 5.1–5.3. The simpler partitioning strategies considered for ML produced very simi-

Table 5.3: Partition scheme comparisons. Optimal partitioning strategy (PF16 scheme in bold font) found by PartitionFinder, compared to several simpler partitioning strategies using the Bayesian information criterion (BIC). Gene region refers to either the single copy (SC) or inverted repeat (IR) regions of the plastid genome.

	No. of (sub)sets	No. of parameters	Log-likelihood	BIC
Non-partitioned	1	466	-281879	568154
Gene region (SC or IR)	2	476	-281370	567231
Codon position	3	486	-278113	560811
Gene region & codon position	6	516	-277351	559569
PF16	**16**	**616**	**-275800**	**557412**

lar tree topologies (4–9 different branches; not shown here), despite having much poorer log-likelihoods and BIC scores (Table 5.3). MP analysis yielded two most parsimonious trees (tree score = 53 402); these were also generally congruent with the PF16 tree (26–27 different branches; not shown here), as was the BI PF16 (3 different branches; not shown here). Branch support is generally high for most branches across the different analyses; branches that were dissimilar among the analyses were typically poorly supported. I focus the results on the ML PF16 tree. *Amborella*, Nymphaeales and Austrobaileyales are successive sister groups to the mesangiosperms, although the position of *Amborella* as the sister group to the remaining angiosperms is only poorly supported. Relationships among the major lineages of mesangiosperms are poorly supported (Figures 5.1–5.3).

Within monocots, ordinal relationships outside of commelinids are moderately to well supported (Figures 5.1–5.2). Acorales, Alismatales, Petrosaviales, Dioscoreales–Pandanales (a clade comprising these two orders), Liliales, and Asparagales are successive sister groups to the commelinids. Most inter-ordinal relationships are poorly supported within the commelinids, with several conflicting topologies seen among the different analyses; the only exception is the strongly supported sister group relationship between

Commelinales and Zingiberales (Figure 5.3). Below I focus on describing the placement of newly added taxa or taxa that differ from the earlier MP analysis of Graham et al. (2006). Several lineages of commelinids (Arecales, Dasypogonales, and Zingiberales) are sparsely sampled, and so intra-ordinal discussion is omitted for them.

Within Alismatales, Tofieldiaceae and Araceae are successive sister groups to the remaining families (the alismatids) with moderate to high support (Figure 5.1). Relationships among the alismatids are congruent with those discussed in Chapter 4. Note that I added *Spirodela* (Araceae subfamily Lemnoideae) and *Tetroncium* (Juncaginaceae) here. These additional taxa placed close to other members of their family or subfamily. Families within Alismatales are generally well supported, with a few exceptions (see Chapter 4). I also find that Petrosaviales, comprising the autotrophic *Japonolirion* and the mycoheterotrophic *Petrosavia*, are strongly supported as monophyletic (Figure 5.2).

Within Dioscoreales the relationship among the three families sampled here (Burmanniaceae, Dioscoreaceae, and Nartheciaceae) are moderately to strongly supported. The Burmanniaceae tribe Thismieae, a group of obligate mycoheterotrophs that is sometimes recognized as the family Thismiaceae (see Merckx et al. 2006), was not sampled here. Within Dioscoreaceae the relationships among the sampled genera (*Dioscorea*, *Tacca*, and *Trichopus*; *Stenomeris* was not sampled) are poorly supported. Within Nartheciaceae *Aletris* and *Metanarthecium* are moderately to well supported as a

---

Figure 5.1 *(following page)*: Monocot phylogram, part 1. Portion of the maximum likelihood phylogram of monocots and outgroups, focused on outgroups and the monocot orders Acorales and Alismatales, based on the coding regions of 17 plastid genes for 230 samples, and a partitioned (PF16 scheme) GTR + Γ model of evolution. Support values adjacent to branches are maximum parsimony bootstrap values, maximum likelihood bootstrap values, and Bayesian posterior probabilities. Support values are presented as percentages. A dash represents support <50%, a filled dot represents support = 100%. Scale represents expected substitution per site.

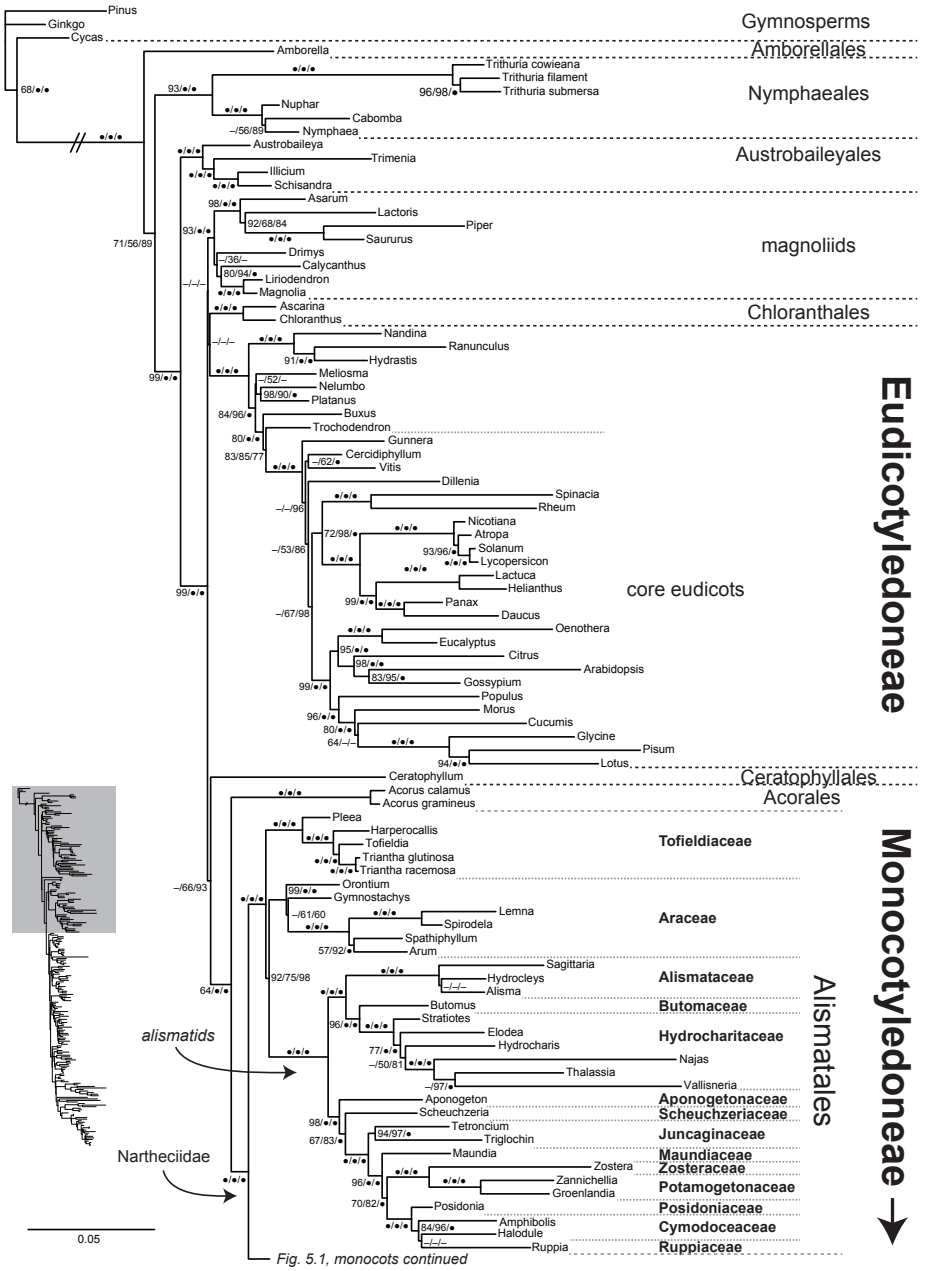


Fig. 5.1, monocots continued

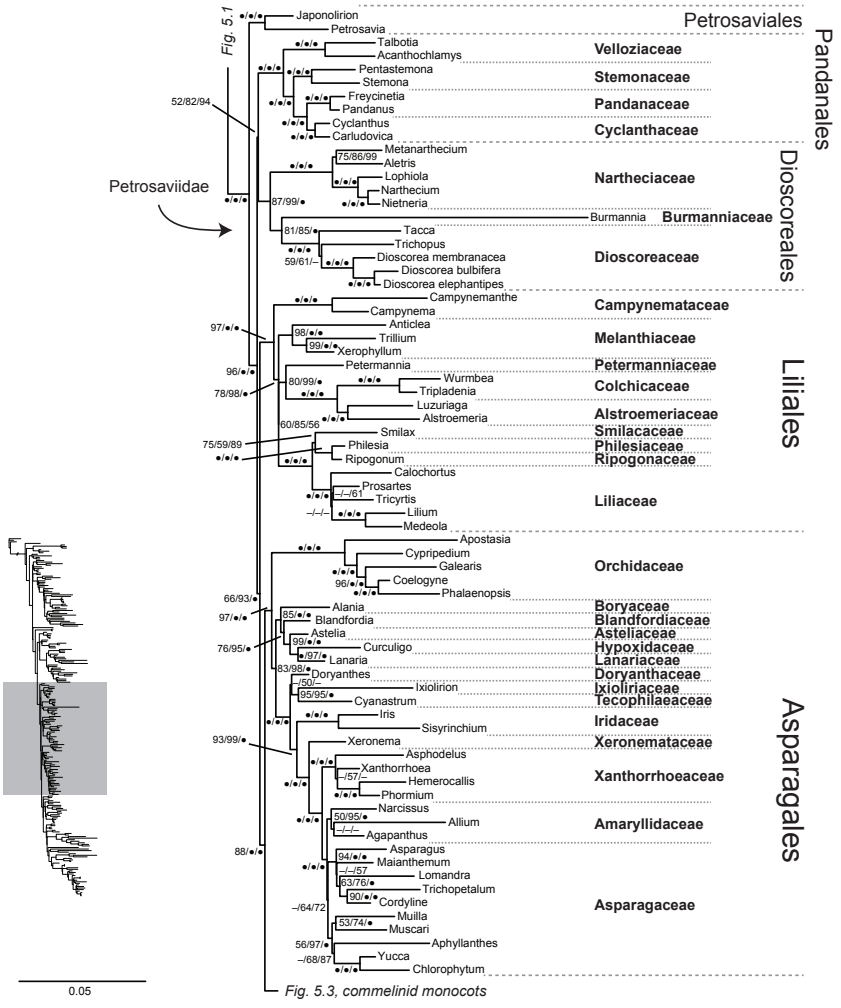


Figure 5.2: Monocot phylogram, part 2. A continuation of the phylogram from Figure 5.1, showing monocot orders Pandanales, Dioscoreales, Liliales and Asparagales.

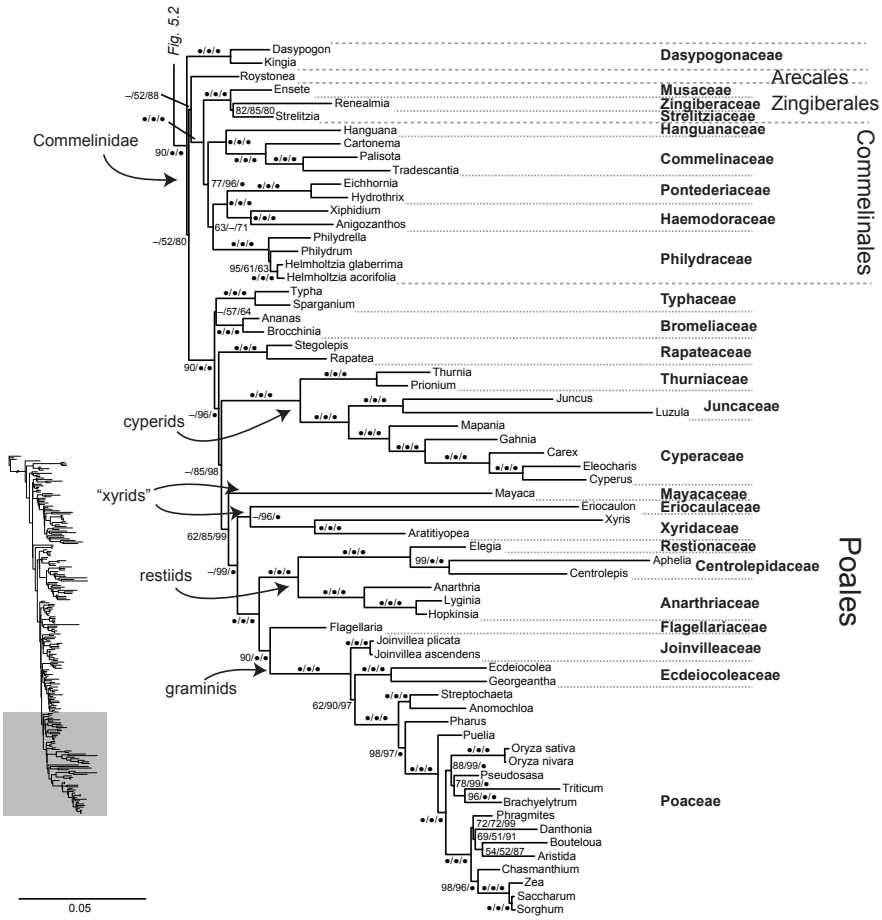


Figure 5.3: Monocot phylogram, part 3. A continuation of the phylogram from Figure 5.2, showing commelinid monocots.

clade that is the sister group to the remaining members of the family (*Lophiola*, *Narthecium*, and *Nietneria*) (Figure 5.2). The relationships among the four sampled families of Pandanales (Cyclanthaceae, Pandanaceae, Stemonaceae, and Velloziaceae; the mycoheterotrophic Triuridaceae was not sampled) are strongly supported, as is the monophyly of each of these families (Figure 5.2).

Within sampled families of Liliales (only the mycoheterotrophic Corsiaceae is not sampled here), Campynemataceae are moderately to well supported as the sister group of the rest of the order. Melanthiaceae are the sister taxon of the remaining families with moderate support in the ML analysis, but poor support from MP and BI. The branch that subtends these remaining families (Alstroemeriaceae, Colchicaceae, Liliaceae, Petermanniaceae, Philesiaceae, Ripogonaceae, Smilacaceae) is very short (not clearly visible on Figure 5.2). Petermanniaceae are the sister group of a clade comprising Colchicaceae and Alstroemeriaceae with moderate to high support. The remaining families comprise a well-supported clade. However, except for a strongly supported relationship between Philesiaceae and Ripogonaceae, relationships among these families are only poorly supported (Figure 5.2). The monophyly of all Liliales families is well supported at the current taxon sampling. Within Liliaceae, relationships among the sampled genera are poorly supported, except for a strongly supported relationship between *Medeola* and *Clintonia* (Figure 5.2).

Within Asparagales, inter-family support is almost uniformly high. One exception is for Doryanthaceae, which is poorly supported as the sister group of a clade comprising Ixioliriaceae and Tecophilaeaceae in the ML analysis (Figure 5.2). An alternative position found in the MP and BI analyses depicts Doryanthaceae as the sister group of a large clade comprising Amaryllidaceae, Asparagaceae, Iridaceae, Xanthorrhoeaceae, and Xeronemataceae; however, this arrangement has poor support in both sets of analyses (data not shown). There is a very short branch subtending crown Asparagaceae, and the support for the monophyly of Asparagaceae is poorly supported (Figure 5.2). The monophyly of the remaining families of Asparagales is generally well supported at the current taxon sampling.

Within Poales, inter-family relationships are mostly moderately to strongly supported (Figure 5.3). One exception concerns the root of the order: Typhaceae and Bromeliaceae are poorly supported as a clade that is the sister group of the remaining members of Poales. Rapateaceae and a clade comprising Cyperaceae, Juncaceae, and Thurniaceae (the 'cyperids') are successive sister groups to the remaining families (comprising the 'graminids,' the 'restiids,' and the 'xyrids'; see Figure 5.3). Model-based methods (i.e., BI and ML) find a different position than MP concerning the placement of the xyrid family Mayacaceae. For the BI and ML analyses, Mayacaceae are moderately to strongly supported as the sister group of a clade comprising the remaining members of xyrids (Eriocaulaceae and Xyridaceae), the restiids, and the graminids (Figure 5.3). In contrast, the shortest MP trees depicts a placement of Mayacaceae embedded within the xyrids as the sister group to Eriocaulaceae, though this placement receives poor support (data not shown). The remaining members of the xyrids, and the restiids (= Anarthriaceae, Centrolepidaceae, and Restionaceae) are successive sister groups of the graminids (Figure 5.3). Within the graminids, Flagellariaceae, Joinvilleaceae, and Ecdeiocoleaceae are successive sister groups of Poaceae with mostly high support (Figure 5.3). Relationships within Poaceae and Commelinales mirror those of Saarela and Graham (2010) and Saarela et al. (2008) respectively, which employed similar taxon sampling, and so they are not discussed here.

5.4.2 Dating analysis

Model statistics were slow to converge under the random local clocks model. After burnin and combination, posterior, likelihood and prior values had relatively low ESSs (134, 93, and 110 respectively). Model parameter values had ESSs that ranged from below 100 to above 200. Despite several low ESSs, visual inspection of most estimated parameters and estimated branches lengths appeared to converge on similar values, and I believe that the search depth is adequate here.

After burnin and thinning (5-fold thinning, leaving 10 000 pre-burnin

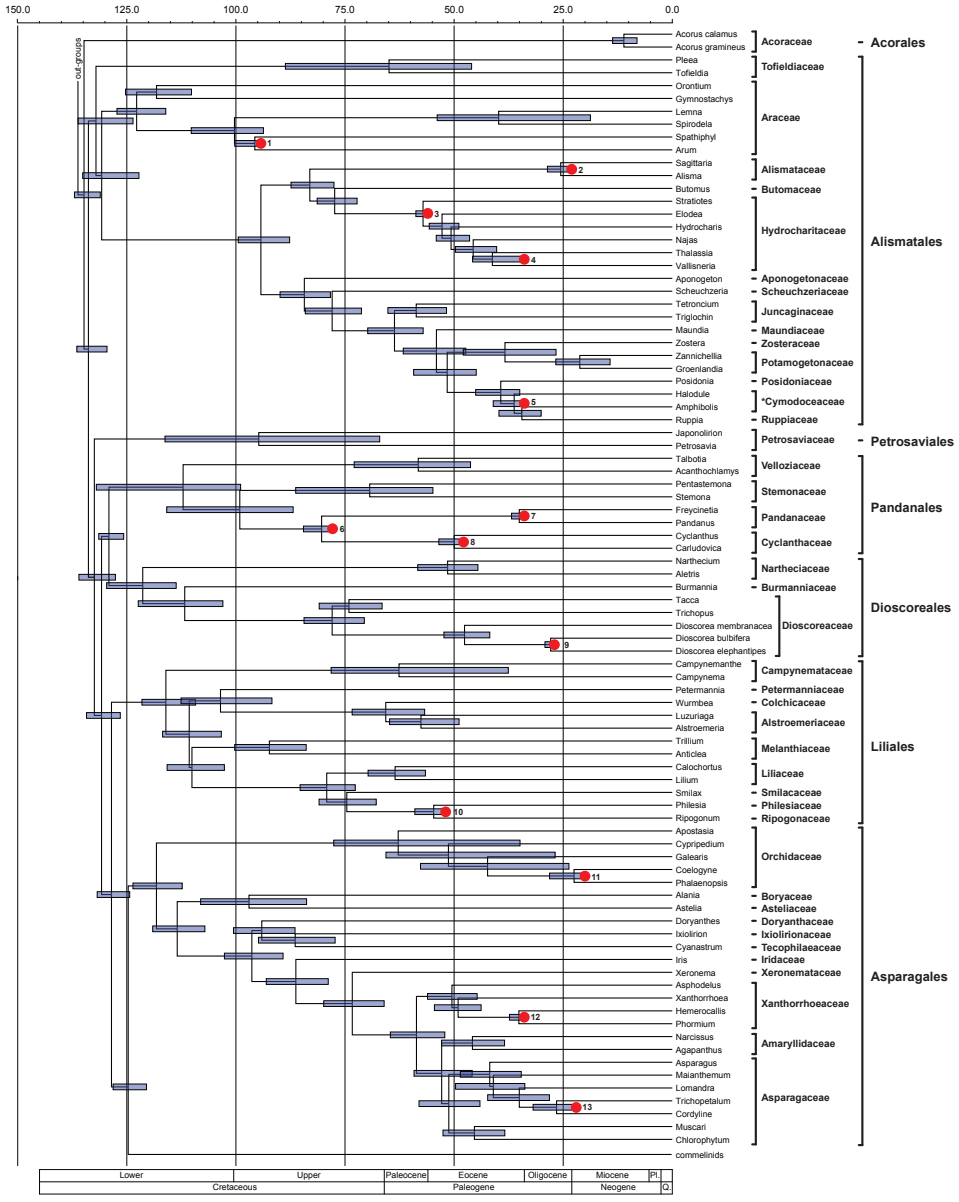
trees per run) trees were combined using TreeAnnotator version 1.7.4 (part of the BEAST package; Drummond and Rambaut 2007) and a maximum clade credibility tree was computed with mean node heights (Figures 5.4–5.5, Table 5.4; note that only monocots are depicted in these figures). The phylogenetic relationships are very similar to the ML tree (compare Figures 5.1–5.3 vs 5.4–5.5) and posterior support values are generally high across the tree (data not shown). The Bayesian chronogram (Figures 5.4–5.5, see also Table 5.4) indicates that all monocot orders originated by the end of the Lower Cretaceous. The crown clades of most orders also date from the Lower Cretaceous (except Acorales, Petrosaviales, and Zingiberales; Upper Cretaceous). Only one family, Araceae, has a crown age dating from the Lower Cretaceous (Figures 5.4–5.5, Table 5.4).

5.5 Discussion

Broad-scale molecular phylogenetic studies of monocots have been carried out for two decades. The earliest include studies by Duvall et al. (1993) and Chase et al. (1995). The phylogenetic results reported here are generally congruent with comparably sampled studies (Figures 5.1–5.3; Chase et al. 2006; Davis et al. 2006; Givnish et al. 2010; Graham et al. 2006), although support levels vary (summarized in Figures 5.6–5.7). Relatively few studies have attempted molecular dating analyses that consider the major lineages of monocots (e.g., Bremer, 2000; Givnish et al., 2005; Janssen and Bremer, 2004), although a number of more focused studies have been performed

---

Figure 5.4 *(following page)*: Monocot chronogram, part 1. Portion of the Bayesian chronogram inferred by BEAST with the local random clock model of rate evolution; this focuses on non-commelinid monocots (outgroups are not shown). This is based on a thinned dataset (153 taxa; only single-copy region genes sampled) with no partitions, under the GTR + Γ model of evolution. The time scale is in Ma; blue bars at nodes represent the 95% highest posterior density. Labelled red dots indicate the location and age of monocot fossil constraints (see Table 5.2).


150.0
125.0
100.0
75.0
50.0
25.0
0.0
out-groups
Acorus calamus
Acorus gramineus
Acoraceae
Acorales
Pleea
Tofieldia
Tofieldiaceae
Orontium
Gymnostachys
Lemna
Spirodela
Spathiphyl
Arum
Araceae
Sagittaria
Alisma
Alismataceae
Butomus
Butomaceae
Stratiotes
Elodea
Hydrocharis
Najas
Thalassia
Vallisneria
Hydrocharitaceae
Alismatales
Aponogeton
Aponogetonaceae
Scheuchzeria
Scheuchzeriaceae
Tetroncium
Triglochin
Juncaginaceae
Maundia
Maundiaceae
Zostera
Zosteraceae
Zannichellia
Groenlandia
Potamogetonaceae
Posidonia
Posidoniaceae
Halodule
Amphibolis
*Cymodoceaceae
Ruppia
Ruppiaceae
Japonolirion
Petrosavia
Petrosaviaceae
Petrosaviales
Talbotia
Acanthochlamys
Velloziaceae
Pentastemona
Stemona
Stemonaceae
Freycinetia
Pandanus
Pandanaceae
Pandanales
Cyclanthus
Carludovica
Cyclanthaceae
Nartheciaceae
Aletris
Burmanniaceae
Tacca
Trichopus
Dioscorea membranacea
Dioscorea bulbifera
Dioscorea elephantipes
Dioscoreaceae
Dioscoreales
Campynemanthe
Campynema
Campynemataceae
Petermannia
Petermanniaceae
Wurmbea
Colchicaceae
Luzuriaga
Alstroemeria
Alstroemeriaceae
Trillium
Anticlea
Melanthiaceae
Calochortus
Lilium
Liliaceae
Smilax
Smilacaceae
Philesia
Philesiaceae
Ripogonum
Ripogonaceae
Liliales
Apostasia
Cypripedium
Galearis
Coelogyne
Phalaenopsis
Orchidaceae
Alania
Boryaceae
Astelia
Asteliaceae
Doryanthes
Doryanthaceae
Ixiolirion
Ixiolirionaceae
Cyanastrum
Tecophilaeaceae
Iris
Iridaceae
Xeronema
Xeronemataceae
Asphodelus
Xanthorrhoea
Hemerocallis
Xanthorrhoeaceae
Phormium
Narcissus
Agapanthus
Amaryllidaceae
Asparagus
Maianthemum
Lomandra
Trichopetalum
Cordyline
Muscari
Chlorophytum
Asparagaceae
commelinids
Asparagales
1
2
3
4
5
6
7
8
9
10
11
12
13
Lower
Upper
Paleocene
Eocene
Oligocene
Miocene
Pl
Q
Cretaceous
Paleogene
Neogene

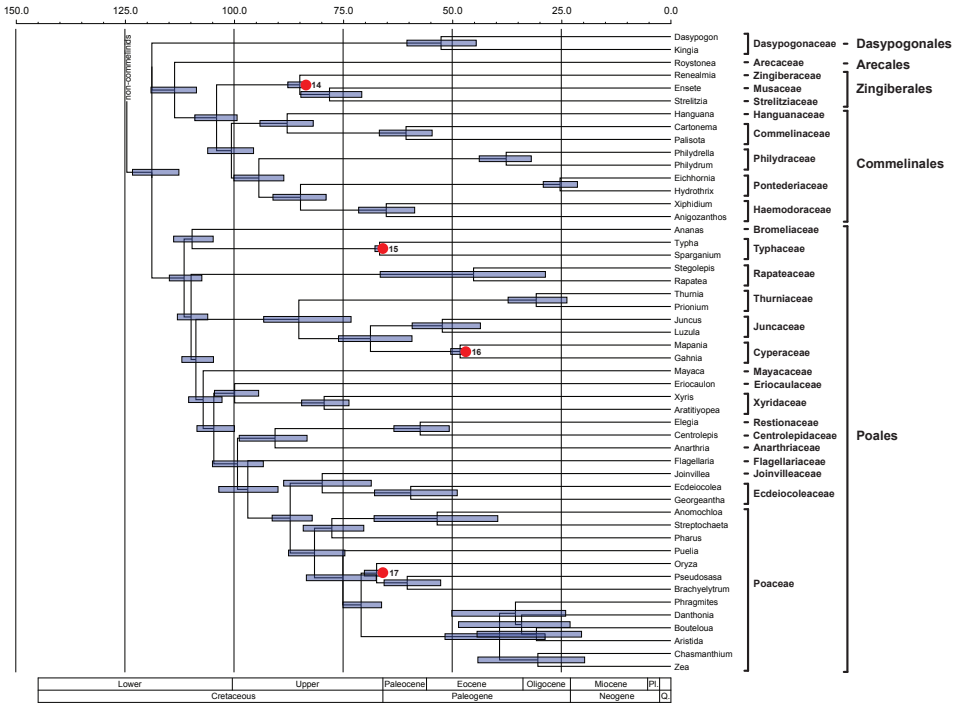


Figure 5.5: Monocot chronogram, part 2. A continuation of Figure 5.4, showing the Bayesian chronogram of commelinid monocots.

(see below). The overall picture is one of general consistency between my results and these earlier studies, particularly with regards to the age of the oldest lineages (Figure 5.8), but there are also some major deviations. In the following sections I discuss the implications of the current study for our understanding of phylogenetic relationships and clade ages in the monocots.

5.5.1 Monocot phylogeny

Relationships among the major angiosperm lineages

Relationships within the angiosperms are generally similar to previous iterations of this data set or subsets of it (Graham and Olmstead, 2000; Graham and Iles, 2009; Graham et al., 2006; Saarela et al., 2007), and are largely consistent with other studies (e.g., Jansen et al., 2007; Moore et al., 2007). They agree that *Amborella* and Nymphaeales represent the two earliest splits in angiosperm phylogeny (though their exact relationship to each and to other angiosperms is uncertain; Graham and Iles 2009), with Austrobaileyales then the sister group of mesangiosperms (Cantino et al., 2007). The ML and BI results depict Ceratophyllales as the sister group to monocots, but this relationship is poorly supported (Figure 5.1), as is a different relationship in the shortest MP trees, which instead recovered monocots as the sister group to a clade comprising Ceratophyllales and eudicots. Relationships among the major lineages of mesangiosperms continue to resist satisfactory resolution (see also Jansen et al., 2007; Moore et al., 2007, 2011; Soltis et al., 2011).

Relationships among the orders of monocots

I focus here on progress made since 2000, with key papers and support for them summarized in Figures 5.6 and 5.7. My results support Acorales, Alismatales, and Petrosaviales as successive sister groups of the remaining monocots (Figures 5.1–5.2) generally with strong support, which is consistent with several recent studies (Figure 5.6A). I also confirm the placement of *Japonolirion* as the sister taxon of *Petrosavia* (Petrosaviaceae, the sole family in Petrosaviales; Cameron et al. 2003; Davis et al. 2006; Fuse and Tamura 2000; Soltis et al. 2000). The stem of this taxon defines one of the

deepest splits in monocot phylogeny, Petrosaviideae (Graham et al., 2013). A few studies have suggested that Acorales belong within the Alismatales as the sister group of the alismatids (e.g., Davis et al., 2006, 2004; Petersen et al., 2006b) or even within Liliales (Duvall et al., 2008), however these results may be artefacts of one or more of the sampled mitochondrial sequences in these studies, which are often are inferred to have heavy RNA editing (Petersen et al., 2006b), and which show very high levels of heterotachy (Mower et al., 2007; Petersen et al., 2006a). The sister-group relationship between Dioscoreales and Pandanales, found here with only moderate support (Figure 5.2), is well supported by plastome-scale data (Givnish et al., 2010) and in one other study (Soltis et al., 2011).

Figure 5.6 *(following page)*: Monocot summary cladograms, part 1. Summary cladograms of major studies of monocots and of orders within monocots. The underlying cladogram is based on the ML PF16 phylogeny (Figures 5.1–5.3) and support values (indicated by branch thickness) correspond to ML bootstrap support values. Branches with <70% support were collapsed. Tick marks correspond to specific studies with thickness indicating support values. A grey branch (for the study here) or an inverted triangle indicates that support for this branch could not be evaluated, i.e., either no taxon was sampled or only one taxon was sampled. (A) Relationships among the orders of monocots. Order of tick marks, figure number and analysis type in parentheses for each study: 1—Chase et al. 2000 (Fig. 2, MP); 2—Davis et al. 2004 (Fig. 1, MP); 3—Tamura et al. 2004 (Fig. 1, MP); 4—Chase et al. 2006 (Fig. 2 and 3, MP); 5—Davis et al. 2006 (Fig. 2 and 3, MP); 6—Givnish et al. 2006 (Fig. 1, MP); 7—Graham et al. 2006 (Fig. 1B, MP); 8—Soltis et al. 2007 (Fig. 3 and 4, BI); 9—Givnish et al. 2010 (Fig. 3, ML); 10—Soltis et al. 2011 (Fig. 2, ML). (B) Relationship among the families of Alismatales. Order of tick marks: 1—Les et al. 1997 (Fig. 2, MP); 2—Chase et al. 2006 (Fig. 2, MP); 3—Petersen et al. 2006 (Fig. 3C, MP), 4—Von Mering & Kadereit 2010 (Fig. 3, ML). (C) Relationships among the families of Liliales. Order of tick marks: 1—Rudall et al. 2000 (Fig. 6, MP); 2—Vinnersten and Bremer 2001 (Fig. 2 and Table 3, MP); 3—Chase et al. 2006 (Fig. 2, MP); 4—Fay et al. 2006 (Fig. 1, MP); 5—Givnish et al. 2006 (Fig. 1, MP); 6—Petersen et al. 2012 (Fig. 1, MP). (D) Relationships among the families of Dioscoreales and Pandanales. Order of tick marks: 1—Chase et al. 2000 (Fig. 2, MP); 2—Caddick et al. 2002; 3—Chase et al. 2006 (Fig. 2, MP); 4—Davis et al. 2006 (Fig. 3, MP); 5—Givnish et al. 2006 (Fig. 1, MP); 6—Merckx et al. 2009 (Fig. 3, ML).

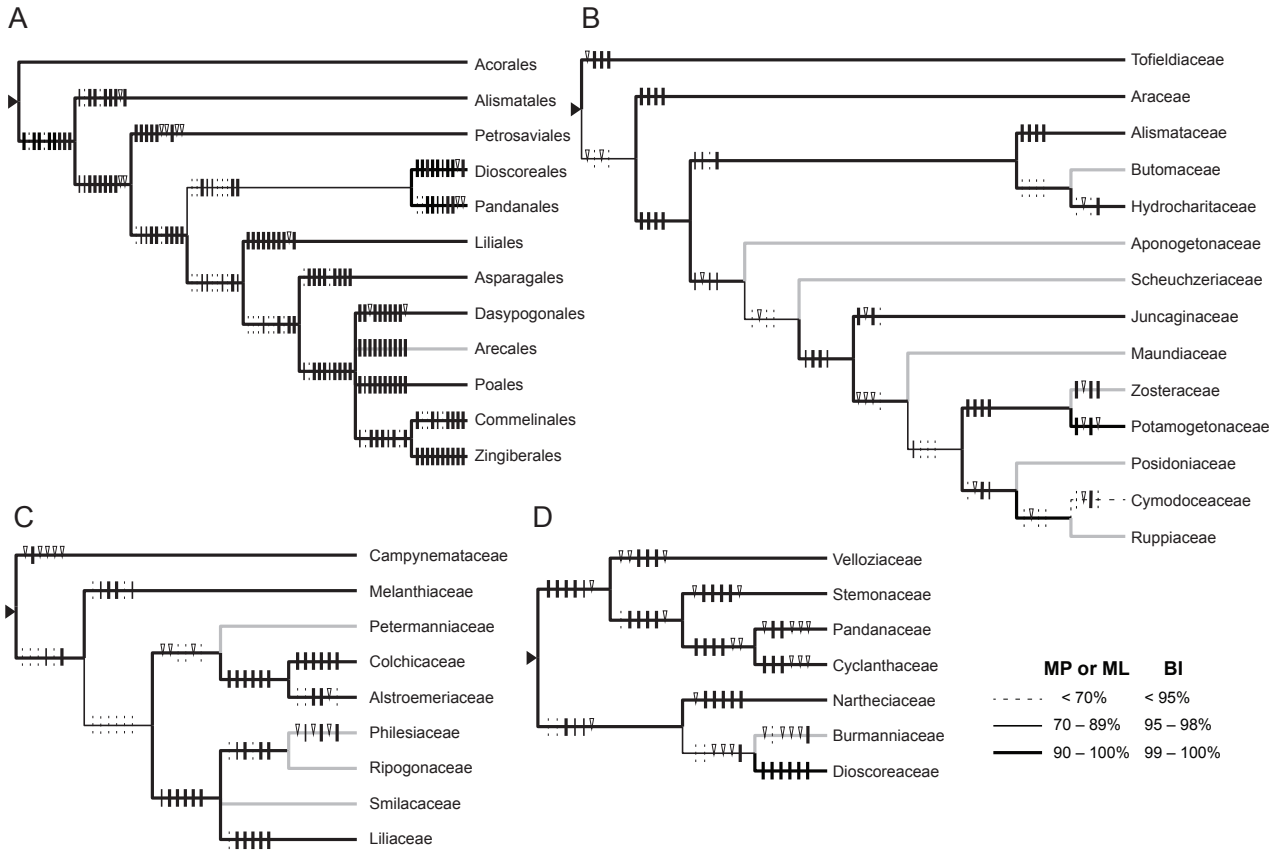
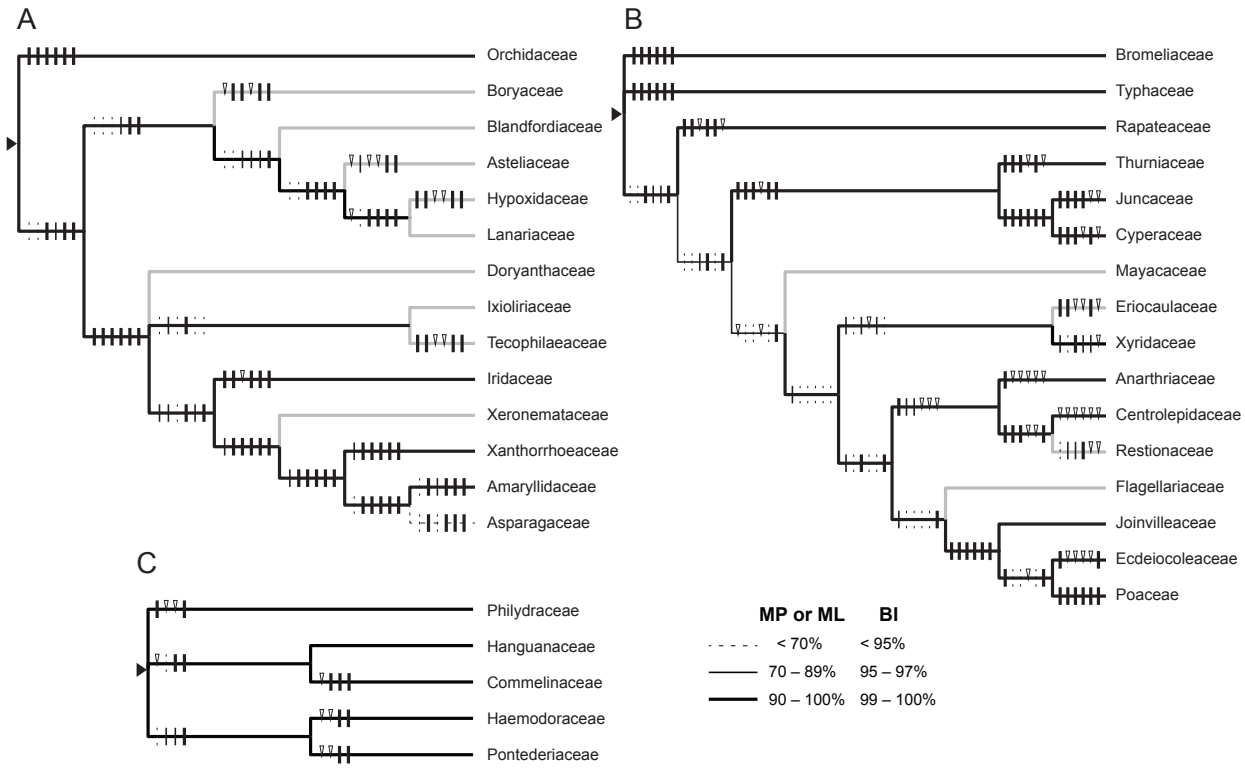

A
Acorales
Alismatales
Petrosaviales
Dioscoreales
Pandanales
Liliales
Asparagales
Dasypogonales
Arecales
Poales
Commelinales
Zingiberales
B
Tofieldiaceae
Araceae
Alismataceae
Butomaceae
Hydrocharitaceae
Aponogetonaceae
Scheuchzeriaceae
Juncaginaceae
Maundiaceae
Zosteraceae
Potamogetonaceae
Posidoniaceae
Cymodoceaceae
Ruppiaceae
C
Campynemataceae
Melanthiaceae
Petermanniaceae
Colchicaceae
Alstroemeriaceae
Philesiaceae
Ripogonaceae
Smilacaceae
Liliaceae
D
Velloziaceae
Stemonaceae
Pandanaceae
Cyclanthaceae
Nartheciaceae
Burmanniaceae
Dioscoreaceae
MP or ML
BI
< 70%
< 95%
70 – 89%
95 – 98%
90 – 100%
99 – 100%

Figure 5.7 *(following page)*: Monocot summary cladograms, part 2. A continuation of Figure 5.6 showing summary cladograms. (A) Relationships among the families of Asparagales. Order of tick marks: 1—Chase et al. 2000 (Fig. 2, MP); 2—Fay et al. 2000 (Fig. 2 and text, MP); 3—Chase et al. 2006 (Fig. 3, MP); 4—Graham et al. 2006 (Fig. 1B, MP); 5—Pires et al. 2006 (Fig. 2, MP) ; 6—Seberg et al. 2012 (Fig. 2, ML). (B) Relationships among the families of Poales. Order of tick marks: 1—Bremer 2002 (Fig. 1, MP); 2—Michelangeli et al. 2003 (Fig. 3, MP); 3—Chase et al. 2006 (Fig. 3, MP); 4—Davis et al. 2006 (Fig. 3, MP); 5—Givnish et al. 2006 (Fig. 1, MP); 6—Givnish et al. 2010 (Fig. 3, ML).(C) Relationships among the families of Commelinales. Order of tick marks: 1—Chase et al. 2006 (Fig. 3, MP); 2—Davis et al. 2006 (Fig. 3, MP); 3—Givnish et al. 2006 (Fig. 1, MP); 4—Saarela et al. 2008 (Fig. 2, ML).


A
Orchidaceae
Boryaceae
Blandfordiaceae
Asteliaceae
Hypoxidaceae
Lanariaceae
Doryanthaceae
Ixioliriaceae
Tecophilaeaceae
Iridaceae
Xeronemataceae
Xanthorrhoeaceae
Amaryllidaceae
Asparagaceae
B
Bromeliaceae
Typhaceae
Rapateaceae
Thurniaceae
Juncaceae
Cyperaceae
Mayacaceae
Eriocaulaceae
Xyridaceae
Anarthriaceae
Centrolepidaceae
Restionaceae
Flagellariaceae
Joinvilleaceae
Ecdeiocoleaceae
Poaceae
C
Philydraceae
Hanguanaceae
Commelinaceae
Haemodoraceae
Pontederiaceae
MP or ML
BI
< 70%
< 95%
70 – 89%
95 – 97%
90 – 100%
99 – 100%

The precise relationships among the three remaining major lineages, Asparagales, Liliales, and commelinids have been more contentious. A sister-group relationship between Asparagales and commelinids is strongly supported here by the two model-based methods (with moderate support from MP); Liliales is also well supported as the sister group of these two (with weak support from MP). These results were also recovered in several recent studies (e.g., Chase et al., 2006; Givnish et al., 2010; Soltis et al., 2011) with moderate to strong support (Figure 5.6A), and in Graham et al. (2006) with similar levels of MP support to the current study (70% vs 66%). Studies that did not recover this set of relationships generally had alternative relationships that were relatively poorly supported (Davis et al., 2006, 2004; Soltis et al., 2007). A recent plastome-scale study (Liu et al., 2012) addressed the placement of Liliales with different data partitions and concluded that the best supported position was as the sister group of the clade comprising Asparagales and commelinids (i.e., the result found here). However, this placement was not significantly better than one with Liliales sister to a clade comprising Dioscoreales and Pandanales, according to the Shimodaira-Hasegawa test (SH; Shimodaira and Hasegawa 1999). Currently, plastome-scale studies tend to only have sparse taxon sampling (e.g., Givnish et al., 2010; Liu et al., 2012). Improved taxon sampling for data sets of this scale may help further in this situation (Heath et al., 2008; Hedtke et al., 2006; Hillis, 1998; Zwickl and Hillis, 2002).

Relationships among the commelinids have proved to be exceptionally difficult to resolve, except for the relationship between Commelinales and Zingiberales (Chase et al., 2000; Davis et al., 2006, 2004; Givnish et al., 2006; Graham et al., 2006; Soltis et al., 2007, 2011; Tamura et al., 2004b). The model-based analyses here suggest a relationship in which Dasypogonales are the sister group of the remaining members of the commelinids, with Arecales then the sister group of a clade comprising Commelinales and Zingiberales. However these sets of relationships are only poorly supported (Figure 5.3). An alternative relationship is suggested in the shortest MP trees (not shown), in which Dasypogonales are the sister group of Poales, and Arecales are the sister group of a clade comprising Commelinales and Zingib-

erales; however this arrangement is also poorly supported (data not shown). Several other studies have also shown conflicting sets of relationships that are moderately well supported, depicting, for example: Dasypogonales as the sister group of Poales (Davis et al., 2006), and Poales as the sister group of a clade comprising Commelinales and Zingiberales (Tamura et al., 2004b). A recent plastome-scale study by Givnish et al. (2010) grouped Arecales and Dasypogonales together with moderate support in an ML analysis, and found Poales to be the sister group of a clade comprising Commelinales and Zingiberales, with strong support. However, the MP analysis from the same study found a conflicting set of relationships, with the depicted relationships were poorly supported (except for Commelinales-Zingiberales). An expanded version of this data set was used to explore relationships among the commelinids (Barrett et al., 2013). They found the same set of relationships as the Givnish et al. (2010) study with similar support values. They were unable to reject seven of the fifteen possible alternative arrangements of the five commelinid orders using SH tests (Barrett et al., 2013).

Relationships within Alismatales

As Alismatales phylogeny is the topic of Chapter 4 I will not discuss it further here, except to note that the new sample of Juncaginaceae (*Tetroncium*) groups with the other sampled species of the family (*Triglochin*), and that this addition does not affect the placement of Maundiaceae or modify the support to a great extent (a segregate family of Juncaginaceae; Figures 4.2, 5.1 and 5.6B). This result is similar to that depicted in von Mering and Kadereit (2010) and Les and Tippery (2013).

Relationships within Dioscoreales-Pandanales

The relationships among the autotrophic members of Pandanales are uncontroversial (Figure 5.6D): Velloziaceae (including *Acanthochlamys*) and Stemonaceae (including *Pentastemona*) are consecutive sister groups of a clade comprising Cyclanthaceae and Pandanaceae in several studies. These relationships are also well supported in the current study, as are support

values for individual families (Figure 5.2). The mycoheterotrophic family Triuridaceae has been associated with Pandanales with moderate levels of support, although its position within the order is poorly understood (Chase et al., 2006, 2000; Davis et al., 2004; Lemaire et al., 2011; Vergara-Silva et al., 2003). This family was not sampled in the current study.

Nartheciaceae are moderately supported as the sister group of a clade comprising Burmanniaceae and Dioscoreaceae in Dioscoreales (Figure 5.2). This set of relationships has been occasionally recovered before (see especially Merckx et al. 2006, 2009), although many monocot-wide studies have not sampled the mycoheterotrophic family Burmanniaceae (e.g., Chase et al., 2006; Fuse and Tamura, 2000; Tamura et al., 2004b), or Nartheciaceae (e.g., Graham et al. 2006). The current data also has complete genus-level coverage of Nartheciaceae. Within Nartheciaceae *Aletris* and *Metanarthecium* are relatively distantly related sister taxa that comprise the sister group of *Lophiola*, *Narthecium*, and *Nietneria*; these relationships are all moderately to strongly supported, as is the sister-group relationship between *Narthecium* and *Nietneria* (Figure 5.2). Previous studies have supported similar relationships, except for the placement of *Metanarthecium*, which has been variable (Caddick et al., 2002; Fuse et al., 2012; Fuse and Tamura, 2000; Merckx et al., 2008b; Tamura et al., 2004b; Zhao et al., 2012). Partly this was due to a mis-identified *Aletris* specimen (labeled as *Metanarthecium luteoviride* Inoue s.n., voucher location unknown, available from the Royal Botanic Gardens Kew DNA Bank, `http://apps.kew.org/dnabank/homepage.html`; Fuse et al. 2012; Zhao et al. 2012; data not shown) which appears to have erroneously placed *Metanarthecium* within *Aletris* or very closely related to it (Caddick et al., 2002; Merckx et al., 2008b). The current circumscription of Burmanniaceae is broad (Angiosperm Phylogeny Group 2009) and includes the segregate family Thismiaceae J.Agardh, a group of obligate mycoheterotrophs, as the tribe Thismieae. Burmanniaceae s.s. includes both chlorophyllous and completely mycoheterotrophic lineages (Merckx et al., 2006) . The species of Burmanniaceae sampled here (*Burmannia capitata*) is chlorophyllous. A suite of studies by Merckx and colleagues (Merckx et al., 2009, 2010, 2006) have

suggested that in its current broad circumscription, Burmanniaceae are not monophyletic and that tribe Thismieae is near or within Dioscoreaceae as one or possibly two distinct lineages (most of the tribe vs. *Afrothismia*). Members of Thismieae were not included here.

The most recent APG classifications lumped Taccaceae and Trichopodaceae in Dioscoreaceae (APG 2003, 2009). I attempted to include all currently accepted genera of Dioscoreaceae (*Dioscorea* [including e.g., *Epipetrum*, *Rajania*, *Tamus*], *Stenomeris*, *Tacca*, and *Trichopus* [including *Avetra*]). A specimen of *Stenomeris* that I considered including (*S. dioscoreifolia* Risdale 550, available from the Royal Botanic Gardens Kew DNA Bank, `http://apps.kew.org/dnabank/homepage.html`) is likely a mis-identified species of *Dioscorea*, as a study that used this specimen found it nested within *Dioscorea* with strong support (Merckx et al., 2006). Previous molecular studies placed *Stenomeris* firmly outside of *Dioscorea* (Caddick et al., 2002; Wilkin et al., 2005) a position that is also supported by morphological data (Caddick et al., 2002; Caddick and Wilkin, 1998; Wilkin et al., 2005). Relationships among the genera of Dioscoreaceae therefore require additional attention.

Relationships among Liliales and Asparagales

The relationships among the deepest splitting lineages of Liliales have been difficult to resolve, making it unclear which taxa define the root split in the order (Chase et al., 2006; Givnish et al., 2006; Rudall et al., 2000)). However two studies focused on the order found moderate to high support for Campynemataceae as the sister group of all other members of Liliales (Figure 5.6C). My data also strongly support this root split (Figure 5.2). The relationships among the four remaining well supported lineages (Melanthiaceae; monotypic Petermanniaceae; a clade comprising Alstroemeriaceae and Colchicaceae; and a clade comprising Liliaceae, Philesiaceae, Ripogonaceae, and Smilacaceae) have also proved resistant to satisfactory resolution. Earlier studies placing Petermanniaceae within Colchicaceae (Chase et al., 2000; Rudall et al., 2000) were a result of mis-labelled *Tripladenia*

(Colchicaceae; sampled here Figure 5.2; see also Chase et al. 2006). Genuine material of *Petermannia* was weakly supported as the sister group of a clade comprising Alstroemeriaceae and Colchicaceae (Chase et al., 2006; Fay et al., 2006; Petersen et al., 2013). I find moderate to strong support for this relationship, for the first time (Figures 5.2 and 5.6C). Even with this advance, relationships among the now three well supported groups are contentious. The ML results recover moderate support for Melanthiaceae as the sister group to the remaining two groups (a clade comprising Alstroemeriaceae, Colchicaceae, and Petermanniaceae; and a clade comprising Liliaceae, Philesiaceae, Ripogonaceae, Smilacaceae), however MP and BI support for this arrangement is substantially lower (85% vs 60% and 56% respectively) and the branch subtending this clade is also very short (Figure 5.2). Relationships within the last group have also proven difficult to resolve. My data suggest a weakly to moderately supported relationship among three lineages that share a predominately vining habit: Philesiaceae, Ripogonaceae and Smilacaceae. However, Fay et al. (2006) and Petersen et al. (2013) found a strongly supported relationship between Liliaceae and Smilacaceae, which appears to be driven by the plastid component of the data set (the other component represented by mitochondrial sequences), although in the Petersen et al. (2013) study the plastid data is a subset of the current data (*rbc*L and *ndh*F), which do not support this arrangement.

Recent studies with broad taxonomic sampling in Liliaceae have identified four major clades: (i) *Lilium* and its allies (*Amana*, *Cardiocrinum*, *Clintonia*, *Erythronium*, *Fritillaria*, *Gagea*, *Lloydia*, *Medeola*, *Nomocharis*, *Notholirion* and *Tulipa*), (ii) *Prosartes* and its allies (*Scoliopus* and *Streptopus*), (iii) *Calochortus*, and (iv) *Tricyrtis* (Fay et al., 2006; Patterson and Givnish, 2002; Petersen et al., 2013). Although Patterson and Givnish (2002) found moderate support for *Prosartes* and allies being the sister group to a clade comprising *Calochortus* and *Tricyrtis*, later studies (Fay et al., 2006; Petersen et al., 2013) and the present study (Figure 5.2) fail to recover any moderately to well supported groupings among these clades.

Relationships within Asparagales are strongly congruent with those of Graham et al. (2006), and have similar levels of support (Figure 5.7A).

The major difference between the current study and that of Graham et al. (2006) is the support and placement of Doryanthaceae. Doryanthaceae are poorly supported here as the sister group of a clade comprising Ixioliriaceae and Tecophilaeaceae (Figure 5.2), whereas Graham et al. (2006) strongly supported the family as the sister group of a large clade comprising Amaryllidaceae, Asparagaceae, Iridaceae, Xanthorrhoeaceae, and Xeronemataceae. Although the taxon composition in Asparagales in these two studies is nearly identical (except for the addition of *Apostasia* in Orchidaceae, and *Cordyline* and *Trichopetalum* in Asparagaceae), the genomic composition considered here are a subset of those in Graham et al. (2006), who included additional multiple noncoding plastid regions. Other studies have not fully resolved the position of Doryanthaceae (Chase et al., 2006; Fay et al., 2000; Pires et al., 2006; Seberg et al., 2012), and its precise placement within Asparagales remains problematic. The low support seen for the crown clade of Asparagaceae is probably due to the inclusion of *Aphyllanthus* (see Graham et al., 2006), and the very short length of the family's stem lineage.

Relationships within Poales

Recent studies on Poales phylogeny have recognized several distinct lineages within the order (Figure 5.7B). These include of several early-splitting families (Bromeliaceae, Rapateaceae, and Typhaceae) and several informally recognized groupings: the cyperids (Cyperaceae, Juncaceae, and Thurniaceae), graminids (Ecdeiocoleaceae, Flagellariaceae, Joinvilleaceae, and Poaceae), restiids (Anarthriaceae, Centrolepidaceae, and Restionaceae), and xyrids (Eriocaulaceae, Mayacaceae, and Xyridaceae). The relationship here (Figure 5.3), in which Bromeliaceae and Typhaceae comprise a clade (poorly supported) has been found in other recent studies (Bremer, 2002; Davis et al., 2006; Givnish et al., 2006), with varying support levels. The plastome-scale study of Givnish et al. (2010) found Bromeliaceae and Typhaceae to be successive sister groups to the remaining families, with moderate to strong support. The placement of Rapateaceae has also proved resistant to resolution (Figure 5.7B), but my results strongly support Rapateaceae as the sister

group of all other Poales (Figure 5.3). I recover the cyperids, Mayacaceae (xyrid p.p.), Eriocaulaceae and Xyridaceae (xyrids p.p.), and restiids as successive sister groups to the graminids, with moderate to strong support in the model-based analyses (Figure 5.3). Maximum parsimony found a similar set of relationships, except that *Mayaca* is embedded in the xyrids as the sister group to Eriocaulaceae, with moderate support (78%). However, support across the broad backbone of Poales phylogeny is poorly supported in the MP analysis here. Similar relationships were found by others (Bremer, 2002; Givnish et al., 2010), but none of these studies recovered Mayacaceae as the sister group of a clade comprising the rest of the xyrids, the restiids, and the graminids. Bremer (2002) found the same set of relationships with varying levels of support (MP analysis; Figure 5.7B) but did not include Mayacaceae. Givnish et al. (2010) sampled Mayacaceae and found that it too grouped with Eriocaulaceae; this clade was in turn the sister group of a clade comprising Xyridaceae and the graminids, a set of relationships with strong support (ML analysis). The MP analysis of Givnish et al. (2010) mirrored my MP analysis in recovering a poorly supported xyrid clade (68% vs 68%). Other studies found arrangements the relationship recovered here in which cyperids and xyrids comprise a clade that is the sister group of a clade comprising the restiids and graminids (Chase et al. 2006; and in the MP analyses of Givnish et al. 2010, 2006). However, these had only moderate levels of support (Figure 5.7B).

Within the graminids, Flagellariaceae, Joinvilleaceae, and Ecdeiocoleaceae are successive sister groups of the Poaceae, with support high support for nearly all branches except for the clade of Ecdeiocoleaceae and Poaceae, which had poor support from MP and moderate support from BI (Figure 5.3). This is similar to recent studies (Figure 5.7B), especially the taxonomically broadly sampled Bremer (2002) and the genomically densely sampled study by Givnish et al. (2010). Surprisingly, this result somewhat conflicts with a previous iteration of the current data set (Saarela and Graham, 2010), which did not resolve intra-graminid relationships with high support using MP or ML. A Bayesian analysis by them did recover most of the relationships seen here with high support, but grouped Joinvilleaceae and

Ecdeiocoleaceae together with poor support (Saarela and Graham, 2010). These mild conflicts might be due to sparser taxon sampling in Ecdeiocoleaceae, Poaceae and the restiids in the older study (Saarela and Graham, 2010). Intra-family relationships in Poaceae remain essentially unchanged since Saarela and Graham (2010), and the placement of Puelioideae (*Puelia* added here) is that seen in other large-scale Poaceae phylogenetic projects (Figure 5.3; e.g., Grass Phylogeny Working Group 2001; Grass Phylogeny Working Group II 2012).

Table 5.4: Ages of major monocot clades. Ages inferred for taxonomic groups (Figures 5.4–5.5) using the random local clocks method in BEAST (Drummond and Rambaut 2007; Drummond and Suchard 2010). Ranks above family (both formal and informal) are in bold font. Ages and the 95% highest posterior density (HPD) are given for crown and stem ages where possible.

Taxon	Crown age (95% HPD)	Stem age (95% HPD)
Angiosperms	140 (139–140)	NA
Nymphaeales (excl. Hydatellaceae)	113.2 (100.5–132.7)	139.8 (138.6–140)
Austrobaileyales	115.5 (105.9–122.3)	138.0 (134.9–139.1)
Mesangiosperms	136.6 (133.4–138.2)	138.0 (134.9–139.1)
Eudicots	116.9 (105.1–135.2)	136.6 (133.4–138.2)
Magnoliids	129.7 (126.5–132.4)	136.6 (133.4–138.2)
Monocots	135.2 (131.4–137.4)	136.6 (133.4–138.2)
Acorales (= Acoraceae)	11.5 (8.5–14.0)	135.2 (131.4–137.4)
Alismatales	132.4 (123.9–136.5)	134.2 (129.9–137.4)
Alismataceae	25.9 (23.0–29.0)	83.4 (77.9–87.7)
Aponogetonaceae	NA	84.7 (78.7–90.2)
Araceae	123.1 (116.4–127.6)	131.1 (122.6–135.5)
Butomaceae	NA	77.7 (72.6–81.7)

Taxon	Crown age (95% HPD)	Stem age (95% HPD)
Cymodoceaceae (incl. Ruppiaceae)	36.6 (33.9–41.4)	39.7 (35.3–45.4)
Hydrocharitaceae	57.5 (56.0–59.1)	77.7 (72.6–81.7)
Juncaginaceae	59 (52.1–65.5)	64 (57.4–70.2)
Maundiaceae	NA	54.4 (47.7–62.0)
Posidoniaceae	NA	39.7 (35.3–45.4)
Potamogetonaceae	21.6 (14.6–27.0)	38.7 (27.0–48.3)
Scheuchzeriaceae	NA	78.3 (71.6–84.5)
Tofieldiaceae	65.3 (46.4–89.0)	132.4 (129.9–137.4)
Zosteraceae	NA	38.7 (27.0–48.3)
Petrosaviales (= Petrosaviaceae)	95.1 (67.4–116.6)	132.8 (128.0–136.3)
Dioscoreales	121.7 (114.0–130.0)	129.5 (126.1–131.8)
Burmanniaceae	NA	112.1 (103.3–122.7)
Dioscoreaceae	78.3 (70.9–84.8)	112.1 (103.3–122.7)
Nartheciaceae	51.9 (44.9–58.7)	121.7 (114.0–130.0)
Pandanales	112.5 (99.3–132.3)	129.5 (126.1–131.8)
Cyclanthaceae	51.9 (47.8–53.9)	80.7 (77.8–84.9)
Pandanaceae	35.4 (33.9–37.2)	80.7 (77.8–84.9)
Stemonaceae	69.7 (55.2–86.7)	99.5 (87.2–116.2)
Velloziaceae	58.6 (46.6–73.3)	112.5 (99.3–132.3)
Liliales	116.4 (109.6–121.9)	128.9 (124.7–132.2)
Alstroemeriaceae	58 (49.2–65.2)	66 (57.1–73.7)
Campynemataceae	63 (37.9–78.5)	116.4 (109.6–121.9)
Colchicaceae	NA	66 (57.1–73.7)
Liliaceae	63.9 (56.9–70.1)	79.5 (73.0–85.7)
Melanthiaceae	92.7 (84.3–100.6)	110.4 (103.0–116.1)
Petermanniaceae	NA	103.9 (92.1–112.9)
Philesiaceae	NA	55 (51.9–59.4)
Ripogonaceae	NA	55 (51.9–59.4)
Smilacaceae	NA	75 (68.2–81.3)

Taxon	Crown age (95% HPD)	Stem age (95% HPD)
Asparagales	118.6 (112.7–124.0)	125 (120.9–128.5)
Amaryllidaceae	46.2 (38.8–53.2)	53.2 (46.2–59.5)
Asparagaceae	51.6 (44.4–58.4)	53.2 (46.2–59.5)
Asteliaceae	NA	97.4 (84.2–108.4)
Boryaceae	NA	97.4 (84.2–108.4)
Doryanthaceae	NA	94.4 (86.8–100.9)
Iridaceae	NA	86.6 (79.2–93.4)
Ixioliriaceae	NA	86.8 (77.6–95.2)
Orchidaceae	63.2 (35.3–77.9)	118.6 (112.7–124.0)
Tecophilaeaceae	NA	86.8 (77.6–95.2)
Xanthorrhoeaceae	50.9 (45.1–56.4)	59 (52.5–65.0)
Xeronemataceae	NA	73.7 (66.4–80.2)
Commelinidae	119.3 (113.1–123.7)	125 (120.9–128.5)
Arecales	NA	113.7 (108.7–119.1)
(= Arecaceae)		
Dasypogonales	52.7 (44.6–60.5)	119.3 (113.1–123.7)
(= Dasypogonaceae)		
Commelinales	100.7 (95.6–106.2)	104.1 (99.4109.1)
Commelinaceae	60.7 (54.7–66.8)	87.9 (82.0–94.2)
Haemodoraceae	65.2 (58.7–71.6)	84.9 (79.0–91.2)
Hanguanaceae	NA	87.9 (82.0–94.2)
Philydraceae	37.7 (32.0–43.9)	94.4 (88.7–100.1)
Pontederiaceae	25.4 (21.4–29.3)	84.9 (79.0–91.2)
Zingiberales	85 (83.6–87.8)	104.1 (99.4–109.1)
Musaceae	NA	78.2 (70.8–84.8)
Strelitziaceae	NA	78.2 (70.8–84.8)
Zingiberaceae	NA	85 (83.6–87.8)
Poales	111.9 (107.9–115.3)	119.3 (113.1–123.7)
Anarthriaceae	NA	91.7 (83.7–99.3)
Bromeliaceae	NA	110.1 (105.2–114.3)
Centrolepidaceae	NA	57.8 (51.2–63.8)

Taxon	Crown age (95% HPD)	Stem age (95% HPD)
Cyperaceae	52.7 (47.0–50.8)	69.2 (59.7–76.5)
Ecdeiocoleaceae	60 (49.3–68.3)	80.3 (69.0–89.1)
Eriocaulaceae	NA	100.3 (94.8–105.0)
Flagellariaceae	NA	97.3 (90.4–104.0)
Joinvilleaceae	NA	80.3 (69.0–89.1)
Juncaceae	48.7 (44.0–59.7)	69.2 (59.7–76.5)
Mayacaceae	NA	107.5 (103.2–110.9)
Poaceae	82 (75.1–88.0)	87.6 (82.6–91.7)
Rapateaceae	45.6 (29.1–67.0)	109.2 (106.5–113.5)
Restionaceae	NA	57.8 (51.2–63.8)
Thurniaceae	31.2 (24.2–37.7)	85.6 (73.7–93.7)
Typhaceae	67.1 (66.0–68.2)	110.1 (105.2–114.3)
Xyridaceae	79.8 (74.1–85.0)	100.3 (94.8–105.0)

5.5.2 Dating analysis

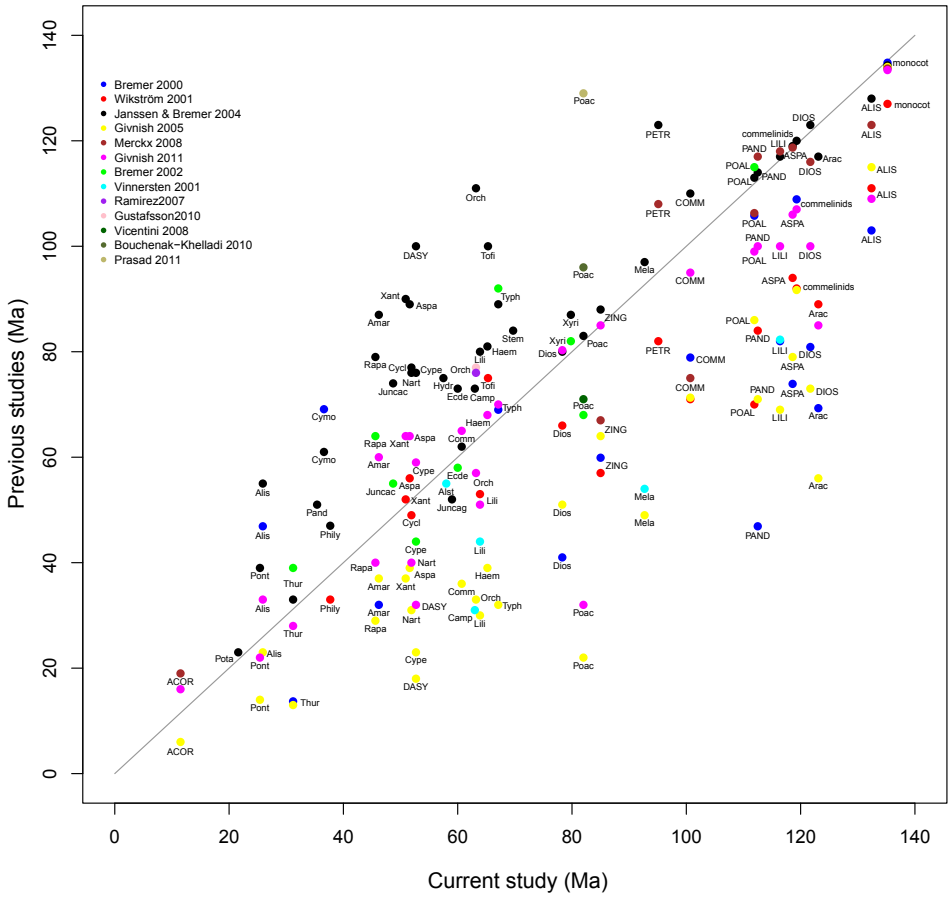
The crown age of monocots inferred here (136 Ma; Figure 5.4) is very similar to other studies (Bell et al., 2005; Bremer, 2000; Magallón, 2010; Magallón and Sanderson, 2005; Wikström et al., 2001, see Figure 5.8) which indicate an age range from 127–137 Ma. The earliest monocot or monocot-like fossils (in addition to those discussed above, including the fossil pollen *Liliacidites* and the vegetative shoot *Acaciaephyllum*; Doyle et al. 2008) are thought to be somewhat younger than these ages (∼10–20 Ma, Aptian vs Hauterivian or Valanginian, see Doyle 2012; Doyle et al. 2008). A few studies have proposed much older ages (156–177 Ma) for the monocot crown clade (Magallón and Castillo, 2009; Smith et al., 2010a). These older ages are more difficult to reconcile with the fossil record, as they suggest a gap of ∼30–60 Ma between the crown clade and any subsequent fossilization.

Within monocots the most extensively sampled study of monocot divergences is that of Janssen and Bremer (2004) who assembled a *rbc*L data-set

of 878 genera and used the resulting phylogeny (based on a constrained phylogenetic backbone and a crown constraint of 134 Ma) to date the diversification of monocots using non-parametric rate smoothing (Sanderson, 1997). They inferred that a large number of families have crown clades that date from the Cretaceous. I find that a majority of sampled families (55 of 68) originated in the Cretaceous (stem ages). However, my crown ages are generally substantially younger than those seen by Janssen and Bremer (2004), as indicated by the black dots above the grey line (Figure 5.8). Only eight families sampled in my study had crown clades extending back to the Cretaceous (Figures 5.4, 5.5, and 5.8; Table5.4). The largest disparities, for age estimates from Janssen and Bremer (2004) that are nearly twice my estimates, are for crown Orchidaceae (111 vs 63 Ma), and crown Dasypogonales (100 vs 53 Ma). Although our taxon samplings were not comparable (especially for Orchidaceae, for which Janssen and Bremer sampled 148 samples vs 5 here) my smaller sampling likely includes the root split in both cases, and so this age-estimate difference likely does not reflect taxon sampling differences (see Cameron, 2004; Górniak et al., 2010; Rudall and Conran, 2012).

Several later studies, particularly those of Givnish et al. (2011, 2005)

Figure 5.8 *(following page)*: Scatter plot of monocot crown clade ages. The x-axis represents mean crown clade ages inferred here; the y-axis represents clade ages from several previous studies. Most studies are monocot-wide or focused on orders (Bremer 2000, 2002; Vinnersten and Bremer 2001; Wikstrm et al. 2001; Janssen and Bremer 2004; Givnish et al. 2005, 2011; Merckx et al. 2008a), except for several focused on large families: Orchidaceae (Ramrez et al. 2007; Gustafsson et al. 2010) and Poaceae (Vicentini et al. 2008; Bouchenak-Khelladi et al. 2010; Prasad et al. 2011). Grey line represents y = x. If sampling was likely to omit the deepest split in a taxon based on the current literature, it was omitted from representation (both for the current study and reviewed studies). Taxon names are truncated to the first four letters in most cases, in a few cases it was necessary to add one or two letters to distinguish taxa with similar names (e.g., Philesiaceae = Phile; Philydraceae = Phily). Standard capitalization = family names; all letters capitalized = order names.


Bremer 2000
Wikström 2001
Janssen & Bremer 2004
Givnish 2005
Merckx 2008
Givnish 2011
Bremer 2002
Vinnersten 2001
Ramirez2007
Gustafsson2010
Vicentini 2008
Bouchenak−Khelladi 2010
Prasad 2011
Previous studies (Ma)
Current study (Ma)
0
20
40
60
80
100
120
140

(magenta and yellow dots, respectively) yielded younger crown-age estimates than Janssen and Bremer (2004) for a wide range of families (Figure 5.8). The earlier study (Givnish et al., 2005), found ages younger than those found here (yellow dots below grey line; Figure 5.8). That study employed a fixed age at the root of the monocots (134 Ma) and no additional fossil constraints. The later study of Givnish et al. (2011), added six constraints from Bremer (2000), and the crown ages for many families and orders are relatively congruent with those inferred here (Figure 5.8, magenta dots), despite the differences in methodology (PL vs BI), taxon sampling, and fossil constraints between our studies (only the *Spirematospermum* and *Typha* fossils are shared).

Several recent studies have focused on the species-rich families Orchidaceae and Poaceae. Recent fossil finds in Orchidaceae (Conran et al., 2009a; Ramírez et al., 2007) allow for internal calibration of the family, which should result in more accurate dating. The ages inferred for the crown clade of Orchidaceae by Ramírez et al. (2007) and Gustafsson et al. (2010) are nearly identical (∼76 Ma) and overlap with my results (Figure 5.8, Table 5.4). Inferred ages for the crown of Poaceae are quite divergent in recent studies (e.g., Bouchenak-Khelladi et al., 2010; Prasad et al., 2011; Vicentini et al., 2008, Figure 5.8). The results here (Figure 5.8) are intermediate between those of other studies that also have Poaceae calibration points (Bouchenak-Khelladi et al., 2010; Prasad et al., 2011; Vicentini et al., 2008); studies lacking Poaceae calibrations (Givnish et al., 2011, 2005) generally have younger ages (but see Janssen and Bremer 2004). This may be due to different fossil choice (or lack of fossils in Poaceae). However, one of the most notable aspects of Poales plastid evolution is the high molecular rate seen in some core poalean families (e.g., Figure 5.3; Gaut et al. 1992). Rate variation like this, combined with poor model fit, may be responsible for some of these and other divergent results seen for monocot families across different studies (Figure 5.8; Ho 2009; Lepage et al. 2007).

5.6 Conclusions

There are significant improvements in resolution and support across the monocot tree here compared to earlier studies using sub-sets of the current data (Figures 5.1–5.3, 5.6A; Graham et al. 2006; Saarela and Graham 2010; Saarela et al. 2008, 2007) these improvement are concentrated in Alismatales (Figure 5.6B, see Chapter 4), Dioscoreales (Figure 5.6D), Pandanales (Figure 5.6D), Liliales (Figure 5.6C), and Poales (Figure 5.7B). This may reflect the increased sampling in these orders. Overall inter-relationships of monocot orders and families found here are similar to recent studies (e.g., Chase et al., 2006; Givnish et al., 2010) while improving our understanding of key lineages (discussed above). The use of 25 local fossil calibrations (17 in monocots) across the tree along with a molecular dating method to model rate heterogeneity (Drummond and Suchard, 2010) should allow for a more accurate estimation of divergence times, than studies employing fewer fossil constraints and less sophisticated analytical methods (e.g., Bremer, 2000; Givnish et al., 2011, 2005; Janssen and Bremer, 2004; Merckx et al., 2008a; Wikström et al., 2001).

My study more consistently and uniformly provided strong support for major monocot clades than any single earlier study (see Figures 5.6 and 5.7), and refined several relationships that were previously unclear, with moderate to strong support here (generally at least matching other studies). Future studies that further expand the amount of genomic data collected per taxon (such at the plastome- and transcriptome focused MonAToL study, see Givnish et al. 2010), may further improve clade support values and satisfactorily resolve the remaining weaknesses in our understanding of higher-order monocot relationships. Further technical advances may also improve our ability to infer divergence times for monocots in the face of substantial rate heterogeneity. The field is still in its infancy (e.g., Lartillot and Delsuc, 2012; Ronquist et al., 2012a; Tamura et al., 2012; Wilkinson et al., 2011). Nonetheless, my current age estimates provide a future benchmark for further refining our understanding of the ages of the major clades of monocot phylogeny.

Chapter 6

Conclusion

The goals of my thesis were to characterize the phylogeny and evolution of the, until recently, obscure angiosperm family Hydatellaceae, and to improve our understanding of phylogenetic relationships and divergence times in the better studied monocots, especially those of the order Alismatales. In this chapter I review the major findings of my thesis, comment on possible limitations of the approaches used here, and end by proposing future avenues of research.

6.1 Overview

In Chapter 2, I evaluated the phylogeny and evolution of Hydatellaceae. I used four plastid markers (*atp*B, *mat*K, *ndh*F, and *rbc*L) and the nuclear ITS region (primarily ITS1 and ITS2) from 41 Hydatellaceae samples, covering all 12 species (Sokoloff et al., 2008a), to address phylogenetic relationships. The two resulting gene trees were congruent across most branches. I used a multi-species coalescent approach (Heled and Drummond, 2010) to construct a species-level phylogeny from these two gene trees. The resulting phylogeny was for the most part well supported and I used it to examine patterns of morphological evolution, focusing on characters that have traditionally been considered to be important for classification (pericarp ribs/dehiscence, sexual system) and characters more recently highlighted as being phylogenetically informative (other fruit and seed characters; Sokoloff et al. 2008a). I found that most characters of the fruit and seed are phylogenetically informative, while pericarp ribs/dehiscence and sexual system are highly homoplasious. I also found that perenniality is a derived character state for one small clade in the family. The ancestral-state recon-

structions for sexual system showed multiple shifts between cosexuality and dioecy, but in most cases the direction was unclear. Sokoloff et al. (2008a) posited that cosexual and dioecious species had quantitative differences in their reproductive structures. I therefore examined this using the phylogenetic ANOVA method of Garland et al. (1993), and found statistically significant correlations between sexual system, and staminal and involucral phyllome characters.

In Chapter 3, I explored the age of Hydatellaceae, its biogeography and niche evolution. Hydatellaceae lack a reliable fossil record, and as a result, fossil calibration must come from outside the family. I assembled a suite of seven seed-plant fossils and added sequences of one additional species of Hydatellaceae (*Trithuria cowieana*) to an existing seed-plant matrix of 17 plastid genes (Graham and Iles, 2009; Saarela et al., 2007) in order to date the deepest splits within Hydatellaceae. I used a Bayesian molecular-dating approach that permits local molecular clocks for clades, assigning different clocks per clade according to an internal model-selection procedure (Drummond and Suchard, 2010). The posterior distributions for the two deepest splits in the family were then used as priors on the age of the two deepest splits in the multi-species coalescent phylogeny from Chapter 2. This two-step process allowed for the estimation of intra-familial divergence times. The geographic distribution of the family in Australia, India, and New Zealand suggested a Gondwanan origin for the crown clade (Gaikwad and Yadav, 2003; Yadav and Janarthanam, 1994, 1995). However, the breakup of East Gondwana ($\sim$132 Ma, when Australia-Antarctica rifted from the India-Madagascar; Ali and Aitchison 2008) is much older than the crown age of Hydatellaceae ($\sim$17.5 Ma) and is several orders of magnitude older than the age of the split estimated between *Trithuria konkanensis* in India and its closest relative, *T. lanterna* in northern Australia ($\sim$0.76 Ma). Considerations of the timing of this and other biogeographic splits suggests that, except for a fundamental vicariance event between north Australian and south Australian clades (defining the root split of the family), the biogeographic patterns in Hydatellaceae are the result of multiple long distance dispersal events (for further examples see Les et al. 2003). Finally, I char-

acterized abiotic niche evolution in the family. I estimated abiotic climate niches using the method of Evans et al. (2009) to extract, for each species, the preference distribution for climate variables from the maximum entropy probability surface (Phillips et al., 2006). The maximum entropy model was derived from herbarium records for individual species and 19 bioclim variables constructed from the interpolated global climate surface (Hijmans et al., 2005). Niche evolution of these climate parameters was evaluated across the phylogeny by considering progressively more complex models (Brownian motion vs "fitness optima" models with one to three optima; Butler and King 2004). I found that model-selection procedures generally preferred models with one or two optima, suggesting a high degree of niche conservatism, the first time that this has been demonstrated for aquatic plants, which are often considered to be highly homoplasious (Santamaría, 2002).

In Chapter 4, I evaluated the phylogenetic relationships of Alismatales. Higher-order evolutionary relationships of most major monocot groups are quite well understood (Chase et al., 2006). However, multiple inter-family relationships among major lineages in Alismatales, especially in the "core" alismatid clade, remain poorly understood, with little progress made since the study by Les et al. (1997). I addressed these short-comings by sampling 17 plastid genes for at least one taxon in each of the recognized families of Alismatales. The resulting phylogeny provided a generally well supported framework for understanding inter-familial relationships in Alismatales. I found evidence that the Australian endemic *Maundia triglochinoides* deserves family status, as its inferred placement renders Juncaginaceae paraphyletic. This mirrors earlier but less well supported results reported by von Mering and Kadereit (2010). I also characterized the convergent loss of genes involved in NADH dehydrogenase function (*ndh* genes) in at least three alismatid lineages. Finally, I demonstrated the effect of long-branch attraction in potentially misleading phylogenetic inference within Alismatales for low taxon densities.

In Chapter 5, I performed a large-scale analysis of monocot higher-order relationships and their ages. Specifically, I fleshed out sampling in the mono-

cot orders Dioscoreales, Liliales, and Pandanales, supplementing existing (published and unpublished) sequence data of a 17-gene plastid matrix (Graham et al., 2006; Saarela, 2006; Saarela and Graham, 2010; Saarela et al., 2008, 2007; Zgurski, 2004). The final sampling included 172 monocot exemplar taxa representing 71 of 78 monocot families. The resulting phylogeny generally confirmed the phylogenetic backbone of monocots seen in other recent studies (e.g., Chase et al., 2006; Givnish et al., 2010; Graham et al., 2006), and many family-level relationships (e.g., Chase et al., 2006; Givnish et al., 2006), but generally with comparable or improved branch support. To evaluate divergence times in monocots I used the Bayesian random local clocks model (Drummond and Suchard, 2010) with 17 monocot fossil constraints (and eight additional angiosperm fossil constraints), providing a detailed justification for their use. I found that the majority of monocot family-level crown clades originated by the end of the Cretaceous (66 Ma). In general my inferred ages were younger than those found by Janssen and Bremer (2004) and among recent studies are most similar to those of Givnish et al. (2011).

6.2 Limitations of the methodology

Chapter 2 provides a solid platform for understanding the phylogeny and character evolution of Hydatellaceae but the study has several possible limitations. First, the method that I use to reconcile gene trees (*BEAST; Heled and Drummond 2010) assumes that species boundaries are correctly known, and is unable to accommodate introgression or allopolyploid speciation. *BEAST assumes that all gene-tree incongruence is due to incomplete lineage sorting, which it reconciles using the multi-species coalescent method. The second caveat concerns the low number of loci covered per species (at most two), and for some species, the low number of populations sampled. The multi-species coalescent is most accurate with a dense sampling of unlinked loci and multiple samples per species (Heled and Drummond, 2010).

In Chapter 3 I evaluated the age, biogeography and niche evolution of

Hydatellaceae. These fields are evolving rapidly and there are many debates about what constitutes appropriate methodology. For example, methods for estimating lineage ages from molecular phylogenies are rapidly expanding (e.g., Lartillot and Delsuc, 2012; Ronquist et al., 2012a; Tamura et al., 2012; Wilkinson et al., 2011). These methods have frequently focused on improving individual aspects of inference. For example, the method of Tamura et al. (2012) was aimed at improving computational speed for the dating of large phylogenies. In contrast Ronquist et al. (2012a) implemented a Bayesian method to place fossils within the phylogeny as part of an integrated process of molecular dating. Explaining the causes of rate variation is one of the major incompletely answered questions of molecular evolution (Gaut et al., 2011). As a result, most methods aim for statistical tractability as opposed to biological reality, and so may sometimes end up fitting real data poorly. To accommodate this, Lartillot and Delsuc (2012) suggest co-estimating divergence times and character evolution, since if character change is associated with change in molecular rate, such as lifespan or body size (e.g., Smith and Donoghue, 2008), this method should improve the estimation of both (Lartillot and Delsuc, 2012). However, I used the random local clocks model (Drummond and Suchard 2010), which allows subclades to have unique molecular clocks. With this method, new local clocks are proposed and accepted or rejected according to an internal model selection procedure. If molecular rates change due to correlated changes in organismal biology (such as life span, Smith and Donoghue 2008) this model should accommodate that variation. However, for a given local clock the molecular rates are (by definition) assumed to be constant, a condition that may not always be met.

Understanding the geographic history of biological lineages has been a focus of biological research since the time of Linnaeus (Briggs and Humphries, 2004). The first quantitative methods in historical biogeography were parsimony based, with model-based methods only appearing in the last five years or so (Ronquist and Sanmartín, 2011). I evaluated the biogeography of Hydatellaceae using a parsimony-based dispersal-vicariance method (Ronquist, 1997), an ML based dispersal-extinction-cladogenesis model (Ree

et al., 2005; Ree and Smith, 2008), and an ML ancestral state reconstruction model (Sanmartín et al., 2010, 2008). These three methods recovered similar historical biogeographic patterns. Possible problems with evaluating the model-based methods are the large number of parameters to be estimated compared to the low number of species considered. An additional issue is their tendency to estimate zero extinction rates, which has been demonstrated by analysing data simulated under high extinction rates (Ree and Smith, 2008). I addressed this issue by considering a variety of extinction rates in my analyses.

There is ongoing controversy about what constitutes phylogenetic niche conservatism and what its importance may be (Losos, 2008a,b; Wiens, 2008; Wiens and Graham, 2005). I consider niche conservatism to be a similarity in ecological variables that is greater than expected under a Brownian motion model of character evolution (Losos, 2008a). How to test for phylogenetic niche conservatism remains contentious, although model testing is often performed by comparing a Brownian motion model to models that consider several "fitness optima" for a trait (Butler and King, 2004; Cooper et al., 2010). This is what I do here. Other options are available such as Pagel's λ, Pagel's δ, and Blomberg's K, where the estimated statistic can be tested against the expected value for Brownian motion (i.e., the null hypothesis; reviewed in Cooper et al. 2010). However, these latter methods do not explicitly test where changes occur in the phylogeny. In the case of testing models of "fitness optima" the location of regime changes (where one fitness optima switches to a secondary one) must be specified a priori, and although this can be based on the ancestral-state reconstruction of some other trait, such as biogeography or a morphological character (e.g., Boucher et al. 2012), this adds an additional layer of subjectivity. Finally, another issue is the controversy over how widespread the pattern of phylogenetic niche conservatism is in nature (Losos, 2008a; Wiens and Graham, 2005), and how unusual or noteworthy its detection might be for any given clade. It is also known that evidence of phylogenetic niche conservatism can change depending upon the phylogenetic scale involved (Losos, 2008a); choosing the phylogenetic scale will always be somewhat arbitrary.

In Chapter 4 I addressed phylogenetic relationships within Alismatales and aspects of molecular evolution in a subset of examined genes. Perhaps the biggest scope for error involves identifying the loss of *ndh* genes with loss of NADH dehydrogenase function. Three different scenarios could explain the two *ndh* pseudogenes observed in several independent lineages: (1) I consider it most likely that the loss of *ndh*B and *ndh*F indicates the loss of the entire NADH dehydrogenase complex. Indeed, in whole plastid genome studies the entire complex is either lost (e.g., Blazier et al., 2011; Braukmann et al., 2009; Chang et al., 2006) or retained, suggesting all subunits are required for enzyme function (Martín and Sabater, 2010; Wicke et al., 2011). In addition, preliminary results in Hydrocharitaceae (*Thalassia*—Gregory Ross, University of British Columbia, pers. comm.; *Najas*—Don Les, University of Connecticut, pers. comm.), two alismatid species that I inferred to have lost *ndh*B and *ndh*F, confirm both losses from the plastid genome and suggest that most or all remaining *ndh* genes are pseudogenes. (2) Functional gene migration from the plastid to the nucleus is known to occur, with subsequent targeting of one or more subunits products back to the plastid (e.g., Kleine et al. 2009), and this could explain the presence of pseudogenes recovered in the plastid, even as the functional complex is retained. Considering the three lineages that appear to lack two subunits, *ndh*B and *ndh*F, which are located in different parts of the plastid genome, this scenario would require at least three parallel sets of migration events of the same gene pair from the plastid to nucleus (with subsequent protein re-targeting to the plastid organelle), an overall scenario that may be improbable. (3) Finally, since only two *ndh* genes were sampled it is possible that the NADH dehydrogenase complex has remained functional through the retention of the remaining *ndh* genes (11 *ndh* genes encoded by the plastid in total, Wicke et al. 2011), although, considering the highly conserved nature of this and other plastid protein complexes, this also seems unlikely. Another possible caveat in this study concerns the possible long-branch effects in misleading phylogenetic inference. I demonstrated that this occurs when considering a reduced taxon set in Alismatales. While the dense sampling I used should minimize the problem (Hillis et al., 2003), long-branch effects

may still have influenced the phylogeny of the Alismatales in this larger data set, and visual inspection of the phylogeny suggests fairly substantial rate elevation in the order compared to other monocots.

Chapter 5 addressed broad-scale monocot phylogeny and molecular dating. I used 25 fossil taxa (17 from the monocots) to constrain molecular-dating analyses. Interpretation of some of the fossils remains contentious (Nixon, 2008; Yoo et al., 2005), and errors in their taxonomic assignments would result in errors in my calibration and dating. Another possible source of error is the prior age distribution chosen to serve as the calibration. I erred on the side of being conservative in choosing minimum ages, generally by using the age at the top of a stratigraphic unit (the youngest possible in the stratum) for a fossil record, or using the suggested ages of Clarke et al. (2011) when the stratigraphy of a region was poorly calibrated (see Chapter 5). The choice of molecular dating method can also influence results (see discussion for Chapter 3). One issue of concern for Chapter 5 (but not for Chapter 3) is the poor convergence of the random local clocks model (Drummond and Suchard, 2010) in the monocot data set. The poor convergence manifested itself here as low estimated sample sizes for the model statistics. This indicates that some aspects of model inference were relatively poorly estimated, such as some clade ages and perhaps tree topology (although it should be noted that the maximum clade credibility tree was very similar to the 'PF16' ML tree). This could be ameliorated using longer runs, although the time needed could be considerable for a data set of this size as the current run time was six weeks.

6.3 Future directions

While the species tree of Hydatellaceae is relatively well supported, it is based on only two linkage groups. A more accurate species tree may be achieved by using additional nuclear loci (Heled and Drummond, 2010), which may also resolve poorly supported regions of the current phylogeny, such as inter-species relationships between *Trithuria bibracteata*, *T. occidentalis*, and *T. submersa*. It may also be beneficial to sample additional

populations per species. In some cases this would be relatively straightforward to accomplish, as (for example) *T. submersa* is widely distributed in Western Australia, Victoria, and Tasmania. Even the recently described *T. austinensis* and *T. australis* are known from at least a half-dozen localities. It would also be useful to use a population structure approach to better understand species relationships and/or species boundaries. This could potentially be done using microsatellites, or SNPs derived from RADSeq, for example (Davey et al., 2010; Falush et al., 2003, 2007; Pritchard et al., 2000).

Our ecological understanding of Hydatellaceae is still relatively poor (reviewed in Sokoloff et al. 2011). Strides are just starting to be made in understanding the family's reproductive ecology (Taylor et al., 2010) and seed germination behaviour (Tuckett et al., 2010a,b). However, most other aspects of its ecology have not been studied quantitatively. In Chapter 2 I used comparative approaches (Harvey and Pagel, 1991; Schluter et al., 1997) to understand macro-evolutionary aspects of sexual system evolution in Hydatellaceae. Further insights will come from field and greenhouse studies of the reproductive ecology within populations or among closely related species such as *T. bracteata*, *T. occidentalis*, and *T. submersa*. In Hydatellaceae, as in *Mercurialis annua*, genetic and chromosomal studies would be valuable (Obbard et al., 2006; Pannell et al., 2004). In Chapter 3 I constructed species distribution models using Maxent (Phillips et al., 2006; Phillips and Dudík, 2008; Phillips et al., 2004) to evaluate aspects of climate niche evolution. However, these are preliminary evaluations, and more finely-tuned and species-specific analyses could be done to provide a more accurate and nuanced understanding of the effects of climate change and relative conservation risk (Elith and Leathwick, 2009).

The work presented in Chapter 5, based on ∼10% of the plastid genome, provides an overview of which parts of the monocot phylogeny are well understood and supported, and which remain poorly understood. Improving support for recalcitrant branches of monocot phylogeny will likely require additional sequence data. While nuclear data undoubtedly represent a valuable and independent source of sequence variation, repeated bouts of polyploidy throughout angiosperm evolution (Barker et al., 2008; Wood et al., 2009)

have made identification of orthologs for phylogenetic analysis a non-trivial task (Duarte et al., 2010). An alternative approach to the identification and use of single-copy nuclear genes would be to use multi-gene family data to construct a species phylogeny based on gene-tree parsimony (Zmasek and Eddy, 2001), although this has only occasionally been done at deep phylogenetic levels in plants (Burleigh et al., 2010; Ness et al., 2011; Sanderson and McMahon, 2007). Whole plastid genome based phylogenies are also rapidly becoming a new standard for evaluating deep phylogenetic relationships in plants (Barrett et al., 2013; Givnish et al., 2010; Jansen et al., 2007; Moore et al., 2007, 2010). In monocots, sequencing of plastid genomes is currently being funded by a NSF grant for MonAToL projects (Barrett et al., 2013; Givnish et al., 2010), with the aim of sampling plastid genomes of nearly all monocot families. MonAToL projects also aim to improve sampling of nuclear genes from transcriptomes from across monocot phylogeny. My study shows the improvement that increased taxon sampling and increased sequence per taxon can bring to resolution of relationships, and it leads me to hope that in the near future, rich new sources of data will fill many of the remaining gaps in our understanding of monocot phylogeny.

6.4 Some conclusions

An accurate phylogeny and temporal scale of a group are both essential for fully understanding multiple aspects of its evolutionary history, such as its speciation modes and rates, extinction patterns, morphological and molecular evolution, and biogeography. In my thesis I studied the phylogeny and evolution of two ancient and distantly related groups of aquatic flowering plants. As a result of my work, a molecular phylogeny is now available for Hydatellaceae, which serves as the framework for testing systematic and evolutionary hypotheses. I hope that my work on Hydatellaceae will provide a platform for further studies of this enigmatic family, and that the analyses on some aspects of its biology, including sexual-system evolution, will promote additional studies on these phenomena.The second group, the monocots, has been the subject of extensive molecular phylogenies for over two

decades, and yet there were still branches that have remained poorly supported phylogenetically, such as multiple ones within the order Alismatales. Other aspects of monocot evolution, such as estimating divergence times, have been relatively meagerly explored (although some groups have been more extensively studied, such as the grasses). My work on Alismatales and (more generally) monocot phylogeny and divergence times should also add to and help further clarify the more substantial body of knowledge available for these major flowering-plant lineages.

Bibliography

Ackerly, D. 2009. Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. *Proceedings of the National Academy of Sciences, USA* 106 Suppl:19699–19706.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *In* B. N. Petrov and F. Csaki (eds.), Proceedings of the Second International Symposium on Information Theory, pp. 267–281, Budapest. Akademiai Kiado.

Alfaro, M. E. and Holder, M. T. 2006. The posterior and the prior in Bayesian phylogenetics. *Annual Review of Ecology, Evolution, and Systematics* 37:19–42.

Ali, J. R. and Aitchison, J. C. 2008. Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian subcontinent from the Middle Jurassic through latest Eocene (166–35Ma). *Earth-Science Reviews* 88:145–166.

Ambrose, B. A., Espinosa-Matías, S., Vázquez-Santana, S., Vergara-Silva, F., Martínez, E., Márquez-Guzmán, J., and Alvarez-Buylla, E. R. 2006. Comparative developmental series of the Mexican triurids support a euanthial interpretation for the unusual reproductive axes of *Lacandonia schismatica* (Triuridaceae). *American Journal of Botany* 93:15–35.

Angiosperm Phylogeny Group 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85:531–553.

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.

Arber, A. 1920. Water Plants: A Study of Aquatic Angiosperms. Cambridge University Press, Cambridge.

Azuma, H. and Tobe, H. 2011. Molecular phylogenetic analyses of Tofieldiaceae (Alismatales): family circumscription and intergeneric relationships. *Journal of Plant Research* 124:349–357.

Backlund, A. and Bremer, K. r. 1998. To be or not to be—Principles of classification and monotypic plant families. *Taxon* 47:391–400.

Baker, H. G. 1955. Self-compatibility and establishment after 'long-distance' dispersal. *Evolution* 9:347–349.

Barker, M. S., Kane, N. C., Matvienko, M., Kozik, A., Michelmore, R. W., Knapp, S. J., and Rieseberg, L. H. 2008. Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Molecular Biology and Evolution* 25:2445–2455.

Barrett, C. F., Davis, J. I., Leebens-Mack, J., Conran, J. G., and Stevenson, D. W. 2013. Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics* 29:65–87.

Barrett, S. C. H. 2002. The evolution of plant sexual diversity. *Nature Reviews Genetics* 3:274–284.

Barrett, S. C. H. 2010. Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365:99–109.

Barrett, S. C. H. and Graham, S. W. 1997. Adaptive radiation in the aquatic plant family Pontederiaceae: insights from phylogenetic analysis, pp. 225–258. *In* T. J. Givnish and K. J. Sytsma (eds.), Molecular Evolution and Adaptive Radiation. Cambridge University Press, Cambridge.

Beaulieu, J. M., Leitch, I. J., Patel, S., Pendharkar, A., and Knight, C. a. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* 179:975–986.

Behnke, H.-D. 2000. Forms and sizes of sieve-element plastids and evolution of the monocotyledons, pp. 163–188. *In* K. L. Wilson and D. A. Morrison (eds.), Monocots: Systematics and Evolution. CSIRO, Collingwood, Australia.

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

Benzecry, A. and Brack-Hanes, S. D. 2008. A new hydrocharitacean seagrass from the Eocene of Florida. *Botanical Journal of the Linnean Society* 157:19–30.

Blazier, J. C., Guisinger, M. M., and Jansen, R. K. 2011. Recent loss of plastid-encoded *ndh* genes within *Erodium* (Geraniaceae). *Plant Molecular Bology* 76:263–272.

Borsch, T., Hilu, K. W., Wiersema, J. H., Lohne, C., Barthlott, W., and Wilde, V. 2007. Phylogeny of *Nymphaea* (Nymphaeaceae): evidence from substitutions and microstructural changes in the chloroplast *trn*T–*trn*F region. *International Journal of Plant Sciences* 168:639–671.

Borsch, T., Löhne, C., Mbaye, M. S., and Wiersema, J. 2011. Towards a complete species tree of *Nymphaea*: shedding further light on subg. *Brachyceras* and its relationships to the Australian water-lilies. *Telopea* 13:193–217.

Borsch, T., Löhne, C., and Wiersema, J. 2008. Phylogeny and evolutionary patterns in Nymphaeales: integrating genes, genomes and morphology. *Taxon* 57:1052–1081.

Bouchenak-Khelladi, Y., Verboom, G. A., Savolainen, V., and Hodkinson, T. R. 2010. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical. *Botanical Journal of the Linnean Society* 162:543–557.

Boucher, F. C., Thuiller, W., Roquet, C., Douzet, R., Aubert, S., and Alvarez, N. 2012. Reconstructing the origins of high-alpine niches and cushion life form in the genus *Androsace* s.l. (Primulaceae). *Evolution* 66:1255–1268.

Braukmann, T. W. A., Kuzmina, M., and Stefanović, S. 2009. Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. *Current Genetics* 55:323–337.

Bremer, K. r. 2000. Early Cretaceous lineages of monocot flowering plants. *Proceedings of the National Academy of Sciences, USA* 97:4707–4711.

Bremer, K. r. 2002. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* 56:1374–1387.

Briggs, J. C. and Humphries, C. J. 2004. Early classics, pp. 5–13. *In* M. V. Lomolino, D. F. Sax, and J. H. Brown (eds.), Foundations of Biogeography: Classic Papers with Commentaries. University of Chicago Press, Chicago.

Bull, J. J. and Charnov, E. L. 1985. On irreversible evolution. *Evolution* 39:1149–1155.

Burleigh, J. G., Bansal, M. S., Eulenstein, O., Hartmann, S., Wehe, A., and Vision, T. J. 2010. Genome-scale phylogenetics: inferring the plant tree of life from 18,896 gene trees. *Systematic Biology* 60:117–125.

Butler, M. A. and King, A. A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *American Naturalist* 164:683–695.

Cabrera, L. I., Salazar, G. A., Chase, M. W., Mayo, S. J., Bogner, J., and Dávila, P. 2008. Phylogenetic relationships of aroids and duckweeds (Araceae) inferred from coding and noncoding plastid DNA. *American Journal of Botany* 95:1153–1165.

Caddick, L. R., Rudall, P. J., Wilkin, P., Hedderson, T. A. J., and Chase, M. W. 2002. Phylogenetics of Dioscoreales based on combined analyses of morphological and molecular data. *Botanical Journal of the Linnean Society* 138:123–144.

Caddick, L. R. and Wilkin, P. 1998. A revision of the genus *Stenomeris* (Dioscoreaceae). *Kew Bulletin* 53:703–712.

Cameron, K. M. 2004. Utility of plastid *psa*B gene sequences for investigating intrafamilial relationships within Orchidaceae. *Molecular Phylogenetics and Evolution* 31:1157–1180.

Cameron, K. M., Chase, M. W., and Rudall, P. J. 2003. Recircumscription of the monocotyledonous family Petrosaviaceae to include *Japonolirion*. *Brittonia* 55:214–225.

Cantino, P. D., Doyle, J. A., Graham, S. W., Judd, W. S., Olmstead, R. G., Soltis, D. E., Soltis, P. S., and Donoghue, M. J. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* 56:822–846.

Carlquist, S. 1974. Island Biology. Columbia University Press, New York.

Carlquist, S. and Schneider, E. L. 2009. Do tracheid microstructure and the presence of minute crystals link Nymphaeaceae, Cabombaceae and Hydatellaceae? *Botanical Journal of the Linnean Society* 159:572–582.

Case, A. L., Graham, S. W., Macfarlane, T. D., and Barrett, S. C. H. 2008. A phylogenetic study of evolutionary transitions in sexual systems in Australasian *Wurmbea* (Colchicaceae). *International Journal of Plant Sciences* 169:141–156.

Chang, C.-C., Lin, H.-C., Lin, I.-P., Chow, T.-Y., Chen, H.-H., Chen, W.-H., Cheng, C.-H., Lin, C.-Y., Liu, S.-M., Chang, C.-C., and Chaw, S.-M. 2006. The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Molecular Biology and Evolution* 23:279–291.

Charlesworth, D. 1999. Theories of the evolution of dioecy, pp. 33–60. *In* M. A. Geber, T. E. Dawson, and L. F. Delph (eds.), Gender and Sexual Dimorphism in Flowering Plants. Springer, Berlin.

Charlesworth, D. 2002. Plant sex determination and sex chromosomes. *Heredity* 88:94–101.

Charnov, E. L. 1982. The Theory of Sex Allocation. Princeton University Press, Princeton.

Chase, M. W. 2004. Monocot relationships: an overview. *American Journal of Botany* 91:1645–1655.

Chase, M. W., Duvall, M. R., Hills, H. G., Conran, J. G., Cox, A. V., Eguiarte, L. E., Hartwell, J., Fay, M. F., Caddick, L. R., Cameron, K. M., and Hoot, S. B. 1995. Molecular phylogenetics of Lilianae, pp. 109–137. *In* P. J. Rudall, P. J. Cribb, D. F. Cutler, and C. J. Humphries (eds.), Monocotyledons: Systematics and Evolution. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.

Chase, M. W., Fay, M. F., Devey, D. S., Maurin, O., Rø nsted, N., Davies, T. J., Pillon, Y., Petersen, G., Seberg, O., Tamura, M. N., Asmussen, C. B., Hilu, K., Borsch, T., Davis, J. I., Stevenson, D. W., Pires, J. C., Givnish, T. J., Sytsma, K. J., McPherson, M., Graham, S. W., and Rai, H. S. 2006. Multigene analyses of monocot relationships: a summary, pp. 63–75. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales), number Apg 1998. Rancho Santa Ana Botanic Garden, Claremont, California.

Chase, M. W., Soltis, D. E., Soltis, P. S., Rudall, P. J., Fay, M. F., Hahn, W. J., Sullivan, S., Joseph, J., Molvray, M., Kores, P. J., Givnish, T. J., Sytsma, K. J., and Pires, J. C. 2000. Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification, pp. 3–16. *In* K. L. Wilson and D. A. Morrison (eds.), Monocots: Systematics and Evolution. CSIRO, Collingwood, Australia.

Chen, J.-M., Chen, D., Gituru, W. R., Wang, Q.-F., and Guo, Y.-H. 2004a. Evolution of apocarpy in Alismatidae using phylogenetic evidence from chloroplast *rbc*L gene sequence data. *Botanical Bulletin of Academia Sinica* 45:33–40.

Chen, J.-M., Robert, G. W., and Wang, Q.-F. 2004b. Evolution of aquatic life-forms in Alismatidae: Phylogenetic estimation from chloroplast *rbc*L gene sequence data. *Israel Journal of Plant Sciences* 52:323–329.

Clarke, J. T., Warnock, R. C. M., and Donoghue, P. C. J. 2011. Establishing a time-scale for plant evolution. *New Phytologist* 192:266–301.

Collinson, M. E. 1983. Palaeofloristic assemblages and palaeoecology of the Lower Oligocene Bembridge marls, Hamstead Ledge, Isle of Wight. *Botanical Journal of the Linnean Society* 86:177–225.

Conran, J. 1997. *Paracordyline kerguelensis*, an Oligocene monocotyledon macrofossil from the Kerguelen Islands. *Alcheringa* 21:129–140.

Conran, J. and Christophel, D. 1998. *Paracordyline aureonemoralis* (Lomandraceae): an Eocene monocotyledon from South Australia. *Alcheringa* 22:349–357.

Conran, J. G., Bannister, J. M., and Lee, D. E. 2009a. Earliest orchid macrofossils: early Miocene *Dendrobium* and *Earina* (Orchidaceae: Epidendroideae) from New Zealand. *American Journal of Botany* 96:466–474.

Conran, J. G., Carpenter, R. J., and Jordan, G. J. 2009b. Early Eocene *Ripogonum* (Liliales: Ripogonaceae) leaf macrofossils from southern Australia. *Australian Systematic Botany* 22:219–228.

Conran, J. G., Christophel, D. C., and Cunningham, L. 2003. An Eocene moncotyledon from Nelly Creek, Central Australia, with affinities to Hemerocallidaceae (Lilianae: Asparagales). *Alcheringa* 27:107–115.

Cook, C. D. K. 1990. Aquatic Plant Book. SPB Academic Publishing, Amsterdam.

Cooke, D. A. 1987. Hydatellaceae, pp. 1–5. *In* A. S. George (ed.), Flora of Australia Volume 45, Hydatellaceae to Liliaceae. Australian Government Publishing Service, Canberra.

Cooper, N., Jetz, W., and Freckleton, R. P. 2010. Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology* 23:2529–2539.

Council of Heads of Australasian Herbaria 2012. Australia's Virtual Herbarium. `http://avh.chah.org.au`.

Crane, P. R., Friis, E. M., and Pedersen, K. R. 1986. Lower Cretaceous angiosperm flowers: fossil evidence on early radiations of dicotyledons. *Science* 232:852–854.

Crane, P. R., Pedersen, K. R., Friis, E. M., and Drinnan, A. N. 1993. Early Cretaceous (early to middle Albian) platanoid inflorescences associated with *Sapindopsis* leaves from the Potomac Group of Eastern North America. *Systematic Botany* 18:328–344.

Crepet, W. L. and Feldman, G. D. 1991. The earliest remains of grasses in the fossil record. *American Journal of Botany* 78:1010–1014.

Crepet, W. L., Nixon, K. C., and Gandolfo, M. A. 2004. Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *American Journal of Botany* 91:1666–1682.

Crisp, M., Cook, L., and Steane, D. 2004. Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359:1551–1571.

Crisp, M. D. and Cook, L. G. 2007. A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution* 43:1106–1117.

Cronquist, A. 1968. The Evolution and Classification of Flowering Plants. Houghton Mifflin, Boston.

Cronquist, A. 1988. The Evolution and Classification of Flowering Plants. New York Botanical Garden, New York, 2nd edition.

Cuenca, A., Petersen, G., Seberg, O., Davis, J. I., and Stevenson, D. W. 2010. Are substitution rates and RNA editing correlated? *BMC Evolutionary Biology* 10:349.

Cusimano, N., Bogner, J., Mayo, S. J., Boyce, P. C., Wong, S. Y., Hesse, M., Hetterscheid, W. L. A., Keating, R. C., and French, J. C. 2011. Relationships within the Araceae: comparison of morphological patterns with molecular phylogenies. *American Journal of Botany* 98:654–668.

Dahlgren, R. M. T. and Clifford, H. T. 1982. The Monocotyledons: A Comparative Study. Academic Press, London.

Dahlgren, R. M. T., Clifford, H. T., and Yeo, P. F. 1985a. Order Hydatellales, pp. 398–400. *In* The Families of the Monocotyledons - Structure, Evolution, and Taxonomy. Springer-Verlag, Berlin.

Dahlgren, R. M. T., Clifford, H. T., and Yeo, P. F. 1985b. The Families of the Monocotyledons: Structure, Evolution, and Taxonomy. Springer-Verlag, Berlin.

Dahlgren, R. M. T. and Rasmussen, F. N. 1983. Monocotyledon evolution: characters and phylogenetic estimation. *Evolutionary Biology* 16:255–395.

Davey, J. W., Davey, J. L., Blaxter, M. L., and Blaxter, M. W. 2010. RADSeq: next-generation population genetics. *Briefings in Functional Genomics* 9:416–423.

Davis, J. I., Petersen, G., Seberg, O., Stevenson, D. W., Hardy, C. R., Simmons, M. P., Michelangeli, F. A., Goldman, D. H., Campbell, L. M., Specht, C. D., and Cohen, J. I. 2006. Are mitochondrial genes useful for the analysis of monocot relationships? *Taxon* 55:857–870.

Davis, J. I., Simmons, M. P., Stevenson, D. W., and Wendel, J. F. 1998. Data decisiveness, data quality, and incongruence in phylogenetic analysis: an example from the monocotyledons using mitochondrial *atp*A sequences. *Systematic Biology* 47:282–310.

Davis, J. I., Stevenson, D. W., Petersen, G., Seberg, O., Campbell, L. M., Freudenstein, J. V., Goldman, D. H., Hardy, C. R., Michelangeli, F. A., Simmons, M. P., Specht, C. D., Vergara-Silva, F., and Gandolfo, M. 2004. A phylogeny of the monocots, as inferred from *rbc*L and *atp*A sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. *Systematic Botany* 29:467–510.

de Lange, P. J., Murray, B. G., and Datson, P. M. 2004. Contributions to a chromosome atlas of the New Zealand flora 38. Counts for 50 families. *New Zealand Journal of Botany* 42:873–904.

Deil, U. 2005. A review on habitats, plant traits and vegetation of ephemeral wetlands - a global perspective. *Phytocoenologia* 35:533–705.

Delph, L. F. 2009. Sex allocation: evolution to and from dioecy. *Current Biology* 19:R249–R251.

DePamphilis, C. W. and Palmer, J. D. 1990. Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. *Nature* 348:337–339.

Diels, L. 1906. Die Pflanzenwelt von West-Australien südlich des Wendekreises mit einer Einleitung über die Pflanzenwetl Gesamt-Australien in Grundzügen. *Vegetation der Erde* 7:1–413.

Dilcher, D. L. and Crane, P. R. 1984. *Archaeanthus*: an early angiosperm from the Cenomanian of the western interior of North America. *Annals of the Missouri Botanical Garden* 71:351–383.

Dinis, J. L., Rey, J., Cunha, P. P., Callapez, P., and Pena dos Reis, R. 2008. Stratigraphy and allogenic controls of the western Portugal Cretaceous: an updated synthesis. *Cretaceous Research* 29:772–780.

Dorken, M. E., Friedman, J., and Barrett, S. C. H. 2002. The evolution and maintenance of monoecy and dioecy in *Sagittaria latifolia* (Alismataceae). *Evolution* 56:31–41.

Douzery, E. J. P., Snell, E. a., Bapteste, E., Delsuc, F., and Philippe, H. 2004. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proceedings of the National Academy of Sciences, USA* 101:15386–15391.

Doyle, J. A. 2005. Early evolution of angiosperm pollen as inferred from molecular and morphological phylogenetic analyses. *Grana* 44:227–251.

Doyle, J. A. 2008. Integrating molecular phylogenetic and paleobotanical evidence on origin of the flower. *International Journal of Plant Sciences* 169:816–843.

Doyle, J. A. 2009. Evolutionary significance of granular exine structure in the light of phylogenetic analyses. *Review of Palaeobotany and Palynology* 156:198–210.

Doyle, J. A. 2012. Molecular and fossil evidence on the origin of angiosperms. *Annual Review of Earth and Planetary Sciences* 40:301–326.

Doyle, J. A. and Endress, P. K. 2010. Integrating early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. *Journal of Systematics and Evolution* 48:1–35.

Doyle, J. A., Endress, P. K., and Upchurch, G. R. 2008. Early Cretaceous monocots: a phylogenetic evaluation. *Acta Musei Nationalis Pragae* 64:59–87.

Doyle, J. A., Hotton, C. L., and Ward, J. V. 1990. Early Cretaceous tetrads, zonasulculate pollen, and Winteraceae. I. Taxonomy, morphology, and ultrastructure. *American Journal of Botany* 77:1544–1557.

Doyle, J. J. and Doyle, J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11–15.

Dransfield, J., Uhl, N. W., Asmussen, C. B., Baker, W. J., Harley, M. M., and Lewis, C. E. 2008. Genera Palmarum: The Evolution and Classification of Palms. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.

Drinnan, A. N., Crane, P. R., Friis, E. M., and Pedersen, K. R. 1991. Angiosperm flowers and tricolpate pollen of buxaceous affinity from the Potomac Group (mid-Cretaceous) of Eastern North America. *American Journal of Botany* 78:153–176.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., and Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4:e88.

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

Drummond, A. J. and Suchard, M. A. 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biology* 8:114.

Duarte, J. M., Wall, P. K., Edger, P. P., Landherr, L. L., Ma, H., Pires, J. C., Leebens-Mack, J., and DePamphilis, C. W. 2010. Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC Evolutionary Biology* 10:61.

Duvall, M. R., Clegg, M. T., Chase, M. W., Clark, W. D., Kress, W. J., Hills, H. G., Eguiarte, L. E., Smith, J. F., Gaut, B. S., Zimmer, E. A., and Learn, G. H. J. 1993. Phylogenetic hypotheses for the monocotyledons constructed from *rbc*L sequence data. *Annals of the Missouri Botanical Garden* 80:607–619.

Duvall, M. R., Robinson, J. W., Mattson, J. G., and Moore, A. 2008. Phylogenetic analyses of two mitochondrial metabolic genes sampled in parallel from angiosperms find fundamental interlocus incongruence. *American Journal of Botany* 95:871–884.

Eckhart, V. M. 1999. Sexual specialization in flowers and inflorescences, pp. 123–148. *In* M. A. Geber, T. E. Dawson, and L. F. Delph (eds.), Gender and Sexual Dimorphism in Flowering Plants. Springer, Berlin.

Edgar, E. 1966. The male flowers of *Hydatella inconspicua* (Cheesem.) Cheesem. (Centrolepidaceae). *New Zealand Journal of Botany* 4:153–158.

Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40:677–697.

Endress, P. K. 2004. Structure and relationships of basal relictual angiosperms. *Australian Systematic Botany* 17:343–366.

Endress, P. K. 2008. Perianth biology in the basal grade of extant angiosperms. *International Journal of Plant Sciences* 169:844–862.

Endress, P. K. 2010. The evolution of floral biology in basal angiosperms. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 365:411–21.

Endress, P. K. and Doyle, J. A. 2009. Reconstructing the ancestral angiosperm flower and its initial specializations. *American Journal of Botany* 96:22–66.

Engler, A. 1892. Syllabus der Vorlesungen über spezielle und medizinisch-pharmoceutische Botanik. Bornträger, Berlin.

Erwin, D. M. and Stockey, R. A. 1989. Permineralized monocotyledons from the Middle Eocene Princeton chert (Allenby Formation) of British Columbia: Alismataceae. *Botany* 67:2636–2645.

ESRI 2009. ArcGIS Desktop.

Evans, M. E. K., Smith, S. A., Flynn, R. S., and Donoghue, M. J. 2009. Climate, niche evolution, and diversification of the "bird-cage" evening primroses (*Oenothera*, sections *Anogra* and *Kleinia*). *American Naturalist* 173:225–240.

Falush, D., Stephens, M., and Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.

Falush, D., Stephens, M., and Pritchard, J. K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular ecology notes* 7:574–578.

Fay, M. F., Chase, M. W., Rø nsted, N., Devey, D. S., Pillon, Y., Pires, J. C., Petersen, G., Seberg, O., and Davis, J. I. 2006. Phylogenetics of Liliales: summarized evidence from the combined analyses of five plastid and one mitochondrial loci, pp. 559–565. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California.

Fay, M. F., Rudall, P. J., Sullivan, S., Stobart, K. L., de Bruijn, A. Y., Reeves, G., Qamaruz-Zaman, F., Hong, W.-P., Joseph, J., Hahn, W. J., Conran, J. G., and Chase, M. W. 2000. Phylogenetic studies of Asparagales based on four plastid DNA regions, pp. 360–371. *In* K. L. Wilson and D. A. Morrison (eds.), Monocots: Systematics and Evolution. CSIRO, Collingwood, Australia.

Feild, T. S., Chatelet, D. S., and Brodribb, T. J. 2009. Ancestral xerophobia: a hypothesis on the whole plant ecophysiology of early angiosperms. *Geobiology* 7:237–64.

Felsenstein, J. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology* 22:240–249.

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

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

Felsenstein, J. 1985b. Phylogenies and the comparative method. *American Naturalist* 125:1–15.

Fischer, T. C., Butzmann, R., Meller, B., Rattei, T., Newman, M., and Hölscher, D. 2009. The morphology, systematic position and inferred biology of *Spirematospermum*—An extinct genus of Zingiberales. *Review of Palaeobotany and Palynology* 157:391–426.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology* 20:406–416.

FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* 3:1084–1092.

Freudenstein, J. V., van der Berg, C., Goldman, D. H., Kores, P. J., Molvray, M., and Chase, M. W. 2004. An expanded plastid DNA phylogeny of Orchidaceae and analysis of jackknife branch support strategy. *American Journal of Botany* 91:149–157.

Friedman, J. and Barrett, S. C. H. 2009. Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Annals of Botany* 103:1515–1527.

Friedman, W. E. 2008. Hydatellaceae are water lilies with gymnospermous tendencies. *Nature* 453:94–97.

Friedman, W. E., Bachelier, J. B., and Hormaza, J. I. 2012. Embryology in *Trithuria submersa* (Hydatellaceae) and relationships between embryo, endosperm, and perisperm in early-diverging flowering plants. *American Journal of Botany* 99:1083–1095.

Friis, E., Pedersen, K., von Balthazar, M., Grimm, G., and Crane, P. 2009. *Monetianthus mirus gen. et sp. nov.*, a nymphaealean flower from the Early Cretaceous of Portugal. *International Journal of Plant Sciences* 170:1086–1101.

Friis, E. M. 1988. *Spirematospermum chandlerae sp. nov.*, an extinct species of Zingiberaceae from the North American Cretaceous. *Tertiary Research* 9:7–12.

Friis, E. M., Crane, P., and Pedersen, K. R. 1997. *Anacostia*, a new basal angiosperm from the Early Cretaceous of North America and Portugal with trichotomocolpate/monocolpate pollen. *Grana* 36:225–244.

Friis, E. M., Crane, P. R., and Pedersen, K. R. 2011. Early Flowers and Angiosperm Evolution. Cambridge University Press, Cambridge.

Friis, E. M., Pedersen, K. R., and Crane, P. R. 2001. Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. *Nature* 410:357–360.

Friis, E. M., Pedersen, K. R., and Crane, P. R. 2004. Araceae from the Early Cretaceous of Portugal: evidence on the emergence of monocotyledons. *Proceedings of the National Academy of Sciences, USA* 101:16565–16570.

Friis, E. M., Pedersen, K. R., and Crane, P. R. 2010. Diversity in obscurity: fossil flowers and the early history of angiosperms. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365:369–382.

Frumin, S. and Friis, E. M. 1999. Magnoliid reproductive organs from the Cenomanian-Turonian of north-western Kazakhstan: Magnoliaceae and Illiciaceae. *Plant Systematics and Evolution* 216:265–288.

Fuse, S., Lee, N. S., and Tamura, M. N. 2012. Biosystematic studies on the family Nartheciaceae (Dioscoreales) I . Phylogenetic relationships, character evolution and taxonomic re-examination. *Plant Systematics and Evolution* 298:1575–1584.

Fuse, S. and Tamura, M. N. 2000. A phylogenetic analysis of the plastid *mat*K gene with emphasis on Melanthiaceae sensu lato. *Plant Biology* 2:415–427.

Gaikwad, S. P. and Yadav, S. R. 2003. Further morphotaxonomical contribution to the understanding of family Hydatellaceae. *Journal of the Swamy Botanical Club* 20:1–10.

Gandolfo, M. A., Nixon, K. C., and Crepet, W. L. 2000. Monocotyledons: a review of their early Cretaceous record, pp. 44–51. *In* K. L. Wilson and D. A. Morrison (eds.), Monocots: Systematics and Evolution. CSIRO, Collingwood, Australia.

Gandolfo, M. A., Nixon, K. C., and Crepet, W. L. 2002. Triuridaceae fossil flowers from the Upper Cretaceous of New Jersey. *American Journal of Botany* 89:1940–1957.

Gandolfo, M. A., Nixon, K. C., and Crepet, W. L. 2004. Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. *Proceedings of the National Academy of Sciences, USA* 101:8056–60.

Gandolfo, M. A., Nixon, K. C., Crepet, W. L., Stevenson, D. W., and Friis, E. M. 1998. Oldest known fossils of monocotyledons. *Nature* 394:532–533.

Garland, T., Dickerman, A. W., Janis, C. M., and Jones, J. A. 1993. Phylogenetic analysis of covariance by computer simulation. *Systematic Biology* 42:265–292.

Gaut, B., Yang, L., Takuno, S., and Eguiarte, L. E. 2011. The patterns and causes of variation in plant nucleotide substitution rates. *Annual Review of Ecology, Evolution, and Systematics* 42:245–266.

Gaut, B. S., Muse, S. V., Clark, W. D., and Clegg, M. T. 1992. Relative rates of nucleotide substitution at the *rbc*L locus of monocotyledonous plants. *Journal of Molecular Evolution* 35:292–303.

Gilg-Benedict, C. 1930. Centrolepidaceae, pp. 15a: 27–33. *In* A. Engler and K. A. E. Prantl (eds.), Natürlichen Pflanzenfamilien. W. Engelmann, Leipzig, 2 edition.

Givnish, T. J., Ames, M., McNeal, J. R., McKain, M. R., Steele, P. R., DePamphilis, C. W., Graham, S. W., Pires, J. C., Stevenson, D. W., Zomlefer, W. B., Briggs, B. G., Duvall, M. R., Moore, M. J., Heaney, J. M., Soltis, D. E., Soltis, P. S., Thiele, K., and Leebens-Mack, J. H. 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales. *Annals of the Missouri Botanical Garden* 97:584–616.

Givnish, T. J., Barfuss, M. H. J., Van Ee, B., Riina, R., Schulte, K., Horres, R., Gonsiska, P. a., Jabaily, R. S., Crayn, D. M., Smith, J. A. C., Winter, K., Brown, G. K., Evans, T. M., Holst, B. K., Luther, H., Till, W., Zizka, G., Berry, P. E., and Sytsma, K. J. 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: insights from an eight-locus plastid phylogeny. *American Journal of Botany* 98:872–895.

Givnish, T. J., Evans, T. M., Pires, J. C., and Sytsma, K. J. 1999. Polyphyly and convergent morphological evolution in Commelinales and Commelinidae: evidence from *rbc*L sequence data. *Molecular Phylogenetics and Evolution* 12:360–85.

Givnish, T. J., Millam, K. C., Evans, T. M., Hall, J. C., Chris, J., Berry, P. E., and Sytsma, K. J. 2004. Ancient vicariance or recent long-distance dispersal? Inferences about phylogeny and South American–African disjunctions in Rapteaceae and Bromeliaceae based on *ndh*F sequence data. *International Journal of Plant Sciences* 165:S35–S54.

Givnish, T. J., Pires, J. C., Graham, S. W., Mcpherson, M. A., Prince, L. M., Patterson, T. B., Rai, H. S., Roalson, E. H., Evans, T. M., Hahn, W. J., Hall, C., Kress, W. J., and Sytsma, K. J. 2006. Phylogenetic relationships of monocots based on the highly informative plastid gene *ndh*F: evidence for widespread concerted convergence, pp. 28–51. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales), volume 22. Rancho Santa Ana Botanic Garden, Claremont, California.

Givnish, T. J., Pires, J. C., Graham, S. W., McPherson, M. a., Prince, L. M., Patterson, T. B., Rai, H. S., Roalson, E. H., Evans, T. M., Hahn, W. J., Millam, K. C., Meerow, A. W., Molvray, M., Kores, P. J., O'Brien, H. E., Hall, J. C., Kress, W. J., and Sytsma, K. J. 2005. Repeated evolution of net venation and fleshy fruits among monocots in shaded habitats confirms a priori predictions: evidence from an *ndh*F phylogeny. *Proceedings of the Royal Society B: Biological Sciences* 272:1481–1490.

Golovneva, L. B. 1997. Morphology, systematics and distribution of the genus *Haemanthophyllum* in the Paleogene floras of the Northern Hemisphere. *Paleontologischeskii Zhurnal* 31:197–207.

Górniak, M., Paun, O., and Chase, M. W. 2010. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, *Xdh*: congruence with organellar and nuclear ribosomal DNA results. *Molecular Phylogenetics and Evolution* 56:784–795.

Gradstein, F. M., Ogg, J. G., and Smith, A. G. 2004. A Geologic Time Scale 2004, volume 13. Cambridge University Press, Cambridge.

Graham, S. and Olmstead, R. 2000. Utility of 17 chloroplast genes for inferring the phylogeny of the basal Angiosperms. *American Journal of Botany* 87:1712–1730.

Graham, S. W. and Iles, W. J. D. 2009. Different gymnosperm outgroups have (mostly) congruent signal regarding the root of flowering plant phylogeny. *American Journal of Botany* 96:216–227.

Graham, S. W., Iles, W. J. D., and Judd, W. S. 2013. Petrosaviideae. *In* Phylonyms: A Companion to the PhyloCode. University of California Press, Berkeley.

Graham, S. W., Kohn, J. R., Morton, B. R., Eckenwalder, J. E., and Barrett, S. C. H. 1998. Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae. *Systematic Biology* 47:545–567.

Graham, S. W., Reeves, P. A., Burns, A. C. E., and 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.

Graham, S. W., Zgurski, J. M., McPherson, M. A., Cherniawsky, D. M., Saarela, J. M., Horne, E. F. C., Smith, S. Y., Wong, W. A., O'Brien, H. E., Biron, V. L., Pires, J. C., Olmstead, R. G., Chase, M. W., and Rai, H. S. 2006. Robust inference of monocot deep phylogeny using an expanded multigene plastid data set, pp. 3–20. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California.

Grass Phylogeny Working Group 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* 88:373–457.

Grass Phylogeny Working Group II 2012. New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *The New phytologist* 193:304–312.

Greenwood, D. R. and Conran, J. G. 2000. The Australian Cretaceous and Tertiary monocot fossil record, pp. 52–59. *In* K. L. Wilson and D. A. Morrison (eds.), Monocots: Systematics and Evolution. CSIRO, Collingwood, Australia.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59:307–321.

Guindon, S. and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52:696–704.

Gustafsson, A. L. S., Verola, C. F., and Antonelli, A. 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evolutionary Biology* 10:177.

Haggard, K. and Tiffney, B. 1997. The flora of the Early Miocene Brandon lignite, Vermont, USA. VIII. *Caldesia* (Alismataceae). *American Journal of Botany* 84:239–252.

Hamann, U. 1976. Hydatellaceae a new family of Monocotyledoneae. *New Zealand Journal of Botany* 14:193–196.

Hamann, U. 1998. Hydatellaceae, pp. 231–234. *In* K. Kubitzki (ed.), The Families and Genera of Vascular Plants. IV. Flowering Plants - Monocotyledons - Alismatanae and Commelinanae. Springer, Berlin.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51:1341–1351.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics (Oxford, England)* 24:129–31.

Harvey, P. H. and Pagel, M. D. 1991. The Comparative Method in Evolutionary Biology. Oxford University Press, Oxford.

Heath, T. A., Hedtke, S. M., and Hillis, D. M. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution* 46:239–257.

Hedtke, S. M., Townsend, T. M., and Hillis, D. M. 2006. Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Systematic Bology* 55:522–529.

Heibl, C. 2011. Phyloclim: integrating phylogenetics and climatic niche modelling.

Heled, J. and Drummond, A. J. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27:570–580.

Hendy, M. D. and Penny, D. 1989. A framework for the quantitative study of evolutionary trees. *Systematic Zoology* 38:297–309.

Henslow, G. 1893. A theoretical origin of endogens from exogens, through self-adaptation to an aquatic habit. *Journal of the Linnean Society of London, Botany* 29:485–528.

Herendeen, P. S. and Crane, P. R. 1995. The fossil history of the monocotyledons, pp. 1–21. *In* P. J. Rudall, P. J. Cribb, D. F. Cutler, and C. J. Humphries (eds.), Monocotyledons: Systematics and Evolution. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.

Herman, A. B. and Kvaček, J. 2010. Late Cretaceous Grünbach Flora of Austria. Naturhistorisches Museum Wien, Wien.

Hieronymus, G. 1888. Centrolepidaceae, pp. 11–16. *In* A. Engler and K. Prantl (eds.), Die Natürlichen Pflanzenfamilien, Vol. II, Part 4. W. Engelmann, Leipzig.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.

Hillis, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology* 47:3–8.

Hillis, D. M., Pollock, D. D., McGuire, J. a., and Zwickl, D. J. 2003. Is sparse taxon sampling a problem for phylogenetic inference? *Systematic Biology* 52:124–126.

Hilu, K. W., Borsch, T., Müller, K., Soltis, D. E., Soltis, P. S., Savolainen, V., Chase, M. W., Powell, M. P., Alice, L. A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T. A. B., Rohwer, J. G., Campbell, C. S., and Chatrou, L. W. 2003. Angiosperm phylogeny based on *mat*K sequence information. *American Journal of Botany* 90:1758–1776.

Ho, S. Y. W. 2009. An examination of phylogenetic models of substitution rate variation among lineages. *Biology Letters* 5:421–424.

Hochuli, P. A., Heimhofer, U., and Weissert, H. 2006. Timing of early angiosperm radiation: recalibrating the classical succession. *Journal of the Geological Society* 163:587–594.

Hofmann, C.-C. and Zetter, R. 2010. Upper Cretaceous sulcate pollen from the Timerdyakh Formation, Vilui Basin (Siberia). *Grana* 49:170–193.

Hooker, J. D. 1860. The Antarctic Voyage of H.M. Discovery Ships Erebus and Terror, Part III, Flora Tasmaniae, Vol. 2. L. Reeve & Co., London.

Hoot, S. B., Culham, A., and Crane, P. R. 1995. The utility of *atp*B gene sequences in resolving phylogenetic relationships: comparison with *rbc*L and 18S ribosomal DNA sequences in the Lardizabalaceae. *Annals of the Missouri Botanical Garden* 82:194–207.

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

Hughes, N. F. 1994. The enigma of angiosperm origins. Cambridge University Press, Cambridge.

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

Hurvich, C. M. and Tsai, C.-L. 1989. Regression and time series model selection in small samples. *Biometrika* 76:297–307.

Huwaldt, J. A. 2012. Plot Digitizer. `http://plotdigitizer.sourceforge.net`.

Iida, S., Kosuge, K., and Kadono, Y. 2004. Molecular phylogeny of Japanese species *Potamogeton* in light of noncoding chloroplast sequences. *Aquatic Botany* 80:115–127.

Iles, W. J. D., Rudall, P. J., Sokoloff, D. D., Remizowa, M. V., Macfarlane, T. D., Logacheva, M. D., and Graham, S. W. 2012. Molecular phylogenetics of Hydatellaceae (Nymphaeales): sexual-system homoplasy and a new sectional classification. *American Journal of Botany* 99:663–676.

International Commission on Stratigraphy 2012. International Chronostratigraphic Chart.

Ito, M. 1987. Phylogenetic systematics of the Nymphaeales. *The Botanical Magazine* 100:17–35.

Ito, Y., Ohi-Toma, T., Murata, J., and Tanaka, N. 2010. Hybridization and polyploidy of an aquatic plant, *Ruppia* (Ruppiaceae), inferred from plastid and nuclear DNA phylogenies. *American Journal of Botany* 97:1156–1167.

Ivany, L. C., Portell, R. W., and Jones, D. S. 1990. Animal-plant relationships and paleobiogeography of an Eocene seagrass community from Florida. *Palaios* 5:244–258.

Jacobson, A. and Hedrén, M. 2007. Phylogenetic relationships in *Alisma* (Alismataceae) based on RAPDs, and sequence data from ITS and *trn*L. *Plant Systematics and Evolution* 265:27–44.

Jansen, R. K., Cai, Z., Raubeson, L. a., Daniell, H., Depamphilis, C. W., Leebens-Mack, J., Müller, K. F., Guisinger-Bellian, M.,

Haberle, R. C., Hansen, A. K., Chumley, T. W., Lee, S.-B., Peery, R., McNeal, J. R., Kuehl, J. V., and Boore, J. L. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences, USA* 104:19369–74.

Janssen, T. and Bremer, K. 2004. The age of major monocot groups inferred from 800+ *rbc*L sequences. *Botanical Journal of the Linnean Society* 146:385–398.

Kato, Y., Aioi, K., Omori, Y., Takahata, N., and Satta, Y. 2003. Phylogenetic analyses of Zostera species based on *rbc*L and *mat*K nucleotide sequences: implications for the origin and diversification of seagrasses in Japanese waters. *Genes and Genetic Systems* 78:329–342.

Keener, B. R. 2005. Molecular systematics and revision of the aquatic monocot genus *Sagittaria* (Alismataceae). PhD thesis, University of Alabama.

Kim, K.-J. and Jansen, R. K. 1995. *Ndh*F sequence evolution and the major clades in the sunflower family. *Proceedings of the National Academy of Sciences, USA* 92:10379–10383.

Kleine, T., Maier, U. G., and Leister, D. 2009. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annual Review of Plant Biology* 60:115–138.

Knight, C. a., Clancy, R. B., Götzenberger, L., Dann, L., and Beaulieu, J. M. 2010. On the relationship between pollen size and genome size. *Journal of Botany* 2010:612017.

Knobloch, E. and Mai, D. H. 1986. Monographie der Früchte und Samen in der Kreide von Mitteleuropa. *Rozpravy Ustredniho Ustavu Geologickeho* 47:1–219.

Kozak, K. H. and Wiens, J. J. 2010. Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *American Naturalist* 176:40–54.

Kubitzki, K., Rudall, P. J., and Chase, M. W. 1998. Systematics and Evolution, pp. 23–33. *In* K. Kubitzki (ed.), The Families and Genera of Vascular Plants. III. Flowering Plants - Monocotyledons - Lilianae (except Orchidaceae). Springer, Berlin.

Kvaček, J. and Herman, A. B. 2004. Monocotyledons from the Early Campanian (Cretaceous) of Grünbach, Lower Austria. *Review of Palaeobotany and Palynology* 128:323–353.

Labra, A., Pienaar, J., and Hansen, T. F. 2009. Evolution of thermal physiology in *Liolaemus* lizards: adaptation, phylogenetic inertia, and niche tracking. *American Naturalist* 174:204–220.

Lande, R. and Schemske, D. W. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* 39:24–40.

Lanfear, R., Calcott, B., Ho, S. Y. W., and Guindon, S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29:1695–1701.

Lartillot, N. and Delsuc, F. 2012. Joint reconstruction of divergence times and life-history evolution in placental mammals using a phylogenetic covariance model. *Evolution* 66:1773–1787.

Lehtonen, S. 2006. Phylogenetics of *Echinodorus* (Alismataceae) based on morphological data. *Botanical Journal of the Linnean Society* 150:291–305.

Lehtonen, S. 2009. Systematics of the Alismataceae—A morphological evaluation. *Aquatic Botany* 91:279–290.

Lehtonen, S. and Myllys, L. 2008. Cladistic analysis of *Echinodorus* (Alismataceae): simultaneous analysis of molecular and morphological data. *Cladistics* 24:218–239.

Lemaire, B., Huysmans, S., Smets, E., and Merckx, V. 2011. Rate accelerations in nuclear 18S rDNA of mycoheterotrophic and parasitic angiosperms. *Journal of Plant Research* 124:561–576.

Lemmon, A. R. and Moriarty, E. C. 2004. The importance of proper model assumption in bayesian phylogenetics. *Systematic Biology* 53:265–277.

Lepage, T., Bryant, D., Philippe, H., and Lartillot, N. 2007. A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution* 24:2669–80.

Les, D. H., Cleland, M. A., and Waycott, M. 1997. Phylogenetic studies in Alismatidae, II: evolution of marine angiosperms (seagrasses) and hydrophily. *Systematic Botany* 22:443–463.

Les, D. H., Crawford, D. J., Kimball, R. T., Moody, M. L., and Landolt, E. 2003. Biogeography of discontinuously distributed hydrophytes: a molecular appraisal of intercontinental disjunctions. *International Journal of Plant Sciences* 164:917–932.

Les, D. H., Crawford, D. J., Landolt, E., Gabel, J. D., and Kimball, R. T. 2002. Phylogeny and systematics of Lemnaceae, the duckweed family. *Systematic Botany* 27:221–240.

Les, D. H., Garvin, D. K., and Wimpee, C. F. 1993. Phylogenetic studies in the monocot subclass Alismatidae: evidence for a reappraisal of the aquatic order Najadales. *Molecular Phylogenetics and Evolution* 2:304–314.

Les, D. H. and Haynes, R. R. 1995. Systematics of subclass Alismatidae: a synthetic approach, pp. 353–377. *In* P. J. Rudall, P. J. Cribb, D. F. Cutler, and C. J. Humphries (eds.), Monocotyledons: Systematics and Evolution. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.

Les, D. H. and Schneider, E. L. 1995. The Nymphaeales, Alismatidae, and the theory of an aquatic monocotyledon origin, pp. 23–42. *In* P. J. Rudall, P. J. Cribb, D. F. Cutler, and C. J. Humphries (eds.), Monocotyledons: Systematics and Evolution. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.

Les, D. H., Schneider, E. L., Padgett, D. J., Soltis, P. S., Soltis, D. E., and Zanis, M. J. 1999. Phylogeny, classification and floral evolution of water lilies (Nymphaeaceae; Nymphaeales): a synthesis of nonmolecular, *rbc*L, *mat*K, and 18S rDNA data. *Systematic Botany* 24:28–46.

Les, D. H. and Tippery, N. P. 2013. In time and in water... the systematics of alismatid monocotyledons, pp. 118–165. *In* P. Wilkin and S. J. Mayo (eds.), Early Events in Monocot Evolution. Cambridge University Press, Cambridge.

Levin, D. A. 1975. Minority cytotype exclusion in local plant populations. *Taxon* 24:35–43.

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

Li, X. and Zhou, Z. 2009. Phylogenetic studies of the core Alismatales inferred from morphology and rbcL sequences. *Progress in Natural Science* pp. 931–945.

Lindqvist, C., De Laet, J., Haynes, R. R., Aagesen, L., Keener, B. R., and Albert, V. a. 2006. Molecular phylogenetics of an aquatic plant lineage, Potamogetonaceae. *Cladistics* 22:568–588.

Liu, J., Qi, Z.-C., Zhao, Y.-P., Fu, C.-X., and Jenny Xiang, Q.-Y. 2012. Complete cpDNA genome sequence of *Smilax china* and phylogenetic placement of Liliales—influences of gene partitions and taxon sampling. *Molecular Phylogenetics and Evolution* 64:545–562.

Lloyd, D. G. 1980. Sexual strategies in plants III. A quantitative method for describing the gender of plants. *New Zealand Journal of Botany* 18:103–108.

Lloyd, D. G. 1982. Selection of combined versus separate sexes in seed plants. *American Naturalist* 120:571–585.

Löhne, C., Borsch, T., and Wiersema, J. H. 2007. Phylogenetic analysis of Nymphaeales using fast-evolving and noncoding chloroplast markers. *Botanical Journal of the Linnean Society* 154:141–163.

Löhne, C., Yoo, M.-j., Borsch, T., Wiersema, J., Wilde, V., Bell, C. D., Barthlott, W., Soltis, D. E., and Soltis, P. S. 2008. Biogeography of Nymphaeales: extant patterns and historical events. *Taxon* 57:1123–1146.

Losos, J. B. 2008a. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters* 11:995–1003.

Losos, J. B. 2008b. Rejoinder to Wiens (2008): Phylogenetic niche conservatism, its occurence and importance. *Ecology Letters* 11:1005–1007.

Lumbert, S. H., den Hartog, C., Phillips, R. C., and Olsen, F. S. 1984. The occurence of fossil seagrasses in the Avon Park Formation (Late Middle Eocene), Levy County, Florida (U.S.A.). *Aquatic Botany* 20:121–129.

Macphail, M. K., Alley, N. F., Truswell, E. M., and Sluiter, I. R. K. 1994. Early Tertiary vegetation: evidence from spores and pollen,

pp. 189–161. *In* R. S. Hill (ed.), History of the Australian vegetation: Cretaceous to recent. Cambridge University Press, Cambridge.

Maddison, W. P. 1997. Gene trees in species trees. *Systematic Biology* 46:523–536.

Maddison, W. P. and Maddison, D. R. 2009. Mesquite: a modular system for evolutionary analysis.

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

Magallón, S. and Sanderson, M. J. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–80.

Magallón, S. A. 2010. Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms. *Systematic Biology* 59:384–399.

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

Manchester, S. R. 2001. Update on the megafossil flora of Florissant, Colorado. *Denver Museum of Nature and Science* 4:137–216.

Martin, H. A. 2006. Cenozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments* 66:533–563.

Martín, M. and Sabater, B. 2010. Plastid *ndh* genes in plant evolution. *Plant Physiology and Biochemistry* 48:636–645.

Martínez, E. and Ramos, C. H. 1989. Lacandoniaceae (Triuiridales): una nueva familia de Mexico. *Annals of the Missouri Botanical Garden* 76:128–135.

Mathews, S., Clements, M. D., and Beilstein, M. A. 2010. A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 365:383–395.

Matsen, F. A. and Steel, M. 2007. Phylogenetic mixtures on a single tree can mimic a tree of another topology. *Systematic Biology* 56:767–775.

Merckx, V., Bakker, F. T., Huysmans, S., and Smets, E. 2009. Bias and conflict in phylogenetic inference of myco-heterotrophic plants: a case study in Thismiaceae. *Cladistics* 25:64–77.

Merckx, V., Chatrou, L. W., Lemaire, B., Sainge, M. N., Huysmans, S., and Smets, E. F. 2008a. Diversification of myco-heterotrophic angiosperms: evidence from Burmanniaceae. *BMC Evolutionary Biology* 8:178.

Merckx, V., Huysmans, S., and Smets, E. 2010. Cretaceous origins of myco-heterotrophic lineages in Dioscoreales, pp. 39–53. *In* O. Seberg, G. Petersen, A. S. Barfod, and J. I. Davis (eds.), Diversity, Phylogeny, and Evolution in the Monocotyledons. Aarhus University Press, Denmark, Aarhus.

Merckx, V., Schols, P., Geuten, K., Huysmans, S., and Smets, E. 2008b. Phylogenetic relationships in Nartheciaceae (Dioscoreales), with a focus on pollen and orbicule morphology. *Belgian Journal of Botany* 141:64–77.

Merckx, V., Schols, P., Maas-van de Kamer, H., Maas, P., Huysmans, S., and Smets, E. 2006. Phylogeny and evolution of Burmanniaceae (Dioscoreales) based on nuclear and mitochondrial data. *American Journal of Botany* 93:1684–1698.

Michelangeli, F. A., Davis, J. I., and Stevenson, D. W. 2003. Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *American Journal of Botany* 90:93–106.

Miller, J. S. and Venable, D. L. 2000. Polyploidy and the evolution of gender dimorphism in plants. *Science* 289:2335–2338.

Miller, M. A., Pfeiffer, W., and Schwartz, T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *In* 2010 Gateway Computing Environments Workshop (GCE), pp. 1–8. IEEE.

Mohr, B. A. R. and Bernardes-de Oliveira, M. E. C. 2004. *Endressinia brasiliana*, a magnolialean angiosperm from the Lower Cretaceous Crato Formation (Brazil). *International Journal of Plant Sciences* 165:1121–1133.

Mohr, B. A. R., Bernardes-de oliveira, M. E. C., and Taylor, D. W. 2008. *Pluricarpellatia*, a nymphaealean angiosperm from the Lower Cretaceous of northern Gondwana (Crato Formation, Brazil). *Taxon* 57:1147–1158.

Moore, B. R. and Donoghue, M. J. 2007. Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. *American Naturalist* 170 Suppl:S28–55.

Moore, M. J., Bell, C. D., Soltis, P. S., and Soltis, D. E. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences, USA* 104:19363–8.

Moore, M. J., Hassan, N., Gitzendanner, M. A., Bruenn, R. A., Croley, M., Vandeventer, A., Horn, J. W., Dhingra, A., Brockington, S. F., Latvis, M., Ramdial, J., Alexandre, R., Piedrahita, A., Xi, Z., Davis, C. C., Soltis, P. S., and Soltis, D. E. 2011. Phylogenetic analysis of the plastid inverted repeat for 244 species: insights into deeper-level angiosperm relationships from a long, slowly evolving sequence region. *International Journal of Plant Sciences* 172:541–558.

Moore, M. J., Soltis, P. S., Bell, C. D., Burleigh, J. G., and Soltis, D. E. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences, USA* 107:4623–8.

Mower, J. P., Touzet, P., Gummow, J. S., Delph, L. F., and Palmer, J. D. 2007. Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *BMC Evolutionary Biology* 7:135.

Nadot, S., Bittar, G., Carter, L., Lacroix, R., and Lejeune, B. 1995. A phylogenetic analysis of monocotyledons based on the chloroplast gene *rps*4, using parsimony and new numerical phenetics method. *Molecular Phylogenetics and Evolution* 4:257–282.

Nee, S., Holmes, E. C., May, R. M., and Harvey, P. H. 1994. Extinction rates can be estimated from molecular phylogenies. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 344:77–82.

Ness, R. W., Graham, S. W., and Barrett, S. C. H. 2011. Reconciling gene and genome duplication events: using multiple nuclear gene families to infer the phylogeny of the aquatic plant family Pontederiaceae. *Molecular Biology and Evolution* 28:3009–3018.

Neubig, K. M., Whitten, W. M., Carlsward, B. S., Blanco, M. a., Endara, L., Williams, N. H., and Moore, M. 2008. Phylogenetic utility of *ycf*1 in orchids: a plastid gene more variable than *mat*K. *Plant Systematics and Evolution* 277:75–84.

Neyland, R. and Urbatsch, L. E. 1996. Phylogeny of subfamily Epidendroideae (Orchidaceae) inferred from *ndh*F chloroplast gene sequences. *American Journal of Botany* 83:1195–1206.

Niklas, K. J. 1985. The aerodynamics of wind pollination. *Botanical Review* 51:328–386.

Nixon, K. C. 2008. Paleobotany, evidence, and molecular dating: an example from the Nymphaeales. *Annals of the Missouri Botanical Garden* 95:43–50.

Obbard, D. J., Harris, S. A., Buggs, R. J. A., and Pannell, J. R. 2006. Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). *Evolution* 60:1801–1815.

Olmstead, R. G. and Sweere, J. a. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* 43:467–481.

Ornduff, R. 1969. Reproductive biology in relation to systematics. *Taxon* 18:121–133.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology* 48:612–622.

Pagel, M., Meade, A., and Barker, D. 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* 53:673–684.

Pan, A. D. 2007. The late oligocene (28-27 Ma) Guang River flora from the northwestern plateau of Ethiopia. PhD thesis, Southern Methodist University.

Pannell, J. R., Obbard, D. J., and Buggs, R. J. A. 2004. Polyploidy and the sexual system: what can we learn from Mercurialis annua? *Biological Journal of the Linnean Society* 82:547–560.

Patterson, T. B. and Givnish, T. J. 2002. Phylogeny, concerted convergence, and phylogenetic niche conservatism in the core Liliales: insights from *rbc*L and *ndh*F sequence data. *Evolution* 56:233–252.

Petersen, G., Seberg, O., and Davis, J. I. 2013. Phylogeny of the Liliales (monocotyledons) with special emphasis on data partition congruence and RNA editing. *Cladistics* p. missing.

Petersen, G., Seberg, O., Davis, J. I., Goldman, D. H., Stevenson, D. W., Campbell, L. M., Michelangeli, F. A., Specht, C. D., Chase, M. W., Fay, M. F., Pires, J. C., Freudenstein, J. V., Hardy, C. R., and Simmons, M. P. 2006a. Mitochondrial data in monocot phylogenetics, pp. 52–62. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California.

Petersen, G., Seberg, O., Davis, J. I., and Stevenson, D. W. 2006b. RNA editing and phylogenetic reconstruction in two monocot mitochondrial genes. *Taxon* 55:871–886.

Phillips, S. J., Anderson, R. P., and Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231–259.

Phillips, S. J. and Dudík, M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175.

Phillips, S. J., Dudík, M., and Schapire, R. E. 2004. A maximum entropy approach to species distribution modeling. *In* C. E. Brodley (ed.), Proceedings of the 21st International Conference on Machine Learning, pp. 655–662, New York. ACM Press.

Pires, J. C., Maureira, I. J., Givnish, T. J., Sytsma, K. J., Seberg, O., Petersen, G., Davis, J. I., Stevenson, D. W., Rudall, P. J., Fay, M. F., and Chase, M. W. 2006. Phylogeny, genome size, and chromosome evolution of Asparagales, pp. 287–304. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.),

Monocots: Comparative Biology and Evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California.

Pledge, D. H. 1974. Some observations on *Hydatella inconspicua* (Cheesem.) Cheesem. (Centrolepidaceae). *New Zealand Journal of Botany* 12:559–561.

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

Posluszny, U. and Charlton, W. 1993. Evolution of the helobial flower. *Aquatic Botany* 44:303–324.

Prasad, V., Strömberg, C. A. E., Alimohammadian, H., and Sahni, A. 2005. Dinosaur coprolites and the early evolution of grasses and grazers. *Science* 310:1177–1180.

Prasad, V., Strömberg, C. A. E., Leaché, A. D., Samant, B., Patnaik, R., Tang, L., Mohabey, D. M., Ge, S., and Sahni, A. 2011. Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. *Nature Communications* 2:480.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Prychid, C. J., Sokoloff, D. D., Remizowa, M. V., Tuckett, R. E., Yadav, S. R., and Rudall, P. J. 2011. Unique stigmatic hairs and pollen-tube growth within the stigmatic cell wall in the early-divergent angiosperm family Hydatellaceae. *Annals of Botany* 108:599–608.

Qiu, Y.-L., Li, L., Wang, B., Xue, J.-Y., Hendry, T. a., Li, R.-Q., Brown, J. W., Liu, Y., Hudson, G. T., and Chen, Z.-D. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *Journal of Systematics and Evolution* 48:391–425.

R Development Core Team 2011. R: a language and environment for statistical computing. `http://www.r-project.org` .

Rai, H. S., Reeves, P. A., Peakall, R., Olmstead, R. G., and Graham, S. W. 2008. Inference of higher-order conifer relationships from a multi-locus plastid data set. *Botany* 86:658–669.

Raine, J. C., Mildenhall, D. C., and Kennedy, E. M. 2011. New Zealand fossil spores and pollen: an illustrated catalogue. http://data.gns.cri.nz/sporepollen/index.htm.

Rambaut, A. 1998. Se-Al (sequence alignment editor version 10 alpha 1). http://tree.bio.ed.ac.uk/software/seal/.

Rambaut, A. 2002. Se-Al version 2.0a11. http://tree.bio.ed.ac.uk/software/seal/.

Rambaut, A. and Drummond, A. J. 2009. Tracer v1.5. http://tree.bio.ed.ac.uk/software/tracer/.

Ramírez, S. R., Gravendeel, B., Singer, R. B., Marshall, C. R., and Pierce, N. E. 2007. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* 448:1042–1045.

Ree, R. H., Moore, B. R., Webb, C. O., and Donoghue, M. J. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.

Ree, R. H. and Smith, S. A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57:4–14.

Remizowa, M. V., Sokoloff, D. D., Macfarlane, T. D., Yadav, S. R., Prychid, C. J., and Rudall, P. J. 2008. Comparative pollen morphology in the early-divergent angiosperm family Hydatellaceae reveals variation at the infraspecific level. *Grana* 47:81–100.

Remizowa, M. V., Sokoloff, D. D., and Rudall, P. J. 2010. Evolutionary history of the monocot flower. *Annals of the Missouri Botanical Garden* 97:617–645.

Renner, S. S. and Ricklefs, R. E. 1995. Dioecy and its correlates in the flowering plants. *American Journal of Botany* 82:596–606.

Riley, M. G. and Stockey, R. a. 2004. *Cardstonia tolmanii gen. et sp. nov.* (Limnocharitaceae) from the Upper Cretaceous of Alberta, Canada. *International Journal of Plant Sciences* 165:897–916.

Ronquist, F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* 46:195–203.

Ronquist, F. and Huelsenbeck, J. P. 2003. MrBayes3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D. L., and Rasnitsyn, A. P. 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* 61:973–999.

Ronquist, F. and Sanmartín, I. 2011. Phylogenetic methods in historical biogeography. *Annual Review of Ecology, Evolution, and Systematics* 42:441–464.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. a., and Huelsenbeck, J. P. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61:539–542.

Rothwell, G. W., van Atta, M. R., Ballard, H. E. J., and Stockey, R. A. 2004. Molecular phylogenetic relationships among Lemnaceae and Araceae using the chloroplast *trn*L–*trn*F intergenic spacer. *Molecular Phylogenetics and Evolution* 30:378–385.

Rudall, P., Sokoloff, D., Remizowa, M., Conran, J., Davis, J., Macfarlane, T., and Stevenson, D. 2007. Morphology of Hydatellaceae, an anomalous aquatic family recently recognized as an early-divergent angiosperm lineage. *American Journal of Botany* 94:1073–1092.

Rudall, P. J. and Bateman, R. M. 2010. Defining the limits of flowers: the challenge of distinguishing between the evolutionary products of simple versus compound strobili. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 365:397–409.

Rudall, P. J. and Conran, J. G. 2012. Systematic placement of Dasypogonaceae among commelinid monocots: evidence from flowers and fruits. *Botanical Review* 78:398–415.

Rudall, P. J., Eldridge, T., Tratt, J., Ramsay, M. M., Tuckett, R. E., Smith, S. Y., Collinson, M. E., Remizowa, M. V., and Sokoloff, D. D. 2009a. Seed fertilization, development, and germination in Hydatellaceae (Nymphaeales): implications for endosperm evolution in early angiosperms. *American Journal of Botany* 96:1581–1593.

Rudall, P. J., Remizowa, M. V., Beer, A. S., Bradshaw, E., Stevenson, D. W., Macfarlane, T. D., Tuckett, R. E., Yadav, S. R., and Sokoloff, D. D. 2008. Comparative ovule and megagametophyte development in Hydatellaceae and water lilies reveal a mosaic of features among the earliest angiosperms. *Annals of Botany* 101:941–956.

Rudall, P. J., Remizowa, M. V., Prenner, G., Prychid, C. J., Tuckett, R. E., and Sokoloff, D. D. 2009b. Nonflowers near the base of extant angiosperms? Spatiotemporal arrangement of organs in reproductive units of Hydatellaceae and its bearing on the origin of the flower. *American Journal of Botany* 96:67–82.

Rudall, P. J., Stobart, K. L., Hong, W.-P., Conran, J. G., and Furness, C. A. 2000. Consider the lilies: systematics and Liliales, pp. 347–359. *In* K. L. Wilson and D. A. Morrison (eds.), Monocots: Systematics and Evolution. CSIRO, Collingwood, Australia.

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

Saarela, J. M. 2006. Molecular systematic studies in commelinid monocots. PhD thesis, University of British Columbia.

Saarela, J. M. and Graham, S. W. 2010. Inference of phylogenetic relationships among the subfamilies of grasses (Poaceae: Poales) using meso-scale exemplar-based sampling of the plastid genome. *Botany* 88:65–84.

Saarela, J. M., Prentis, P. J., Rai, H. S., and Graham, S. W. 2008. Phylogenetic relationships in the monocot order Commelinales, with a focus on Philydraceae. *Botany* 86:719–731.

Saarela, J. M., Rai, H. S., Doyle, J. A., Endress, P. K., Mathews, S., Marchant, A. D., Briggs, B. G., and Graham, S. W. 2007. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. *Nature* 446:312–315.

Sakai, A. K. and Weller, S. G. 1999. Gender and sexual dimorphism in flowering plants: a review of terminology, biogeographic patterns, ecological correlates, and phylogenetic approaches, pp. 1–31. *In* M. A. Geber, T. E. Dawson, and L. F. Delph (eds.), Gender and Sexual Dimorphism in Flowering Plants. Springer, Berlin.

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.

Sanderson, M. J. and McMahon, M. M. 2007. Inferring angiosperm phylogeny from EST data with widespread gene duplication. *BMC Evolutionary Biology* 7 Suppl 1:S3.

Sanmartín, I., Anderson, C. L., Alarcon, M., Ronquist, F., and Aldasoro, J. J. 2010. Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters* 6:703–707.

Sanmartín, I., van der Mark, P., and 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.

Santamaría, L. 2002. Why are most aquatic plants widely distributed? Dispersal, clonal growth and small-scale heterogeneity in a stressful environment. *Acta Oecologica* 23:137–154.

Sarkissian, T. S., Barrett, S. C. H., and Harder, L. D. 2001. Gender variation in *Sagittaria latifolia* (Alismataceae): is size all that matters? *Ecology* 82:360–373.

Sauquet, H., Ho, S. Y. W., Gandolfo, M. a., Jordan, G. J., Wilf, P., Cantrill, D. J., Bayly, M. J., Bromham, L., Brown, G. K., Carpenter, R. J., Lee, D. M., Murphy, D. J., Sniderman, J. M. K., and Udovicic, F. 2012. Testing the impact of calibration on molecular divergence times Using a fossil-rich group: the case of *Nothofagus* (Fagales). *Systematic Biology* 61:289–313.

Savolainen, V., Chase, M. W., Hoot, S. B., Morton, C. M., Soltis, D. E., Bayer, C., Fay, M. F., de Bruijn, a. Y., Sullivan, S., and Qiu, Y. L. 2000. Phylogenetics of flowering plants based on combined

analysis of plastid *atp*B and *rbc*L gene sequences. *Systematic Biology* 49:306–362.

Schaefer, H. and Renner, S. S. 2010. A three-genome phylogeny of *Momordica* (Cucurbitaceae) suggests seven returns from dioecy to monoecy and recent long-distance dispersal to Asia. *Molecular Phylogenetics and Evolution* 54:553–60.

Schluter, D., Price, T., Mooers, A. O., and Ludwig, D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51:1699–1711.

Schneider, E. L. and Carlquist, S. 2009. Xylem of early angiosperms: novel microstructure in stem tracheids of *Barclaya* (Nymphaeaceae). *Aquatic Botany* 91:61–66.

Schneider, H., Schuettpelz, E., Pryer, K. M., Cranfill, R., Magallón, S., and Lupia, R. 2004. Ferns diversified in the shadow of angiosperms. *Nature* 428:553–557.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461–464.

Sculthorpe, C. D. 1967. The Biology of Aquatic Vascular Plants. Edward Arnold, London.

Seberg, O., Petersen, G., Davis, J. I., Pires, J. C., Stevenson, D. W., Chase, M. W., Fay, M. F., Devey, D. S., Jø rgensen, T., Sytsma, K. J., and Pillon, Y. 2012. Phylogeny of the Asparagales based on three plastid and two mitochondrial genes. *American Journal of Botany* 99:875–889.

Sharma, P. P. and Giribet, G. 2012. Out of the Neotropics: Late Cretaceous colonization of Australasia by American arthropods. *Proceedings of the Royal Society B: Biological Sciences* 279:3501–3509.

Shimodaira, H. and Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* 16:1114–1116.

Sicard, A. and Lenhard, M. 2011. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany* 107:1433–1443.

Sille, N. P., Collinson, M. E., Kucera, M., and Hooker, J. J. 2006. Morphological evolution of *Stratiotes* through the Paleogene in England: an example of microevolution in flowering plants. *Palaios* 21:272–288.

Simpson, D. A., Muasya, A. M., Alves, M. V., Bruhl, J. J., Dhooge, S., Chase, M. W., Furness, C. A., Ghamkhar, K., Goetghebeur, P., Hodkinson, T. R., Marchant, A. D., Reznicek, A. A., Nieuwborg, R., Roalson, E. H., Smets, E., Starr, J. R., Thomas, W. W., Wilson, K. L., and Zhang, X. 2006. Phylogeny of Cyperaceae based on DNA sequence data—a new *rbc*L analysis, pp. 72–83. *In* Monocots: Comparative Biology and Evolution (excluding Poales).

Slocum, P. D. and Robinson, P. 1996. Water Gardening: Water Lilies and Lotuses. Timber Press, Portland, Oregon, USA.

Smith, S. A., Beaulieu, J. M., and Donoghue, M. J. 2010a. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences, USA* 107:5897–5902.

Smith, S. A. and Donoghue, M. J. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* 322:86–89.

Smith, S. Y. 2013. The fossil record of non-commelinid monocots. *In* P. Wilkin and S. J. Mayo (eds.), Early Events in Monocot Evolution. Cambridge University Press, Cambridge.

Smith, S. Y., Collinson, M. E., and Rudall, P. J. 2008. Fossil Cyclanthus (Cyclanthaceae, Pandanales) from the Eocene of Germany and England. *American Journal of Botany* 95:688–699.

Smith, S. Y., Collinson, M. E., Rudall, P. J., and Simpson, D. A. 2010b. Cretaceous and Paleogene fossil record of Poales: Review and current research. *In* O. Seberg, G. Petersen, A. S. Barfod, and J. I. Davis (eds.), Diversity, Phylogeny and Evolution in the Monocotyledons. Aarhus University Press, Denmark, Aarhus.

Smith, S. Y., Collinson, M. E., Simpson, D. a., Rudall, P. J., Marone, F., and Stampanoni, M. 2009. Elucidating the affinities and habitat of ancient, widespread Cyperaceae: *Volkeria messelensis gen. et sp. nov.*, a fossil mapanioid sedge from the Eocene of Europe. *American Journal of Botany* 96:1506–1518.

Sokoloff, D. D., Remizowa, M. V., Briggs, B. G., and Rudall, P. J. 2009a. Shoot architecture and branching pattern in perennial Hydatellaceae (Nymphaeales). *International Journal of Plant Sciences* 170:869–884.

Sokoloff, D. D., Remizowa, M. V., Linder, H. P., and Rudall, P. J. 2009b. Morphology and development of the gynoecium in Centrolepidaceae: The most remarkable range of variation in Poales. *American Journal of Botany* 96:1925–1940.

Sokoloff, D. D., Remizowa, M. V., Macfarlane, T. D., and Rudall, P. J. 2008a. Classification of the early-divergent angiosperm family Hydatellaceae: one genus instead of two, four new species and sexual dimorphism in dioecious taxa. *Taxon* 57:179–200.

Sokoloff, D. D., Remizowa, M. V., Macfarlane, T. D., Tuckett, R. E., Ramsay, M. M., Beer, A. S., Yadav, S. R., and Rudall, P. J. 2008b. Seedling diversity in Hydatellaceae: implications for the evolution of angiosperm cotyledons. *Annals of Botany* 101:153–164.

Sokoloff, D. D., Remizowa, M. V., Macfarlane, T. D., Yadav, S. R., and Rudall, P. J. 2011. Hydatellaceae: a historical review of systematics and ecology. *Rheedea* 21:115–138.

Sokoloff, D. D., Remizowa, M. V., Yadav, S. R., and Rudall, P. J. 2010. Development of reproductive structures in the sole Indian species of Hydatellaceae, *Trithuria konkanensis*, and its morphological differences from Australian taxa. *Australian Systematic Botany* 23:217–228.

Soltis, D. E., Gitzendanner, M. A., and Soltis, P. S. 2007. A 567 taxon data set for angiosperms: the challenges posed by Bayesian analyses of large data sets. *International Journal of Plant Sciences* 168:137–157.

Soltis, D. E., Smith, S. a., Cellinese, N., Wurdack, K. J., Tank, D. C., Brockington, S. F., Refulio-Rodriguez, N. F., Walker, J. B., Moore, M. J., Carlsward, B. S., Bell, C. D., Latvis, M., Crawley, S., Black, C., Diouf, D., Xi, Z., Rushworth, C. a., Gitzendanner, M. a., Sytsma, K. J., Qiu, Y.-L., Hilu, K. W., Davis, C. C., Sanderson, M. J., Beaman, R. S., Olmstead, R. G., Judd, W. S., Donoghue, M. J., and Soltis, P. S. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98:704–730.

Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savolainen, V., Hahn, W. H., Hoot, S. B., Fay, M. F., Axtell, M., Swensen, S. M., Prince, L. M., Kress, W. J., Nixon, K. C., and Farris, J. S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbc*L, and *atp*B sequences. *Botanical Journal of the Linnean Society* 133:381–461.

Stamatakis, A. 2006a. Phylogenetic models of rate heterogeneity: a high performance computing perspective. *In* Proceedings 20th IEEE International Parallel & Distributed Processing Symposium, p. 8 pp. IEEE.

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

Stamatakis, A., Hoover, P., and Rougemont, J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Systematic Biology* 57:758–771.

Stebbins, G. L. 1957. Self fertilization and the population variability in the higher plants. *American Naturalist* 91:337–354.

Steele, P. R., Hertweck, K. L., Mayfield, D., McKain, M. R., Leebens-Mack, J., and Pires, J. C. 2012. Quality and quantity of data recovered from massively parallel sequencing: Examples in Asparagales and Poaceae. *American Journal of Botany* 99:330–348.

Stefanović, S. and Olmstead, R. G. 2005. Down the slippery slope: plastid genome evolution in Convolvulaceae. *Journal of Molecular Evolution* 61:292–305.

Stevens, P. F. 2001+. Angiosperm Phylogeny Website. `http://www.mobot.org/MOBOT/research/APweb/`.

Stockey, R. A. 2006. The fossil record of basal monocots, pp. 91–106. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California.

Stone, J. L. 2002. Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. *The Quarterly Review of Biology* 77:17–32.

Strömberg, C. A. E. 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences* 39:517–544.

Sugiura, N. 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics Theory and Methods A7* 7:13–26.

Sun, G., Ji, Q., Dilcher, D. L., Zheng, S., Nixon, K. C., and Wang, X. 2002. Archaefructaceae, a new basal angiosperm family. *Science* 296:899–904.

Swenson, U., Nylinder, S., and Wagstaff, S. J. 2012. Are Asteraceae 1.5 billion years old? A reply to Heads. *Systematic Biology* 61:522–532.

Swofford, D. L. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4. Sinauer.

Sytsma, K. J., Smith, J. F., and Berry, P. E. 1991. The use of chloroplast DNA to assess biogeography and evolution of morphology, breeding systems, and flavonoids in *Fuchsia* sect. *Skinnera* (Onagraceae). *Systematic Botany* 16:257–269.

Takashima, R., Kawabe, F., Nishi, H., Moriya, K., Wani, R., and Ando, H. 2004. Geology and stratigraphy of forearc basin sediments in Hokkaido, Japan: Cretaceous environmental events on the north-west Pacific margin. *Cretaceous Research* 25:365–390.

Takhtajan, A. L. 1997. Diversity and Classification of Flowering Plants. Columbia University Press, New York.

Tamura, K., Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., and Kumar, S. 2012. Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences, USA* 109:19333–19338.

Tamura, M. N. 1998. Nartheciaceae, pp. 381–392. *In* K. Kubitzki (ed.), The Families and Genera of Vascular Plants. III. Flowering Plants - Monocotyledons - Lilianae (except Orchidaceae). Springer, Berlin.

Tamura, M. N., Azuma, H., Yamashita, J., Fuse, S., and Ishii, T. 2010. Biosystematic studies of the family Tofieldiaceae II. Phylogeny of species of *Tofieldia* and *Triantha* inferred from plastid and nuclear DNA sequences. *Acta Phytotaxonomica et Geobotanica* 60:131–140.

Tamura, M. N., Fuse, S., Azuma, H., and Hasebe, M. 2004a. Biosystematic studies on the family Tofieldiaceae I. Phylogeny and circumscription of the family inferred from DNA sequences of *mat*K and *rbc*L. *Plant Biology* 6:562–567.

Tamura, M. N., Yamashita, J., Fuse, S., and Haraguchi, M. 2004b. Molecular phylogeny of monocotyledons inferred from combined analysis of plastid *mat*K and *rbc*L gene sequences. *Journal of Plant Research* 117:109–20.

Tanaka, N., Kuo, J., Omori, Y., Nakaoka, M., and Aioi, K. 1997. Phylogeny of the family Hydrocharitaceae inferred from *rbc*L and *mat*K gene sequence data. *Journal of Plant Research* 110:329–337.

Tanaka, N., Kuo, J., Omori, Y., Nakaoka, M., and Aioi, K. 2003. Phylogenetic relationships in the genera *Zostera* and *Heterozostera* (Zosteraceae) based on *mat*K sequence data. *Journal of Plant Research* 116:273–279.

Tank, D. C. and Olmstead, R. G. 2008. From annuals to perennials: phylogeny of subtribe Castillejinae (Orobanchaceae). *American Journal of Botany* 95:608–625.

Taylor, D. W. 2008. Phylogenetic analysis of Cabombaceae and Nymphaeaceae based on vegetative and leaf architectural characters. *Taxon* 57:1082–1095.

Taylor, D. W., Brenner, G. J., and Basha, S. H. 2008. *Scutifolium jordanicum gen. et sp. nov.* (Cabombaceae), an aquatic fossil plant from the Lower Cretaceous of Jordan, and the relationships of related leaf fossils to living genera. *American Journal of Botany* 95:340–352.

Taylor, M. L. 2011. Developmental evolution of the progamic phase in Nymphaeales. PhD thesis, University of Tennessee.

Taylor, M. L., Hudson, P. J., Rigg, J. M., Strandquist, J. N., Green, J. S., Thiemann, T. C., and Osborn, J. M. 2012. Tapetum structure and ontogeny in *Victoria* (Nymphaeaceae). *Grana* 51:37–41.

Taylor, M. L., Macfarlane, T. D., and Williams, J. H. 2010. Reproductive ecology of the basal angiosperm *Trithuria submersa* (Hydatellaceae). *Annals of Botany* 106:909–920.

Taylor, T. N. and Taylor, E. L. 1993. The Biology and Evolution of Fossil Plants. Prentice Hall, Englewood Cliffs, New Jersey.

Terachi, T., Ogihara, Y., and Tsunewaki, K. 1987. The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops*. VI. Complete nucleotide sequences of the rbcL genes incoding H- and L-type Rubisco large subunits in common wheat and *Ae. crassa* 4x. *Japanese Journal of Genetics* 62:375–387.

Thien, L. B., Azuma, H., and Kawano, S. 2000. New perspectives on the pollination biology of basal angiosperms. *International Journal of Plant Sciences* 161 (Suppl:225–235.

Thien, L. B., Sage, T. L., Jaffre, T., Bernhardt, P., Pontieri, V., Weston, P. H., Malloch, D., Azuma, H., Graham, S. W., McPherson, M. a., Rai, H. S., Sage, R. F., and Dupre, J.-L. 2003. The population structure and floral biology of *Amborella trichopoda* (Amborellaceae). *Annals of the Missouri Botanical Garden* 90:466–490.

Thiers, B. 2013. Index Herbariorum: a global directory of public herbaria and associated staff. `http://sweetgum.nybg.org/ih/`.

Thomson, J. D. and Brunet, J. 1990. Hypotheses for the evolution of dioecy in seed plants. *Trends in Ecology and Evolution* 5:11–16.

Thorne, R. F. 1992. Classification and geography of the flowering plants. *Botanical Review* 58:225–348.

Tillich, H.-J., Tuckett, R., and Facher, E. 2007. Do Hydatellaceae belong to the monocotyledons or basal angiosperms? Evidence from seedling morphology. *Willdenowia* 37:399–406.

Tomlinson, P. B. 1982. Anatomy of the Monocotyledons. VII. Helobiae (Alismatidae). Clarendon Press, Oxford.

Tratt, J., Prychid, C. J., Behnke, H.-D., and Rudall, P. J. 2009. Starch-accumulating (S-type) sieve-element plastids in Hydatellaceae: implications for plastid evolution in flowering plants. *Protoplasma* 237:19–26.

Tuckett, R. E., Merritt, D. J., Hay, F. R., Hopper, S. D., and Dixon, K. W. 2010a. Comparative longevity and low-temperature storage of seeds of Hydatellaceae and temporary pool species of south-west Australia. *Australian Journal of Botany* 58:327–334.

Tuckett, R. E., Merritt, D. J., Rudall, P. J., Hay, F., Hopper, S. D., Baskin, C. C., Baskin, J. M., Tratt, J., and Dixon, K. W. 2010b. A new type of specialized morphophysiological dormancy and seed storage behaviour in Hydatellaceae, an early-divergent angiosperm family. *Annals of Botany* 105:1053–1061.

Vamosi, J. C., Otto, S. P., and Barrett, S. C. H. 2003. Phylogenetic analysis of the ecological correlates of dioecy in angiosperms. *Journal of evolutionary biology* 16:1006–18.

van den Berg, C., Goldman, D. H., Freudenstein, J. V., Pridgeon, A. M., Cameron, K. M., and Chase, M. W. 2005. An overview of the phylogenetic relationships within Epidendroideae inferred from multiple DNA regions and recircumscription of Epidendreae and Arethuseae (Orchidaceae). *American Journal of Botany* 92:613–624.

van der Ham, R., van Konijnenburg-van Cittert, J., and Indeherberge, L. 2007. Seagrass foliage from the Maastrichtian type area (Maastrichtian, Danian, NE Belgium, SE Netherlands). *Review of Palaeobotany and Palynology* 144:301–321.

Vergara-Silva, F., Ambrose, B. M., Davis, J. I., Vasquez-Lobo, A., and Gandolfo, M. A. 2003. Triuridaceae within Pandanales: a cladistic analysis based on molecular and morphological matrices. *In* Monocots III: 3rd International Conference on the Comparative Biology of Monocotyledons — Abstract Book, p. 92, Claremont, California.

Vialette-Guiraud, A. C. M., Alaux, M., Legeai, F., Finet, C., Chambrier, P., Brown, S. C., Chauvet, A., Magdalena, C., Rudall, P. J., and Scutt, C. P. 2011. *Cabomba* as a model for studies of early angiosperm evolution. *Annals of Botany* 108:589–598.

Vicentini, A., Barber, J. C., Aliscioni, S. S., Giussani, L. M., and Kellogg, E. a. 2008. The age of the grasses and clusters of origins of C 4 photosynthesis. *Global Change Biology* 14:2963–2977.

Vinnersten, A. and Bremer, K. r. 2001. Age and biogeography of major clades in Liliales. *American Journal of Botany* 88:1695–1703.

Volz, S. M. and Renner, S. S. 2008. Hybridization, polyploidy, and evolutionary transitions between monoecy and dioecy in *Bryonia* (Cucurbitaceae). *American Journal of Botany* 95:1297–1306.

von Balthazar, M., Pedersen, K. R., Crane, P. R., and Friis, E. M. 2008. *Carpestella lacunata gen. et sp. nov.*, a new basal angiosperm flower from the Early Cretaceous (Early to Middle Albian) of Eastern North America. *International Journal of Plant Sciences* 169:890–898.

von Mering, S. and Kadereit, J. W. 2010. Phylogeny , systematics , and recircumscription of Juncaginaceae —A cosmopolitan wetland family, pp. 55–79. *In* O. Seberg, G. Petersen, Barfod, and J. I. Davis (eds.), Diversity, Phylogeny, and Evolution in the Monocotyledons. Aarhus University Press, Denmark.

Wagner, W. L. and Lorence, D. H. 1999. A revision of *Trimenia* Seem. (Trimeniaceae) in the Marquesas Islands with description of a new species, *Trimenia nukuhivensis*. *Adansonia* 21:225–230.

Wakeley, J. 2008. Coalescent Theory: An Introduction. Roberts and Co., Greenwood Village, Colorado, USA.

Wang, B. and Qiu, Y.-L. 2006. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* 16:299–363.

Wang, Q. D., Zhang, T., and Wang, J. B. 2007. Phylogenetic relationships and hybrid origin of *Potamogeton* species (Potamogetonaceae) distributed in China: insights from the nuclear ribosomal internal transcribed spacer sequence (ITS). *Plant Systematics and Evolution* 267:65–78.

Warner, K. A., Rudall, P. J., and Frohlich, M. W. 2008. Differentiation of perianth organs in Nymphaeales. *Taxon* 57:1096–1109.

Waycott, M., Freshwater, D. W., York, R. A., Calladine, A., and Kenworthy, W. J. 2002. Evolutionary trends in the seagrass genus *Halophila* (Thouars): insights from molecular phylogeny. *Bulletin of Marine Science* 71:1299–1308.

Waycott, M., Procaccini, G., Les, D. H., and Reusch, T. B. H. 2006. Seagrass evolution, ecology and conservation: a genetic perspective, pp. 25–50. *In* A. W. D. Larkum, R. J. Orth, and C. M. Duarte (eds.), Seagrasses: Biology, Ecology and Conservation. Springer.

Wells, R., Clayton, J., and de Winton, M. 1998. Submerged vegetation of Lakes Te Anau, Manapouri, Monowai, Hauroko, and Poteriteri, Fiordland, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 32:621–638.

Wertheim, J. O., Fourment, M., and Kosakovsky Pond, S. L. 2012. Inconsistencies in estimating the age of HIV-1 subtypes due to heterotachy. *Molecular Biology and Evolution* 29:451–456.

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

Wicke, S., Schneeweiss, G. M., DePamphilis, C. W., Müller, K. F., and Quandt, D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology* 76:273–297.

Wiens, J. J. 2008. Commentary on Losos (2008): niche conservatism déjà vu. *Ecology Letters* 11:1004–1005.

Wiens, J. J. and Graham, C. H. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics* 36:519–539.

Wiens, J. J. and Morrill, M. C. 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Systematic biology* 60:719–731.

Wikström, N., Savolainen, V., and Chase, M. W. 2001. Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society B: Biological Sciences* 268:2211–2220.

Wilkin, P., Schols, P., Chase, M. W., Chayamarit, K., Furness, C. a., Huysmans, S., Rakotonasolo, F., Smets, E., and Thapyai, C. 2005. A plastid gene phylogeny of the yam genus, *Dioscorea*: roots, fruits and Madagascar. *Systematic Botany* 30:736–749.

Wilkinson, R. D., Steiper, M. E., Soligo, C., Martin, R. D., Yang, Z., and Tavaré, S. 2011. Dating primate divergences through an integrated analysis of palaeontological and molecular data. *Systematic Biology* 60:16–31.

Wolf, D. E. and Takebayashi, N. 2004. Pollen limitation and the evolution of androdioecy from dioecy. *The American Naturalist* 163:122–137.

Wolfe, K. H. 2001. Yesterday's polyploids and the mystery of diploidization. *Nature Reviews Genetics* 2:333–341.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., and Rieseberg, L. H. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* 106:13875–13879.

Yadav, S. R. and Janarthanam, M. K. 1994. Hydatellaceae: a new family to Indian flora with a new species. *Rheedea* 4:17–20.

Yadav, S. R. and Janarthanam, M. K. 1995. *Trithuria konkanensis* (Hydatellaceae), eine neue Art aus Indien. *Aqua-Planta* 3:91–97.

Yamada, T., Nishida, H., Umebayashi, M., Uemura, K., and Kato, M. 2008. Oldest record of Trimeniaceae from the Early Cretaceous of northern Japan. *BMC Evolutionary Biology* 8:135.

Yamashita, J. and Tamura, M. 2000. Molecular phylogeny of the Convallariaceae (Asparagales), pp. 387–400. *In* K. Wilson and D. Morrison (eds.), Monocots: systematics and evolution. CSIRO, Melbourne.

Yang, Z. 2006. Computational Molecular Evolution. Oxford University Press, Oxford.

Yang, Z. and Rannala, B. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. *Molecular Biology and Evolution* 14:717–724.

Yoo, M.-J., Bell, C. D., Soltis, P. S., and Soltis, D. E. 2005. Divergence times and historical biogeography of Nymphaeales. *Systematic Botany* 30:693–704.

Yu, Y., Harris, A. J., and He, X. 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution* 56:848–850.

Yu, Y., Harris, A. J., and He, X.-J. 2011. RASP (Reconstructed Ancestral State in Phylogenies). `http://mnh.scu.edu.cn/soft/blog/RASP`.

Zgurski, J. M. 2004. Analysis of multiple levels of vascular-plant phylogeny using nuclear RNA polymerase subunit gene (RPB2) and a large plastid data set: Studies of the tracheophytes, cycads and the lilies and relatives.

Zhang, T., Wang, Q., Li, W., Cheng, Y., and Wang, J. 2008. Analysis of phylogenetic relationships of *Potamogeton* species in China based on chloroplast *trn*T–*trn*F sequences. *Aquatic Botany* 89:34–42.

Zhao, Y.-M., Wang, W., and Zhang, S.-R. 2012. Delimitation and phylogeny of *Aletris* (Nartheciaceae) with implications for perianth evolution. *Journal of Systematics and Evolution* 50:135–145.

Zhou, Q. and Fu, D. 2008. Reproductive morphology of *Nuphar* (Nymphaeaceae), a member of basal angiosperms. *Plant Systematics and Evolution* 272:79–96.

Zmasek, C. M. and Eddy, S. R. 2001. A simple algorithm to infer gene duplication and speciation events on a gene tree. *Bioinformatics* 17:821–828.

Zomlefer, W. B., Judd, W. S., Whitten, W. M., and Williams, N. H. 2006. A synopsis of Melanthiaceae (Liliales) with a focus on character evolution in tribe Melanthieae, pp. 566–578. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson (eds.), Monocots: Comparative Biology and Evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California.

Zuckerkandl, E. and Pauling, L. B. 1962. Molecular disease, evolution, and genetic heterogeneity, pp. 189–225. *In* M. Kasha and B. Pullman (eds.), Horizons in Biochemistry. Academic Press, New York.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis, The University of Texas at Austin.

Zwickl, D. J. and Hillis, D. M. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology* 51:588–598.

Appendix A

Voucher information for Chapter 2

Specimen voucher and GenBank accession details. Collection information, herbarium in parentheses (following abbreviations in Thiers (2013), gender of plant (if dioecious), GenBank accession: *atp*B, *mat*K, *ndh*F, *rbc*L and ITS (ITS1–5.8S–ITS2). For outgroups collection information is omitted unless new sequences were generated; in cases where all sequences came from a whole plastid-genome sequence, only that GenBank accession number is provided. ITS data were not included for outgroup taxa. Newly determined sequences have GenBank accessions starting with JQ. NA: indicates that the region was not recovered.

Trithuria austinensis **D. D. Sokoloff, Remizowa, T. D. Macfarl. & Rudall: Australia: Western Australia:** *Keighery B.J. & Gibson s.n.* (PERTH), ♀ –NA, JQ284121, JQ284084, JQ284234, JQ284161; *Macfarlane 4586* (PERTH), ♀ 1–JQ284197, JQ284122, JQ284085, JQ284235, JQ284162, ♀ 3–JQ284199, JQ284124, JQ284087, JQ284237, JQ284164,♀ 5–JQ284201, JQ284126, JQ284089, JQ284239, JQ284165, ♂ 2–JQ284198, JQ284123, JQ284086, JQ284236, JQ284163, ♂ 4–JQ284200, JQ284125, JQ284088, JQ284238, NA, ♂ 6–JQ284202, JQ284127, JQ284090, JQ284240, JQ284166; *Macfarlane 4163 & Hearn* (PERTH), ♀ 6–JQ284204, JQ284129, JQ284092, JQ284242, JQ284168, ♂ 1–JQ284203, JQ284128, JQ284091, JQ284241, JQ284167, ♂ 7–JQ284205, JQ284130, JQ284093, JQ284243, NA. ***Trithuria australis*** **(Diels) D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall: Australia: Western Australia:** *Keighery G.J. & Gibson 2584* (PERTH), JQ284207, JQ284132, JQ284095, JQ284245, JQ284170; *Rudall s.n.* (MW), JQ284206, JQ284131, JQ284094, JQ284244, JQ284169; *Taylor s.n.* (TENN), JQ284209, JQ284134, JQ284097, JQ284247, JQ284172; *Taylor 62* (TENN), JQ284208, JQ284133, JQ284096, JQ284246, JQ284171.

Trithuria bibracteata **Stapf ex D.A. Cooke: Australia: Western Australia:** *Gunness & al. 13/37* (PERTH), JQ284215, JQ284142, JQ284102, JQ284253, JQ284178; *Keighery B.J. & Gibson 801* (PERTH), JQ284217, JQ284144, JQ284104, JQ284255, JQ284180; *Taylor 60* (TENN), JQ284216, JQ284143, JQ284103, JQ284254, JQ284179. ***Trithuria cookeana*** **D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall: Australia: Northern Territory:** *Cowie 5934* (DNA), ♀ , JQ284186, JQ284110, JQ284073, JQ284223, JQ284151. ***Trithuria cowieana*** **D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall: Australia: Northern Territory:** *Cowie & Dixon s.n.* (DNA), JQ284189, JQ284113, JQ284076, JQ284226, JQ284154; *Cowie & Jacka 9995* (DNA), JQ284190, JQ284114, JQ284077, JQ284227, NA; *Macfarlane & al. 4217* (MW), JQ284187, JQ284111, JQ284074, JQ284224, JQ284152; *Macfarlane & al. 4238* (MW), JQ284188, JQ284112, JQ284075, JQ284225, JQ284153. ***Trithuria filamentosa*** **Rodway: Australia: Tasmania:** *Briggs 9775* (NSW), ♀ , JQ284210, JQ284135, NA, JQ284248, JQ284173; *Buchanan 12328* (HO), ☿ , JQ284211, JQ284136, JQ284098, JQ284249, JQ284174; *Feild 210* (TENN), ♀ , FJ514801, JQ284137, FJ514806, FJ514807, JQ284175. ***Trithuria inconspicua*** **Cheesem.: New Zealand: North Island:** *Chapman s.n.* (NSW), ♀ , JQ284212, JQ284138, JQ284099, JQ284250, NA. **South Island:** *Clayton 1145* (CHR), ♀ , JQ284213, JQ284139, JQ284100, JQ284251, JQ284176; *Forester & Goh s.n.* (AK), ♀ , JQ284214, JQ284140, JQ284101, JQ284252, JQ284177. ***Trithuria konkanensis*** **Yadav & Jarnarthanam: India: Maharashtra:** *Yadav s.n.* (MW), 2006, JQ284193, JQ284117, JQ284080, JQ284230, JQ284157; *Yadav s.n.* (MW), 2007, JQ284192, JQ284116, JQ284079, JQ284229, JQ284156. ***Trithuria lanterna*** **D.A. Cooke: Australia: Northern Territory:** *Egan 4816 & Knox* (DNA), JQ284196, JQ284120, JQ284083, JQ284233, JQ284160; *Macfarlane & al. 4268* (MW), JQ284194, JQ284118, JQ284081, JQ284231, JQ284158; *Macfarlane & al. 4321* (MW), JQ284195, JQ284119, JQ284082, JQ284232, JQ284159. ***Trithuria occidentalis*** **Benth.: Australia: Western Australia:** *Keighery G.J. 4204* (PERTH), ♀ , NA, JQ284141, NA, NA, NA. ***Trithuria polybracteata*** **D.A. Cooke ex D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall. Australia: Western Australia:** *Willis s.n.* (MEL), ♀ , JQ284191, JQ284115, JQ284078, JQ284228, JQ284155. ***Trithuria submersa*** **Hook.f.: Australia: South Australia:** *Conran 961 & Rudall* (ADU), JQ284218, JQ284145, NA, DQ915188, NA; *Doust et al. 1123* (MELU), AJ419142, JQ284146, AF547020, DQ915187, JQ284181. **Tasmania:** *Moscal 20272* (HO), JQ284219, JQ284147, JQ284105, JQ284256, JQ284182. **Western**

Australia: *Sandiford 902* (PERTH), JQ284220, JQ284148, JQ284106, JQ284257, JQ284183; *Taylor 61* (TENN), JQ284222, JQ284150, JQ284108, JQ284259, JQ284185; *Taylor 63* (TENN), JQ284221, JQ284149, JQ284107, JQ284258, JQ284184. **Outgroups:** ***Amborella trichopoda*** **Baill.**, AF235041, AJ506156, AJ506156, L12628; ***Austrobaileya scandens*** **C.T. White**, AF092107, DQ182344, NA, L12632; ***Cabomba caroliniana*** **A. Gray**, *Les s.n.* (CONN), AF187058, JQ284109, JQ284072, M77027; ***Chloranthus spicatus*** **Makino**, EF380352; ***Illicium oligandrum*** **Merr. & Chun**, NC_009600; ***Nuphar advena*** **Ait.**, NC_008788; ***Nymphaea alba*** **L.**, AJ627251; ***Schisandra chinensis*** **(Turcz.) Baill.**, AF239790, DQ185526, NA, AF238061; ***Trimenia moorei*** **(Oliv.) Philipson**, *Floyd s.n.* (NSW), AY116653, DQ401360, JQ284071, AY116658.

Appendix B

Trithuria classification for Chapter 2

A new sectional treatment for *Trithuria*.

Trithuria Hook. f., Fl. Tasman. 2: 78 (1858).

Typus: *Trithuria submersa* Hook. f.

(1) *Trithuria* sect. *Hamannia* D.D. Sokoloff, Iles, Rudall & S.W. Graham, *sect. nov.*

Typus: *Trithuria lanterna* D.A. Cooke.

Description: Fruits with three longitudinal ribs. Papillae and distinct epicuticular wax deposits on fruit surface absent. Fruits usually dehiscent, but specialized endocarp cells facilitating fruit dehiscence absent. Fruit apex without thickened endocarp cells. Seeds smooth, with firm cuticle. Seedlings without a cotyledonary sheath.

Etymology: The section is named in honour of Ulrich Hamann whose extensive work demonstrated the distinctiveness of the family Hydatellaceae and its unclear position in the classification of angiosperms.

Included species: *T. lanterna*, *T. konkanensis* Yadav & Janarthanam, *T. polybracteata* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall.

(2) *Trithuria* sect. *Altofinia* D.D. Sokoloff, Iles, Rudall & S.W. Graham, *sect. nov.*

Typus: *Trithuria cowieana* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall.

Description: Fruits without longitudinal ribs. Exocarp cells each typically form two papillae. Distinct epicuticular wax deposits on fruit surface absent. Fruits indehiscent. Fruit apex with thickened endo-

carp cells. Seeds smooth, with firm cuticle.

Etymology: The name is derived from the Latin words *altus* (high, elevated, tall) and *finis* (boundary, limit, border, end), reflecting the currently known distribution of representatives of the section in the Top End region of Northern Territory, Australia.

Included species: *T. cowieana*, *T. cookeana* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall.

(3) *Trithuria* sect. *Hydatella* (Diels) D.D. Sokoloff, Iles, Rudall & S.W. Graham, *comb. & stat. nov.*

Basionym: *Hydatella* Diels, Bot. Jahrb. Syst. 35: 93. (1904).

Lectotypus (designated by Sokoloff et al., Taxon 57: 192, 2008): *Hydatella australis* Diels. (≡ *T. australis* (Diels) D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall).

Included species: *T. australis*, *T. austinensis* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall, *T. filamentosa* Rodway, *T. inconspicua* Cheesem.

(4) *Trithuria* sect. *Trithuria*

Typus: *Trithuria submersa* Hook. f.

Included species: *T. submersa*, *T. occidentalis* Benth., *T. bibracteata* Stapf ex D.A. Cooke.

Appendix C

Supplementary material for Chapter 3

C.1 Dating seed-plant phylogeny

I sampled early lineages of angiosperms and seed-plant relatives for which I could confidently assign representative fossils (see below). I included 17 lineages comprising three gymnosperms, *Amborella*, three water lilies (Cabombaceae and Nymphaeaceae), three Austrobaileyales, and four members of Mesangiospermeae (Cantino et al. 2007; see Appendix C.5). I included representatives of the two deepest splits in Hydatellaceae (Chapter 2) by adding *Trithuria cowieana* to an earlier alignment (Graham and Olmstead, 2000; Graham and Iles, 2009; Saarela et al., 2007). Genomic sampling focused on 13 single-copy plastid genes. I amplified and sequenced the targeted markers that lie in six distinct plastid regions (comprising four multi-gene clusters, *psb*D-*psb*C, *psb*E-*psb*F-*psb*L-*psb*J, *psb*B-*psb*T-*psb*N-*psb*H, and three single-gene regions, *ndh*F, *rbc*L and *atp*B). I considered only protein-coding regions in analyses. Details of DNA extraction, amplification, sequencing, contig assembly and alignment are described elsewhere (Graham and Olmstead 2000; Graham et al. 2000; Saarela et al. 2008; Chapter 2).

Fossils with unequivocal affinity to Hydatellaceae are unknown (see Chapter 3) and as a result information about when Hydatellaceae diverged from water lilies, and when subsequent diversification within the lineage took place, must come from fossil evidence outside of the family. I used the earliest appearance of Cordaitales (Taylor and Taylor, 1993) as a constraint on the age of the crown seed plants (Doyle, 2008; Magallón, 2010;

Mathews et al., 2010), and used glossopterids as a constraint on the age of the stem angiosperms (Doyle, 2008; Hughes, 1994; Magallón, 2010; Mathews et al., 2010; Taylor and Taylor, 1993). Two fossils provided constraints on the crown clade of water lilies (Friis et al., 2009, 2001; Mohr et al., 2008; Taylor et al., 2008). The stem age of Trimeniaceae was constrained using a recently discovered seed fossil (Yamada et al., 2008). To be conservative I considered well characterized Lower Cretaceous tricolpate pollen (Hughes and McDougall, 1990) as a constraint for the stem age of the eudicots, rather than their crown age (Clarke et al., 2011). A suite of fossil Platanaceae from the Upper Cretaceous (Doyle and Endress, 2010) provide a constraint on the crown age of eudicots. The crown age of monocots was constrained using the aroid fossil *Mayoa portugallica* (Friis et al., 2004). I used minimum ages of fossil localities noted in the references in Table C.1; note that in some cases these are later adjustments of Clarke et al. (2011). Geological ages follow Gradstein et al. (2004)).

To test for and accommodate non-clocklike behaviour in the seed-plant data set I used the Bayesian random local clocks method (Drummond and Suchard, 2010). This accommodates molecular rate variation by allowing different parts of the tree to have unique molecular clocks. An internal model-selection procedure is used to determine the number of local clocks (it therefore automatically tests against the hypothesis of one global clock; Drummond and Suchard 2010). The method is implemented in BEAST version 1.6.1 (Drummond and Rambaut, 2007) and co-estimates phylogeny, divergence times and the number and location of local clocks using Bayesian inference. I used a GTR + Γ model of sequence evolution, with default priors (or those suggested by `http://code.google.com/p/beast-mcmc/wiki/ParameterPriors` if not automatically implemented). I present results in which cycads were constrained to be the sister group of angiosperms among extant seed plants, consistent with some recent studies (e.g., Doyle, 2008; Graham and Iles, 2009; Magallón, 2010; Mathews et al., 2010). In all analyses, I also constrained Nymphaeaceae s.s. to be monophyletic based on molecular and morphological analyses (Löhne et al., 2007; Soltis et al., 2011; Taylor, 2008). The BEAST analysis requires that each fossil calibrations

Table C.1: Calibration fossils for seed-plant phylogeny. Fossil taxa and associated taxonomic assignments are noted for labelled nodes in Figure 3.1. Fossil ages and assigned lognormal priors are noted with references (some ages are based on Clarke et al., 2011, noted with asterisk).

Fossil, higher taxon (node)	Age of fossil(s) (Ma)	Prior parameters: 95% prior density interval (log-mean, log-standard deviation).	References
Cordaites spp., Cordaitales (A)	315	316.0367.5 (2, 1)	Taylor and Taylor (1993)
Gangamopteris spp., glossopterids (C)	293.8	294.8346.3 (2, 1)	Taylor and Taylor (1993), Hughes (1994)
Pluricarpellatia peltata, Cabombaceae (G)	98.7*	99.1118.0 (1, 1)	Mohr et al. (2008), Taylor et al. (2008)
Unnamed seed, Trimeniaceae (M)	98.7	98.9110.4 (0.5, 1)	Yamada et al. (2008)
Tricolpate pollen, Eudicotyledonae (N)	124*	124.2135.7 (0.5, 1)	Hughes and McDougall (1990)
Mayoa portugallica, Araceae (O)	96.1*	96.5115.4 (1, 1)	Friis et al. (2004)
West Brothers platanoid and *Sapindopsis*, Platanaceae (P)	92.8*	93.2112.1 (1, 1)	Doyle and Endress (2010)

have an associated prior. I used lognormal priors with 95% highest prior density intervals of ∼10–40 Ma, with larger intervals for older fossils (Appendix Table C.1), except for Trimeniaceae and Eudicotyledonae. I ran seven runs of 4.0×10^7 generations, and considered four that converged on the same posterior and likelihood scores after 10% burnin. The estimated sample sizes (ESS) of run statistics (posterior, prior, likelihood, parameter estimates) were all over 200 when these runs were pooled. Details of inferred node ages and credibility intervals (CI, posterior probability interval) are shown graphically in Figure 3.1 are provided in Table C.2. I also explored alter-

native possible sister groups in which other gymnosperms were constrained as the sister group (conifers alone, *Ginkgo* along, or pairwise combinations of conifers, cycads and *Ginkgo*) or no outgroup was constrained. This had minimal effect on the two ages within Hydatellaceae (<1 Ma difference; data not shown).

Table C.2: Inferred ages of splits in seed-plant phylogeny based on a Bayesian random local clocks analysis. Mean and 95% credibility intervals are indicated for individual nodes (Figure 3.1), which correspond to the crown clade indicated. MRCA, most recent common ancestor; N/A, not applicable (due to constrained split or root node).

Node	Clade	Mean age (Ma), 95% CI
A	Seed plants	325.87, 315.21–349.37
B	MRCA *Ginkgo* + *Pinus*	301.56, 274.79–330.71
C	MRCA angiosperms + *Cycas*	302.16, 294.18–314.42
D	Angiosperms	158.67, 151.05–167.72
E	MRCA *Amborella* + Nymphaeales	152.44, 146.25–158.91
F	Nymphaeales	126.69, 120.58–133.21
G	MRCA *Cabomba* + Nymphaeaceae	102.11, 98.84–107.07
H	Nymphaeaceae	95.52, 92.94–99.61
I	Hydatellaceae	19.07, 15.72–23.36
J	MRCA *T. filamentosa* + *T. submersa*	16.21, 13.41–19.28
K	MRCA Austrobaileyales + Mesangiosperms	140.22, 135.32–145.58
L	Austrobaileyales	118.27, 109.87–130.83
M	MRCA *Schisandra* + *Trimenia*	102.12, 98.77–107.93
N	Mesangiosperms	125.11, 124.06–126.82
O	Monocots	104.24, 97.98–110.47
P	Eudicots	96.28, 92.98–100.39

C.2 Dating the species tree of Hydatellaceae

I used a Bayesian implementation of the multi-species coalescent (Heled and Drummond, 2010) to date divergences within Hydatellaceae. This method finds the best species tree given individual gene trees (based on unlinked loci) using a coalescent process (Heled and Drummond, 2010). The method, which is implemented in BEAST version 1.6.1 (Drummond and Rambaut, 2007) as *BEAST (Heled and Drummond, 2010), co-estimates the species tree and divergences. Currently this method cannot accommodate reticulation events (hybridization); incongruence among gene-trees is assumed to be due to incomplete lineage sorting (Heled and Drummond, 2010). I used the molecular data-set of Chapter 2, which includes all extant species, and comprises four concatenated four plastid regions and a separate nuclear-based data set comprising the ribosomal internal transcribed spacers 1 and 2. I followed the methods of Chapter 2 to reconstruct the multi-species coalescent species tree, except that I assigned two Hydatellaceae posterior distributions (mean and 95% CI) determined from the seed-plant dating analysis (Figure 3.1, Table C.2) as normal priors for the corresponding splits. These priors were only applied to the plastid loci, using the rooting of Hydatellaceae determined in the seed-plant analysis.

C.3 Biogeographic analysis of Hydatellaceae

I identified five discrete geographic areas reflecting the current distribution of the family, considering breaks that reflect natural barriers to regular dispersal (e.g., the extremely arid Nullarbor Plain of southern Australia). These areas are: northern Australia (NA), southeastern Australia (including Tasmania) (SEA), southwestern Australia (SWA), India, and New Zealand (Figure 3.2A; Saarela et al. 2008). I treated SEA as one area despite it being climatically variable, because species ranges (*Trithuria filamentosa* and *T. submersa*) are essentially contiguous in Tasmania, and there is no evidence of a strong biogeographic barrier between them (Sokoloff et al. 2008). The dispersal-vicariance (DIVA) of Ronquist (1997) and the dispersal-extinction-

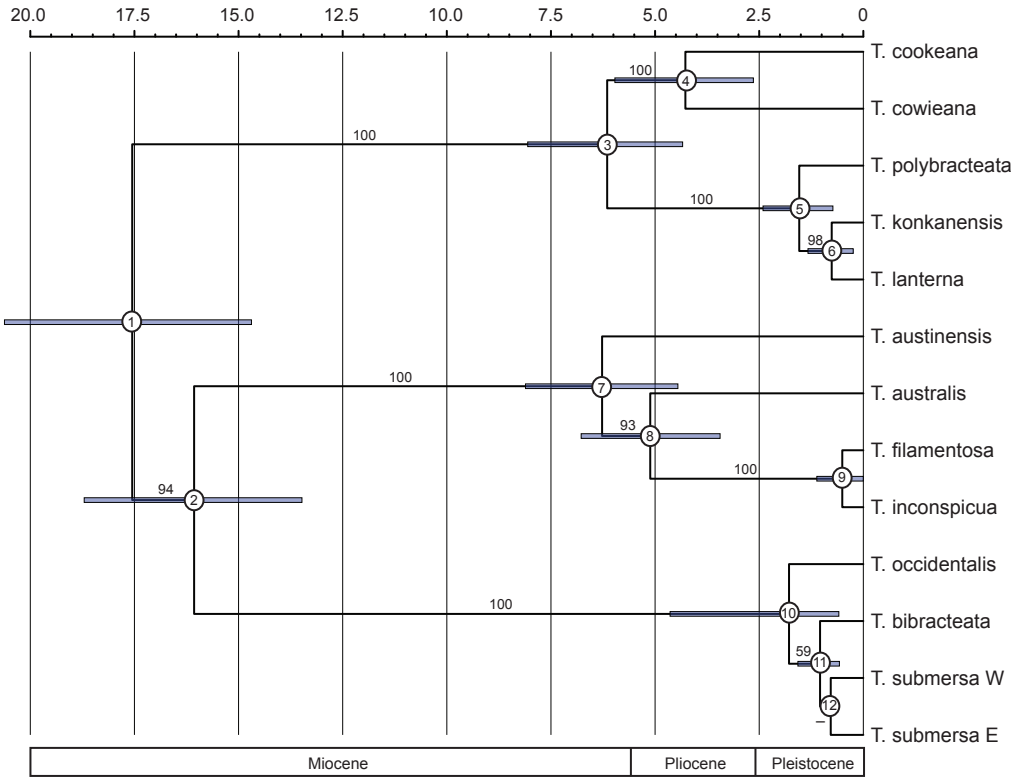


Figure C.1: Dated species tree of Hydatellaceae inferred using a multispecies coalescent analysis using uncorrelated lognormal clocks, based on four plastid genes and nuclear ITS. Labeled nodes are referred to in Appendix Table C.3. Numbers beside branches are posterior probabilities (PP) expressed as percentages; dash indicates <50% support. 95% credibility intervals are noted. The time scale is in Ma, the geological scale is shown. Letters adjacent to tips represent: E = east, T. = *Trithuria*, W = west.

cladogenesis (DEC) method of Ree et al. (2005) and Ree and Smith (2008) allow extant species and reconstruction of internal nodes (ancestral species) to encompass multiple areas. Examples of these at internal nodes facilitate identification of dispersal, extinction and vicariance events within these frameworks. In contrast, maximum likelihood ancestral state reconstructions (ML ASR Lewis, 2001; Schluter et al., 1997) only implicitly consider dispersal, and restrict each species range and internal node to a single area. In all biogeographic analyses, the range of each extant species comprises one of the five areas (I treated the SEA and SWA members of *Trithuria submersa* as distinct species, following Chapter 2). ML ASR may be suitable when the group of interest is younger than tectonic events separating areas, or where areas were never geologically connected; the former seems to apply here given the ages determined in the dating analyses (Figure 3.2, Appendix Figure C.1). It also allows for direct estimation of dispersal rates between areas, and provides a straightforward framework for model comparison.

Table C.3: Ages of nodes in Hydatellaceae based on a Bayesian multi-species coalescent analysis of plastid and nuclear data. Mean and 95% credibility intervals are indicated for individual nodes (Appendix Figure C.1). NA, not applicable because split occurred in less than 50% of bipartitions. Sectional classification according to Chapter 2.

Node	(taxon if applicable)	Mean age (Ma), 95% CI
1	crown *Trithuria*	17.55, 14.69–20.62
2		16.07, 13.48–18.71
3		6.15, 4.34–8.06
4	crown sect. *Altofinia*	4.27, 2.64–5.97
5	crown sect. *Hamannia*	1.54, 0.73–2.41
6		0.76, 0.24–1.33
7	crown sect. *Hydatella*	6.27, 4.45–8.11
8		5.12, 3.44–6.77
9		0.51, 0.00–1.12
10	crown sect. *Trithuri*a	1.78, 0.59–4.64
11		1.04, 0.57–1.57
12		0.78, NA

I used the maximum clade credibility species tree (Appendix Figure C.1) as the reference tree for all biogeographic reconstructions. I mostly used default settings in the DIVA analysis (using RASP version 2.0; Yu et al. 2010, 2011), and for the DEC analysis (using Lagrange version 20110117; Ree and Smith 2008), but in both cases limited the connectivity (dispersal or vicariance) of India to NA, and that of New Zealand to SEA. Although the models used by DIVA and DEC permit area extinction (extirpation), in practice this is rarely or never estimated (Ree and Smith, 2008; Ronquist and Sanmartín, 2011). I therefore estimated the lineage extinction rate using a Bayesian tree-shape based method (Nee et al. 1994a, 1994b) implemented in Diversitree version 0.7-2 (FitzJohn, 2012); note that area and lineage extinction concepts are only partly related (see Chapter 3). The Diversitree analysis gave a mean speciation rate of 0.430 Ma^{-1} (0.107–0.881; 95% Bayesian credibility intervals, CI) and a mean lineage extinction rate of 0.446 Ma^{-1} (0.003–0.955; 95% CI). Because of the large uncertainty in lineage extinction rates, I decided to examine the effect of a broad range of area extinction rates on the DEC analysis. I considered six area extinction rates (0.001, 0.01, 0.05, 0.1, 0.5, and 1.0 Ma^{-1}) spanning the range of lineage extinction rates. I iteratively changed values of the dispersal rate in Lagrange to find its optimal value for each extinction rate. I plotted the best geographic range splits of individual nodes in the DEC analyses based on their relative likelihoods, as a function of area extinction rate (Appendix Figure C.2). I also plotted the range outcomes of for each descendent lineage (represented by cumulative relative likelihoods) as a function of the area extinction rate (Appendix Figure C.3). I performed ML ASR of geographic ranges using BayesTraits version 1.0 (Pagel 1999; `www.evolution.rdg.ac.uk`), for the three models described in Chapter 3 (simple ML ASR, continental ML ASR, and full ML ASR). Root frequencies were set to empirical values. The full ML ASR had the best AICc score (differences between best and alternative models: simple ML ASR $\Delta = 1.06$; continental ML ASR $\Delta = 2.135$).

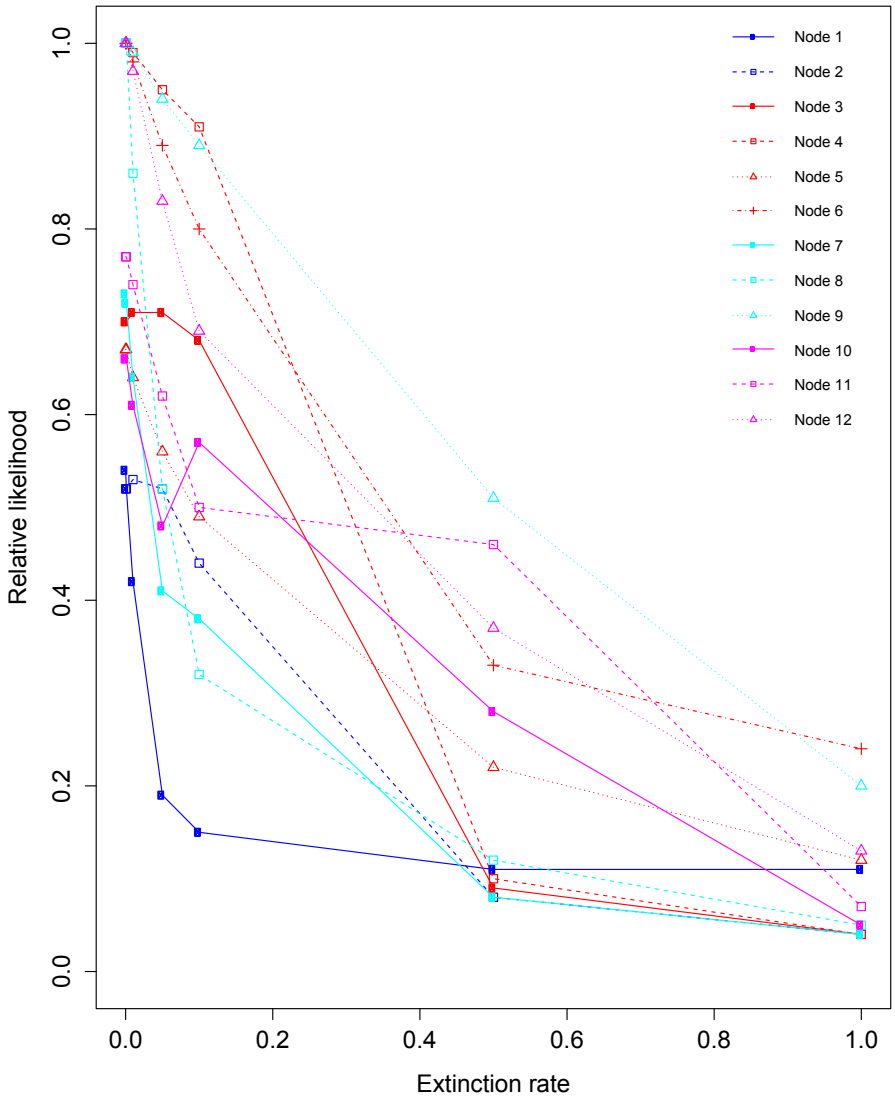


Figure C.2: The effect of extinction on best geographic range split. Relationship between area extinction rate and the relative likelihood of the best geographic range split in the DEC analysis. All except the first extinction value was assigned, with dispersal rate then optimized relative to them. The first extinction value is the auto-optimized estimate (4.285×10^{-9}). Node numbers correspond to those in Appendix Figure C.1.

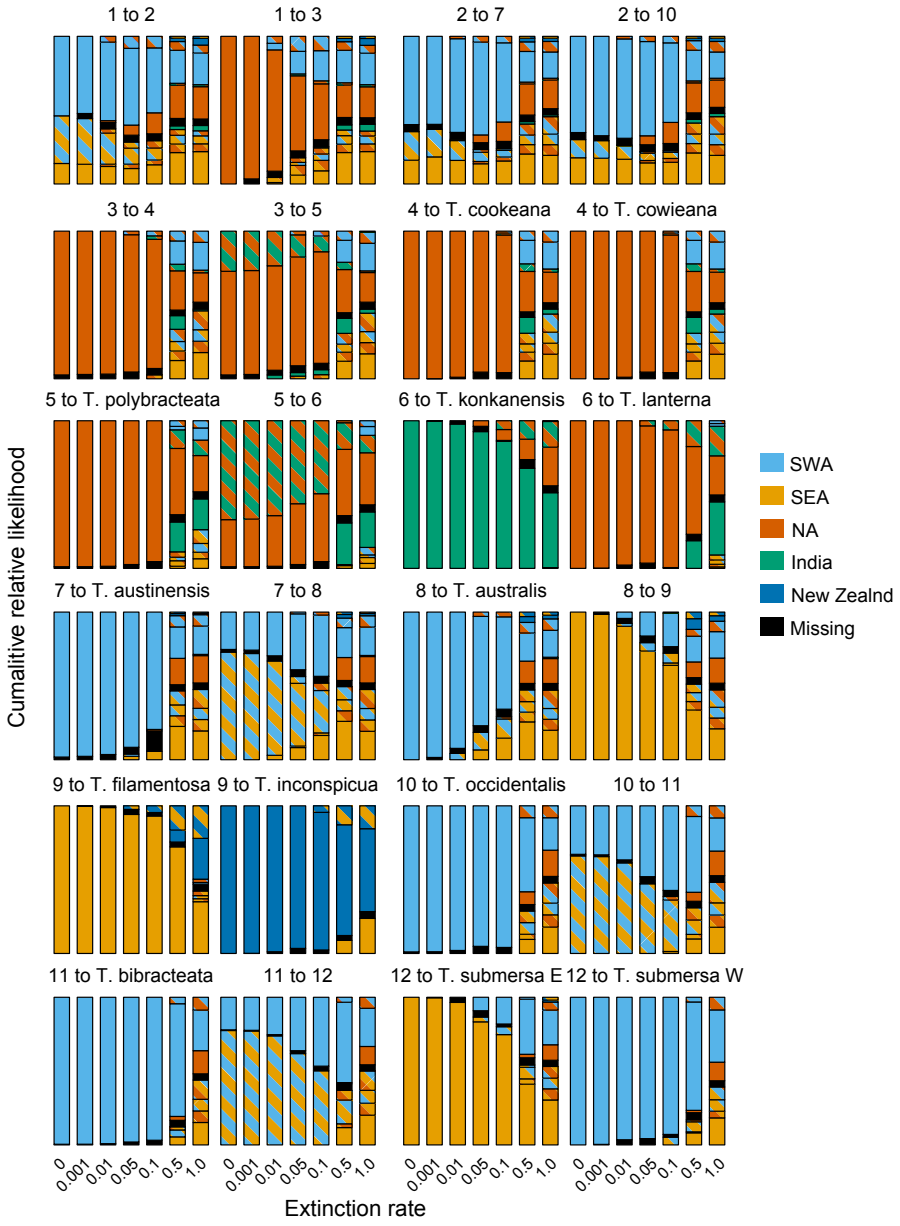
C.4 Niche evolution in Hydatellaceae

Species occurrence data were collated from online herbaria sources (Council of Heads of Australasian Herbaria, 2012) augmented by recent collections (Appendix Table C.4). We georeferenced herbarium collections lacking explicit latitude and longitude coordinates using locality information where this was sufficiently precise (1 km or less). Accurate niche modeling is not possible with limited data, and so I excluded three of four dioecious species (*Trithuria cookeana*, *T. cowieana*, and *T. polybracteata*), as these are currently known from single localities (Sokoloff et al., 2008a).

To estimate the climate niche of extant species I used a maximum entropy approach (Maxent version 3.3.3e; Phillips and Dudík 2008; Phillips et al. 2004, which allows the use of presence-only data such as herbarium records. I considered the Bioclim data set of 19 derived climate variables (Hijmans et al. 2005; `http://www.worldclim.org`); the climate niche is integrated across all variables. Map grids were formatted using ArcGIS version 9.3 ESRI (2009). For each species I ran the Maxent analysis on each species, considering 100 bootstrap replicates. I used 20% of the data for training and the remainder for testing. The accuracy of the model was evaluated using the area-under-the-curve method. Only scores >0.95 were considered adequate: all analyzed species displayed adequate model accuracy. Niche profiles for individual species were output as raw probability surfaces. To recover occupancy profiles (probability surface) of individual climate variables from the integrated niche of each species I employed the

---

Figure C.3 *(following page)*: The effect of extinction on the cumulative geographic range. The relationship between area extinction rate and the cumulative relative likelihood of each geographic range in the DEC analysis for individual branches (node to node) noted above each graph (see also Appendix Figure C.2). Stacked bar graphs sum to unity; subdivisions represent the proportion for each inferred range. The 'zero' extinction rate is the auto-optimized estimate (4.285×10^{-9}). Hashed ranges comprise more than one area. NA, northern Australia; SEA, southeastern Australia; SWA, southwestern Australia.


1 to 2
1 to 3
2 to 7
2 to 10
3 to 4
3 to 5
4 to T. cookeana
4 to T. cowieana
5 to T. polybracteata
5 to 6
6 to T. konkanensis
6 to T. lanterna
7 to T. austinensis
7 to 8
8 to T. australis
8 to 9
9 to T. filamentosa
9 to T. inconspicua
10 to T. occidentalis
10 to 11
11 to T. bibracteata
11 to 12
12 to T. submersa E
12 to T. submersa W
Cumalitive relative likelihood
0
0.001
0.01
0.05
0.1
0.5
1.0
Extinction rate
SWA
SEA
NA
India
New Zealnd
Missing

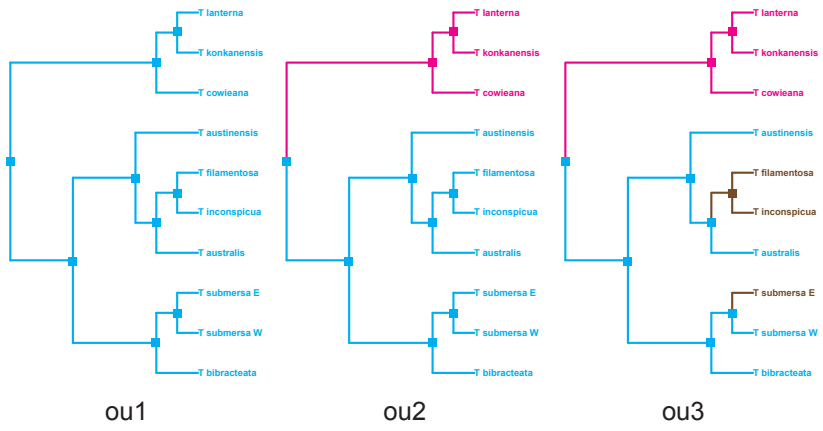


Figure C.4: Cladograms showing assigned locations of bioclimatic regimes assumed for the three different Ornstein-Uhlenbeck models (ou1–3). Node assignments are based on the biogeographic shifts predicted by the full ML ASR (Figure 3.2B).

method of Evans et al. (2009), using the program Phyloclim version 0.8.1 (Heibl, 2011). This surface relates the probability of spatial occupancy of a species to a particular climate value. These profiles were used to calculate mean values of climate preferences for individual species.

I looked for evidence of phylogenetic niche conservatism in individual climate variables by model comparison of Brownian motion and Ornstein-Uhlenbeck processes. The Ornstein-Uhlenbeck model is a Brownian-motion like process with a central tendency. It has three parameters: the strength of the tendency (α), the value of the optimum (θ), and the variability of the process (σ^2). When α is zero the model reduces to classical Brownian motion (Butler and King, 2004; Hansen, 1997). The Ornstein-Uhlenbeck process was introduced in evolutionary biology to model stabilizing selection in quantitative traits (Hansen, 1997). It has also been used to infer phylogenetic niche conservatism by testing against a Brownian motion model

(Boucher et al., 2012; Cooper et al., 2010; Kozak and Wiens, 2010; Labra et al., 2009). The three Ornstein-Uhlenbeck process models examined here relate to varying degrees of conservatism (Appendix Figure C.4; see Chapter 3). Mean values of climate preferences for individual species represent point estimates of niche. To understand how the occupancy profiles affects calculation of parameter estimates and model fit I sampled 1000 replicates for each climate variable proportionately from its surface (Table 3.1). Intervals reported in the tables are 95% bounds inferred from these replicates, and give a sense of the underlying variability in parameter estimates and model choice.

Table C.4: Specimen accessions with latitude and longitude (decimal degrees) used for Maxent analyses. Herbarium acronyms follow Thiers (2013).

Species	Collector	Herbarium	Latitude	Longitude
Trithuria austinensis	Gibson & Lyons 2387	PERTH	-32.6169	115.7759
	Hearn s.n.	PERTH, K	-34.3599	116.7245
	Hopper & Yates 8681	PERTH, K	-33.9869	123.187
	Keighery B.J. & Gibson 85	PERTH	-32.6156	115.7867
	Keighery B.J. & Gibson s.n.	PERTH	-34.4068	116.7218
	Keighery G.J. 15707	PERTH	-31.3263	116.4035
	Macfarlane & Hearn 4163	PERTH	-34.3609	116.7237
	Macfarlane & Tuckett 3988	PERTH	-34.3317	116.7332
	Macfarlane 4586	PERTH	-34.3481	116.7236
Trithuria australis	Annels & Hearn 5347	PERTH	-34.2731	116.6988
	Fitzgerald s.n.	NSW	-31.8676	115.9996
	Fitzgerald s.n.	B, K	-31.8635	116.0062
	Hearn s.n.	PERTH, K	-34.4043	116.7198
	Keighery G.J. & Gibson 2584	PERTH	-34.4558	116.8345
	Keighery G.J. et al 7251	PERTH	-34.9175	116.4619

Species	Collector	Herbarium	Latitude	Longitude
	Macfarlane & Annels 2283	PERTH	-34.2731	116.6988
	Macfarlane & Hearn 3357	NSW, PERTH	-34.266	116.7102
	Macfarlane & Hearn 3361	PERTH	-34.3366	116.7876
	Taylor MLT 62	TENN	-34.2686	116.705
	Taylor s.n.	TENN	-34.3358	116.7866
Trithuria bibracteata	Annels & Hearn 4626	CANB, PERTH	-34.8238	116.973
	Annels & Hearn 5350	PERTH, K	-34.1755	116.5899
	Annels 5349	PERTH	-34.3195	116.6426
	Archer 311051	PERTH	-33.8306	122.0933
	Clifford s.n.	BRI	-32.022	115.9347
	Fitzgerald s.n.	NSW	-31.8676	115.9996
	Fitzgerald s.n.	K	-31.8635	116.0062
	Gunness et al. 13/37	PERTH	-31.1412	116.8713
	Keighery B.J. & Gibson 682	PERTH	-33.72	115.4343
	Keighery B.J. & Gibson 801	PERTH	-33.7329	115.3902
	Keighery G.J. & Gibson 2798	PERTH	-34.3735	116.8678
	Keighery G.J. 10496	PERTH	-33.5154	116.8093
	Keighery G.J. 10904	PERTH	-32.0291	115.9719
	Keighery G.J. 4125	PERTH	-30.7716	115.4738
	Keighery G.J. 4280	PERTH	-34.1167	116.2055
	Keighery G.J. 6719	PERTH	-34.0535	117.6448
	Kelly 90/50	PERTH	-32.0303	115.9755
	Kelly 90/62	PERTH	-31.7535	116.0352
	Macfarlane 3479	PERTH	-34.1755	116.5899
	Macfarlane 3903	PERTH	-34.1333	116.2172
	Macfarlane and Hearn 3366	PERTH	-34.3735	116.8678
	Macfarlane and Hearn s.n.	PERTH	-34.3363	116.8378
	Staer s.n.	E	-31.8635	116.0062
	Taylor 60	TENN	-34.3147	116.5267
Trithuria cowieana	Cowie & D.J. Dixon s.n.	DNA	-12.8203	132.0128
	Cowie & S. Jacka 9995	DNA	-12.9781	130.7597
	Macfarlane et al. 4217	MW	-12.9761	130.7583
	Macfarlane et al. 4238	DNA	-12.4364	131.1339

Species	Collector	Herbarium	Latitude	Longitude
	Macfarlane et al. 4261	MW	-12.8564	131.7865
	Macfarlane et al. 4267	MW	-12.8239	132.0117
	Macfarlane et al. 4336	DNA	-12.7308	131.0378
	Michell 3863	DNA	-13.9199	132.552
Trithuria filamentosa	Briggs 9774	NSW	-41.6569	145.9581
	Briggs 9775	NSW	-41.6542	145.9638
	Brown 340	HO	-42.9046	145.8524
	Buchanan 1194b	HO	-42.6584	146.5668
	Buchanan 12328	HO	-42.01	146.272
	Buchanan 13009	HO	-42.5513	146.2877
	Buchanan 8092	HO	-43.0222	145.7363
	Buchanan 8094	HO	-43.0222	145.7363
	Buchanan 9938	HO	-41.9742	145.8448
	Burns 438	HO	-41.6542	145.9638
	Curtis s.n.	K	-42.6845	146.5928
	Curtis s.n.	HO	-41.8603	146.5456
	de Malahide & Curtis s.n.	K	-42.6845	146.5928
	de Malahide s.n.	K	-41.6655	145.965
	Field 210	TENN	-41.6655	145.965
	Gibbs 6674	BM, K	-42.1141	146.1718
	Gibbs 6809	K	-43.2386	146.7591
	Jarman s.n.	HO	-42.2943	145.8401
	Jarman s.n.	HO	-42.9333	145.7333
	Jarman s.n.	HO	-42.9	146.4
	Jarman s.n.	HO	-42.8501	145.8281
	Jarman s.n.	HO	-42.2957	145.8316
	Moscal 13940	K	-43.5335	146.0852
	Moscal 13942b	HO	-43.5335	146.0852
	Moscal 1426	HO	-41.8431	146.2794
	Moscal 1517	HO	-41.7969	146.0682
	Moscal 1522	HO	-41.7824	146.0728
	Moscal 2008	HO	-41.7387	145.9201
	Moscal 20292	HO	-42.0897	146.8359
	Moscal 2052	HO	-41.7419	145.9489
	Moscal 6881	HO	-41.7858	146.5404
	Moscal 7003	HO	-41.7993	146.5209
	Rodway 907a	HO, MEL	-42.6687	146.5853
	Rodway s.n.	HO	-42.6674	146.5865
	Rodway s.n.	HO	-42.6355	146.5866
	Sauquet 128	UBC	-42.6764	146.6181

Species	Collector	Herbarium	Latitude	Longitude
	Tyler s.n.	HO	-42.5507	146.2862
	Wells, de Winton & Clayton VM31	HO	-42.1167	146.1833
	Wells, de Winton & Clayton VM9	HO	-42.685	146.593
Trithuria inconspicua	Anonymous s.n.	CHR	-35.7727	173.6265
	Butcher s.n.	NSW	-34.9424	173.183
	Carse 1796	K	-35.0295	173.1952
	Carse 2131	K	-35.0295	173.1952
	Chapman s.n.	NSW	-34.882	173.3115
	Clayton & P.N. Johnson 1138	CHR	-45.9986	167.3808
	Clayton 1145	CHR	-45.399	167.6726
	Clayton s.n.	CHR	-46.0987	167.1457
	Cooper s.n.	K	-34.9424	173.183
	de Lange 1146	CHR	-36.3471	174.148
	de Lange 2504	CHR	-35.816	173.6595
	Forester & J. Goh s.n.	AK	-35.8139	173.6599
	Matthews s.n.	B	-35.0295	173.1952
	Pledge s.n.	CHR	-35.8046	173.6515
	Powell s.n.	CHR	-34.9503	173.1814
	Tanner s.n.	CHR	-35.7973	173.6431
	Wells s.n.	CHR	-45.5149	167.5963
Trithuria konkanensis	Yadav 1	SUK	17.01115	73.33918
	Yadav 10	SUK	13.3378	74.88028
	Yadav 2	SUK	16.48462	73.71788
	Yadav 3	SUK	16.43847	73.76258
	Yadav 4	SUK	16.08517	73.47045
	Yadav 5	SUK	15.76895	73.66538
	Yadav 6	SUK	15.6732	74.11702
	Yadav 7	SUK	15.67218	74.26835
	Yadav 8	SUK	15.45445	73.8345
	Yadav 9	SUK	14.93858	74.06662
Trithuria lanterna	Bean 16480	DNA	-20.0176	148.1091
	Dunlop 9345 & G. Leach	DNA	-13.825	136.2184
	Egan 4781 & S. Knox	DNA	-12.5978	132.9908
	Egan 4816 & S. Knox	DNA	-12.5917	132.9883

Species	Collector	Herbarium	Latitude	Longitude
	Egan 5290 & Michell	DNA	-12.4644	131.13
	Graven & G.M. Wightman 8331	DNA	-13.05	133.0833
	Harwood 370	DNA	-12.6558	131.7828
	Johnson 4555	DNA	-12.6092	133.0495
	Johnson 4597	DNA	-12.6092	133.0495
	Latz 3699	NT	-12.724	131.015
	Latz 3816	NT	-12.3167	132.6167
	Macfarlane et al. 4218	MW	-12.9761	130.7583
	Macfarlane et al. 4237	MW	-12.4361	131.1347
	Macfarlane et al. 4242	MW	-12.4439	131.1153
	Macfarlane et al. 4257	MW	-12.5983	131.2156
	Macfarlane et al. 4258	MW	-12.6978	131.3972
	Macfarlane et al. 4262	MW	-12.8239	132.0225
	Macfarlane et al. 4268	MW	-12.8239	132.0117
	Macfarlane et al. 4272	MW	-12.7325	132.3647
	Macfarlane et al. 4275	MW	-12.4333	132.9703
	Macfarlane et al. 4278	MW	-12.4764	132.9036
	Macfarlane et al. 4282	MW	-12.6439	132.8178
	Macfarlane et al. 4287	MW	-12.7372	132.7672
	Macfarlane et al. 4292	MW	-12.9036	132.6283
	Macfarlane et al. 4297	MW	-12.9403	132.5408
	Macfarlane et al. 4304	MW	-13.5656	132.2636
	Macfarlane et al. 4307	MW	-13.5667	132.2622
	Macfarlane et al. 4312	MW	-13.5797	131.8569
	Macfarlane et al. 4315	MW	-14.1867	132.0639
	Macfarlane et al. 4319	MW	-13.7428	131.7342
	Macfarlane et al. 4321	MW	-12.7263	131.016
	Macfarlane et al. 4329	MW	-12.7308	131.0389
	McKee 9231	NSW	-16.45	144.65
	M.D. Barrett 1930	PERTH	-16.3398	124.776
	M.D. Barrett 2054	PERTH	-14.8064	126.5464
	M.D. Barrett 3503b	PERTH	-14.5169	126.4444
	Michell & B. Deichmann 3400	DNA	-14.2378	132.4353
	Michell 2698	DNA	-14.1614	132.0389
	R.L. Barrett 4082	PERTH	-15.7485	125.374
	R.L. Barrett 4083	PERTH	-15.6321	125.546
	R.L. Barrett 6824	PERTH	-14.6892	125.9591
	Specht 413	MEL	-13.8597	136.8006
	Specht 566	MEL	-13.7715	136.1943
	Symon 7806	DNA	-12.7832	136.5256

Species	Collector	Herbarium	Latitude	Longitude
Trithuria submersa southeast	Beauglehole & Courtney 69473	MEL	-37.0494	143.5694
	Beauglehole & L.K.M. Elmore 71248	MEL	-38.1728	145.1735
	Beauglehole 29726	MEL	-36.7635	141.813
	Beauglehole 43325	MEL	-37.3415	140.96
	Beauglehole 5922	MEL	-37.3373	140.9835
	Beauglehole 61240	MEL	-37.2164	143.0116
	Beauglehole 66275	MEL	-36.5674	141.3358
	Beauglehole 75708A	MEL	-37.2881	142.1033
	Beauglehole 76028	MEL	-37.5492	142.417
	Beauglehole 76145	MEL	-37.2216	141.8331
	Beauglehole 76554	MEL	-37.5546	141.1346
	Beauglehole & N. Krachenbuel 19268	MEL	-37.4046	140.7305
	Clarke 2671	MEL	-37.5764	142.3166
	Corrick 4862	MEL	-36.5	141.2167
	Corrick 729	MEL	-36.6934	141.1566
	Curtis s.n.	HO, MEL	-42.0762	147.2094
	Curtis s.n.	HO	-41.9903	147.3306
	Curtis s.n.	HO, MEL	-41.9675	146.8741
	Curtis s.n.	HO	-41.814	147.4179
	Curtis s.n.	MEL	-41.8136	147.4178
	Gibbs 6809	K	-43.2386	146.7591
	Gunn s.n.	HO	-41.743	147.0689
	Hicks s.n	MEL	-36.5674	141.3358
	Hunt s.n.	NSW	-36.9833	140.2944
	Marshall s.n.	K	-36.1077	145.2355
	McBarron 5403	NSW	-35.966	146.8704
	McBarron 5688	NSW	-35.5169	147.0231
	McBarron 5932	NSW	-35.966	146.8704
	Moscal 20272	HO	-42.0858	146.8229
	Robinson s.n.	MEL	-37.8132	144.7878
	Rodway 153	HO	-41.9426	147.4741
	Rudall	ADU	-36.5971	140.9393
	Short & Corrick 6770	MEL	-36.6194	141.6495
	Smith s.n.	HO	-41.6007	147.0534
	Strudwick 0073	MEL	-36.617	145.9141
	Strudwick 604	MEL	-36.6142	145.2838
	Tepper 1247	MEL	-35.9006	137.4031
	von Mueller 20369	BM	-38.0533	142.896
	Williamson s.n.	MEL	-38.3209	142.0521

Species	Collector	Herbarium	Latitude	Longitude
Trithuria submersa south-west	Annels 5930	PERTH	-34.2669	116.5398
	Clifford s.n.	BRI	-31.9953	115.9807
	Hearn s.n.	K, PERTH	-34.405	116.7169
	Keighery B.J. & Gibson 109	PERTH	-32.6829	115.9222
	Keighery B.J. & Gibson 201	PERTH	-31.3392	115.784
	Keighery G.J. & Gibson 2396	PERTH	-34.4539	116.8333
	Keighery G.J. & Gibson 2584	PERTH	-34.4539	116.8333
	Keighery G.J. 10962	PERTH	-34.4519	116.6731
	Keighery G.J. 5667	PERTH	-33.4842	115.5875
	Keighery G.J. 6718	PERTH	-34.0594	117.6547
	Keighery G.J. et al 7251	PERTH	-34.9175	116.4619
	Macfarlane & Hearn 3358	K, PERTH	-34.2731	116.6988
	Macfarlane & Hearn 3362	K, PERTH	-34.3374	116.7891
	Macfarlane & Hearn 3374	PERTH	-34.3904	116.6547
	Macfarlane 3902	PERTH	-34.1333	116.2172
	McCallum Webster 640	PERTH	-35.018	117.7451
	R.L. Barrett 5305	PERTH	-33.5095	115.6443
	Royce 3866	PERTH	-33.5229	115.6215
	Royce 4356	PERTH	-33.5229	115.6215
	Sandiford 902	PERTH	-34.5	117.2333
	Staer s.n.	E	-31.8635	116.0062
	Taylor 61	TENN	-34.3222	116.7792
	Taylor 63	TENN	-34.2686	116.705

C.5 Accessions for seed-plant dating

List of species and GenBank accessions used in seed-plant molecular dating. Taxon, collection and herbarium (Thiers, 2013) if newly sequenced, Gen-

Bank accession numbers for *atp*B, *ndh*F, *psb*BTNH, *psb*DC, *psb*EFLJ, and *rbc*L unless otherwise noted. Newly sequenced regions are indicated by an asterisk.

***Acorus calamus* L.**, AJ235381, AY007647, AF123843, AF123813, AF-123828, D28865. ***Amborella trichopoda* Baill.**, AF235041, AF235046, AF235042, AF235043, AF235044, L12628. ***Austrobaileya scandens* C. T. White**, AF092107, AF238052, AY007460, AF239777, AY007475, L12632. ***Cabomba caroliniana* A. Gray**, AF187058, AF123801, AF123845, AF123815, AF123830, M77027. ***Cycas revoluta* Thunb.**, AF469657, AF469695, AF469707, AF462403, AF469716, AF462411. ***Ginkgo biloba* L.**, AJ235481, AF123807, AF123851, AF123821, AF123836, D10733. ***Nandina domestica* Thunb.**, complete plastid genome NC_-008336. ***Nuphar advena* Ait.**, complete plastid genome NC_008788. ***Nymphaea odorata* Ait.**, AJ235544, AF188853, AF188851, AF188854, AF188852, M77034. ***Pinus thunbergii* Parl.**, complete plastid genome D17510. ***Platanus occidentalis* L.**, complete plastid genome NC_-008335. ***Schisandra chinensis* (Turcz.) Baill.**, AF239790, AF238062, AY007470, AF239791, AY007485, AF238061. ***Spathiphyllum wallisii* Regel**, AJ235606, AY007658, AY007471, AF239794, AY007487, AJ235807. ***Trimenia moorei* (Oliv.) Philipson**, AY116653, AY116655, AY116656, AY116657, AY116652, AY116658. ****Trithuria cowieana* D. D. Sokoloff, Remizowa, T. D. Macfarl. & Rudall**, *Macfarlane et al. 4217* (MW), atpB, ndhF, psbBTNH, psbDC, psbEFLJ, rbcL. ***Trithuria filamentosa* Rodway**, FJ514801, FJ514806, FJ514802, FJ514803, FJ514804, FJ514807. ***Trithuria submersa* Hook.f.**, AJ419142, AF547020, DQ915189, EF153940, EF153946, DQ915188.

Appendix D

Voucher information for Chapter 4

List of specimens and associated GenBank accessions used in this study. Taxon; collection information, herbarium (acronyms according to Thiers [continuously updated]); GenBank accessions for *atp*B, *ndh*F, *psb*BTNH, *psb*DC, *psb*EFLJ, *rbc*L, *rpl*2, 3'*rps*12-*rps*7-*ndh*B-*trn*L(CAA). Missing regions are indicated by "N/A". Sequence data of *Vallisneria spiralis* L. (Hydrocharitaceae) were provided by Mike Moore (Oberlin College, Ohio). An asterisk "*" indicates a sequence published previously. Whole plastid genomes were sampled from GenBank for the outgroup taxa *Nandina domestica* Thunb. (Berberidaceae, DQ923117), *Nuphar advena* Aiton (Nymphaeaceae, NC_008788), *Phalaenopsis aphrodite* Rchb. f. (Orchidaceae, NC_007499) and *Platanus occidentalis* L. (Platanaceae, NC_008335). Details of previously published Alismatales and outgroup taxa can be found in Graham & Olmstead (2000), Graham et al. (2000, 2006).

Acorales. Acoraceae. *Acorus gramineus* Sol. ex W. Aiton; *G. A. Rothwell & Williams s.n.*, ALTA; HQ901511, HQ901538, HQ901404, HQ901484, HQ901457, HQ901561, HQ901430, HQ901377. Alismatales. Alismataceae. *Alisma triviale* Pursh; textitS. Y. Smith 47, ALTA; HQ901513, HQ901541, HQ901405, HQ901486, HQ901459, HQ901563, HQ901432, HQ901380. *Hydrocleys martii* Seub.; *R. A. Stockey & G. W. Rothwell 86*, no voucher, Botanischer Garten München-Nymphenburg living collection; HQ901514, HQ901542, HQ901406, HQ901487, HQ901460, HQ901564, HQ901433, HQ901381. Aponogetonaceae. *Aponogeton distachyos* L. f.; *R. A. Stockey & G. W. Rothwell 6*, no voucher, Botanischer Garten München-Nymphenburg living collection; HQ901529, HQ901552, HQ901421, HQ901502, HQ901475, HQ901579, HQ901448, HQ901395. Araceae. *Arum italicum* Mill.; *W.*

J. D. Iles 2010-001, UBC; HQ901533, HQ901556, HQ901425, HQ901506, HQ901479, HQ901583, HQ901452, HQ901399. *Gymnostachys anceps* R. Br.; *M. W. Chase 3841*, K; HQ901532, HQ901555, HQ901424, HQ901505, HQ901478, HQ901582, HQ901451, HQ901398. *Lemna trisulca* L.; *R. A. Stockey & G. W. Rothwell 82*, ALTA; HQ901530, HQ901553, HQ901422, HQ901503, HQ901476, HQ901580, HQ901449, HQ901396. *Orontium aquaticum* L.; *R. A. Stockey & G. W. Rothwell 40*, ALTA; HQ901531, HQ901554, HQ901423, HQ901504, HQ901477, HQ901581, HQ901450, HQ901397. Cymodoceaceae. *Amphibolis griffithii* (J. M. Black) Hartog; *Hopper 8539*, KPBG; HQ901524, HQ901548, HQ901416, HQ901497, HQ901470, HQ901574, HQ901443, HQ901390. *Halodule wrightii* Asch.; *D. A. Kolterman & I. López 1003*, ALTA; HQ901525, HQ901549, HQ901417, HQ901498, HQ901471, HQ901575, HQ901444, HQ901391. Hydrocharitaceae. *Elodea canadensis* Michx.; *S. Y. Smith 55*, ALTA; HQ901516, HQ901544, HQ901408, HQ901489, HQ901462, HQ901566, HQ901435, HQ901383. *Hydrocharis morsus-ranae* L.; *S. Y. Smith 51*, ALTA; HQ901517, HQ901545, HQ901409, HQ901490, HQ901463, HQ901567, HQ901436, HQ901384. *Najas flexilis* (Willd.) Rostk. & W. L. E. Schmidt; *S. Y. Smith 30*, ALTA; HQ901519, N/A, HQ901411, HQ901492, HQ901465, HQ901569, HQ901438, HQ901590. *Stratiotes aloides* L.; *Bogner s.n.*, ALTA; HQ901515, HQ901543, HQ901407, HQ901488, HQ901461, HQ901565, HQ901434, HQ901382. *Thalassia testudinum* Banks & Sol. ex K. D. Koenig; *D. A. Kolterman & I. López 1001*, ALTA; HQ901518, N/A, HQ901410, HQ901491, HQ901464, HQ901568, HQ901437, HQ901385. Juncaginaceae. *Maundia triglochinoides* F. Meull.; *L. Stanberg & G. Sainty LS 80*, NSW; HQ901527, HQ901551, HQ901419, HQ901500, HQ901473, HQ901577, HQ901446, HQ901393. *Triglochin maritima* L.; *M. Buzgo 1011*, K DNA 10463; HQ901528, AF546998*, HQ901420, HQ901501, HQ901474, HQ901578, HQ901447, HQ901394. Posidoniaceae. *Posidonia australis* Hook.f.; *M. van Keulen s. n.*, ALTA; HQ901523, N/A, HQ901415, HQ901496, HQ901469, HQ901573, HQ901442, HQ901389. Potamogetonaceae. *Groenlandia densa* (L.) Fourr.; *Bogner s. n.*, ALTA; HQ901521, HQ901546, HQ901413, HQ901494, HQ901467, HQ901571, HQ901440, HQ901387. *Zannichellia palustris* L.; *Bruinsma s.n.*, UBC; HQ901522, HQ901547, HQ901414, HQ901495, HQ901468, HQ901572, HQ901441, HQ901388. Ruppiaceae. *Ruppia maritima* L.; *D. A. Kolterman, G. J. Breckon, J. Vlez-Gaviln & A. R. Lewis 1005*, ALTA; HQ901526, HQ901550, HQ901418, HQ901499, HQ901472, HQ901576, HQ901445, HQ901392. Tofieldiaceae. *Harperocallis flava* McDaniel; *M. W. Chase 306*, NCU; HQ901536, HQ901559, HQ901428, HQ901509, HQ901482, HQ901586, HQ901455, HQ901402. *Pleea tenuifolia* Michx.; *W. Zomle-*

fer 789, GA; HQ901537, HQ901560, HQ901429, HQ901510, HQ901483, HQ901587, HQ901456, HQ901403. *Tofieldia coccinea* Richardson; *M. J. Waterway 2006-241*, UBC; HQ901535, HQ901558, HQ901427, HQ901508, HQ901481, HQ901585, HQ901454, HQ901401. *Triantha racemosa* (Walter) Small; *W. Zomlefer 801*, GA; HQ901534, HQ901557, HQ901426, HQ901507, HQ901480, HQ901584, HQ901453, HQ901400. Zosteraceae. *Zostera angustifolia* (Hornem.) Rchb.; *M. W. Chase 2795 W2*, K; HQ901520, AF547022*, HQ901412, HQ901493, HQ901466, HQ901570, HQ901439, HQ901386. Dioscoreales. Nartheciaceae. *Narthecium ossifragum* L.; *R. A. Stockey & G. A. Rothwell 59*, ALTA; AY147597, AY147763, AY147503, AY147642, AY147550, AY149348, AY147689, AY147454. Liliales. Philesiaceae. *Philesia magellanica* J. F. Gmel.; *M. W. Chase 545*, K; AY465551, AY465656, AY465578 & AY465744, AY465682, AY465605, AY465707, AY465734, AY465633. Rhipogonaceae. *Rhipogonum elseyanum* F. Muell.; *M. W. Chase 187*, NCU; AY465553, AY465658, AY465580 & AY465745 , AY465684, AY465607, AY465709, AY465736, AY465635. Poales. Bromeliaceae. *Brocchinia micrantha* (Baker) Mez; no voucher, U. Wisconsin Botany Greenhouse living collection; EU832849, EU832884, EU832899, EU832915, EU832935, EU832951, EU832964, EU832867 & EU832982.

Appendix E

Voucher information for Chapter 5

Voucher information and GenBank accession details of newly used material. **Bolded samples** are new to this thesis. Other new taxa represent GenBank downloads, or previously unpublished material. Voucher information is omitted for sequences downloaded from GenBank. Gene order is as follows: *atp*B, *ndh*B, *ndh*F, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*J, *psb*L, *psb*N, *psb*T, *rbc*L, *rpl*2, *rps*7, 3'*rps*12. Where downloaded sequences are from an entire plastid genome, only that accession number is given. Accession details on previously published material can be found in: Graham and Olmstead (2000); Graham et al. (2006); Saarela (2006); Saarela and Graham (2010); Saarela et al. (2008, 2007); Zgurski (2004). Herbarium abbreviations follow Thiers (2013).

Monocotyledonae: Alismatales: Araceae: *Spirodela polyrhiza* (L.) Schleid., NC_015891. Juncaginaceae: ***Tetroncium magellanicum*** Willd., Alvarez s.n. (CONN): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Asparagales: Asparagaceae: *Cordyline australis* (G.Forst.) Endl.: JQ273624, JQ276512, JQ276801, JQ275924, JQ275875, JQ275826, JQ275777, JQ275728, JQ275679, JQ275581, JQ275483, psbN, JQ275385, JQ273919, JQ275287, JQ274263, JQ274214. *Trichopetalum plumosum* (Ruiz & Pav.) J.F.Macbr.: JQ273625, JQ276513, JQ276802, JQ275925, JQ275876, JQ275827, JQ275778, JQ275729, JQ275680, JQ275582, JQ275484, JQ275435, JQ275386, JQ273920, JQ275288, JQ274264, JQ274215. Orchidaceae: *Apostasia wallichii* R.Br.: HQ180445, HQ180934, HQ181103, HQ181862, HQ181904, HQ181947, HQ181990, HQ182034, HQ182076, HQ182160, HQ182245, HQ182329, HQ182373, HQ182416, HQ182458,

HQ183105, HQ183231. Dioscoreales: Dioscoreaceae: *Dioscorea elephantipes* (L'Hér.) Engl., NC_009601. ***Dioscorea membranacea*** Pierre ex Prain & Burkill, M.W.Chase 21050 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. ***Tacca leontopetaloides*** (L.) Kuntze, P.Wilkin 817 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. ***Trichopus sempervirens*** (H.Perrier) Caddick & Wilkin, Wilkin et al. 948 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Nartheciaceae: ***Aletris glabra*** Bureau & Franch., M.W.Chase 517 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. ***Lophiola aurea*** Ker Gawl., Whitten 95028 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. ***Metanarthecium luteoviride*** Maxim., I.V.Tatarenko s.n. (MW): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. ***Nietneria paniculata*** Steyerm., O.Hokche & P.J.M.Maas 849 (U): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Liliales: Alstroemeriaceae: *Luzuriaga radicans* Ruiz & Pav., XXXXX (X): AY465548, AY465630, AY465653, AY465575, AY465679, AY465679, AY465602, AY465602, AY465742, AY465602, AY465602, AY465742, AY465575, AY465705, AY465731, AY465630, AY465630. Campynemataceae: *Campynema lineare* Labill., XXXXX (X): atpB, AY465629, ndhF, AY465574, AY465678, AY465678, AY465601, AY465601, AY465574, AY465601, AY465601, AY465574, AY465574, rbcL, AY465730, AY465629, AY465629. ***Campynemanthe viridiflora*** Baill., Pillon et al. 24 (NOU): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Colchicaceae: *Tripladenia cunninghamii* D.Don, XXXXX (X): AY465550, AY465632, AY465655, AY465577, AY465681, AY465681, AY465604, AY465604, AY465577, AY465604, AY465604, AY465577, AY465577, rbcL, AY465733, AY465632, AY465632. Liliaceae: *Calochortus apiculatus* Baker, XXXXX (X): AY465547, AY465628, AY465652, AY465573, AY465677, AY465677, AY465600, AY465600, AY465573, AY465600, AY465600, AY465573, AY465573, AY465704, AY465729, AY465628, AY465628. *Medeola virginiana* L., XXXXX (X): AY465549, AY465631, AY465654, AY465576, AY465680, AY465680, AY465603, AY465603, AY465576, AY465603, AY465603, AY465576, AY465576, AY465706, AY465732, AY465631, AY465631. *Prosartes trachycarpa* S.Watson, XXXXX (X): AY465552, AY465634, AY465657, AY465579, AY465683, AY465683, AY465606,

AY465606, AY465579, AY465606, AY465606, AY465579, AY465579, AY465708, AY465735, AY465634, AY465634. *Tricyrtis* sp., Waterway s.n. (X): AY465555, AY465637, AY465660, AY465582, AY465686, AY465686, AY465609, AY465609, AY465582, AY465609, AY465609, AY465582, AY465582, AY465711, AY465738, AY465637, AY465637. Melanthiaceae: ***Xerophyllum tenax*** (Pursh) Nutt., J.M.Saarela 181 (UBC): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Pandanales: Cyclanthaceae: ***Cyclanthus bipartitus*** Poit. ex A.Rich., M.W.Chase 1237 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Pandanaceae: ***Freycinetia*** sp., XXXX (X): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Stemonaceae: ***Pentastemona sumatrana*** Steenis, XXXXX (X): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Velloziaceae: ***Acanthochlamys bracteata*** P.C.Kao, P.C.Kao 1993 (K): atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbJ, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'rps12. Petrosaviales: Petrosaviaceae: *Petrosavia* sp., S.W.Graham 2004 (UBC): NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, AY465690, AY465715, AY465613, AY465613. Poales: Anarthriaceae: *Anarthria scabra* R.Br., Briggs 9581 (NSW): NA, EU832864, EU832882, NA, EU832928, EU832928, EU832932, EU832932, NA, EU832932, EU832932, NA, NA, NA, EU832961, EU832980, EU832980. *Hopkinsia anoectocolea* (F.Muell.) D.F.Cutler, Briggs s.n. (NSW): NA, EU832865, EU832845, EU832898, EU832914, EU832914, EU832933, EU832933, EU832898, EU832933, EU832933, EU832898, EU832898, NA, EU832962, EU832981, NA. *Lyginia imberbis* R.Br., Briggs 98477 (NSW): NA, EU832866, EU832883, NA, NA, NA, EU832934, EU832934, NA, EU832934, EU832934, NA, NA, NA, EU832963, rps7, 3'rps12. Cyperaceae: *Carex cordillerana* Saarela & B.A.Ford, Saarela 196 (ALTA): EU832850, EU832868, EU832885, EU832900, EU832916, EU832916, EU832936, EU832936, EU832900, EU832936, EU832936, EU832900, EU832900, NA, EU832965, NA, NA. *Eleocharis palustris* (L.) Roem. & Schult., Saarela 258 (CAN): EU832851, EU832869, EU832886, EU832901, EU832917, EU832917, EU832937, EU832937, EU832901, EU832937, EU832937, EU832901, EU832901, EU832952, EU832966, EU832844, EU832844. *Gahnia baniensis* Benl, Simpson s.n. (K): EU832852, EU832870, EU832887, EU832902, EU832918, EU832918, EU832938, EU832938, EU832902, EU832938, EU832938, EU832902, EU832902, EU832953, EU832967, EU832983, EU832983. *Mapania* cf. *pubisquama*, Walters et al. 563 (MO): EU832853, EU832871,

ndhF, EU832903, EU832919, EU832919, EU832939, EU832939, EU832903, EU832939, EU832939, EU832903, EU832903, rbcL, EU832968, rps7, EU832847. Eriocaulaceae: *Eriocaulon compressum* Lam., Unwin 241 (MU): EU832854, EU832873, EU832890, EU832905, EU832921, EU832-921, EU832941, EU832941, EU832905, EU832941, EU832941, EU832905, EU832905, EU832954, EU832970, EU832985, EU832985. Joinvilleaceae: *Joinvillea ascendens* Gaudich. ex Brongn. & Gris, Weston 2501 (NSW): EU832855, EU832874, EU832891, EU832906, EU832922, EU832922, EU832942, EU832942, EU832906, EU832942, EU832942, EU832906, EU832906, EU832955, EU832971, EU832986, EU832986. Juncaceae: *Juncus effusus* L., Rai 1004 (ALTA): EU832856, EU832875, EU832892, EU832907, EU832923, EU832923, EU832943, EU832943, EU832907, EU832943, EU832943, EU832907, EU832907, EU832956, EU832972, EU832987, EU832987. *Luzula parviflora* (Ehrh.) Desv., Peterson et al. 18634 (US): EU832857, EU832876, NA, EU832908, EU832924, EU832924, EU832944, EU832944, EU832908, EU832944, EU832944, EU832908, EU832908, EU832957, EU832973, EU832988, EU832988. Poaceae: *Oryza rufipogon* Griff.: NC_005973. *Puelia olyriformis* (Franch.) Clayton, Bradley et al. 1060 (MO): EU832858, EU832877, EU832893, EU832909, EU832925, EU832925, EU832945, EU832945, EU832909, EU832945, EU832945, EU832909, EU832909, NA, EU832974, EU832989, EU832989. Rapateaceae: *Rapatea* sp., M.W.Chase 195 (K): EU832859, EU832878, EU832894, EU832910, EU832929, EU832929, EU832946, EU832946, NA, EU832946, EU832946, NA, NA, EU832958, EU832975, EU832990, EU832990. *Stegolepis* sp., Kubitzki et al. 97-30 (HBG): EU832860, EU832879, EU832895, EU832995, EU832930, EU832930, EU832947, EU832947, EU832995, EU832947, EU832947, EU832995, EU832995, NA, EU832976, EU832991, EU832991. Thurniaceae: *Prionium serratum* (L.f.) Drège, National Botanic Garden of Belgium living collection (accession 19880003): EU832861, EU832880, EU832896, EU832911, EU832926, EU832926, EU832948, EU832948, EU832911, EU832948, EU832948, EU832911, EU832911, EU832959, EU832977, EU832992, EU832992. *Thurnia sphaerocephala* (Rudge) Hook.f., Kelloff et al. 1335 (US): EU832862, EU832843, NA, EU832912, EU832931, EU832931, EU832949, EU832949, EU832912, EU832949, EU832949, EU832912, EU832912, NA, EU832978, EU832993, EU832993. Xyridaceae: *Aratitiyopea lopezii* (L.B.Sm.) Steyerm. & P.E.Berry, van der Werff et al. 16131 (MO): EU832863, EU832881, EU832897, EU832913, EU832927, EU832927, EU832950, EU832950, EU832913, EU832950, EU832950, EU832913, EU832913, EU832960, EU832979, EU832994, EU832994. Zingiberales: Zingiberaceae: *Renealmia*

alpinia (Rottb.) Maas: HQ180478, HQ180968, HQ181133, HQ181895, NA, NA, HQ182024, HQ182067, HQ182109, HQ182193, HQ182278, HQ182363, HQ182406, HQ182449, HQ182491, HQ183139, HQ183258. Eudicotyledonae: Asterales: Asteraceae: *Lactuca sativa* L.: NC_007578. Buxales: Buxaceae: *Buxus microphylla* Siebold & Zucc.: NC_009599. Dilleniales: Dilleniaceae: *Dillenia indica* L.: GQ997134, GQ997145, GQ997149, GQ997167, GQ997168, GQ997169, GQ997170, GQ997171, GQ997172, GQ997174, GQ997176, GQ997178, GQ997179, GQ997181, GQ997184, GQ997205, GQ997196. Fabales: Fabaceae: *Glycine max* (L.) Merr.: XXXX. Malphigiales: Salicaceae: *Populus* XXXX. Malvales: Malvaceae: *Gossypium* XXXX. Myrtales: Myrtaceae: Eucalyptus XXXXX. Proteales: Nelumbonaceae: *Nelumbo nucifera* Gaertn.: GQ997549, GQ997560, GQ997564, GQ997582, GQ997583, GQ997584, GQ997585, GQ997586, GQ997587, GQ997589, GQ997591, GQ997593, GQ997594, GQ997596, GQ997599, GQ997620, GQ997611. Sabiales: Sabiaceae: *Meliosma* aff. *cuneifolia*: GQ997466, GQ997477, GQ997481, GQ997499, GQ997500, GQ997501, GQ997502, GQ997503, GQ997504, GQ997506, GQ997508, GQ997510, GQ997511, GQ997513, GQ997516, GQ997537, GQ997528. Solanales: Solanaceae: *Atropa belladonna* L.: NC_004561. *Solanum lycopersicum* L.: NC_007898. *Solanum tuberosum* L.: NC_008096