DR SEAN W GRAHAM (Orcid ID : 0000-0001-8209-5231)

Article type : Article

A bi-organellar phylogenomic study of Pandanales: inference of higher-order relationships and unusual rate-variation patterns

Marybel Soto Gomez[1,2], Qianshi Lin[1,2], Eduardo da Silva Leal[3], Timothy Gallaher[4], David Scherberich[5], Constantijn B. Mennes[6], Selena Y. Smith[7], Sean W. Graham[1,2,*]

[1] *Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia, V6T 1Z4, Canada*

[2] *UBC Botanical Garden & Centre for Plant Research, University of British Columbia, 6804 Marine Drive SW, Vancouver, British Columbia, V6T 1Z4*

[3] *Universidade Federal Rural da Amazônia, Campus Capanema, Rua João Pessoa, 121, Capanema, PA CEP 68700-030, Brazil*

[4] *Department of Biology, University of Washington, Box 351800, Seattle, Washington, 98195-1800*

[5] *Jardin Botanique de la Ville de Lyon, Mairie de Lyon, 69205 Lyon Cedex 01, France*

[6] *Naturalis Biodiversity Center, Vondellaan 55, 2332 AA, Leiden, The Netherlands*

[7] *Department of Earth & Environmental Sciences and Museum of Paleontology, Universit of Michigan, Ann Arbor, Michigan, 48109, USA*

***Corresponding author**: *Sean W. Graham, Department of Botany, University of British Columbia, Vancouver, Canada, phone +1 604 822 4816, e-mail: swgraham@interchange.ubc.ca*

Short running title: Bi-organellar phylogenomics of Pandanales

ABSTRACT

We used a bi-organellar phylogenomic approach to address higher-order relationships in Pandanales, including the first molecular phylogenetic study of the panama-hat family, Cyclanthaceae. Our genus-level study of plastid and mitochondrial gene sets includes a comprehensive sampling of photosynthetic lineages across the order, and provides a framework for investigating clade ages, biogeographic hypotheses and organellar molecular evolution. Using multiple inference methods and both genomes, we recovered mostly congruent and strongly supported relationships within and between families, including the placement of fully mycoheterotrophic Triuridaceae. Cyclanthaceae and Pandanaceae plastomes have slow substitution rates, contributing to weakly supported plastid-based relationships in Cyclanthaceae. While generally slowly evolving, mitochondrial genomes exhibit sporadic rate elevation across the order. However, we infer well-supported relationships even for slower evolving mitochondrial lineages in Cyclanthaceae. Clade age estimates across photosynthetic lineages are largely consistent with previous studies, are well correlated between the two organellar genomes (with slightly younger inferences from mitochondrial data), and support several biogeographic hypotheses. We show that rapidly evolving non-photosynthetic lineages may bias age estimates upwards at neighbouring photosynthetic nodes, even using a relaxed clock model. Finally, we uncovered new genome structural variants in photosynthetic taxa at plastid inverted repeat boundaries that show promise as interfamilial phylogenetic markers.

Keywords: Arborescent monocots; comparative genome evolution; monocot phylogenomics; organellar phylogenomics; substitution rate heterogeneity; substitution rate shifts; Stemonaceae; Triuridaceae; Velloziaceae

TABLE OF CONTENTS

INTRODUCTION

The modern circumscription of the monocot order Pandanales (Angiosperm Phylogeny Group, APG, 2003, 2009, 2016) was an early surprise of plant molecular systematics, as its component families (Cyclanthaceae, Pandanaceae, Stemonaceae, Triuridaceae and Velloziaceae) had been aligned with a very diverse array of higher-order taxa in previous classification schemes (reviewed in Dahlgren and Clifford, 1982). The modern conception of the order was developed primarily using molecular systematic data sets comprising two plastid and one nuclear loci (e.g., Chase et al., 1993, 1995, 2000). These molecular systematic studies culminated in the finding that the non-photosynthetic family Triuridaceae belongs in Pandanales, based on analyses that incorporated sequence data of the nuclear 18S rDNA locus from *Sciaphila*, a genus of Triuridaceae, in a broad study of monocot relationships (Chase et al., 2000). The family Triuridaceae is a fully mycoheterotrophic clade that lacks photosynthetic genes such as *rbc*L and *atp*B (Lam et al., 2015; Petersen et al., 2018), two of the three loci that broadly underpin the APG classification systems. The assignment of Triuridaceae to Pandanales has since been supported in diverse studies using additional sources of data, including morphology, mitochondrial genes, a subset of non-photosynthetic plastid genes, and whole plastid genomes (Rudall and Bateman, 2006; Mennes et al., 2013; Lam et al., 2015, 2016, 2018; Givnish et al., 2018).

Classification schemes based on early morphology-based cladistic inferences by Dahlgren and Rasmussen (1983) have been highly influential in monocot classification, but are also at odds with current schemes for Pandanales (e.g., APG, 2016). Dahlgren and Clifford (1982) and Dahlgren et al. (1985) also dispersed component families across multiple orders in their influential monographs on monocot classification. Specifically,

they inferred that Cyclanthaceae and Pandanaceae are closely related to the palm order Arecales, Stemonaceae to the yam order Dioscoreales, Velloziaceae to bromeliads (now part of the grass order, Poales), and also tentatively aligned mycoheterotrophic Triuridaceae with taxa now recognized as Alismatales (Dahlgren and Clifford, 1982; Dahlgren et al., 1985). These discrepancies between morphological and molecular phylogenetic inferences may reflect the remarkable diversity of growth forms in the order, in addition to difficulties in interpreting reproductive structures in higher-order comparisons. Pandanales are a relatively small order, with some 1300 species in ~34 genera, but its growth forms include trees, shrubs, woody and herbaceous climbers, and photosynthetic and mycoheterotrophic herbs (Eriksson, 1994; Harling et al., 1998; Kao and Kubitzki, 1998; Kubitzki, 1998a,b,c; Stone et al., 1998; Rudall et al., 2005; Mabberley, 2008; Mello-Silva et al., 2011; Gallaher et al., 2015). Flowers of Cyclanthaceae and Pandanaceae are often densely aggregated and can be highly reduced, making them harder to compare to other monocot flowers (e.g., Dahlgren et al., 1985; Harling et al., 1998; Stone et al., 1998; Rudall and Bateman, 2006; Sajo et al., 2014), those of Stemonaceae deviate substantially from 'canonical' trimerous monocot flowers in being dimerous or pentamerous (Dahlgren et al., 1985; Rudall et al., 2005), and those of Triuridaceae appear to blur boundaries between flowers and inflorescences (e.g., Rudall and Bateman, 2006).

Higher-order relationships in Pandanales

Relationships inferred among the five component families of Pandanales have been examined in pre-phylogenomic studies that employed various combinations of plastid loci (*atp*B, *acc*D, *clp*P, *mat*K, *rbc*L), mitochondrial regions (*atp*A, *mat*R, *nad*1b-c intron), the nuclear 18S rDNA locus, and updated interpretations of morphology (Chase et al., 2000; Caddick et al., 2002; Davis et al., 2004; Rudall and Bateman, 2006; Mennes et al., 2013; Lam et al., 2016). Setting aside Triuridaceae, multiple studies have recovered a close relationship between Cyclanthaceae and Pandanaceae, with Stemonaceae as their sister group, and Velloziaceae as the sister group of the other families in the order (Caddick et al., 2002; Davis et al., 2004; Graham et al., 2006; Chase et al., 2006; Mennes

et al., 2013; Lam et al., 2016). However, the mycoheterotrophic family Triuridaceae has drifted to different higher-order positions when included in phylogenetic analyses. For example, various molecular studies based on mitochondrial and nuclear markers recovered Triuridaceae as the sister group of Pandanaceae (Chase et al., 2000, Caddick et al., 2002), the sister group of Velloziaceae (Davis et al., 2004), and the sister group of a clade comprising Stemonaceae, Cyclanthaceae and Pandanaceae (Davis et al., 2004; Mennes et al., 2013); a morphological analysis inferred it to be a lineage nested within Stemonaceae (Rudall and Bateman, 2006). Mennes et al. (2013) subsequently inferred strong support for the respective monophyly of Triuridaceae and Stemonaceae in phylogenetic estimates that included several mitochondrial and nuclear genes, and a broad sampling of taxa in both families.

More recently, plastid data have become available for Triuridaceae, which were once predicted to lack plastid genomes (Cronquist, 1988; Merckx et al., 2009). Lam et al. (2016) retrieved up to three retained plastid genes (*acc*D, *clp*P, *mat*K) for four members of *Sciaphila* using Sanger sequencing, and placed Triuridaceae as the sister group of Cyclanthaceae and Pandanaceae in a monocot-wide phylogenetic analysis, although with very weak support for this arrangement. Plastid phylogenomic studies used all 18 non-photosynthetic protein-coding genes and four rDNA genes retained in the plastid genome of a representative of Triuridaceae (*Sciaphila densiflora*) and added those sequences to full gene sets representing the unreduced plastid genomes of photosynthetic monocots (Lam et al., 2015, 2018; Givnish et al. 2018). Parsimony analyses of these phylogenomic data either yielded weakly supported placements of Triuridaceae as the sister group of the Velloziaceae genus *Xerophyta* in a study that examined a subset of genera in Pandanales (Supplementary Fig. S7 in Lam et al., 2015), or inferred it to be part of a bizarre clade comprising other rapidly evolving mycoheterotrophic lineages belonging to multiple monocot orders (i.e., members of Burmanniaceae, Corsiaceae, Thismiaceae and Orchidaceae) in a broader angiosperm-wide analysis (Fig. 5 in Lam et al., 2018). The latter result conflicts substantially with our current understanding of monocot phylogeny, suggesting that both analyses were misled by long-branch artefacts concerning Triuridaceae (e.g., Felsenstein, 1978; Hendy and Penny, 1989). In contrast, likelihood analyses considering DNA, amino-acid and codon-based substitution models all placed

Triuridaceae as the sister-group of Cyclanthaceae-Pandanaceae, with generally strong support (Lam et al., 2015, 2018; Givnish et al., 2018). However, plastid genome (plastome)-based studies have not yet sampled broadly across the photosynthetic genera in the order, and no molecular studies of relationships within Cyclanthaceae have been conducted to date. Moreover, a broad phylogenomic study of Pandanales that includes full mitochondrial gene sets has yet to be performed.

Divergence time and molecular rate estimation

Substitutional rate shifts related to changes in life history have been postulated for clades that include a mixture of tree-like vs. herbaceous taxa (e.g., for the monocot orders Arecales vs. Poales, respectively; Bousquet et al., 1992; Gaut et al., 1996; Barrett et al., 2016; De la Torre et al., 2017). This mix of growth forms is also found in Pandanales, but previous studies have not addressed in detail whether the order also exhibits similar rate shifts. Order-wide divergence times are also of interest as current age estimates for Pandanales have focused either on specific families (e.g., Pandanaceae: Gallaher et al., 2015; Velloziaceae: Mello-Silva et al., 2011, Alcantara et al., 2018; Triuridaceae: Mennes et al., 2013), or were part of broader monocot- or angiosperm-wide inferences (e.g., Janssen and Bremer, 2004; Bell et al., 2010; Hertweck et al., 2015; Magallón et al., 2015; Eguchi and Tamura, 2016, Foster et al., 2017, Givnish et al., 2018). The most thorough molecular dating analysis of Pandanales to date included multiple taxa of Triuridaceae, which are all fully mycoheterotrophic/non-photosynthetic (Mennes et al., 2013). Elevated evolutionary rates in all three genomic compartments of mycoheterotrophs (e.g., Merckx et al., 2009; Lemaire et al., 2011; Lam et al., 2018) may make it more challenging to incorporate these taxa in molecular dating analyses, as postulated by Iles et al. (2015). However, this hypothesis, and (more broadly) order-wide age estimates using data derived from both organelles remain to be addressed in detail.

Genome structural evolution

Plant mitochondrial genomes are challenging to assemble because of their rapid structural evolution and substantial structural variation within species, a contrast with their very slow evolution at the DNA substitutional level (reviewed in Gualberto and Newton, 2017). A substantial difference between mitochondrial and plastid genomes is that the latter are generally highly conserved structurally (e.g., Palmer, 1991; Raubeson and Jansen, 2005; Wicke et al., 2011). Rare plastome structural changes such as inversions, gene and intron loss, and shifting boundaries of the inverted repeat (IR) region, may therefore have utility as phylogenetic markers of major clades (e.g., Raubeson and Jansen, 1992; Kim and Jansen, 1994; Plunkett and Downie, 2000; Graham et al., 2000; McPherson et al., 2004; Wang et al., 2008; Davis and Soreng, 2010; Steele et al., 2012; Guisinger et al., 2010; Harris et al., 2013; Zhu et al., 2016). The increasing availability of full plastid genomes has further facilitated comparative studies on plastid genome structural evolution (e.g., Guisinger et al., 2011; Downie and Jansen, 2015; McKain et al., 2016). While comparative plastid genome evolution studies in Pandanales have been conducted within Stemonaceae (Lu et al., 2018), and for mycoheterotrophic *Sciaphila densiflora* and *S. thaidanica* (Triuridaceae) relative to each other and to photosynthetic *Carludovica palmata* (Cyclanthaceae; Lam et al., 2015; Petersen et al., 2018), a comparative family-level study has yet to be performed across Pandanales as a whole.

Study goals

We assembled organellar phylogenomic data sets comprising plastid and mitochondrial genes sets for all currently recognized photosynthetic genera in Pandanales. The non-photosynthetic family Triuridaceae was represented by one or two species from *Sciaphila,* the only genus in this family for which plastid genes have been recovered to date. We use these data sets to: (i) infer genus-level relationships of photosynthetic taxa in Pandanales, including the first molecular study of higher-order relationships in the panama-hat family Cyclanthaceae; (ii) investigate clade ages, consequences of these estimates for biogeographic hypotheses, and address possible biases in molecular dating analyses due to inclusion of substantially rate-elevated mycoheterotrophic taxa; (iii) characterize substitution rate shifts across photosynthetic members of Pandanales for both

organellar genomes; and (iv) investigate the evolution of overall plastid genome structure by comparing several completed genomes for each family.

MATERIALS AND METHODS

Taxon sampling and library preparation

We used genome skims to generate organellar genes sets for phylogenetic analyses by adding new plastid and mitochondrial data for 18 and 28 species, respectively, to published sequences in final matrices that included 28–29 Pandanales taxa for mitochondrial and plastid data sets, respectively (Table S1, S2). All 26 currently recognized genera in the four photosynthetic families of Pandanales (Cyclanthaceae, Pandanaceae, Stemonaceae, Velloziaceae) are included, with the non-photosynthetic family Triuridaceae represented by a single genus, *Sciaphila* (one or two species). For plastid data we added newly sequenced species to a previously published matrix (Lam et al., 2018) that included data from 10 Pandanales taxa; we also retrieved sequences for *Sciaphila thaidanica* (Triuridaceae) from GenBank (MG757197.1), yielding a 96-taxon plastid matrix. This includes 85 taxa from all 12 monocot orders (sensu Givnish et al., 2018), and 11 taxa representing outgroups from eudicots, magnoliids, and the orders Amborellales, Nymphaeales and Austrobaileyales. For the mitochondrial matrix, we generated data from 28 Pandanales taxa and seven additional monocots. The resulting 37-taxon matrix includes representatives from six monocot orders and a eudicot outgroup (*Arabidopsis thaliana;* GenBank accession NC_037304) (Table S1, S2), but excludes *Sciaphila thaidanica* as mitochondrial data were not available for this taxon.

Genome skims for Pandanales taxa were generated following approaches outlined in Lam et al. (2015). Briefly, we prepared genomic libraries using two additional library preparation kits (Bioo NEXTflex Rapid DNA Sequencing Kit, Bioo Scientific Corp., Austin, USA and NEBNext Ultra II DNA Library Prep Kit, New England Biolabs, Ipswich, USA) and size-selected all libraries using magnetic beads. We sequenced 17 libraries as 100-bp or 125-bp paired-end reads respectively on multiplexed lanes of HiSeq 2000 and HiSeq 2500 platforms (Illumina, San Diego, USA) that included 20–41

samples. We also sequenced a library for one taxon (*Barbacenia involucrata*) as 300-bp reads on a MiSeq platform using the v3 kit in a multiplexed lane that included four samples. The genome skims used to obtain mitochondrial sequences for additional monocot taxa were produced using methods outlined in Ross et al. (2015; Alismatales), Givnish et al. (2016; Liliales) and Lam et al. (2018; Asparagales, Petrosaviales).

Contig assembly, plastid gene annotation and plastome reconstruction

We assembled full circularized plastome maps for *Acanthochlamys bracteata* (Velloziaceae), *Benstonea copelandii* (Pandanaceae) and *Croomia japonica* (Stemonaceae) following procedures described in Lam et al. (2015). We used Sanger sequencing to bridge gaps between contigs and to verify contig overlap, employing primers designed using Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012). We also retrieved the published complete plastome maps of five Stemonaceae taxa (*Croomia japonica, C. heterosepala, C. pauciflora, Stemona japonica, S. mairei*; Lu et al., 2018), one Cyclanthaceae taxon (*Carludovica palmata*; Lam et al., 2015) and two mycoheterotrophic Triuridaceae taxa (*Sciaphila densiflora*; Lam et al., 2015 and *S. thaidanica*; Petersen et al., 2018), and the sequences from the junction of the inverted repeat (IR) and large single copy (LSC) region of one Pandanaceae taxon (*Pandanus odoratissimus*; Wang et al., 2008). These allow for comparative examination of plastome structure for at least one representative from each of the five families in Pandanales. We therefore considered a total of 11 entire plastomes and one set of IR-LSC junction sequences; note that the five Stemonaceae taxa represented by plastome maps were not included in organellar gene sets, as mitochondrial data were not available for them, and all genera in the family were already represented for both organelles. We used approaches described in Lam et al. (2015) to assemble plastid contigs and retrieve plastid gene sets, with exceptions noted here. We selected contigs with at least 30x average coverage, and used *Carludovica palmata* (Pandanales; NC_026786.1) and *Dioscorea elephantipes* (Dioscoreales; NC_009601.1) as reference taxa for manually adjusting plastid gene start/stop codons and exon boundaries. We selected contigs with at least 10x average coverage for assembling mitochondrial contigs and retrieved mitochondrial

genes using the blastn program from BLAST+ NCBI (version NCBI-BLAST-2.2.30+, Camacho et al., 2009), using *Arabidopsis thaliana* as a query.

Sequence alignment and data matrix construction

We generated 82 individual plastid gene files, representing 78 protein-coding genes and four rDNA genes, each with 96 taxon terminals; mycoheterotrophic taxa *Sciaphila thaidanica* and *S. densiflora* were respectively represented by 14 and 18 protein-coding genes and four rDNA genes retained in their plastid genomes (Lam et al., 2015; Petersen et al., 2018). We retrieved plastid gene *ycf*1 for most photosynthetic taxa, but included this gene only in an additional Pandanales-focused matrix comprising 83 genes from 30 taxa (see below), as we had difficulties aligning it across the broader taxon set. For the mitochondrial data we generated 37 gene files, each representing a protein-coding gene and 37 taxon terminals. Missing genes (i.e., those lost in mycoheterotrophic *Sciaphila* and a few that were not recovered for individual photosynthetic taxa) were coded as missing data. We followed Lam et al. (2015) in constructing plastid and mitochondrial data matrices, with the following modifications: we generated initial DNA-based alignments of plastid genes using MAFFT v. 7.407 (Katoh et al., 2002), implemented in AliView v. 1.18.1 (Larsson, 2014), using the localpair flag to incorporate local pairwise alignment information and 1000 iterative refinement cycles (maxiterate 1000 flag). We then concatenated individual plastid gene files using SequenceMatrix v. 100.0 (Vaidya et al., 2011). For mitochondrial genes, we conducted initial DNA-based alignments using the MUSCLE (Edgar, 2004a,b) online portal with default settings and concatenated individual gene files manually. We used AliView and Mesquite v. 3.04 (Maddison and Maddison, 2015) to perform alignment adjustments on separate plastid and mitochondrial matrices, respectively, using criteria laid out in Graham et al. (2000), Kelchner (2000), and Simmons and Ochoterena (2000). We staggered regions that were difficult to align globally, resulting in offset blocks of locally aligned sequences, which should minimize errors due to misalignment while retaining potentially informative variation (e.g., Steane et al., 1999; Graham et al., 2006; Bell et al., 2020). We verified open reading frames for all protein-coding genes, and checked for matrix compilation errors as described in Lam

et al. (2015). We obtained a 80,183-bp plastid matrix and a 36,807-bp mitochondrial matrix (for reference, these are derived from 68,668 bp and 21,143 bp of unaligned plastid and mitochondrial data in *Acanthochlamys bracteata,* respectively), and translated both into corresponding amino-acid matrices (24,839 and 12,269 residues, respectively). We also assembled a 119-gene combined organellar matrix (i.e., a plastid and mitochondrial dataset) that included all sampled Pandanales (excluding *S. thaidanica*, for which only plastid data were available) and five outgroup taxa. All datasets are available at figshare (https://figshare.com/s/c57739fd6fec7f6ea958).

Phylogenetic inference

We analyzed data matrices using PAUP* v. 4.0a build 164 (Swofford, 2002) for parsimony analyses, and RAxML v7.4.2 (Stamatakis, 2006) under a graphical interface (Silvestro and Michalak, 2012) for likelihood analyses. For parsimony, we conducted separate analyses of the plastid data set (95-taxon version; see below for additional versions of this data set), the mitochondrial data set (37-taxon matrix) and a combined organellar data set (33-taxon matrix). In all three cases we represented Triuridaceae only by *Sciaphila densiflora*, and excluded *Sciaphila thaidanica* (Triuridaceae); the latter is an extremely rapidly evolving species that was not available for the mitochondrial data set, but was included in one plastid-based likelihood analysis (see below). For each matrix, we ran heuristic searches using tree bisection and reconnection (TBR) branch swapping, considering 1000 random stepwise addition replicates, holding up to 100 trees at each step, and otherwise used default settings.

For likelihood analyses, we considered unpartitioned and partitioned analyses of DNA and amino-acid data. We partitioned DNA alignments by gene and codon position (a 'G x C' scheme, e.g., Lam et al., 2015), with additional partitions included for each rDNA gene, starting with 238 initial partitions for the plastid DNA matrix, 111 for the mitochondrial DNA matrix and 349 for the combined organellar matrix. We started with 78 and 37 initial partitions for translated plastid and mitochondrial matrices, respectively, considering one partition per gene for the respective organellar genomes. We used PartitionFinder 2 (Lanfear et al., 2017) to combine partitions that did not have

substantially different substitution models, limiting the latter to models implemented in RAxML, and employing the relaxed hierarchical clustering algorithm (r-cluster) and corrected Akaike Information Criterion (AICc) (see Table S3 for final partitions and models). The best fit DNA substitution models were consistently GTR, GTR+G or GTR+I+G; we used the GTR+G model for all partitions in subsequent analyses, as the invariant sites parameter 'I' may be accommodated by the gamma parameter 'G' (Yang, 2006). We obtained a range of optimal amino-acid substitution models and applied them to corresponding final partitions (Table S3). We estimated per-partition base frequencies and among- and within-site rate parameters independently, while linking branch length estimates across partitions. We also used PartitionFinder 2 to find optimal substitution models for unpartitioned DNA- and amino-acid (AA)-based matrices (Table S3).

We ran 12 likelihood analyses in total. Eight of these were based on one core version of each of the plastid and mitochondrial data sets (i.e., a 95-taxon plastid matrix and a 37-taxon mitochondrial matrix, see below); they correspond to partitioned plus unpartitioned analyses of both DNA and AA versions of each matrix, for a total of four analyses per organelle. For the plastid data we focused most efforts on the 95-taxon matrix as a common reference point for examining the effect of different likelihood approaches. Both core organellar matrices included a single mycoheterotrophic representative, *Sciaphila densiflora*. Three additional partitioned likelihood analyses of plastid data were performed to assess: (i) the effect of including both representatives of mycoheterotrophic *Sciaphila* (in an analysis of a 96-taxon plastid DNA matrix that includes *S. densiflora* and *S. thaidanica*; both are on long branches but the branch for the latter taxon was exceptionally long, see Results and Fig. S7); (ii) the effect of excluding both long-branch *Sciaphila* taxa altogether (in an analysis of a 94-taxon plastid DNA matrix excluding *S. densiflora* and *S. thaidanica*); (iii) the effect of including *ycf*1, a highly variable plastid gene that we were able to align within Pandanales and two outgroups (in a DNA-based analysis of 83 plastid genes for 30 taxa that included one mycoheterotrophic representative, *S. densiflora*). Finally, we ran a combined DNA-based analysis of plastid and mitochondrial data; this was a partitioned analysis that included the 28 Pandanales taxa common to both organellar data sets and five outgroups. For each likelihood analysis, we conducted 20 independent searches for the best tree.

We estimated branch support for parsimony and likelihood analyses using bootstrap analysis (Felsenstein, 1985). We considered poorly and strongly supported branches to have <70% or ≥95% bootstrap support, respectively (e.g., Soltis and Soltis, 2003). For parsimony analyses, we estimated branch support using 500 bootstrap replicates (Felsenstein, 1985), each with 100 random stepwise addition replicates and TBR branch swapping, and otherwise using default settings. For likelihood analyses we used 500 rapid bootstrap replicates, considering GTRGAMMA and PROTGAMMA approximations for the nucleotide and amino-acid substitution models, respectively; each bootstrap analysis used the same partitioning schemes as the searches for best trees.

Divergence time and substitution rate estimation

We used the plastid and mitochondrial DNA-based data sets generated here to estimate divergence times and substitution rate shifts in Pandanales and related monocots using a Bayesian framework in BEAST v. 2.4.8 (Bouckaert et al., 2014). Each matrix excluded two long-branch taxa, *Sciaphila densiflora* and *S. thaidanica*, as their inclusion may distort age estimation (see below). The plastid matrix also excluded seven magnoliid, Amborellales, Nymphaeales and Austrobaileyales taxa, as their inclusion caused analyses to terminate prematurely; we otherwise included 87 and 36 taxa for plastid and mitochondrial data, respectively. We constructed a three-gene plastid matrix (*acc*D, *mat*K, *rps*3) and a 12-gene mitochondrial matrix (*atp*1*, atp*8, *ccm*C*, ccmFc, ccmFn, cox*1*, mat*R, *mtt*B*, nad*5, *rpl*2*, rps*3, *rps*4). Both data sets contain approximately 2000 parsimony informative sites for their sampled taxa; we did not include the full concatenated matrix in either case because previous studies have shown that adding genes tends not to significantly improve age estimates (e.g., Mulcahy et al., 2012; Foster et al., 2017); trial analyses using larger gene sets also proved impractical with available computational resources. We used a random local clock (RLC) model (Drummond and Suchard, 2010) to estimate divergence times and the number of molecular rate shifts. We ran plastid analyses using unpartitioned data, or with the data partitioned by codon position. The age and rate shift estimates were not substantially different across these two variant analyses, and the attempted partitioned analyses did not reach stationarity for

several parameters (data not shown). We ran only unpartitioned analyses for the mitochondrial data. In both cases we constrained the full tree topology (e.g., Givnish et al., 2016, 2018), considering the trees resulting from partitioned analyses of the 95-taxon plastid matrix and the partitioned analyses of the 37-taxon mitochondrial matrix, after pruning mycoheterotrophic taxa from both respective trees and seven non-monocot taxa from the plastid tree. We used the bladj function in Phylocom (Webb et al., 2008) to obtain ultrametric plastid and mitochondrial input trees.

Seven fossil calibrations were considered for the plastid data and two for the mitochondrial data (Table S4), each applied with a conservative uniform prior distribution placing equal probability between the lower and upper age bounds. We set minimum fossil ages as the lower bounds of the uniform distribution, and set upper bounds to 151.8 Ma, the maximum age estimate for the angiosperm crown node in analyses by Silvestro et al. (2015). We used GTR + G with four gamma categories as the substitution model and employed an empirical estimate of base frequency, selecting the estimate clock.rate option, and using a birth-death branching model for the tree prior, a 'birth diff' rate of 0.1, an upper limit of 1.0 for the speciation and extinction rates, and a gamma distribution ($\alpha = 2$, $\beta = 0.5$) for nucleotide exchange rates. We set the mean of the RLC model (ucldMean) as an exponential distribution with a mean of 0.33, leaving all other parameters as default, and used CIPRES (Miller et al., 2010) to run seven and 13 independent analyses using plastid and mitochondrial data, respectively. Each analysis was run for 200 million generations, sampling every 2,000 generations.

We also explored the consequences of conducting molecular dating analyses that include vs. exclude mycoheterotrophic taxa. The latter often have highly elevated rates of evolution in plastid, nuclear, and mitochondrial genes (e.g., Merckx et al., 2009; Lam et al., 2018). Iles et al. (2015) hypothesized that such extremely long branches could have a distorting effect on age estimates. We assessed this by re-examining a published monocot-focused matrix of Pandanales by Mennes et al. (2013), which consists of two loci (one nuclear, one mitochondrial) and includes sequences for 30 mycoheterotrophic taxa in the orders Pandanales, Dioscoreales and Petrosaviales, in addition to 96 photosynthetic monocot taxa. We removed all mycoheterotrophs from the matrix used by Mennes et al. (2013), otherwise retaining their alignment, and ran analyses that replicated

all of the conditions in their study. To parallel their methods, we used BEAST v.1.7.5 (Drummond et al., 2012) to run two independent analyses, each for 200 million generations, sampling every 10,000 generations, and used an uncorrelated lognormal relaxed clock (UCLN), GTR+I+G as the substitution model, a Yule branching model for the tree prior, and four fossil calibrations with lognormal prior distributions using the age of the fossil as the offset value (see Mennes et al., 2013, for complete methods).

For all Bayesian analyses conducted here we assessed stationarity and convergence among independent analyses using Tracer v.1.7.1 (Rambaut et al., 2018). We used LogCombiner (v1.7.5 and v2.4.8; Drummond et al., 2012; Bouckaert et al., 2014) to combine independent analyses that converged (ignoring those that did not converge); this led to us combining five of seven analyses based on plastid data, nine of 13 based on mitochondrial data, and both analyses based on the version of the Mennes et al. (2013) matrix that excluded mycoheterotrophs. All resulting combined analyses reached stationarity with ESS values > 200 for all parameters. We summarized sampled trees from analyses that converged using TreeAnnotator (v1.7.5 and v2.4.8; Drummond et al., 2012; Bouckaert et al., 2014), discarding the first 10–25% of trees as burn-in, resampling at a lower frequency to obtain ~10,000 final trees when necessary, and with consensus ages representing the median of posterior distributions.

RESULTS

Higher-order relationships in Pandanales

Nearly all tree-wide relationships recovered in the three parsimony and 12 likelihood analyses of plastid, mitochondrial and combined organellar datasets are well supported by bootstrap analysis and are congruent with each other, ignoring taxon-sampling differences (Figs. 1–3, S1–S15; support values summarized in Table S5). We use the topology from partitioned likelihood analysis of the combined organellar dataset (Fig. 1, left-hand tree, and Fig. S4) as a general reference point for comparing results of other analyses here; inconsistent relationships relative to the combined organellar ML tree are highlighted in all tree figures. Most disagreements concern branches that were weakly

supported in Cyclanthaceae in individual analyses (see below), in addition to a long-branch taxon (*Sciaphila densiflora*, Triuridaceae) that places differently in a single parsimony analysis of plastid data (Figs. 3, S2).

The monophyly of Pandanales and its sister-group status to Dioscoreales are inferred with moderate to strong bootstrap support across different analyses (Figs. 1–3, S1–S15, Table S5). Setting aside Triuridaceae (for which we included only a single genus, *Sciaphila*), the remaining Pandanales families are nearly all strongly supported as monophyletic across analyses. The only exception is that Stemonaceae monophyly was strongly rejected in parsimony analysis of plastid data (Figs. 3, S2), as *Sciaphila* (Triuridaceae) was inferred to be nested within Stemonaceae, as a well-supported sister-group of *Pentastemona* (labeled as branch 'aa' in Fig. 3, left-hand tree, and see Fig. S2). This represents the only strong conflict between parsimony and likelihood analyses (i.e., cases where both conflicting arrangements are strongly supported). The unusual placement of Triuridaceae also represents a strong conflict between plastid and mitochondrial parsimony analyses: although its precise placement was poorly supported by the mitochondrial parsimony data, the latter analysis strongly rejects a sister-group relationship between *Sciaphila* (Triuridaceae) and *Pentastemona* (Stemonaceae), as Stemonaceae monophyly was strongly supported (Fig. 3, right-hand tree, and Fig. S3). The mitochondrial parsimony placement of Triuridaceae as the sister-group of Cyclanthaceae and Pandanaceae is identical to its placement in all the remaining analyses: all likelihood analyses of mitochondrial data (Figs. 2, S12–S15; see branches 'r' and 'q' in Fig. 1 and Table S5), parsimony and likelihood analyses of combined organellar data (Figs. 1, S1, S4, Table S5), and all likelihood analyses of plastid data (Figs. 2, S5–S11, Table S5). Plastid-based likelihood analyses had the strongest support for the placement of Triuridaceae (90–99% bootstrap support for branch 'q'; Figs. 2, S5–S11; Table S5). Setting aside the unusual Stemonaceae-Triuridaceae clade in plastid-based parsimony analysis, relationships among the five Pandanales families were congruent across parsimony and likelihood analyses, and generally had moderate to strong support.

Most genus-level relationships in Pandanales were consistent and strongly supported across all analyses performed here (Figs. 1–3, S1–S15, Table S5). Considering

the 26 internal branches in each rooted subtree that corresponds to Pandanales for analyses that included all sampled taxa except mycoheterotrophic *S. thaidanica*, 15–22 branches had strong support across likelihood analyses, 1–6 were moderately supported and 1–5 were poorly supported (Figs 1, 2, S4–S15, Table S5). In parsimony analyses there were 15–19 branches with strong support, 1–6 with moderate support and 1–4 with poor support (for the plastid-based parsimony tree we considered 27 branches to take account of the placement of *Sciaphila* within Stemonaceae in that analysis; Figs. 1, 3, S1–S3, Table S5).

Nearly all genus-level relationships in Pandanaceae and Stemonaceae were completely congruent across phylogenetic analyses, and most were strongly supported (Figs. 1–3, S1–S15, Table S5). The sole difference in topology was the strongly supported placement of *Sciaphila* as the sister group of *Pentastemona* within Stemonaceae in plastid-based parsimony analyses (see above; Figs. 3, S2). There were no topological differences within Pandanaceae across analyses, and all inferred relationships had strong bootstrap support. In Velloziaceae, genus-level relationships were congruent and moderately to strongly supported across analyses, except for the placement of *Barbacenia*: there was a moderately supported conflict for the placement of *Barbacenia* as either the sister group of *Xerophyta* (branch 'x' in Fig. 1, Table S5; the highest bootstrap support for this relationship was 71% in the plastid-based likelihood analysis that included gene *ycf*1, Fig. S9) or as the sister-group of all Velloziaceae genera except *Acanthochlamys* (see branch 'bb' in Figs. 3, S3; this relationship had 71–82% bootstrap support in two mitochondrial-based analyses, Figs. 3, S3, S14, Table S5).

We inferred strong support for most relationships in Cyclanthaceae in likelihood analysis of the combined organellar matrix; these were mostly congruent with the other phylogenetic analyses (cf. Figs. 1, S4 vs. Figs 2, 3, S1–S3, S5–S15 and see Table S5). All deviations from the topology obtained from likelihood analysis of the combined organellar matrix (Fig. 1, left-hand tree, and Fig. S4) were poorly supported, except for the placements of *Dianthoveus*, *Evodianthus* and *Thoracocarpus*: most analyses, including the combined organellar likelihood analysis, recovered a *Dianthoveus*-*Evodianthus* clade with poor to moderate support (clade 'e' in Fig. 1, Table S5; see Figs. 2, 3, S1, S3, S4, S10–S15), an arrangement that represents a moderately supported

conflict with the plastid likelihood analysis that included gene *ycf1*, as the latter recovered an *Evodianthus-Thoracocarpus* clade with strong support (clade 'cc' in Fig. S9, Table S5).

Plastid genome structural evolution in Pandanales

We compared the newly assembled complete circular plastomes for three Pandanales taxa here (*Acanthochlamys bracteata*, Velloziaceae, Fig. S16, GenBank accession MN905940; *Benstonea copelandii*, Pandanaceae, Fig. S17, GenBank accession MN905941; *Croomia japonica*, Stemonaceae, Fig. S18, GenBank accession MN905942) to eight additional published plastomes for the order (for a total of 11 plastomes, nine from photosynthetic taxa and two from mycoheterotrophs; Table S6). The nine plastomes from photosynthetic taxa are comparable in size, organization, gene content and gene order to those of other angiosperms (e.g., Hansen et al., 2007); their sizes range from 153,842 bp in *Acanthochlamys* to 158,545 bp in *Carludovica.* They are 91.7–91.9% bigger than *S. thaidanica*, the smallest of the two Triuridaceae plastomes (Table S6). The total number of unique genes in the nine green plastomes ranges from 112–113, with 18–20 of these duplicated in the inverted repeat (IR) region (Table S6). Counting IR duplications, there are 130–133 genes in the plastomes of photosynthetic taxa, with differences due both to gene loss and expansion/contraction of the IR at the junction with the large single copy (LSC) region (Fig. S19). Compared to Stemonaceae plastomes, which have the highest gene content in our sampling, *Sciaphila thaidanica* has a genome that is ~15–18% as large (range considering total and unique genes); *S. densiflora* is ~21–25% as large, respectively (Table S6). These differences reflect loss of photosynthesis-related and other genes, and to some degree the loss of the IR in the Triuridaceae taxa (Petersen et al., 2018 for *S. thaidanica*; note that *S. densiflora* may have small cryptic repeats, Lam et al., 2015).

We did not recover the two exons for the *rps*16 locus in the LSC of *Acanthochlamys* (Fig. S16); BLAST-based searches of intergenic space between *trn*Q-UUG and *trn*K-UUU, the two genes that usually flank *rps*16, also did not recover these exons. Stemonaceae taxa considered here (Table S6) have the largest IR (20 genes,

including *rps*19 and *rpl*22; Figs. S18, S19c, Table S6). In *Acanthochlamys* both *rps*19 and *rpl*22 are part of the LSC, and so the IR of this taxon contains 18 genes (Figs. S16, S19d, Table S6). The IR of *Carludovica* has 19 genes; it includes *rps*19 but not *rpl*22 (Fig. S19a, Table S6). In *Benstonea*, the 5'-end of *rps*19 straddles the IR_A junction and this gene portion is duplicated in the IR_B; thus its IR encompasses 18 intact genes and one short partial reading frame (Figs. S17, S19b, Table S6). *Sciaphila thaidanica* and *S. densiflora* respectively retain eight and nine of the 18–20 genes found in the IRs of the photosynthetic taxa sampled here; these genes are colinear across sampled photosynthetic and non-photosynthetic taxa, except for an inversion of the *rps*7 and 3'-*rps*12 gene block in *S. densiflora* (Fig. 3, Lam et al., 2015). Outside the IR regions, there were additional inversions inferred in *S. densiflora* (*rps*18-*acc*D gene block; Fig. 3 in Lam et al., 2015) and *S. thaidanica* (*rps*12-*rps*2 gene block; Fig. 2 in Petersen et al., 2018). Pandanales taxa were otherwise all fully colinear, except for a single inversion of the *cem*A-*pet*A gene block in the two *Stemona* taxa (Lu et al., 2018).

Divergence times and substitution rate shifts in Pandanales and other monocots

All 95% highest posterior density (HPD) intervals from analyses of the plastid and mitochondrial date sets overlapped for nodes of interest, but plastid data produced narrower intervals (mean interval length of 20 Ma) than mitochondrial data (mean of 31 Ma; Table 1, Figs. 4, S20, S21); note that the mycoheterotrophic family Triuridaceae was excluded from these analyses (see below for rationale). Median ages of major clades estimated from plastid and mitochondrial data were highly correlated (Pearson correlation coefficient, $r = 0.97$); ages based on the mitochondrial data are somewhat younger (slope, $m = 0.94$) by an average of 15 Ma (0–25 Ma, SD = 8 Ma) for the 13 nodes that overlap between the two data sets (Table 1, Figs. 4, S20, S21). Outside Pandanales, the 95% HPD intervals for node ages overlapped across plastid and mitochondrial analyses (Table S7).

The estimated stem age of Pandanales (and thus the date of separation from its sister group, Dioscoreales) is 106 (98–117) Ma, and it began diversifying into its constituent families (= order crown age) at 93 (85–104) Ma based on plastid data; mitochondrial age estimates for these two splits are more recent: 84 (70–109) and 71 (58–

94) Ma, respectively (Table 1, Figs. 4, S20, S21). Plastid and mitochondrial data yielded similar estimates for the dates by which family-level splits in Pandanales were complete: 61 (53–71) and 56 (49–74) Ma, respectively. Stemonaceae have the oldest family-level crown age based on plastid data (65 [53–75] Ma), and Cyclanthaceae considering mitochondrial data (49 [47–54] Ma); Pandanaceae have the youngest family-level crown age based on both plastid (35 [23–51] Ma) and mitochondrial data (26 [20–35] Ma) (Table 1; Figs. 4, S20, S21).

Bayesian analyses estimated 31 and 18 molecular rate shifts in the time-calibrated phylogenies inferred from plastid and mitochondrial data, respectively (note that these have different taxon samplings). Considering monocots, the plastid-based analyses yielded inferences of four rate shifts in Pandanales and 25 in the remaining monocot taxa (Figs. 5, S22; full tree shown in the latter figure). The fastest overall plastid rates were observed in Zingiberaceae (3.2 x 10^{-3} substitutions per site per million years) and Poaceae (3.1 x 10^{-3} substitutions per site per million years). The slowest plastid rates were recovered in Arecaceae after the divergence of *Calamus* (0.1 x 10^{-4} substitutions per site per million years), and in Cyclanthaceae, *Sararanga* (Pandanaceae) and *Japonolirion* (Petrosaviaceae) (the latter all 0.2 x 10^{-4} substitutions per site per million years). In Pandanales, successive rate decelerations in plastid genes occurred in each of the common ancestors of Stemonaceae, Pandanaceae-Cyclanthaceae, Pandanaceae, *Sararanga* (Pandanaceae), and Cyclanthaceae, with the slowest overall rates in the latter family. Only one rate acceleration was inferred in Pandanales plastid data, after the divergence of *Sararanga* from all other Pandanaceae (Figs. 5, S22).

In analyses of mitochondrial data, 12 of the 18 estimated rate shifts occurred in Pandanales, where taxon sampling is heaviest, and the remaining six in other monocots (Figs. 5, S23; full tree shown in the latter figure). The fastest overall mitochondrial rates were observed in *Benstonea* (Pandanaceae; 1 x 10^{-3} substitutions per site per million years), and in *Ludovia* (Cyclanthaceae) and *Stratiotes* (Hydrocharitaceae) (the latter both 9 x 10^{-4} substitutions per site per million years). The slowest mitochondrial rate (1 x 10^{-4} substitutions per site per million years) was recovered for *Tofieldia* (Tofieldiaceae) and multiple Stemonaceae, Pandanaceae and Cyclanthaceae lineages (Figs. 5, S23). In Pandanales, there were five rate decelerations and seven accelerations. Independent

decelerations occurred once in Velloziaceae (in the *Barbacenia* terminal lineage), Stemonaceae (after the divergence of *Pentastemona*) and Pandanaceae, and three times in Cyclanthaceae (in the terminal lineages leading to *Cyclanthus* and *Stelestylis,* and the lineage leading to the common ancestor of all Cyclanthaceae excluding *Cyclanthus*). Accelerations occurred in each of the terminal lineages leading to *Stichoneuron* (Stemonaceae), and *Benstonea* (Pandanaceae), the lineage leading to the common ancestor of Pandanaceae-Cyclanthaceae, and four Cyclanthaceae taxa (terminal lineages leading to *Asplundia*, *Dianthoveus*, *Ludovia*, and the lineage leading to the common ancestor of a clade comprising *Chorigyne, Ludovia, Sphaeradenia* and *Stelestylis*) (Figs. 5, S23).

Divergence time estimates excluding mycoheterotrophic taxa

We re-analyzed a published dating analysis by Mennes et al. (2013) to explore the possible age biasing effect of very rapidly evolving taxa, by excluding mycoheterotrophs (in the orders Dioscoreales, Pandanales, Petrosaviales) and comparing nodes of interest in Pandanales and other monocots to the original results. There were multiple differences in median age estimates for nodes in the vicinity of where mycoheterotroph lineages were placed in the original study. Median age estimates from the original and modified Mennes et al. (2013) analyses were within the 95% HPD intervals of each other, except for the Pandanales crown node (= Velloziaceae stem node); this node was 62 (38–90) Ma with mycoheterotrophs excluded vs. 91 (69–110) Ma with them included (Table 1, cf. Fig. S24 here vs. Fig. 4 in Mennes et al., 2013). The remaining two largest age differences across analyses were for the crown node of Dioscoreales (91 [70–110] Ma with mycoheterotrophs excluded vs. 101 [85–116] Ma with them included), and the crown node of Velloziaceae (31 [10–57] vs. 37 [9–75] Ma, respectively). These differences contrast with negligible age differences for the nine orders that did not include mycoheterotrophic taxa in the original matrix (0–2 Ma).

DISCUSSION

Our study provides a comprehensive and well resolved picture of phylogenetic relationships at the family and genus level in the monocot order Pandanales using phylogenomic data from the plastid and mitochondrial genomes. These findings are well supported for most branches and are congruent across data sets and inference methods, with a key exception that appears to reflect a long-branch artefact involving the local placement of mycoheterotrophic Triuridaceae for one data set and analysis method (i.e., plastid data using parsimony). These new results provide a useful framework for thinking about substitution rate shifts, clade ages, biogeography and plastid genome structural evolution.

Organellar data produce robust family-level relationships in Pandanales

The family-level topology we recovered across all likelihood analyses here (Figs. 1, 2, S4–S15) is congruent with the plastid phylogenomic studies of Lam et al. (2015) and Givnish et al. (2018), which also recovered strongly supported arrangements of the five families of Pandanales: (Velloziaceae, (Stemonaceae, (Triuridaceae, (Cyclanthaceae, Pandanaceae)))). Plastid rates appear to be much more elevated than mitochondrial rates in mycoheterotrophic *Sciaphila* based on the substantially longer branches subtending this taxon in plastid-based likelihood analyses (cf. Fig. S6 vs. S13). These very long branches appear to underlie the unusual placement of *Sciaphila* in plastid-based parsimony analysis, as a strongly supported sister-group of *Pentastemona* within Stemonaceae here; labeled as branch 'aa' (Fig. 3 left-hand tree, Fig. S2, Table S5), or as the sister group of *Xerophyta* (Velloziaceae) in Lam et al. (2015, although with poor support in the latter study). In contrast, the mitochondrial-based parsimony analysis, and all likelihood analyses for both genomes considered individually or together, support Triuridaceae as the sister group of Cyclanthaceae-Pandanaceae. Considered together, the parsimony placement of Triuridaceae for the plastid data, but not for the less rapidly evolving mitochondrial data, likely reflects a long-branch artefact for the former genome, and so based on the current data at least, the placement of Triuridaceae in the order can be considered settled.

Multiple studies have focused on relationships within individual Pandanales families; these generally used comprehensive genus-level samplings of each family and surveyed one to several genes, morphological characters, or a combination of these, resulting in varying degrees of support for relationships. For example, the strongly supported topologies we recovered within Stemonaceae and Pandanaceae across most analyses here are consistent with previous studies that also reconstructed relationships in each family with strong support: for Stemonaceae, Caddick et al. (2002) and Rudall et al. (2005) used combined molecular and morphological data, and Lam et al. (2015) used full plastome data; for Pandanaceae, Gallaher et al. (2015) used nuclear and combined nuclear-plastid data sets. Mennes et al. (2013) performed a detailed analysis of relationships in Triuridaceae. The representative taxon included here (*Sciaphila*) is classified under tribe Sciaphileae (Cheek, 2003). However, the tribe may not be monophyletic (Rudall and Bateman, 2006; Mennes et al., 2013). Additional mitochondrial and nuclear phylogenomic data would be useful to resolve these and other unresolved relationships in Triuridaceae.

Relationships in the remaining two Pandanales families have also remained somewhat unclear. In Velloziaceae, we clarified the placements of *Barbaceniopsis* and *Vellozia* as a strongly supported clade in all analyses here; the placements of these two genera have varied and were poorly supported in previous studies (Menezes et al., 1993; Behnke et al., 2000, 2013; Salatino et al., 2001; Mello-Silva, 2005; Mello-Silva et al., 2011; Alcantara et al., 2018). The sole relationship in the family that remains uncertain in our analyses is that of *Barbacenia,* which we recovered as either the sister-group of *Xerophyta* (branch 'x' in Fig. 1, Table S5, and see Figs. 2, S4–S9) or more commonly as the sister-group of *Xerophyta* plus *Barbaceniopsis*-*Vellozia* (branch 'bb' in Fig. 3, Table S5, and see Figs. 1, 2, S1–S3, S10–S15), with poor to moderate support for each hypothesis. Only the latter arrangement has been recovered with strong support to date, in Bayesian analyses by Alcantara et al. (2018; posterior probability, PP, of 0.97) using five molecular markers. The relatively dense taxon sampling of Alcantara et al. (2018), comprising 150 of the ~310 species in Velloziaceae (Govaerts, 2019), may have

contributed to their increased support for the placement of *Barbacenia* compared to our study and that of Mello-Silva et al. (2011). Here we used a much larger molecular data set, ranging from 37 to 119 organellar genes, focused on a representative genus-level sampling of the family consisting of six taxa; Mello-Silva et al. (2011) used the same five molecular markers as Alcantara et al. (2018) for a sampling of 47 Velloziaceae taxa but recovered *Barbacenia* as the sister group of *Barbaceniopsis-Vellozia* with poor support in parsimony analyses.

Previous morphology-based studies produced a baseline phylogenetic outline for Cyclanthaceae, but multiple poorly supported branches persisted, and molecular studies of this group had not been performed until this study and a parallel study by one of our co-authors (E.S.L. and colleagues, Universidade Federal Rural da Amazônia, unpublished data). Harling (1958) divided the family into subfamilies Cyclanthoideae (comprising *Cyclanthus*) and Carludovicoideae (comprising the remaining genera). The latter subfamily was further divided into the informal Sphaeradenia group (comprising *Chorigyne-Stelestylis, Sphaeradenia, Ludovia*) and Asplundia group (comprising the remaining genera in the subfamily). However, multiple relationships among and within these proposed groups remained unclear (Harling, 1958; Hammel and Wilder, 1989; Eriksson, 1989, 1994). Our analyses confirm previously proposed hypotheses for the placement of *Cyclanthus* and for relationships within the Sphaeradenia group (Harling, 1958; Hammel and Wilder, 1989; Eriksson, 1989, 1994). The strongly supported sister-group relationship of *Cyclanthus* to all other Cyclanthaceae recovered here had been previously assumed by Harling (1958) and Eriksson (1994) as their studies did not include an outgroup; only Hammel and Wilder (1989) tested the latter placement in preliminary analyses using *Freycinetia* (Pandanaceae) as an outgroup, but they did not report on branch support. Our topology for the Sphaeradenia group is consistent with the morphology-based hypotheses of Harling (1958), Hammel and Wilder (1989) and Eriksson (1994), although only the latter study included *Chorigyne* (described by Eriksson, 1989) and assessed branch support.

Our analyses also recovered four strongly supported arrangements in Cyclanthaceae that had not been previously proposed: an *Evodianthus-Thoracocarpus* clade, a *Carludovica-Schultesiophytum* clade, an *Asplundia-Dicranopygium* clade, and an

arrangement with *Asplundia-Dicranopygium* as the sister-group of all Cyclanthaceae except *Carludovica-Schultesiophytum* and *Cyclanthus*. Support for these newly proposed relationships was generally strong from analyses using combined organellar and mitochondrial data, except for the *Evodianthus-Thoracocarpus* clade, which was strongly supported only in plastid-based analyses that included the large and variable plastid locus *ycf*1 (Fig. S9, Table S5). Two portions of *ycf*1 are among the most variable coding regions in land-plant plastomes (Dong et al., 2015), and this gene may therefore have contributed to improved support for the short branch subtending *Evodianthus-Thoracocarpus* relative to the other analyses here that excluded this gene. We found strong support for the non-monophyly of the Asplundia group, as Eriksson (1994) had suspected, as the Sphaeradenia group is nested within it. We thus propose maintaining the two-subfamily classification of Cyclanthaceae, but not the subdivision of the latter into smaller groups.

Fluctuating inverted repeat boundaries and gene loss in Pandanales plastomes

The plastomes of photosynthetic land plants are highly conserved in structural organization, gene content, and gene order (e.g., Palmer, 1991; Raubeson and Jansen, 2005; Wicke et al., 2011). Large clusters of polycistronically transcribed genes impose a constraint on plastome organization, making large structural changes relatively rare, and therefore potentially useful as markers for clades (e.g., Downie and Palmer, 1992; Wang et al., 2008). Plastomes typically have a quadripartite structure with an inverted repeat (IR) region separating a small single-copy (SSC) region and a large single-copy (LSC) region (e.g., Kolodner and Tewari, 1979; Palmer, 1991). The IR of land plants varies somewhat in size as a result of expansions and contractions; these shifts are often lineage-specific and may have phylogenetic significance (Wang et al., 2008; Wicke et al., 2011). For example, changes in the location of the junction between the IR and single-copy regions have been used as synapomorphies at various phylogenetic scales in angiosperms, including at the genus level (e.g., the *Berberis-Mahonia* clade in Berberidaceae; Kim and Jansen, 1994) and subfamily level (e.g., Ehrhartoideae-Pooideae clade in Poaceae; Guisinger et al., 2010). The IR-LSC junction appears to have shifted dynamically in

Pandanales, as its location varies across the four photosynthetic families (Fig. S19). Junctions of the IR and LSC vary less within Pandanales families than between them, at least considering our sampling. For example, *rpl*22 terminates the IR in all six Stemonaceae plastomes examined here (Fig. S19, Table S6); in Pandanaceae, a partial *rps*19 copy terminates the IR in the two species examined here (note that one of them, *Pandanus odoratissimus*, is represented by IR-LSC junction sequences, not by a full plastome in Wang et al., 2008). Therefore, IR-LSC junctions likely act as genome structural synapomorphies for major clades in Pandanales, and additional complete plastomes should be surveyed to address this possibility.

Land-plant plastomes commonly encode 21 small and large ribosomal subunit proteins (*rps* and *rpl* genes) and loss of these genes from the plastome is rare in photosynthetic plants (Wicke et al., 2011). However, independent losses of *rps*16 have been documented for multiple plant lineages, including *Acanthochlamys* here (Table S6, Fig. S16), additional monocots (*Dioscorea* and some Asparagales: Hansen et al., 2007; Steele et al., 2012; McKain et al., 2016), eudicots (*Populus* and some Fabaceae: Doyle et al., 1995; Steane, 2005), gymnosperms (*Pinus*: Tsudzuki et al., 1992) and liverworts (*Marchantia*: Ohyama et al., 1986). The gene *rps*16 may be part of a group of plastid genes that are especially prone to being lost from the plastome following functional transfer to the nucleus (e.g., Martin et al., 1998; Ueda et al., 2008). Loss of *rps*16 may have occurred at least twice more in Velloziaceae, as we did not recover the 5'-exon from *Barbaceniopsis* or the 3'-exon from *Vellozia.* Additional full and completed plastomes would be useful to confirm the latter two cases, and to explore how labile the loss of *rps*16 is among closely related taxa in Pandanales.

Organellar age estimates and consequences for biogeographic hypotheses

Our finding that age estimates from plastid- and mitochondrial-based dating analyses were highly correlated and similar (mean difference per node = 13.09 Ma, SD = 7.54 Ma, Pearson correlation coefficient $r = 0.97$, slope $m = 0.94$; Tables 1, S7), suggests that the relaxed clock models we used are robust to data sets under different substitution rate regimes when the latter are correctly accounted for. These findings are consistent with

those of Magallón et al. (2013), who found that ages estimated from two variant data sets (consisting of four highly conserved plastid genes vs. one faster evolving plastid gene) were highly correlated despite an order of magnitude difference in the absolute rates of each data set. Variation in age estimates across data sets here may instead be more strongly influenced by differences in the number of fossil calibrations used (in turn affected by differences in taxon sampling); we respectively used six and two fossil calibrations for plastid- and mitochondrial-based analyses due to taxon sampling differences. For nodes of interest here, plastid and mitochondrial estimates were most similar in the vicinity of the two calibrated nodes common to the two organellar analyses (i.e., the monocot stem node and the *Cyclanthus* stem node): ages differed by 4 Ma for the monocot stem node, by 5 Ma for the Cyclanthaceae-Pandanaceae stem node, and were the same for the Cyclanthaceae crown node (Table 1, Figs. 4, S20, S21).

Pandanales age estimates from plastid-based analyses here are most similar to the angiosperm-focused study of Magallón et al. (2015) considering also the monocot-wide study of Givnish et al. (2018), the original (and our re-analyzed) versions of Mennes et al. (2013), and family-focused studies on Pandanaceae (Gallaher et al., 2015) and Velloziaceae (Alcantara et al., 2018), summarized in Table 1. Our mitochondrial-based analyses are most similar to those of Magallón et al. (2015) and Givnish et al. (2018) (Table 1). We estimated younger ages for the stem and crown nodes of Pandanaceae and Velloziaceae compared to Gallaher et al. (2015) and Alcantara et al. (2018), respectively. Nevertheless, our estimates add support to the hypotheses presented in both of those studies that neither family follows a Gondwanan vicariance diversification pattern, as previously suggested by Callmander et al. (2003) for Pandanaceae, and by Mello-Silva et al. (2011) for Velloziaceae. For example, the diversification of Velloziaceae genera from Africa (*Xerophyta*) and South America (*Barbacenia, Barbaceniopsis, Vellozia*) began 11–16 Ma based on plastid data, 5–10 Ma based on mitochondrial analyses (data not shown) and 40–55 Ma in analyses by Alcantara et al. (2018), which is well after the separation of West Gondwana into Africa and South America in the early Cretaceous (~105 Ma; McLoughlin, 2001). In Pandanaceae, the Madagascar endemic *Martellidendron* diverged from its common ancestor with *Pandanus* and *Benstonea* 7–21 Ma based on plastid analyses, 4–11 Ma based on mitochondrial analyses (data not

shown), and 15–38 Ma in analyses by Gallaher et al. (2015), which is well after the separation of Madagascar from East Gondwanan landmasses in the late Cretaceous (~65 Ma; McLoughlin, 2001).

Members of Pandanaceae have a paleotropical distribution and the sister taxon of the family, Cyclanthaceae, is exclusively neotropical (Eriksson, 1994; Stone et al., 1998). These families diverged 53–71 Ma based on plastid data and 49–74 Ma based on mitochondrial data, well after the separation of Africa and South America (~105 Ma) (Table 1). Stem ages estimated here for Pandanaceae and Cyclanthaceae, in combination with fossils from modern-day Europe attributed to each family (Iles et al., 2015), may point to a Laurasian origin for both lineages, as previously considered by Gallaher et al. (2015). Stemonaceae are distributed in tropical Asia and northern Australia, with one species (*Croomia pauciflora*) found in the southeastern United States (Rudall and Bateman, 2006). *Croomia* diverged from its sister taxon, *Stichoneuron*, 9–33 My based on plastid analyses and 18–35 My based on mitochondrial analyses (data not shown). North America separated from former Gondwanan landmasses 165–180 Ma (McLoughlin, 2001), and so the current distribution of Stemonaceae may be consistent with an Old World origin and subsequent dispersal to the New World. However, there is no known fossil record for Stemonaceae to further clarify the latter hypothesis (Collinson et al., 1993; Herendeen and Crane, 1995; Smith, 2013). Our analyses excluded mycoheterotrophic Triuridaceae, but Mennes et al. (2013) attributed the pantropical distribution of the family largely to dispersal events that postdate major tectonic events.

Age estimates in the presence of rapidly evolving mycoheterotrophic taxa

Elevated rates in the three genomic compartments of mycoheterotrophs (e.g., Merckx et al., 2009; Lemaire et al., 2011; Lam et al., 2018) led Iles et al. (2015) to hypothesize that these taxa may be challenging to incorporate in molecular dating analyses. Here, we found that inclusion of fast-evolving mycoheterotrophs within a phylogenetic framework of slower-evolving photosynthetic relatives likely moderately biased ages upwards in the vicinity of the former taxa by 2–29 Ma. This finding is consistent with previous studies that have shown that relaxed clock methods, such as the uncorrelated lognormal relaxed clock (UCLN) model used here for reanalyzing the data set of Mennes et al. (2013), may

perform poorly and mislead age estimation when there is substantial among-lineage rate variation (e.g., Dornburg et al., 2011; Beaulieu et al., 2015). For example, Beaulieu et al. (2015), found that substantial lineage-specific rate heterogeneity simulated near the angiosperm crown node produced ages that were ~70 Ma older than the fossil-based age that they used to fix that node, but biases towards older ages were localized in the vicinity of simulated rate shifts and did not have knock-on effects on ages for the rest of the phylogeny. We therefore recommend cautious interpretation of ages estimated in molecular dating analyses that include mycoheterotrophs, particularly for crown ages of the closest photosynthetic relatives. This may be particularly important when using plastid genes, as these may experience substantial relaxation or release of purifying selection after loss of photosynthesis in mycoheterotrophs (Graham et al., 2017), which may further exacerbate lineage-specific rate heterogeneity. In contrast, mitochondrial and nuclear genes may be less prone to this effect, although even in these genomes rates can be moderately to substantially accelerated relative to photosynthetic relatives (e.g., see Fig. S13 here; Merckx et al., 2009; Lemaire et al., 2011).

Rate accelerations and decelerations in Pandanales organellar genomes

Of the three plant genomes, the mitochondrial genome (mitogenome) usually has by far the lowest substitution rate, followed by the plastome and then the nuclear genome (Wolfe et al., 1987; Palmer and Herbon, 1988). Substitution rate regimes correlate with plant life history traits, and so rate shifts commonly occur in parallel across the three plant genomes (Eyre-Walker and Gaut, 1997; Drouin et al., 2008; Zhu et al., 2014). For example, molecular rates are consistently lower in woody trees, arborescent taxa and shrubs with relatively long generation times, compared to higher and more variable rates in herbaceous plants, which tend to have substantially shorter generation times; additional processes that may contribute to rate differences include plant height, population size, speciation rate and paternal inheritance in the case of organelles (Bousquet et al., 1992; Gaut et al., 1996; Laroche et al., 1997; Smith and Donoghue, 2008; Korall et al., 2010; Lanfear et al., 2013). More rarely, rate shifts may be decoupled across the plant genomes, as is the case here, where we detected mitochondrial rate accelerations in *Benstonea*

(Pandanaceae) and multiple Cyclanthaceae taxa, alongside sequential plastid rate decelerations in the common ancestors of the Pandanaceae-Cyclanthaceae clade and Cyclanthaceae itself, where we observed faster mitochondrial than plastid rates for multiple taxa (Figs. 5, S22, S23). Significant mitochondrial-only rate elevations have been observed at synonymous substitution sites in a few angiosperm lineages (Palmer et al., 2000; Cho et al., 2004; Parkinson et al., 2005; Bakker et al., 2006; Mower et al., 2007; Sloan et al., 2008; Zhu et al., 2014; Skippington et al., 2015); here we did not partition rate estimates among synonymous and nonsynonymous sites. Life-history traits in Pandanaceae and Cyclanthaceae may contribute to the slow plastid rates estimated for each family. For example, Pandanaceae comprise arborescent taxa, shrubs and woody climbers (Stone et al., 1998) and Cyclanthaceae taxa are herbaceous, but can be relatively tall (e.g., up to 5 m in *Carludovica palmata* and *Sphaeradenia gigantea*; Harling et al., 1998). In contrast, we estimated relatively faster plastid rates for the smaller and mostly herbaceous taxa in Stemonaceae and Velloziaceae (Kubitzki, 1998a,b,c). Life-history traits may underlie decoupling of organellar rates in some Pandanaceae and Cyclanthaceae taxa leading to the rate decelerations evident in plastid sequences here, but may be obscured in mitochondrial sequences by unknown processes that led to rate accelerations in this genome (Figs. 5, S22, S23). The number of sequenced angiosperm mitogenomes remains limited (Havird et al., 2019), but as more become available, this should help clarify the extent and frequency of lineage-specific mitochondrial rate heterogeneity.

ACKNOWLEDGEMENTS

The authors thank anonymous reviewers for comments that helped to improve the text, Daniel Spalink for advice in running BEAST, the Missouri Botanical Garden for providing material of *Dianthoveus cremnophilus*, and Ruth Stockey, Gar Rothwell and the Munich Botanical Garden for Stemonaceae material. This research was enabled in part by WestGrid (www.westgrid.ca) and Compute Canada (www.computecanada.ca).

This work was supported by an NSERC Postgraduate Fellowship and UBC Four-Year Fellowship to MSG,and an NSERC Discovery Grant to SWG.

LITERATURE CITED

Alcantara, S., Ree, R.H., Mello-Silva, R. 2018. Accelerated diversification and functional trait evolution in Velloziaceae reveal new insights into the origins of the *campos rupestres*' exceptional floristic richness. Ann. Bot. 122, 165–180.

Angiosperm Phylogeny Group. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 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. Bot. J. Linn. Soc. 161, 105–121.

Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181, 1–20.

Bakker, F.T., Breman, F., Merckx, V. 2006. DNA sequence evolution in fast evolving mitochondrial DNA nad1 exons in Geraniaceae and Plantaginaceae. Taxon 55, 887–896.

Barrett, C.F., Baker, W.J., Comer, J.R., Conran, J.G., Lahmeyer, S.C., Leebens-Mack, J.H., Li, J., Lim, G.S., Mayfield-Jones, D.R., Perez, L., Medina, J., Pires, J.C., Santos, C., Stevenson, D.W., Zomlefer, W.B., Davis, J.I. 2016. Plastid genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. New Phytol. 209, 855–870.

Beaulieu, J.M., O'Meara, B.C., Crane, P., Donoghue, M.J. 2015. Heterogeneous rates of molecular evolution and diversification could explain the Triassic age estimate for angiosperms. Syst. Biol. 64, 869–878.

Behnke, H.-D., Treutlein, J., Wink, M., Kramer, K., Schneider, C., Kao, P.C. 2000. Systematics and evolution of Velloziaceae, with special reference to sieve-element

plastids and *rbcL* sequence data. Bot. J. Linn. Soc. 134, 93–129.

Behnke, H.-D., Hummel, E., Hillmer, S., Sauer-Gürth, H., Gonzales, J., Wink, M. 2013. A revision of African Velloziaceae based on leaf anatomy characters and *rbcL* nucleotide sequences. Bot. J. Linn. Soc. 172, 22–94.

Bell, C.D., Soltis, D.E., Soltis, P.S. 2010. The age and diversification of the angiosperms re-visted. Am. J. Bot. 97, 1296–1303.

Bell, D., Lin, Q., Gerelle, W.K., Joya, S., Chang, Y., Taylor, Z.N., Rothfels, C.J., Larsson, A., Villarreal, J.C., Li, F.-W., Pokorny, L., Szövényi, P., Crandall-Stotler, B., DeGironimo, L., Floyd, S.K., Beerling, D.J., Deyholos, M.K., von Konrat, M., Ellis, S., Shaw, A.J., Chen, T., Wong, G.K.-S., Stevenson, D.W., Palmer, J.D., Graham, S.W. 2020. Organellomic data sets confirm a cryptic consensus on (unrooted) land-plant relationships and provide new insights into bryophyte molecular evolution. Am. J. Bot. 107, 1–25.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.

Bousquet, J., Strauss, S.H., Doerksen, A.H., Price, R.A. 1992. Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. Proc. Natl. Acad. Sci. USA. 89, 7844–7848.

Caddick, L.R., Rudall, P.J., Wilkin, P., Hedderson, T.A.J., Chase, M.W. 2002. Phylogenetics of Dioscoreales based on combined analyses of morphological and molecular data. Bot. J. Linn. Soc. 138, 123–144.

Callmander, M.W., Chassot, P., Küpfer, P., Lowry II, P.P. 2003. Recognition of *Martellidendron*, a new genus of Pandanaceae, and its biogeographic implications. Taxon 52, 747–762.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421.

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H, Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D, Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, D., Hedren, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier,

G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn Jr., G.H., Graham, S.W., Barrett, S.C.H., Dayanandan, S., Albert, V.A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Mo. Bot. Gard. 80, 528–580.

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., Hoot, S. 1995. Molecular phylogenetics of Lilianae. In: Rudall, P.J., Cribb, P.J., Cutler, D.F., Humphries, C.J. (Eds.), Monocotyledons: Systematics and Evolution. Royal Botanic Gardens, Kew, Richmond, UK, pp. 109–137.

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

Chase, M.W., Fay, M.F., Devey, D.S., Maurin, O., Rønsted, 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.A., Graham, S.W., Rai, H.S. 2006. Multigene analyses of monocot relationships: a summary. Aliso, 22, 63–75.

Cheek, M. 2003. Kupeaeae, a new tribe of Triuridaceae from Africa. Kew Bull. 58, 939–949.

Cho, Y., Mower, J.P., Qiu, Y.-L., Palmer, J.D. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proc. Natl. Acad. Sci. USA. 101, 17741–17746.

Collinson, M.E., Boulter, M.C., Holmes, P.L. 1993. 45, Magnoliophyta ('Angiospermae'). In: Benton, M.J. (Ed.), The Fossil Record 2. Chapman and Hall, London, UK, pp. 809–841.

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

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

Dahlgren, R.M.T. and Rasmussen, F.N. 1983. Monocotyledon evolution: characters and phylogenetic estimation. In: Hecht, M.K., Wallace, B., Prance, G.T. (Eds.), Evolutionary Biology, Volume 16. Plenum Press, New York, USA, pp. 255–395.

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

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., Gandolfo, M. 2004. A phylogeny of the monocots, as inferred from *rbcL* and *atpA* sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. Syst. Bot. 29, 467–510.

Davis, J.I. and Soreng, R.J. 2010. Migration of endpoints of two genes relative to boundaries between regions of the plastid genome in the grass family (Poaceae). Am. J. Bot. 97, 874–892.

De la Torre, A.R., Li, Z., Van de Peer, Y., Ingvarsson, P.K. 2017. Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. Mol. Biol. Evol. 34, 1363–1377.

Dong, W., Xu, C., Li, C., Sun, J., Zuo, Y., Shi, S., Cheng, T., Guo, J., Zhou, S. 2015. *ycf1*, the most promising plastid DNA barcode of land plants. Sci. Rep. 5, 8348.

Dornburg, A., Brandley, M.C., McGowen, M.R., Near, T.J. 2011. Relaxed clock and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea). Mol. Biol. Evol. 29, 721–736.

Doyle, J.J., Doyle, J.L., Palmer, J.D. 1995. Multiple independent losses of two genes and one intron from legume chloroplast genomes. Syst. Bot. 20, 272–294.

Downie, S.R. and Palmer, J.D. 1992. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis, P.S., Soltis, D.E., Doyle, J.J. (Eds.), Molecular Systematics of Plants, Springer, Boston, USA, pp. 14–35.

Downie, S.R. and Jansen, R.K. 2015. A comparative analysis of whole plastid genomes from the Apiales: expansion and contraction of the inverted repeat, mitochondrial to

plastid transfer of DNA, and identification of highly divergent noncoding regions. Syst. Bot. 40, 336–351.

Drouin, G., Daoud, H., Xia, J. 2008. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. Mol. Phylogenet. Evol. 49, 827–831.

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

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Edgar, R.C. 2004a. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

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

Eguchi, S. and Tamura M.N. 2016. Evolutionary timescale of monocots determined by the fossilized birth-death model using a large number of fossil records. Evolution. 70, 1136–1144.

Eriksson, R. 1989. *Chorigyne*, a new genus of the Cyclanthaceae from Central America. Nord. J. Bot. 9, 31–45.

Eriksson, R. 1994. Phylogeny of the Cyclanthaceae. Plant Syst. Evol. 190, 31–47.

Eyre-Walker, A. and Gaut, B.S. 1997. Correlated rates of synonymous site evolution across plant genomes. Mol. Biol. Evol. 14, 455–460.

Felsenstein, J. 1978 . Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27, 401–410.

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

Foster, C.S.P., Sauquet, H., Van Der Merwe, M., McPherson, H., Rossetto, M., Ho, S.Y.W. 2017. Evaluating the impact of genomic data and priors on Bayesian estimates of the angiosperm evolutionary timescale. Syst. Biol. 66, 338–351.

Gallaher, T., Callmander, M.W., Buerki, S., Keeley, S.C. 2015. A long distance dispersal hypothesis for the Pandanaceae and the origins of the *Pandanus tectorius* complex. Mol. Phylogenet. Evol. 83, 20–32.

Gaut, B.S., Morton, B.R., McCaig, B.C., Clegg, M.T. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proc. Natl. Acad. Sci. USA. 93, 10274–10279.

Givnish, T.J., Zuluaga, A., Marques, I., Lam, V.K.Y., Soto Gomez, M., Iles, W.J.D., Ames, M., Spalink, D., Moeller, J.R., Briggs, B.G., Lyon, S.P., Stevenson, D.W., Zomlefer, W., Graham, S.W. 2016. Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. Cladistics, 0, 1–25.

Givnish, T.J., Zuluaga, A., Spalink, D., Soto Gomez, M., Lam, V.K.Y., Saarela, J.M., Sass, C., Iles, W.J.D., Lima de Sousa, D.J., Leebens-Mack, J., Pires, J.C., Zomlefer, W.B., Gandolfo, M.A., Davis, J.I., Stevenson, D.W., dePamphilis, C., Specht, C.D., Graham, S.W., Barrett, C.F., Ané, C. 2018. Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. Am. J. Bot. 105, 1–23.

Govaerts, R. World Checklist of Velloziaceae. 2019. Facilitated by the Royal Botanic Gardens, Kew. http://wcsp.science.kew.org/ [accessed 22 July 2019].

Graham, S.W., Reeves, P.A., Burns, A.C.E., Olmstead, R.G. 2000. Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int. J. Plant Sci. 161, S83–S96.

Graham, S.W., Zgurski, J.M., McPherson, M.A., Cherniawsky, D.M., Saarela, J.M., Horne, E.S.C., Smith, S.Y., Wong, W.A., O'Brien, H.E., Biron, V.L., Pires, J.C., Olmstead, R.G., Chase, M.W., Rai, H.S. 2006. Robust inference of monocot deep phylogeny using an expanded multigene plastid data set. Aliso, 22, 3–20.

Graham, S.W., Lam, V.K.Y., Merckx, V.S.F.T. 2017. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. New Phytol. 214, 48–55.

Gualberto, J.M. and Newton, K.J. 2017. Plant mitochondrial genomes: dynamics and mechanisms of mutation. Annu. Rev. Plant Biol. 68, 225–252.

Guisinger, M.M., Chumley, T.W., Kuehl, J.V., Boore, J.L., Jansen, R.K. 2010. Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in Poaceae. J. Mol. Evol. 70, 149–166.

Guisinger, M.M., Kuehl, J.V., Boore, J.L., Jansen, R.K. 2011. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. Mol. Biol. Evol. 28, 583–600.

Hammel, B.E. and Wilder, G.J. 1989. *Dianthoveus*: a new genus of Cyclanthaceae. Ann. Mo. Bot. Gard. 76, 112–123.

Hansen, D.R., Dastidar, S.G., Cai, Z., Penaflor, C., Kuehl, J.V., Boore, J.L., Jansen, R.K. 2007. Phylogenetic and evolutionary implications of complete chloroplast genome sequences of four early-diverging angiosperms: *Buxus* (Buxaceae), *Chloranthus* (Chloranthaceae), *Dioscorea* (Dioscoreaceae), and *Illicium* (Schisandraceae). Mol. Phylogenet. Evol. 45, 547–563.

Harling, G. 1958. Monograph of the Cyclanthaceae. Acta Horti Berg. 18, 1–428.

Harling, G., Wilder, G.J., Eriksson, R. 1998. Cyclanthaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Vol. III, Flowering Plants. Monocotyledons: Lilianae (except Orchidaceae). Springer-Verlag, Berlin Heidelberg, pp. 202–215.

Harris, M.E., Meyer, G., Vandergon, T., Oberholzer Vandergon, V. 2013. Loss of the acetyl-CoA carboxylase (*accD*) gene in Poales. Plant Mol. Biol. Rep. 31, 21–31.

Havird, J.C., Noe, G.R., Link, L., Torres, A., Logan, D.C., Sloan, D.B., Chicco, A.J. 2019. Do angiosperms with highly divergent mitochondrial genomes have altered mitochondrial function? Mitochondrion 49, 1–11.

Hendy, M.D. and Penny, D. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38, 297–309.

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

Hertweck, K.L., Kinney, M.S., Stuart, S.A., Maurin, O., Mathews, S., Chase, M.W., Gandolfo, M.A., Pires, J.C. 2015. Phylogenetics, divergence times and diversification from three genomic partitions in monocots. Bot. J. Linn. Soc. 178, 375–393.

Iles, W.J.D., Smith, S.Y., Gandolfo, M.A., Graham, S.W. 2015. Monocot fossils suitable for molecular dating analyses. Bot. J. Linn. Soc. 178, 346–374.

Janssen, T. and Bremer, K. 2004. The age of major monocot groups inferred from 800+ *rbcL* sequences. Bot. J. Linn. Soc. 146, 385–398.

Kao, P.-C. and Kubitzki, K. 1998. Acanthochlamydaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Vol. III, Flowering Plants. Monocotyledons: Lilianae (except Orchidaceae). Springer-Verlag, Berlin Heidelberg, pp. 55–58.

Katoh, K., Misawa, K., Kuma, K., Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kelchner, S.A. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann. Mo. Bot. Gard. 87, 482–498.

Kim, Y.-D. and Jansen, R.K. 1994. Characterization and phylogenetic distribution of a chloroplast DNA rearrangement in the *Berberidaceae*. Plant Syst. Evol. 193, 107–114.

Kolodner, R. and Tewari, K.K. 1979. Inverted repeats in chloroplast DNA from higher plants. Proc. Natl. Acad. Sci. USA 76, 41–45.

Korall, P., Schuettpelz, E., Pryer, K.M. 2010. Abrupt deceleration of molecular evolution linked to the origin of arborescence in ferns. Evolution 64, 2786–2792.

Koressaar, T. and Remm, M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291.

Kubitzki, K. 1998a. Pentastemonaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Vol. III, Flowering Plants. Monocotyledons: Lilianae (except Orchidaceae). Springer-Verlag, Berlin Heidelberg, pp. 404–406.

Kubitzki, K. 1998b. Stemonaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Vol. III, Flowering Plants. Monocotyledons: Lilianae (except Orchidaceae). Springer-Verlag, Berlin Heidelberg, pp. 422–425.

Kubitzki, K. 1998c. Velloziaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Vol. III, Flowering Plants. Monocotyledons: Lilianae (except Orchidaceae). Springer-Verlag, Berlin Heidelberg, pp. 459–467.

Lam, V.K.Y., Soto Gomez, M., Graham, S.W. 2015. The highly reduced plastome of mycoheterotrophic *Sciaphila* (Triuridaceae) is colinear with its green relatives and is under strong purifying selection. Genome Biol. Evol. 7, 2220–2236.

Lam, V.K.Y., Merckx, V.S.F.T., Graham, S.W. 2016. A few-gene plastid phylogenetic framework for mycoheterotrophic monocots. Am. J. Bot. 103, 692–708.

Lam, V.K.Y., Darby, H., Merckx, V.S.F.T., Lim, G., Yukawa, T., Neubig, K.M., Abbott, J.R., Beatty, G.E., Provan, J., Soto Gomez, M., Graham, S.W. 2018. Phylogenomic inference *in extremis*: a case study with mycoheterotroph plastomes. Am. J. Bot. 105, 480–494.

Lanfear, R., Ho, S.Y.W., Davies, J., Moles, A.T., Aarssen, L., Swenson, N.G., Warman, L., Zanne, A.E., Allen, A.P. 2013. Taller plants have lower rates of molecular evolution. Nat. Commun. 4, 1879.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.

Laroche, J., Li, P., Maggia, L., Bousquet, J. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. Proc. Natl. Acad. Sci. USA 94, 5722–5727.

Larsson, A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30, 3276–3278.

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

Lu, Q., Ye, W., Lu, R., Xu, W., Qiu, Y. 2018. Phylogenomic and comparative analyses of complete plastomes of *Croomia* and *Stemona* (Stemonaceae). Int. J. Mol. Sci. 19, 2383.

Mabberley, D.J. 2008. Mabberley's Plant-Book: A Portable Dictionary of Plants, their Classifications and Uses (Third Edition). Cambridge University Press, Cambridge, UK.

Maddison, W.P. and Maddison, D.R. 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04 http://mesquiteproject.org

Magallón, S., Hilu, K.W., Quandt, D. 2013. Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. Am. J. Bot. 100, 556–573.

Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L., Hernández-Hernández, T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol. 207, 437–453.

Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M., Kowallik, K.V. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393, 162–165.

McKain, M.R., McNeal, J.R., Kellar, P.R., Eguiarte, L.E., Pires, J.C., Leebens-Mack, J. 2016. Timing of rapid diversification and convergent origins of active pollination within Agavoideae (Asparagaceae). Am. J. Bot. 103, 1717–1729.

McLoughlin, S. 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Aust. J. Bot. 49, 271–300.

McPherson, M.A., Fay, M.F., Chase, M.W., Graham, S.W. 2004. Parallel loss of a slowly evolving intron from two closely related families in Asparagales. Syst. Bot. 29, 296–307.

Mello-Silva, R. 2005. Morphological analysis, phylogenies and classification in Velloziaceae. Bot. J. Linn. Soc. 148, 157–173.

Mello-Silva, R., Santos, D.Y.A.C., Salatino, M.L.F., Motta, L.B., Cattai, M.B., Sasaki, D., Lovo, J., Pita, P.B., Rocini, C., Rodrigues, C.D.N., Zarrei, M., Chase, M.W. 2011. Five vicarious genera from Gondwana: the Velloziaceae as shown by molecules and morphology. Ann. Bot. 108, 87–102.

Menezes, N.L., Mello-Silva, R., Mayo, S.J. 1993. A cladistic analysis of the Velloziaceae. Kew Bull. 49, 71–92.

Mennes, C.B., Smets, E.F., Moses, S.N., Merckx, V.S.F.T. 2013. New insights in the long-debated evolutionary history of Triuridaceae (Pandanales). Mol. Phylogenet. Evol. 69, 994–1004.

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

Miller, M.A., Pfeiffer, W., Schwartz, T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, USA, pp 1–8.

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

Mulcahy, D.G., Noonan, B.P., Moss, T., Townsend, T.M., Reeder, T.W., Sites Jr., J.W., Wiens, J.J. 2012. Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. Mol. Phylogenet. Evol. 65, 974–991.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., Ozeki, H. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322, 572–574.

Palmer, J.D. and Herbon, L.A. 1988. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J. Mol. Evol. 28, 87–97.

Palmer, J.D. 1991. Plastid chromosomes: structure and evolution. In: Bogorad, L., Vasil, I.K. (Eds.), The Molecular Biology of Plastids. Cell Culture and Somatic Genetics of Plants, Volume 7A. Academic Press, San Diego, USA, pp. 5–53.

Palmer, J.D., Adams, K.L., Cho, Y., Parkinson, C.L., Qiu, Y.-L., Song, K. 2000. Dynamic evolution of plant mitochondrial genomes: Mobile genes and introns and highly variable mutation rates. Proc. Natl. Acad. Sci. USA 97, 6960–6966.

Parkinson, C.L., Mower,J.P., Qiu, Y.-L., Shirk, A.J., Song, K., Young, N.D., dePamphilis, C.W., Palmer, J.D. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. BMC Evol. Biol. 12, 1–12.

Petersen, G., Zervas, A., Pedersen, H.A.E., Seberg, O. 2018. Genome reports: Contracted genes and dwarfed plastome in mycoheterotrophic *Sciaphila thaidanica* (Triuridaceae, Pandanales). Genome Biol. Evol. 10, 976–981.

Plunkett, G.M. and Downie, S.R. 2000. Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. Syst. Bot. 25, 648–667.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904.

Raubeson, L.A. and Jansen, R.K. 1992. Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science 255, 1697–1699.

Raubeson, L.A. and Jansen, R.K. 2005. Chloroplast genomes of plants. In: Henry, R.J. (Ed.), Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants. CABI Publishing, Wallingford, UK, pp. 45–68.

Ross, T.G., Barrett, C.F., Soto Gomez, M., Lam, V.K.Y., Henriquez, C.L., Les, D.H., Davis, J.I., Cuenca, A., Petersen, G., Seberg, O., Thadeo, M., Givnish, T.J., Conran, J., Stevenson, D.W., Graham, S.W. 2015. Plastid phylogenomics and molecular evolution of Alismatales. Cladistics 0, 1–19.

Rudall, P.J., Cunniff, J., Wilkin, P., Caddick, L.R. 2005. Evolution of dimery , pentamery and the monocarpellary condition in the monocot family Stemonaceae (Pandanales). Taxon, 54, 701–711.

Rudall, P.J. and Bateman, R.M. 2006. Morphological phylogenetic analysis of Pandanales: testing contrasting hypotheses of floral evolution. Syst. Bot. 31, 223–238.

Sajo, M.G., Lombardi, J.A., Forzza, R.C., Rudall, P.J. 2014. Comparative anatomy of reproductive structures in Cyclanthaceae (Pandanales). Int. J. Plant Sci. 175, 814–827.

Salatino, A., Salatino, M.L.F., Mello-Silva, R., van Sluys, M.-A., Giannasi, D.E., Price, R.A. 2001. Phylogenetic inference in Velloziaceae using chloroplast *TrnL-F* sequences. Syst. Bot. 26, 92–103.

Silvestro, D. and Michalak, I. 2012. raxmlGUI: a graphical front-end for RAxML. Org. Divers. Evol. 12, 335–337.

Silvestro, D., Cascales-Miñana, B., Bacon, C.D., Antonelli, A. 2015. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. New Phytol. 207, 425–436.

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

Skippington, E., Barkman, T.J., Rice, D.W., Palmer, J.D. 2015. Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proc. Natl. Acad. Sci. USA. 112, E3515–E3524.

Sloan, D.B., Barr, C.M., Olson, M.S., Keller, S.R., Taylor, D.R. 2008. Evolutionary rate variation at multiple levels of biological organization in plant mitochondrial DNA. Mol. Biol. Evol. 25, 243–246.

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

Smith, S.Y. 2013. The fossil record of noncommelinid monocots. In: Wilkin, P. and Mayo, S.J. (Eds.), Early Events in Monocot Evolution. Cambridge University Press, Cambridge, UK, pp. 29–59.

Soltis, P.S. and Soltis, D.E. 2003. Applying the bootstrap in phylogeny reconstruction. Stat. Sci. 18, 256–267.

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

Steane, D.A., Scotland, R.W., Mabberley, D.J., Olmstead, R.G. 1999. Molecular systematics of *Clerodendrum* (Lamiaceae): ITS sequences and total evidence. Am. J. Bot. 86, 98–107.

Steane, D.A. 2005. Complete nucleotide sequence of the chloroplast genome from the Tasmanian Blue Gum, *Eucalyptus globulus* (Myrtaceae). DNA Res. 12, 215–220.

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

Stone, B.C., Huynh, K.-L., Poppendieck, H.-H. 1998. Pandanaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Vol. III, Flowering Plants. Monocotyledons: Lilianae (except Orchidaceae). Springer-Verlag, Berlin Heidelberg, pp. 397–404.

Swofford, D.L. 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, USA.

Thiers, B. Continuously updated. Index Herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/science/ih/ [accessed 19 March 2019].

Tsudzuki, J., Nakashima, K., Tsudzuki, T., Hiratsuka, J., Shibata, M., Wakasugi, T., Sugiura, M. 1992. Chloroplast DNA of black pine retains a residual inverted repeat

lacking rRNA genes: nucleotide sequences of *trnQ*, *trnK*, *psbA*, *trnI* and *trnH* and the absence of *rps16*. Mol. Gen. Genet. 232, 206–214.

Ueda, M., Nishikawa, T., Fujimoto, M., Takanashi, H., Arimura, S.-I., Tsutsumi, N., Kadowaki, K.-I. 2008. Substitution of the gene for chloroplast RPS16 was assisted by generation of a dual targeting signal. Mol. Biol. Evol. 28, 1566–1575.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G. 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Res. 40, e115.

Vaidya, G., Lohman, D.J., Meier, R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180.

Wang, R.-J., Cheng, C.-L., Chang, C.-C., Wu, C.-L., Su, T.-M., Chaw, S.-M. 2008. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. BMC Evol. Biol. 8, 36.

Webb, C.O., Ackerly, D.D., Kembel, S.W. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 24, 2098–2100.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Müller, K.F., Quandt, D. 2011. The evolution of the plastid chromosome in land plants: gene conent, gene order, gene function. Plant Mol. Biol. 76, 273–297.

Wolfe, K.H., Li, W.-H., Sharp, P.M. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA 84, 9054–9058.

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

Zhu, A., Guo, W., Jain, K., Mower, J.P. 2014. Unprecedented heterogeneity in the synonymous substitution rate within a plant genome. Mol. Biol. Evol. 31, 1228–1236.

Zhu, A., Guo, W., Gupta, S., Fan, W., Mower, J.P. 2016. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytol. 209, 1747–1756.

TABLES

Table 1. Estimated ages for Pandanales inferred here and in previous studies. Three variant analyses performed here (columns 2–4) excluded mycoheterotrophic Triuridaceae and were based on (i) a subset of the plastid matrix comprising three genes, analyzed using a Bayesian random local clock (RLC) model; (ii) a subset of the mitochondrial matrix comprising 12 genes, analyzed using a RLC model and; (iii) a version of a matrix used in Mennes et al. (2013) with all mycoheterotrophs excluded here (analyzed with the Bayesian uncorrelated lognormal, UCLN, relaxed clock model as in the original study). Ages are shown for major Pandanales clades only (see Table S7 for other lineages). Ages are median estimates here (or median or mean estimates as noted in previous studies); 95% HPD (highest posterior density) intervals of ages shown in brackets when available (n/a = not applicable; age not reported or unspecified due to taxon sampling differences).

Clade/node	Plastid data (this study)	Mitochondrial data (this study)	Mennes et al. (2013) data, mycoheterotrophs excluded (analyzed here)	Mennes et al. (2013)[a]	Magallón et al. (2015)[b]	Givnish et al. (2018)[c]	Other studies[d, e]
PANDANALES							
Stem	106 (98–117)	84 (70–109)	106 (90–121)	110 (96–123)	110 (97–122)	123 (120–126)	131 (93–175)[d]
Crown	93 (85–104)	71 (58–94)	62 (38–90)	91 (69–110)	72 (53–100)	93 (86–102)	101 (91–118)[d]
Velloziaceae							
Stem	93 (85–104)	71 (58–94)	62 (38–90)	91 (69–110)	72 (53–100)	93 (86–102)	101 (91–118)[d]
Crown	51 (44–58)	34 (27–47)	31 (10–57)	37 (9–75)	49 (24–76)	n/a	

Stemonaceae							
Stem	81 (69–90)	64 (52–86)	43 (25–64)	46 (24–74)	54 (49–89)	76 (68–83)	–
Crown	65 (53–75)	46 (37–61)	27 (11–47)	27 (9–53)	n/a	50 (44–56)	–
Triuridaceae							
Stem	n/a	n/a	n/a	84 (62–103)	n/a	75 (68–82)	–
Crown	n/a	n/a	n/a	69 (50–90)	n/a	n/a	–
Pandanaceae							
Stem	61 (53–71)	56 (49–74)	37 (22–56)	38 (20–64)	n/a	44 (38–50)	85 (79–133)[e]
Crown	35 (23–51)	26 (20–35)	n/a	n/a	n/a	30 (26–35)	65 (42–98)[e]
Cyclanthaceae							
Stem	61 (53–71)	56 (49–74)	37 (22–56)	38 (20–64)	54 (49–89)	44 (38–50)	85 (79–133)[e]
Crown	49 (47–55)	49 (47–54)	22 (10–36)	21 (9–37)	n/a	18 (13–22)	45 (45–72)[e]

[a] Mennes et al. (2013): median ages and 95% HPD intervals from analyses implemented in BEAST using a UCLN model.

[b] Magallón et al. (2015): median ages and 95% HPD intervals from analyses implemented in BEAST using a UCLN model.

[c] Givnish et al. (2018): mean ages and 95% HPD intervals from analyses implemented in BEAST using a UCLN model.

[d] Alcantara et al. (2018): ages and 95% HPD intervals from analyses implemented in BEAST using a UCLN model.

[e] Gallaher et al. (2015): ages and 95% HPD intervals from analyses implemented in BEAST using a UCLN model and two fossil calibrations. Ages for the Cyclanthaceae-Pandanaceae stem node and Cyclanthaceae crown node represent the peak of the 95% HPD interval; the age for the Pandanaceae crown node is the median estimate.

LEGENDS TO FIGURES

Fig. 1. Phylogenetic relationships in Pandanales inferred from combined analysis of 119 organellar genes (37 mitochondrial and 82 plastid genes in photosynthetic taxa; 35 mitochondrial and 22 plastid genes in mycoheterotrophic *Sciaphila densiflora,* Triuridaceae), using partitioned likelihood (left-hand) and parsimony (right-hand) analyses of DNA data (see text for analysis details). Thick lines indicate 100% bootstrap support; values <100% are shown above branches. A circled support value in the parsimony tree indicates a branch that conflicts with the likelihood tree, also indicated by crossed dotted lines between taxa. Branches (labeled 'a-z') are summarized for other analyses in Table S5. The parsimony cladogram is one of ten shortest trees; an arrow indicates a Pandanales branch that collapses in the strict consensus tree (note that two additional branches collapse in outgroup taxa, Fig. S1). Inset phylograms show full taxon samplings of each analysis (Figs. S1, S4); shaded portions represent Pandanales, scale bars indicate estimated substitutions per site or number of changes.

Fig. 2. Phylogenetic relationships in Pandanales inferred from partitioned likelihood analyses of plastid and mitochondrial DNA data, each using a gene-by-codon partitioning scheme (see text for analysis details). The left-hand cladogram is from analysis of 82 plastid genes in photosynthetic taxa (22 in mycoheterotrophic *Sciaphila densiflora,* Triuridaceae), the right-hand cladogram from analysis of 37 mitochondrial genes (35 in *S. densiflora*). Thick lines indicate 100% bootstrap support; values <100% are shown at branches (values above vs. below branches in the plastid-based tree are with *S. densiflora* included or excluded, respectively; n/a = not applicable). Circled support values indicate branches that conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree); crossed lines indicate topological conflict between trees. Inset phylograms show full taxon samplings of each analysis (Figs. S6, S13); shaded portions represent Pandanales, scale bars indicate estimated substitutions per site.

Fig. 3. Phylogenetic relationships in Pandanales inferred from parsimony analyses of plastid and mitochondrial DNA data (see text for analysis details). The left-hand

cladogram shows relationships in the shortest tree from analysis of 82 plastid genes in photosynthetic taxa (22 in mycoheterotrophic *Sciaphila densiflora,* Triuridaceae), the right-hand cladogram shows one of two shortest trees from analysis of 37 mitochondrial genes (35 in *S. densiflora*; an arrow indicates a branch that collapses in the strict consensus tree). Thick lines indicate 100% bootstrap support; values <100% are shown above branches. Circled support values and labels indicate branches that conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree); crossed lines indicate topological conflict between trees. Two branches with moderate to strong support (labeled 'aa', 'bb') are summarized for other analyses in Table S5. Inset phylograms show full taxon samplings of each analysis (Figs. S2, S3); shaded portions represent Pandanales, scale bars indicate estimated number of changes.

Fig. 4. Clade ages of Pandanales estimated using a Bayesian random local clock (RLC) model. The tree topology and node heights (the latter representing median age estimates) are based on the three-gene plastid analysis (Fig. S20); bars represent 95% highest posterior density (HPD) intervals for node ages, derived from analyses of the same three-gene plastid data set (lower bars in dark grey) and a 12-gene mitochondrial data set (upper bars in lighter grey). Plastid and mitochondrial age estimates outside Pandanales are respectively shown in Figs. S20, S21. Scale bar indicates time (millions of years, Ma).

Fig. 5. Substitution rates in Pandanales inferred from a three-gene plastid data set (left-hand tree) and a 12-gene mitochondrial data set (right-hand tree) using a Bayesian random local clock (RLC) model. Branch colors (blue to red) summarize the range of branch rates, with rate shifts also indicated as substitutions per site per million years. Plastid and mitochondrial rates outside Pandanales are shown in Figs. S22 and S23, respectively.

LEGENDS TO SUPPLEMENTARY TABLES AND FIGURES

Table S1. Specimen source information for new data; herbarium abbreviations follow Thiers (continuously updated).

Table S2. GenBank accession numbers for newly sequenced plastid and mitochondrial gene sets (83 and 37 genes, respectively) and three fully circularized plastomes.

Table S3. Optimal substitution models for unpartitioned analyses and for final partitioning schemes inferred using PartitionFinder 2 (see text). DNA analyses were partitioned using a gene-by-codon ('G x C') scheme; amino-acid (AA) analyses were partitioned by gene. (i) Plastid, (ii) mitochondrial and (iii) combined organellar. (i-A) Unpartitioned 95-taxon DNA matrix including a mycoheterotrophic taxon (*Sciaphila densiflora*); (i-B) G x C partitioning scheme for the same 95-taxon DNA matrix; (i-C) G x C partitioning scheme for a 96-taxon DNA matrix that includes two mycoheterotrophic taxa (*S. densiflora* and *S. thaidanica*); (i-D) G x C partitioning scheme for a 94-taxon DNA matrix excluding both mycoheterotrophic taxa (*S. densiflora* and *S. thaidanica*); (i-E) G x C partitioning scheme for a 30-taxon Pandanales-focused DNA matrix that includes the plastid *ycf*1 locus; (i-F) Unpartitioned 95-taxon AA matrix; (i-G) Gene-based partitioning scheme for the same 95-taxon AA matrix; (ii-A) Unpartitioned DNA matrix; (ii-B) G x C partitioning scheme for DNA matrix; (ii-C) Unpartitioned AA matrix; (ii-D) Gene-based partitioning scheme for AA matrix; (iii-A) G x C partitioning scheme for a 33-taxon combined organellar DNA matrix. Genes are shown before the underscore and the 'pos' term after the underscore indicates the codon position (not applicable for *rrn* genes).

Table S4. Fossil calibrations for estimating divergence times and substitution rate shifts using a Bayesian random local clock (RLC) model. Each calibration was implemented using a uniform prior distribution, with minimum fossil ages as the lower bound and an upper bound of 151.8 My (the maximum age estimate for the angiosperm crown node in Silvestro et al., 2015). All calibrations were applied to plastid-based analyses; those

indicated with an asterisk were applied to mitochondrial-based analyses. See Iles et al. (2015) for additional details on fossil taxa.

Table S5. Summary of bootstrap support for Pandanales relationships across 15 phylogenetic analyses conducted here (Figs. 1–3, S1–S15). Listed branches are moderately to strongly supported in at least one analysis (labels 'a-z' indicated in Fig. 1, 'aa' in Figs. 3, S2 and 'bb' in Figs. 3, S3). Abbreviations for analyses: 'P' = parsimony, 'ML-u' = unpartitioned likelihood analysis, 'ML-p' = partitioned likelihood analysis (using a gene-by-codon scheme for DNA data and a gene-based scheme for amino-acid data). Branches with <50% bootstrap support are shown as '–'; n/a = branch not available.

Table S6. Characteristics of full circular plastomes representing the five Pandanales families. Gene counts only include those with intact reading frames; partial genes at IR boundaries are listed separately. Genes missing in *Acanthochlamys* and *Sciaphila* are with respect to the typical gene content of angiosperm plastomes.

Table S7. Estimated monocot ages inferred here and in previous studies. Three variant analyses performed here (columns 2–4) excluded mycoheterotrophic Triuridaceae and were based on (i) a subset of the plastid matrix comprising three genes, analyzed using a Bayesian random local clock (RLC) model; (ii) a subset of the mitochondrial matrix comprising 12 genes, analyzed using a RLC model and (iii) a version of a matrix used in Mennes et al. (2013) with all mycoheterotrophs excluded here (analyzed with the Bayesian uncorrelated lognormal, UCLN, relaxed clock model as in the original study). Ages are median estimates here (or median or mean estimates as noted in previous studies); 95% HPD (highest posterior density) intervals of ages shown in brackets when available (n/a = not applicable; age not reported or unspecified due to taxon sampling). Ages for major Pandanales clades are summarized in Table 1.

Figure S1. Pandanales phylogeny inferred from combined parsimony analysis of 119 organellar genes (37 mitochondrial and 82 plastid genes in photosynthetic taxa; 35

mitochondrial and 22 plastid genes in mycoheterotrophic *Sciaphila densiflora,* Triuridaceae) using DNA data (the right-hand tree in Fig. 1 depicts a subset of this taxon sampling). One of 10 shortest trees found for a Pandanales-focused matrix (excluding mycoheterotrophic *Sciaphila thaidanica*; arrows indicate branches that collapse in the strict consensus tree). Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the same combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated number of changes.

Figure S2. Angiosperm phylogeny inferred from parsimony analysis of 82 plastid genes (22 in mycoheterotrophic *Sciaphila densiflora*, Triuridaceae) using DNA data, showing the shortest tree found for a 95-taxon plastid matrix (excluding *Sciaphila thaidanica*, Triuridaceae). Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree); a strongly supported branch (labeled 'aa') is summarized for other analyses in Table S5. Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated number of changes. The left-hand tree in Fig. 3 depicts a subset of this taxon sampling.

Figure S3. Monocot phylogeny inferred from parsimony analysis of 37 mitochondrial genes (35 in mycoheterotrophic *Sciaphila densiflora*, Triuridaceae) using DNA data, showing one of two shortest trees found (an arrow indicates a branch that collapses in the strict consensus tree). Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree); a moderately supported branch (labeled 'bb') is summarized for other analyses in Table S5. Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated number of changes. The right-hand tree in Fig. 3 depicts a subset of this taxon sampling.

Figure S4. Pandanales phylogeny inferred from combined DNA-based likelihood analysis of 119 organellar genes (37 mitochondrial and 82 plastid genes in photosynthetic taxa; 35

mitochondrial and 22 plastid genes in mycoheterotrophic *Sciaphila densiflora,* Triuridaceae), using a gene-by-codon partitioning scheme for a Pandanales-focused matrix (excluding mycoheterotrophic *Sciaphila thaidanica*). Inset cladogram shows relationships in Cyclanthaceae. Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site. The left-hand tree in Fig. 1 depicts a subset of this taxon sampling.

Figure S5. Angiosperm phylogeny inferred from unpartitioned DNA-based likelihood analysis of 82 plastid genes (22 in mycoheterotrophic *Sciaphila densiflora*, Triuridaceae) for a 95-taxon plastid matrix (excluding *Sciaphila thaidanica*, Triuridaceae). Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S6. Angiosperm phylogeny inferred from partitioned DNA-based likelihood analysis of 82 plastid genes (22 in mycoheterotrophic *Sciaphila densiflora*, Triuridaceae) using a gene-by-codon partitioning scheme for a 95-taxon plastid matrix (excluding *Sciaphila thaidanica*, Triuridaceae). Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site. The left-hand tree in Fig. 2 depicts a subset of this taxon sampling.

Figure S7. Angiosperm phylogeny inferred from partitioned DNA-based likelihood analysis of 82 plastid genes (22 and 18 respectively in mycoheterotrophic *Sciaphila densiflora* and *S. thaidanica*, Triuridaceae) using a gene-by-codon partitioning scheme for a 96-taxon plastid matrix. Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S8. Angiosperm phylogeny inferred from partitioned DNA-based likelihood analysis of 82 plastid genes using a gene-by-codon partitioning scheme for a 94-taxon plastid matrix (excluding mycoheterotrophic *Sciaphila densiflora* and *S. thaidanica*, Triuridaceae). Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S9. Pandanales phylogeny inferred from partitioned DNA-based likelihood analysis of 83 plastid genes, including the *ycf*1 locus, using a gene-by-codon partitioning scheme for a Pandanales-focused matrix (excluding mycoheterotrophic *Sciaphila thaidanica*). Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree); a strongly supported branch (labeled 'cc') is summarized for other analyses in Table S5. Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S10. Angiosperm phylogeny inferred from unpartitioned amino-acid-based likelihood analysis of 78 plastid genes (18 in mycoheterotrophic *Sciaphila densiflora*, Triuridaceae) for a 95-taxon plastid matrix (excluding mycoheterotrophic *Sciaphila thaidanica*). Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S11. Angiosperm phylogeny inferred from partitioned amino-acid-based likelihood analysis of 78 plastid genes (18 in mycoheterotrophic *Sciaphila densiflora*, Triuridaceae) using a gene-based partitioning scheme for a 95-taxon plastid matrix (excluding mycoheterotrophic *Sciaphila thaidanica*). Inset cladogram shows relationships in Cyclanthaceae. Red branches conflict with the likelihood analysis of the combined

organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S12. Monocot phylogeny inferred from unpartitioned DNA-based likelihood analysis of 37 mitochondrial genes. Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S13. Monocot phylogeny inferred from partitioned DNA-based likelihood analysis of 37 mitochondrial genes using a gene-by-codon partitioning scheme. Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site. The right-hand tree in Fig. 2 depicts a subset of this taxon sampling.

Figure S14. Monocot phylogeny inferred from unpartitioned amino-acid-based likelihood analysis of 37 mitochondrial genes. Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S15. Monocot phylogeny inferred from partitioned amino-acid-based likelihood analysis of 37 mitochondrial genes using a gene-based partitioning scheme. Inset cladogram shows relationships in Cyclanthaceae. The red branch conflicts with the likelihood analysis of the combined organellar data (Fig. 1, left-hand tree). Thick lines indicate 100% bootstrap support; values <100% are shown at branches. Scale bar indicates estimated substitutions per site.

Figure S16. Circular plastome map of *Acanthochlamys bracteata* (Velloziaceae). Genes inside the circle are transcribed clockwise and those outside counterclockwise. The gray circle indicates GC content (innermost gray circle shows a 50% threshold). Thicker lines indicate the inverted repeat (IR) copies. Genes with introns are indicated with '*'; a short partial copy of *ycf*1 in the IR_B is indicated with 'ψ'. The junctions between the IR copies and the large single copy (LSC) regions are depicted in Fig. S19D.

Figure S17. Circular plastome map of *Benstonea copelandii* (Pandanaceae). Genes inside the circle are transcribed clockwise and those outside counterclockwise. The gray circle shows GC content (innermost gray circle shows a 50% threshold). Thicker lines indicate the inverted repeat (IR) copies. Genes with introns are indicated with '*'; short partial copies of *ycf*1 and *rps*19 in the IR_B are indicated with 'ψ'. The junctions between the IR copies and the large single copy (LSC) regions are depicted in Fig. S19B.

Figure S18. Circular plastome map of *Croomia japonica* (Stemonaceae). Genes inside the circle are transcribed clockwise and those outside counterclockwise. The gray circle shows GC content (innermost gray circle shows a 50% threshold). Thicker lines indicate the inverted repeat (IR) copies. Genes with introns are indicated with '*'; a short partial copy of *ycf*1 in the IR_B is indicated with 'ψ'. The junctions between the IR copies and the large single copy (LSC) regions are depicted in Fig. S19C.

Fig. S19. Linearized plastomes from each of the five Pandanales families. Differences between plastomes from photosynthetic families Cyclanthaceae, Pandanaceae, Stemonaceae and Velloziaceae (A-D) imply shifting boundaries between the inverted repeat (IR) and large single-copy (LSC) regions, indicated by dotted lines. Homologous regions in the plastome of a non-photosynthetic Triuridaceae (E) are present as single copy, as the IR was likely lost for this taxon (see text). Asterisks (*) indicate genes with introns; scale bar shows plastome size (bp). Full plastomes are depicted in Figs. S16–S18 here (and Figs. 1, 2 in Lam et al., 2015)

Figure S20. Clade ages of monocots estimated using a three-gene plastid data set and a Bayesian random local clock (RLC) model. Node heights represent median age estimates; gray bars the 95% highest posterior density (HPD) intervals for node ages. Labeled nodes indicate fossil calibrations placements ('A–F'; Table S4); scale bar shows time (millions of years, Ma). The tree in Fig. 4 depicts a subset of this taxon sampling.

Figure S21. Clade ages of monocots estimated using a 12-gene mitochondrial data set and a Bayesian random local clock (RLC) model. Node heights represent median age estimates; gray bars the 95% highest posterior density (HPD) intervals for node ages. Labeled nodes indicate fossil calibration placements ('C', 'F'; Table S4); scale bar shows time (millions of years, Ma). Highest posterior density intervals within Pandanales are depicted in the tree in Fig. 4 (upper bars).

Figure S22. Substitution rates in monocots inferred from a three-gene plastid data set using a Bayesian random local clock (RLC) model. Branch colors (blue to red) summarize the range of branch rates, with rate shifts also indicated as substitutions per site per million years. The left-hand tree in Fig. 5 depicts a subset of this taxon sampling.

Figure S23. Substitution rates in monocots inferred from a 12-genes mitochondrial data set using a Bayesian random local clock (RLC) model. Branch colors (blue to red) summarize the range of branch rates, with rate shifts also indicated as substitutions per site per million years. The right-hand tree in Fig. 5 depicts a subset of this taxon sampling.

Figure S24. Clade ages of monocots estimated from a version of the matrix used for dating analyses in Mennes et al. (2013) with all mycoheterotrophs excluded here (analyzed with the Bayesian uncorrelated lognormal, UCLN, relaxed clock model as in the original study). Node heights represent median age estimates; gray bars the 95% highest posterior density (HPD) intervals for node ages. Scale bar shows time (millions of years, Ma).