Phylogenomics resolves deep subfamilial relationships in Malvaceae s.l.

Tijana Cvetković,1,2 Fabiola Areces-Berazain,3,4 Damien D. Hinsinger5,3,5 Daniel C. Thomas,6 Jan J. Wieringa,7 Santhana K. Ganesan,6 and Joeri S. Strijk1,6,8

1Institute of Parasitology, Biology Centre CAS, 37005 České Budějovice, Czech Republic
2Biodiversity Genomics Team, Guangxi University, Nanning, Guangxi 530005, China
3Alliance for Conservation Tree Genomics, 06000 Luang Prabang, Laos
4Herbarium UPRRP, Department of Biology, University of Puerto Rico, San Juan PR 00925-2537, USA
5Université Paris-Saclay, INRAE, Etude du Polyorphismes des Génomes Végétaux (EPGV), 91000 Evry, France
6National Parks Board, Singapore Botanic Gardens, Singapore 259569, Singapore
7Naturalis Biodiversity Center, 2333 CR Leiden, The Netherlands
8Institute for Biodiversity and Environmental Research, Universiti Brunei Darussalam, BE1410 Brunei Darussalam

Abstract

Malvaceae s.l., the most diverse family within Malvales, includes well-known species of great economic importance like cotton, cacao, and durian. Despite numerous phylogenetic analyses employing multiple markers, relationships between several of its nine subfamilies, particularly within the largest lineage /Malvadendrina, remain unclear. In this study, we attempted to resolve the relationships within the major clades of Malvaceae s.l. using plastid genomes of 48 accessions representing all subfamilies. Maximum likelihood and Bayesian analyses recovered a fully resolved and well-supported topology confirming the split of the family into /Byttneriineae (/Grewioideae + /Byttnerioideae) and /Malvadendrina. Within /Malvadendrina, /Helicteroideae occupied the earliest branching position, followed by /Sterculioideae, /Brownlowioideae, /Tilioidae, and /Dombeyioideae formed a clade sister to /Malvatheca (/Malvoideae + /Bombacoideae), a grouping morphologically supported by the lack of androgynophore. Results from dating analyses suggest that all subfamilies originated during hot or warm phases in the Late Cretaceous to Paleocene. This study presents a well-supported phylogenetic framework for Malvaceae s.l. that will aid downstream revisions and evolutionary studies of this economically important plant family.

Keywords: phylogenomics; Malvaceae s.l; next-generation sequencing; plastomes; historical diversification

Introduction

Malvaceae is a diverse and economically important family distributed throughout the tropical and temperate areas of both hemispheres (Bayer and Kubitzki 2003). Members of this family are widely used in agriculture, forestry and horticulture. Well-known examples are cotton (Gossypium spp.), cacao (Theobroma cacao), okra (Abelmoschus esculentus), durian (Durio zibethinus), cola (Cola spp.), baobab (Adansonia spp.), and the highly valued ornamental species of Hibiscus and Alcea (Bayer and Kubitzki 2003). Malvaceae account for about 70% of the diversity of the order Malvales. The family comprises 244 genera and over 4225 accepted species (Stevens 2001; Bayer and Kubitzki 2003; Christenhusz and Byng 2016), with new ones being regularly added (Alverson 1991; De Carvalho-Sobrinho and De Queiroz 2008; Bovini 2016; Areces-Berazain and Vega-Lopez 2019; Ganesan et al. 2020).

Earlier morphology-based studies (e.g., Bentham and Hooker 1867, Schumann 1890, Hutchinson 1926, Edlin 1935a, b) recognized just four of the ten families now included in the order Malvales: Bombacaceae, Malvaceae, Sterculiaceae, and Tiliaceae. However, the boundaries and relationships among the four groups were long considered problematic due to the difficulty in interpreting morphological traits that now appear to be homoplasious (Alverson et al. 1999; Le Péchon and Gigord 2014). Phylogenetic studies based on both morphological (Judd and Manchester 1997) and molecular data (Alverson et al. 1998, 1999; Baum et al. 1998; Bayer et al. 1999) confirmed the Malvanean identity of these families but revealed that each was nonmonophyletic. The traditional families were merged into an expanded Malvaceae, which was subdivided into nine subfamilies: Bombacoideae, /Brownlowioideae, /Byttnerioideae, /Dombeyioideae, /Grewioideae, /Helicteroideae, /Malvoideae, /Sterculioideae, and /Tilioidae (Bayer et al. 1999).

In most molecular phylogenetic analyses of Malvaceae s.l. (e.g., Baum et al. 1998, Alverson et al. 1999; Whitlock et al. 2001; Richardson et al. 2015; Hernández-Gutiérrez and Magallón 2019) subfamilies Grewioideae and Byttnerioideae, which include species previously placed in Tiliaceae and Sterculiaceae, respectively, form a well-supported clade that has been named /Malvadendrina. However, the boundaries and relationships among these subfamilies were long considered problematic due to the difficulty in interpreting morphological traits that now appear to be homoplasious (Alverson et al. 1999; Le Péchon and Gigord 2014). Phylogenetic studies based on both morphological (Judd and Manchester 1997) and molecular data (Alverson et al. 1998, 1999; Baum et al. 1998; Bayer et al. 1999) confirmed the Malvanean identity of these families but revealed that each was nonmonophyletic. The traditional families were merged into an expanded Malvaceae, which was subdivided into nine subfamilies: Bombacoideae, /Brownlowioideae, /Byttnerioideae, /Dombeyioideae, /Grewioideae, /Helicteroideae, /Malvoideae, /Sterculioideae, and /Tilioidae (Bayer et al. 1999).

Received: March 19, 2021. Accepted: April 11, 2021
© The Author(s) 2021. Published by Oxford University Press on behalf of Genetics Society of America.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Byttneriina. The remaining subfamilies are grouped in /Malvadendrina (Baum et al. 1998; Alverson et al. 1999; Nyffeler et al. 2005). Within /Malvadendrina, /Malvoideae, and /Bombacoideae form another well-supported monophyletic group known as /Malvatheca (Baum et al. 1998, 2004). In contrast, the positions of /Brownlowioidae, /Dombeyoideae, /Helicteroideae, and /Sterculioidae vary greatly between different datasets and studies and remain unresolved (Alverson et al. 1999; Bayer and Kubitzki 2003; Nyffeler et al. 2005). The lack of a well-resolved framework in /Malvadendrina has not only limited progress on classification, but also impaired our understanding of the evolution of morphological characters and the family’s biogeographic history (Alverson et al. 1999; Nyffeler et al. 2005).

Molecular studies in Malvaceae s.l. have typically relied on a small number of markers, which have failed to provide an adequate resolution at the subfamily level, particularly within /Malvadendrina. A recent study employing eight markers (six plastid, one nuclear, and one mitochondrial) (Hernández-Gutiérrez and Magallón 2019) retrieved /Helicteroideae sister to the rest of /Malvadendrina with high support, but relationships between the other subfamilies within the clade remained poorly supported. Also, relationships among tribes and early-branching genera within /Malvoideae and /Bombacoideae were notably different from those found in previous phylogenetic analyses (e.g., Baum et al. 2004; Areces-Berazain and Ackerman 2017).

The most straightforward way to address relationships that have not been resolved with small datasets is to increase sequence data. Analyses based on tens to hundreds of genes (i.e., genomic data) have provided resolution among major clades of Angiosperms, and successfully clarified difficult relationships within many orders and families (Soltis et al. 2013). In Malvaceae, the potential of comparative genomics was recently explored by Conover et al. (2019) to infer the number and location of whole-genome multiplication events. Their phylogenetic analysis, based on 67 plastid genes for 8 species and transcriptome data for 11 species representing eight of the nine subfamilies, showed improved resolution within /Malvadendrina, albeit with moderate to low support. The extremely short branches at the base of /Malvadendrina recovered in this and previous studies (e.g., Alverson et al. 1999, Nyffeler et al. 2005) suggest rapid radiation of the subfamilies and contribute to explain prior difficulties in resolving subfamilial relationships (Conover et al. 2019).

Here, we assembled a plastome dataset of 35 species of Malvaceae s.l. to investigate the subfamilial relationships and establish a preliminary temporal framework for the evolution of the family. We used a genome-skimming approach to generate plastome sequences for 28 species in 25 genera, representing all subfamilies in Malvaceae s.l. and two outgroup families, Thymelaeaceae and Dipterocarpaceae.

Our objectives were (1) to assess whether this expanded plastome dataset can clarify the backbone phylogeny and relationships among subfamilies of Malvaceae s.l., especially at the base of /Malvadendrina; (2) to estimate molecular divergence times of subfamilies and major lineages (/Byttneriina, /Malvadendrina, and /Malvatheca).

Materials and methods

Taxon selection

We obtained samples from 28 species representing all nine subfamilies in Malvaceae s.l., plus species from two malvaceous families, Thymelaeaceae and Dipterocarpaceae. Samples were collected during field work in Yunnan (China), Gabon, Cameroon, and Singapore Botanic Gardens, or retrieved from silica-gel dried tissue samples from the Naturalis Biodiversity Centre collections (Leiden, The Netherlands). All vouchers were deposited in our herbarium (BGT, Brunei Darussalam). Additionally, we retrieved 19 plastome sequences from GenBank, 15 belonging to Malvaceae, three to Brassicaceae, and one to Thymelaeaceae (for accession numbers and voucher information see Supplementary Appendix S1).

DNA extraction, sequencing, and phylogenomic analyses

Standard DNA extraction protocols were followed with minor modifications (Doyle and Doyle 1990). Genomic DNA was purified from either fresh/frozen or silica-dried, then a 350 bp paired-end sequencing library was built and sequenced on an Illumina HiSeq2500 platform (San Diego, California, USA) with 150 bp reads. Library construction and sequencing were performed by Novogene (Beijing, China). Raw sequence reads were assembled into circular contigs using NovoPlasty v2.7.0 (Dierckxsens et al. 2016) and ORG.Asm v0.2.05 (ORG.ASM 2016) with default settings. The contigs were then imported in Geneious v10.0.5 (http://www.geneious.com; Kearse et al. 2012), for further assembly and refinement, as described by Cvetković et al. (2019). Plastomes were annotated with cpGAVAS (Liu et al. 2012) and GeSeq (Tillich et al. 2017), followed by manual curation, and aligned using the MAFFT v7 (Kato and Standley 2013) plugin in Geneious R10 with default settings, following Hinsinger and Strijk (2015) and Cvetković et al. (2017). The complete plastome sequences were submitted to DRYAD (https://datadryad.org/stash/share/LkbLwUlzW_GJ5rBDMHYZz69S19HzkNWY6fPOySH9tBQ).

We used ModelTest-NG v0.1.5 (Darriba et al. 2019) to select the best-fitting nucleotide substitution model for the plastome dataset [with one inverted repeat (IR) removed]. No partition was defined because, in our experience with plastome data, we have not observed an effect of partitioning on tree inference (see Areces-Berazain et al. 2020). A maximum likelihood (ML) tree was built with RAxML-NG v0.9.0 (Kozlov et al. 2019) using the GTR+I+G model. We performed an ‘all in one’ analysis with 20 parsimony starting trees and 1000 bootstrap replicates for the evaluation of branch support. The best-scoring ML tree was edited using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Molecular divergence time estimation

We performed a divergence time estimation analysis with the plastome dataset in BEAST v2.6.3 (Drummond and Rambaut 2007; Suchard and Rambaut 2009). We used an uncorrelated relaxed clock model (Drummond et al. 2006) and the Birth–Death model (Nee et al. 1994) as tree prior. The two hyperparameters of the clock model were assigned exponential distributions with an ukdMean.c: mean = 10.0, and ukdStdev.c: mean = 0.333 as specified in Areces-Berazain and Ackerman (2016).

Eight fossil calibrations were selected following a survey of the relevant literature (see Supplementary Appendix S2). The root node (corresponding to the crown node of Malvales) was constrained to a minimum age of 82 Ma based on the malvaceous fossil wood Bombaxoylon langstoni from the late Cretaceous (Campanian) of Texas (Wheeler and Lehman 2000). We applied a log-normal prior to this node age with a mean of 1.0 and SD of 1.25. The other fossil constraints were used to calibrate seven ingroup nodes representing major clades of Malvaceae s.l. Five of these fossils were considered in the study of
Hernández-Gutiérrez and Magallón (2019) although three of them were here assigned to different nodes (Supplementary Appendix S2).

The stem Sterculia (or crown Sterculioideae) was calibrated at 70 Ma based on fossil leaves of several species of Sterculia (S. washburnii and S. patagonica) described from the late Cretaceous of South America (Berry 1938). Together with Sterculiaephiyllum australis from the Cretaceous of Antarctica (Dutra and Batten 2000), these leaf impressions, which resemble those of modern Sterculia, are among the oldest records of Sterculioideae.

Two palynomorphs were used to constrain the age of stems Tilioideae and Brownlowioideae, respectively. For Tilioideae, we used Tilia-type pollen from the Late Cretaceous (72-66 Ma) of Canada (Rouse et al. 1970). The stem Brownlowioideae was calibrated with Discoidites boneensis, a palynomorph from the Paleocene (66-56 Ma) of Malaysia, similar to pollen of modern Brownlowia and Pentax (Muller 1968).

The stem Grewioideae was assigned a minimum age of 65 Ma based on the fossil wood Grewinium canalisum. The combined occurrence of tile cells and radial canals allows us to confidently assign this fossil wood to this subfamily (Wheeler et al. 2017). Another fossil wood, Guazuma santacrucensis from the middle Eocene (39 Ma) of Peru, and which shares many anatomical characteristics with the New-world genus Guazuma (Woodcock et al. 2019), was used to constrain the age of the crown group of Byttnerioideae.

Within /Malvatheca, the age of modern bombacoids was calibrated with Bombaccioides anniae, a fossil pollen from the Mid- to Late Paleocene (66–56 Ma) of Colombia very similar to the Bombax-type pollen of most Bombacoideae (Van Der Hammen 1954; Carvalho et al. 2011). The age of crown eumalvoid was constrained with Malvavichyum maconucis, found in Colombian mid- to late Paleocene (61.6–56 Ma) deposits. Leaves of this species exhibit several architectural features of modern malvoids and appear to be the oldest fossils assignable to this clade (Carvalho et al. 2011).

All calibrated ingroup nodes were assigned exponential priors with a mean of 1.0 and offset values corresponding to the fossil age. Two independent runs were conducted in the CIPRES with a mean of 1.0 and offset values corresponding to the fossil (Rambaut et al. 2018). The combined sampling trees every 4000 generations. Effective sample size values and convergence of the runs were assessed in Tracer v1.7.1 (Drummond and S. washburnii and S. patagonica) described from the late Cretaceous of South America (Berry 1938). Together with Sterculiaephiyllum australis from the Cretaceous of Antarctica (Dutra and Batten 2000), these leaf impressions, which resemble those of modern Sterculia, are among the oldest records of Sterculioideae.

Two palynomorphs were used to constrain the age of stems Tilioideae and Brownlowioideae, respectively. For Tilioideae, we used Tilia-type pollen from the Late Cretaceous (72-66 Ma) of Canada (Rouse et al. 1970). The stem Brownlowioideae was calibrated with Discoidites boneensis, a palynomorph from the Paleocene (66-56 Ma) of Malaysia, similar to pollen of modern Brownlowia and Pentax (Muller 1968).

The stem Grewioideae was assigned a minimum age of 65 Ma based on the fossil wood Grewinium canalisum. The combined occurrence of tile cells and radial canals allows us to confidently assign this fossil wood to this subfamily (Wheeler et al. 2017). Another fossil wood, Guazuma santacrucensis from the middle Eocene (39 Ma) of Peru, and which shares many anatomical characteristics with the New-world genus Guazuma (Woodcock et al. 2019), was used to constrain the age of the crown group of Byttnerioideae.

Within /Malvatheca, the age of modern bombacoids was calibrated with Bombaccioides anniae, a fossil pollen from the Mid- to Late Paleocene (66–56 Ma) of Colombia very similar to the Bombax-type pollen of most Bombacoideae (Van Der Hammen 1954; Carvalho et al. 2011). The age of crown eumalvoid was constrained with Malvavichyum maconucis, found in Colombian mid- to late Paleocene (61.6–56 Ma) deposits. Leaves of this species exhibit several architectural features of modern malvoids and appear to be the oldest fossils assignable to this clade (Carvalho et al. 2011).

All calibrated ingroup nodes were assigned exponential priors with a mean of 1.0 and offset values corresponding to the fossil age. Two independent runs were conducted in the CIPRES with a mean of 1.0 and offset values corresponding to the fossil (Rambaut et al. 2018). The combined sampling trees every 4000 generations. Effective sample size values and convergence of the runs were assessed in Tracer v1.7.1 (Rambaut et al. 2018). The sampled trees were combined with LogCombiner 2.6.0 (Rambaut and Drummond 2017a), but due to the large number of trees generated, we discarded 88% of them. The maximum clade credibility (MCC) tree was constructed with the remaining 60,000 trees using TreeAnnotator v 2.6.0 (Rambaut and Drummond 2017b).

Data availability
Complete plastome sequences are available through the Dryad Digital Repository, at https://datadryad.org/stash/share/LkblWUl2W_GJsrBDMfHYzggz69s19H2kNWiY6jPOySH9tBQ. Supplemental material available at figshare: https://doi.org/10.25387/g3.14245967.

Results
Plastome size and genomic data
We assembled 24 plastomes of Malvaceae s.l. species, three of Thymelaeaceae and one of Dipterocarpaceae. These plastomes varied in length from 151,071 bp (Sterculia lanceolata) to 172,707 bp (Edgeworthia chrysantha), and had the typical organization of a large single copy (~90 kb, 34.5% GC), small single copy (~20 kb, 31.5% GC) and two IRs (~25 kb, 42.7% GC). The length of the alignment, including the 20 sequences retrieved from Genbank, was 182,323 bp, with 43,475 (23.8%) of identical sites and 84.7% pairwise identity.

Phylogenomic analyses
The best-scoring ML tree based on plastomes was highly supported (BS ≥ 99), except for the clade formed by /Malvatheca plus subfamilies /Brownlowioideae, /Dombeyoideae, and /Tilioideae which was moderately supported (BS = 72). Within subfamilies, only two nodes in Bombacoideae had moderate support (BS = 66; 73, Figure 1). Monophyletic /Byttneriina (/Grewioidae and /Byttneriidae) was retrieved as sister to /Malvadendrina, a clade composed of all other Malvaceae s.l. subfamilies (BS = 100; Figure 1). Within /Malvadendrina, /Helicteroideae was retrieved as sister to a clade including the other /Malvadendrina subfamilies (/Sterculioideae, /Brownlowioideae, /Dombeyoideae, /Tilioideae, /Bombacoideae, and /Malvoideae). The latter clade consists of /Sterculioideae which is sister to the rest, and a /Brownlowioideae- /Dombeyoideae- /Tilioideae clade that is sister to the /Malvatheca clade (/Bombacoideae and /Malvoideae).

Dating analyses
The Bayesian MCC tree resulting from the analysis of the plastome dataset in BEAST2 was well-supported and fully congruent with the ML tree (Figure 2). The molecular dating analysis placed the origin of Malvaceae in the Late Jurassic, about 153 Ma, albeit with a very wide highest posterior density (HPD) interval (95–208, 95% HPD; Figure 2 and Supplementary Appendix S3). The split of /Malvatheca and /Byttneriina (corresponding to the crown node of Malvaceae s.l.) was dated to the late early Cretaceous (110 Ma, 86–143; 95% HPD). Diversification of /Malvatheca was estimated to have begun around 93 Ma (79–112; 95% HPD) with the separation of Helicteroideae. Sterculioideae diverged shortly after (about 84 Ma, 75–100; 95% HPD), followed by the split of /Malvatheca from its sister clade formed by /Brownlowioideae- /Dombeyoideae- /Tilioidae about 80 Ma (72–95; 95% HPD). As shown in Figure 2 (see also Supplementary Appendix S3), differentiation of the lineages leading to the nine subfamilies occurred in the late Cretaceous, between 93 and 66 Ma.

Discussion
Genomics has increasingly enabled the use of phylogenomics to address phylogenetic relationships in Malvales over the last decade (Argout et al. 2011; Heckenhauer et al. 2018; Conover et al. 2019; Cvetković et al. 2019; Abdullah et al. 2020). Here, we used a genome skimming strategy for the inference of phylogenomic relationships in Malvales, with a focus on subfamily relationships in Malvaceae s.l. Plastome analyses are largely consistent with the results obtained in previous studies that used a limited number of loci (Alversen et al. 1998, 1999; Bayer et al. 1999; Nyffeler and Baum 2000; Whitlock et al. 2001; Baum et al. 2004; Nyffeler et al. 2005). However, plastome data clarified some subfamilial relationships that remained unresolved in previous analyses (see below) relying on single or few plastid loci (Stevens 2001; Smith and Baum 2003; Le Pechon and Gigord 2014) providing the best estimate of the plastome phylogeny of Malvaceae s.l. to date.

The results of our phylogenomic analyses provided strong support for the monophyly of the three previously recognized major clades within Malvaceae s.l., viz. /Byttneriina, /
Figure 1. ML phylogenomic tree using a plastome dataset inferred by RAxML. Bootstrap branch support shown at the nodes, unlabeled branches = 100 BS.

Figure 2. Chronogram using a plastome dataset inferred using BEAST 2. Node ages (in Ma) shown at nodes, with the 95% HPD intervals (blue bars). All nodes have posterior probabilities (PP) 1, except a node with gray circle (PP = 0.65). Calibrations used in this study (red circles): I. B. langstoni (84–74 Ma, Wheeler and Lehman 2000), II. G. canalisum (67–64 Ma, Wheeler et al. 2017), III. Guazuma santoncrunzais (39 Ma, Woodcock et al. 2019), IV. S. patagonica and S. waiburnii (72–66 Ma, Berry 1938, Dutra and Batten 2000), V. D. borneensis (66–56 Ma, Muller 1968), VI. T. gloriosa (72–66 Ma, Rouse et al. 1970), VII. B. anane (62–56 Ma, Van Der Hammen 1954, Carvalho et al. 2011), VIII. M. macondicus (62–56 Ma, Carvalho et al. 2011). Geological time scale in millions of years shown.

Malvadendrina, and /Malvatheca (Baum et al. 1998; Alverson et al. 1999; Smith and Baum 2003), which are discussed below. The characteristic morphological features and potential synapomorphies of the subfamilies and some major clades, based on our inferred plastome phylogenomic tree were presented in Figure 3.
New subfamilial phylogenetic relationships at the base of /Malvadendrina

Among Malvaceae s.l., the phylogenomic relationships at the base of /Malvadendrina (Brownlowioideae, Dombeyoideae, Helicteroideae, Sterculioideae, and Tilioideae) are arguably the most problematic as morphology-based and molecular classifications differ (Bayer et al. 1999; Whitlock et al. 2001; Nyffeler et al. 2005; Le Péchon and Gigord 2014). The boundaries between the former families of the “Core Malvales”: Sterculiaceae-Tiliaceae, Sterculiaceae-Bombacaceae, and Bombacaceae-Malvaceae “remain nebulous” (Baum et al. 1998), and would require substantial revision to accommodate the molecular-based subfamilies (Baum et al. 1998; Kubitzki and Chase 2003; Baum et al. 2004; Wilkie et al. 2006).

Traditional taxonomic placement of /Sterculioideae was uncertain, as it is a group with a highly variable morphology, anatomy and palynology (Wilkie et al. 1998; Takhtajan 2009). The phylogenetic placement of /Sterculioideae within Malvadendrina was previously unclear (Wilkie et al. 2006); it was retrieved as sister to /Brownlowioideae (Alverson et al. 1999; Bayer et al. 1999) or to /Malvatheca (Nyffeler et al. 2005; Hernández-Gutiérrez and Magallón 2019). In the study by Conover et al. (2019), /Sterculioideae formed a moderately supported sister clade with /Tilioideae, but /Brownlowioideae were not included in their dataset. Our analyses of the plastome dataset indicate that /Sterculioideae may be sister to a clade comprising a /Brownlowioideae-/Dombeyoideae-/Tilioideae subclade, and /Malvatheca (Bombacoideae and /Malvaceae), but this relationship was only weakly supported (BS = 72). The /Brownlowioideae-/Dombeyoideae-/Tilioideae+ /Malvatheca clade can be characterized by the absence of androgynophore (Figure 3), although this structure is uniquely present in Pterospermum (Dombeyoideae) and is rudimentary in a few genera of Brownlowioideae (Bayer and Kubitzki 2003).

/Dombeyoideae were strongly supported as sister to /Tilioideae (BS = 100) in our analysis of the plastid dataset, corroborating previous studies (Nyffeler et al. 2005). In Hernández-Gutiérrez and Magallón (2019), a sister clade of /Dombeyoideae and /Brownlowioideae was unsupported. A putative synapomorphy for this grouping might be the presence of free (or nearly free) sepals, which are typically connate in the rest of /Malvadendrina (Figure 3). The phylogenetic relationship of /Tilioideae within /Malvadendrina has remained elusive despite detailed studies of pollen morphology (e.g., Chambers and Godwin 1961, 1971) and various molecular phylogenetic studies (e.g., basal position in Malvadendrina (Bayer et al. 1999); close to /Malvatheca (Alverson et al. 1999); sister to /Brownlowioideae-/Sterculioideae-/Malvaceae (Nyffeler et al. 2005); moderately supported (BS = 80) sister clade of /Sterculioideae (Conover et al. 2019); poorly supported sister to a clade comprising /Sterculioideae and /Malvatheca (Hernández-Gutiérrez and Magallón 2019]). Previous studies comparing the anatomy of the petiole’s vascular system, placed polyphyletic /Tiliaceae close to /Sterculioideae-/Malvaceae (Conover et al. 2019).

/Brownlowioideae were strongly supported (BS = 99) as sister to the /Dombeyoideae-/Tilioideae clade in our plastome analysis, resolving polytomous or poorly supported relationships from previous studies (e.g., Alverson et al. 1999; Le Péchon and Gigord 2014; Hernández-Gutiérrez and Magallón 2019). The three subfamilies share the presence of five elongate staminodes, although
it is unclear whether these structures are homologous. Staminodes can be free or integrated into the staminal tube (Dombeyoideae) and can vary from linear to ovate or petaloid (Bayer and Kubitzki 2003). Such staminodes are absent in the sister /Malvatheca and in /Sterculioideae.

**Dating analyses**

Our estimates of divergence times based on eight fossil calibrations (Supplementary Appendix S2) revealed a relatively old origin for Malvaceae (152 Ma) compared with the estimate of 98 Ma inferred by Hernández-Gutiérrez and Magallón (2019). Similarly to their study, we assigned the oldest fossil, B. langstoni, to the stem (rather than crown Malvaceae) to provide a minimum age constraint of 82 Ma for the origin of this family. We used a lognormal distribution for this node as it places the highest probability on ages slightly older than the fossil.

Hernández-Gutiérrez and Magallón (2019), however, used a different strategy to calibrate the root of their tree. They employed a uniform prior with maximum and minimum age bounds of 103.2 and 89.44 Ma based on the age of the stem Malvales estimated in a previous study (Magallón et al. 2015). As a result, their time estimate for the origin of Malvaceae was younger than that obtained in this study. Our main reason for not following their approach to calibrate the root of our tree was to avoid the use of secondary calibrations, which has been shown to produce unreliable estimates and “a false impression of precision” (Schenk 2016).

The crown age of Malvaceae inferred in our study was 110 Ma (Figure 2), about 20 Ma older than that estimated by Hernández-Gutiérrez and Magallón (2019). However, the differentiation of the subfamilies, here estimated to have occurred in the upper Cretaceous-Paleocene interval (93–66 Ma; Figure 2 and Supplementary Appendix S3), was relatively consistent with the ages reported in their study (between 82 and 72 Ma). In addition to the different calibration approaches for the tree root, it is also very likely that our scarce taxon sampling plays a role in the different divergence time estimates reported. It is clear that an increased taxon sampling is needed in order to obtain robust age estimates as it has been shown that undersampling affects the accuracy of phylogenetic estimation (in terms of branch lengths) and thus the molecular divergence dates (Heath et al. 2008; Schulte 2013).

The stem group age estimates for the Malvaceae s.l. subfamilies suggest early onset of Malvaceae diversification occurred. Stem groups of most subfamilies originated during humid and hot conditions of the Late Cretaceous or a subsequent warm phase in the Paleocene. Most Malvaceae subfamilies are predominantly tropical, except for /Tilioidae, which is predominantly temperate, and /Malvoideae, which is widespread in both temperate and tropical habitats (Zachos et al. 2001; Huber et al. 2018; Hernández-Gutiérrez and Magallón 2019). Additional analyses incorporating a denser taxon sampling and fossil distributions are needed to infer ancestral geographic ranges and clarify where and how often niche shifts between tropical-humid, tropical-semi-arid, and temperate habitats occurred.

**/Bytterniina**

Our analyses confirmed the placement of /Bytterniina, the clade comprising /Grewioideae and /Bytternioidae, as sister to all remaining Malvaceae s.l. subfamilies (/Malvadendrina; Figures 1 and 2 and Supplementary Appendix S3). The plastome dataset strongly supports the monophyly of /Bytterniina, which is in contrast to some traditional classifications based on morphology [e.g., /Grewioideae included in /Tiliaceae (Hutchinson 1967; Takhtajan 1997)]. No morphological synapomorphies of /Bytterniina have been identified to date.

In this clade, /Bytternioidae can usually be characterized by a reduction in stamen number, stamens in multiple of five, arranged in groups opposite the petals, and by the distinctive hooded petals, often with a broad or cucullate base and an apical appendage (Whitlock et al. 2001; Judd et al. 2008), although the apical appendages may be homoplasious (Whitlock et al. 2001; Le Péchon and Gigord 2014).

/Grewioideae are supported by numerous, mostly free stamens, staminodes that are modified outer stamens and by the nectaries on petal bases or androgynophore (Whitlock et al. 2001; Judd et al. 2008; Brunken and Muellner 2012).

**/Malvadendrina**

No morphological synapomorphies characterizing /Malvadendrina have been identified.

/Helicteroideae is strongly supported as sister clade to the remaining taxa of /Malvadendrina (Alverson et al. 1999; Bayer et al. 1999; Conover et al. 2019; Hernández-Gutiérrez and Magallón 2019; this study). The presence of an androgynophore has been indicated as a potential synapomorphy for the subfamily (Stevens 2001; Judd et al. 2008). The placement of Durio within Bombacaceae based on morphology (Kostermans 1958, Fryxell 1968; Takhtajan 2009) was previously not supported by both chromosome numbers (Baum and Oginuma 1994) and molecular studies (Bayer et al. 1999, Nyffeler and Baum 2000, 2001; Nyffeler et al. 2005; Hernández-Gutiérrez and Magallón 2019). Our plastome-based molecular phylogeny placed Durio among /Helicteroideae.

Monophyly of /Sterculioideae is supported by a suite of unique characters including unisexual or polygamous flowers, coloured calyx of petaloid sepals and a loss of petals, secondary apocarpy, presence of an androgynophore, and woody follicle fruits (Figure 3; Stevens 2001; Bayer and Kubitzki 2003; Judd et al. 2008; Le Péchon and Gigord 2014) and was confirmed in several previous studies (Judd and Manchester 1997; Wilkie et al. 2006).

/Dombeyoideae are characterized by bilobed cotyledons, a short staminode tube, and large-sized, echinate pollen grains with 3-zonoporate apertures (Falque et al. 1992, Stevens 2001; Bayer and Kubitzki 2003; Perveen and Qaiser 2009; Hamdy and Shamsa 2010; Silveira Júnior et al. 2017). /Tilioidae have been characterized by their pollen structure (“tilioid pollen type”; Chambers and Godwin 1961, 1971; Bayer and Kubitzki 2003; Le Péchon and Gigord 2014), and the presence of folded cotyledons and staminodes opposite the petals have also been highlighted as potential synapomorphies (Figure 3; Stevens 2001; Bayer and Kubitzki 2003; Judd et al. 2008).

Potential synapomorphies of /Brownlowioideae include a strongly fused, bell-shaped calyx, and basally dilated and divergent anthers with apically contiguous anther sacs (Figure 3; Stevens 2001; Bayer and Kubitzki 2003; Cheek 2007; Judd et al. 2008).

**/Malvatheca**

The monophyly of /Malvatheca (/Bombacoideae and /Malvoideae) in our plastome analysis is consistent with numerous previous studies (Alverson et al. 1999; Bayer et al. 1999; Baum et al. 2004; Nyffeler et al. 2005; De Carvalho-Sobrinho et al. 2016). Anthers with transversely septate locules (in most species modified to 1-loculed half-anthers) were indicated as synapomorphy in some studies (e.g., Judd et al. 2008), but this character is likely...
homoplasious in /Malvatheca (Le Péchon and Gigord 2014). Le Péchon and Gigord (2014) indicate that the aspects of the androecial development as described in detail by von Balthazar et al. (2004, 2006) may be synapomorph, and Stevens (2001) additionally indicates the absence of a root hypoderms and stamen filaments forming a tube as potential synapomorphies (Figure 3).

/Bombacoideae are characterized by palmately compound leaf laminas, the triangular, nonspiny pollen, a unique chromosome number \(n = 36\), embryo morphology, seed anatomy and the absence of mucilaginous substances (Fuchs 1967; Hutchinson 1967; Fryxell 1968; Singh and Chauhan 1984; Baum and Oginuma 1994; El Naggar 2001; Bayer and Kubitzki 2003; Takhtajan 2009).

/Malvoideae, by far the largest (~1800 species) of the nine subfamilies in Malvaceae s.l., have been characterized by the frequent presence of spiny, usually panatoporous pollen grains and a staminal column with five apical teeth (Bayer and Kubitzki 2003; Judd et al. 2008).

We present a detailed, well-supported phylogenetic framework for Malvaceae s.l. that will aid downstream revisions and evolutionary studies of this iconic and economically important plant family.

Conclusions
Resolving the subfamilial phylogenomic relationships at the base of /Malvadendrina is a crucial step for further evolutionary analyses. The inclusion of additional nuclear data in particular could be of interest for further evolutionary and taxonomic studies in this complex group. Incorporating new techniques [e.g., MIG-seq (Suyama and Matsuki 2015); high-throughput sequencing of tar-seq (Jones and Good 2016; Bossert and Danforth 2018; Johnson et al. 2019)] could provide valuable data from rare or unplaced (Binh et al. 2018), extinct (Feigin et al. 2018), or herbarium specimens (Strijk et al. 2020).

Our study provides a phylogenetic framework with robust support at deep levels, suitable for downstream revisions and evolutionary studies of this iconic and economically important plant family. As such, it serves as a valuable genomic resource to further investigate the complex evolutionary history of the family at lower taxonomic levels. The nine subfamilies were retrieved as previously circumscribed confirming the promising potential of using next-generation sequencing data in studies of Malvales.

Acknowledgments
We would like to acknowledge the horticultural staff of Hangzhou Botanical Garden (Zhejiang, China), Jing Dong Subtropical Botanical Garden (Yunnan, China), Naturalis Biodiversity Center (Netherlands), Qin Xiu Shan Park (China), Singapore Botanic Gardens (Singapore), and Xishuangbanna Tropical Botanical Garden (Yunnan, China) for their assistance in our sampling.

Author contributions
T.C. performed the experiment and analyses and collected samples. D.D.H. and J.S.S. designed the experiment and collected samples. F.A.B. and D.C.T. performed the analyses. J.J.W, S.K.G., and D.C.T. contributed sampling materials. All authors contributed to the writing, editing, and review of the article, and approved the final version for submission.

Funding
This work was supported by grants from the China Scholarship Council (No. 2016GZXS580) to T.C. and from the China Postdoctoral Science Foundation (No. 2015MS82481 and 2016T08022) to D.D.H., and funding from the Guangxi University, the Bagui Scholarship team funding (No. C33600992001) and the provincial government of Guangxi Province (“100 Talents” Program; recruitment of overseas talents for colleges and universities in Guangxi) to J.S.S.

Conflicts of interest: None declared.

Literature cited
Abdullah F, Mehmood I, Shahzadi S, Waseem B, Mirza, et al. 2020. Chloroplast genome of Hibiscus rosa-sinensis (Malvaceae): comparative analyses and identification of mutational hotspots. Genomics. 112:581–591.
Alverson WS. 1991. A synopsis of Phragmotheca (Bombacaceae), with two new species and a new subspecies. Brittonia. 43:73–87.
Alverson WS, Kenneth KG, Baum DA, Chase MW, Swensen SMR, et al. 1998. Circumscription of the malvales and relationships to other rosidae: evidence from rcl-seq data. Am J Bot. 85:876–887.
Alverson WS, Whitlock B, Nyffeler R, Bayer C, Baum DA. 1999. Phylogeny of the core malvales from ndhF sequence data. Am J Bot. 86:1474–1486.
Areces-Berazain F, Ackerman JD. 2016. Phylogenetics, delimitation and historical biogeography of the pantropical tree genus Thespesia (Malvaceae, Gossypieae). Bot J Linn Soc. 181:171–198.
Areces-Berazain F, Ackerman JD. 2017. Diversification and fruit evolution in eumalvoids (Malvaceae). Bot J Linn Soc. 184:401–417.
Areces-Berazain F, Vega-Lopez VJ. 2019. Hamaea lanceolata (Malvaceae, Malvoideae), a new species from Chiapas and Oaxaca, Mexico. Phytotaxa. 404:121–126.
Areces-Berazain F, Wang Y, Hinsinger DD, Strijk JS. 2020. Plastome comparative genomics in maples resolves the infrageneric backbone relationships. PeerJ. e9483.
Argout X, Salse J, Aury JM, Gutilman MJ, Droog G, et al. 2011. The genome of Theobroma cacao. Nat Genet. 43:101–108.
Baum DA, Oginuma K. 1994. A Review of chromosome numbers in Bombacaceae with new counts for Adansonia. Taxon. 43:11–20.
Baum AD, William AS, Nyffeler R. 1998. A Durian by any other name: taxonomy and nomenclature of the core Malvales. Harv Pap Bot. 3:315–330.
Baum AD, Smith SD, Yen A, William AS, Nyffeler R, et al. 2004. Phylogenetic relationships of malvatheca (Bombacoideae and Malvoideae, Malvaceae sensu lato) as inferred from plastid DNA sequences. Am J Bot 91:1863–1871.
Bayer C, Fay MF, Bruijn AY, Savolainen V, Morton CM, et al. 1999. Support for an expanded family concept of Malvaceae within a recircumscribed order Malvales: a combined analysis of plastid atpB and rbcL DNA sequences. Bot J Linn Soc. 129:267–303.
Bayer C, Kubitzki K. 2003. Malvaceae. In: K Kubitzki, C Bayer, editors. Dicotyledons: The Families and Genera of Vascular Plants, Malvales, Capparales and Non-betalain Caryophyllales, Vol. 5, Berlin: Springer, p. 225–312.
Berry EW. 1938. Tertiary flora from the Rio Fichileufu. Argentina Spec Pap Geol Soc Am. 12:1–149.
Bentharn G, Hooker JD. 1867. Genera Plantarum ad Exemplaria Imprimis in Herbaris Kewensibus Servata Definita, vol. 1, part 1. London, UK: Reeve.
platform for the organization and analysis of sequence data. Bioinformatics. 28:1647–1649.
Koestermans AJGH. 1958. The genus Durio Adans. (Bombac.). Reinwardtia. 4:47–153.
Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference, version 0.9.0. Bioinformatics. 35:4453–4455.
Kubitzki K, Chase MW. 2003. Introduction to Malvales. In: K Kubitzki, Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. Kubitzki K, Chase MW. 2003. Introduction to Malvales. In: K Kubitzki, Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019.
Kostermans AJGH. 1958. The genus Durio Adans. (Bombac.). (Bombacaceae-Durioneae). Org Divers Evol. 1:165–178.
Kubitzki K, Chase MW. 2003. Introduction to Malvales. In: K Kubitzki, Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019.
Kostermans AJGH. 1958. The genus Durio Adans. (Bombac.). (Bombacaceae-Durioneae). Org Divers Evol. 1:165–178.
Kubitzki K, Chase MW. 2003. Introduction to Malvales. In: K Kubitzki, Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019.
Kubitzki K, Chase MW. 2003. Introduction to Malvales. In: K Kubitzki, Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019.
Woodcock DW, Meyer HW, Prado Y. 2019. The Piedra Chamana fossil woods (Eocene, Peru), II. IAWA J. 40:551–595.

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

Communicating editor: S. Mathews