Article type : Research Article

RESEARCH ARTICLE

Short Title: Matsunaga and Smith— Fruits reveal deep roots of palm diversity

Fossil palm reading: using fruits to reveal the deep roots of palm diversity

Kelly K.S. Matsunaga[1,2,3] and Selena Y. Smith[1,3]

Manuscript received 7 July 2020; revision accepted 22 October 2020

[1] Department of Earth and Environmental Sciences and Museum of Paleontology, University of Michigan, Ann Arbor, MI 48109, USA

[2] Current address: Department of Ecology and Evolutionary Biology & Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA

[3] Authors for correspondence (e-mails: matsunaga@ku.edu, sysmith@umich.edu); ORCID id for K. K. S. M. 0000-0002-9533-2934

Citation: Matsunaga, K. K. S. and S. Y. Smith. 2021. Fossil palm reading: using fruits to reveal the deep roots of palm diversity. *American Journal of Botany* 108(3): XXXX

DOI: XXXX

PREMISE: Fossils are essential for understanding evolutionary history because they provide direct evidence of past diversity and geographic distributions. However, resolving systematic relationships between fossils and extant taxa, an essential step for many macroevolutionary

studies, requires extensive comparative work on morphology and anatomy. While palms (Arecaceae) have an excellent fossil record that includes numerous fossil fruits, many are difficult to identify due in part to limited comparative data on modern fruit structure.
METHODS: We studied fruits of 207 palm species, representing nearly every modern genus, using X-ray microcomputed tomography. We then developed a morphological data set to test whether the fossil record of fruits can improve our understanding of palm diversity in the deep past. To evaluate the accuracy with which this data set recovers systematic relationships, we performed phylogenetic pseudofossilization analyses. We then used the data set to investigate the phylogenetic relationships of five previously published fossil palm fruits.
RESULTS: Phylogenetic analyses of fossils and pseudofossilization of extant taxa show that fossils can be placed accurately to the tribe and subtribe level with this data set, but node support must be considered. The phylogenetic relationships of the fossils suggest origins of many modern lineages in the Cretaceous and early Paleogene. Three of these fossils are suitable as new node calibrations for palms.
CONCLUSIONS: This work improves our knowledge of fruit structure in palms, lays a foundation for applying fossil fruits to macroevolutionary studies, and provides new insights into the evolutionary history and early diversification of Arecaceae.

KEY WORDS: Arecaceae; paleobotany; comparative morphology; phylogenetics; X-ray micro-CT; morphospace; evolution; plant anatomy; divergence-time analysis; biogeography

Answering many fundamental questions in evolutionary biology requires information from fossils because they provide direct evidence of the diversity and geographic distributions of organisms through time. For studies of extant lineages, fossils cannot usually be incorporated without taxonomic or phylogenetic information. Unfortunately, well-sampled comparative morphological data sets are lacking for many plant clades, hindering the inclusion of fossils in many types of analyses. Comparative work in plant morphology and anatomy is therefore crucial for filling this gap and paving the way for more integrative research in plant macroevolution.

We focus here on the Arecaceae (palms), which are a widespread tropical angiosperm family composed of approximately 2500 species organized into five subfamilies, 28 tribes, 27 subtribes, and 181 genera (Baker and Dransfield, 2016). They exhibit broad morphological and ecological diversity. Ranging in habit from acaulescent understory herbs to canopy trees and

heavily armed lianas, today, palms occupy nearly all terrestrial environments of the tropics from rainforests to arid deserts. Palms have been prominent components of terrestrial environments for the last ~85 million years, during the Late Cretaceous and Cenozoic. Unequivocal macrofossils first appear in Coniacian strata of the Late Cretaceous (~90–86 Ma; Berry 1914) and become widespread by the Maastrichtian (~72–66 Ma), demonstrated by numerous occurrences of mangrove palms (subfamily Nypoideae; Gee, 2001; Harley, 2006; Dransfield et al., 2008) and palm pollen (Harley and Baker, 2001; Pan et al., 2006; Vajda and Bercovici, 2014). Moreover, palms were ubiquitous in many Late Cretaceous and Cenozoic floras around the world and exhibited broad geographic ranges that extended into high latitudes during warm and equable climatic intervals like the Eocene (Eldrett et al., 2009; Pross et al., 2012; Greenwood and West, 2017).

The macrofossil record of palms consists primarily of vegetative organs like leaves and stems. Although readily recognized as palms owing to their distinctive morphology and anatomy, most leaf and stem specimens are difficult to identify and are frequently assigned to broad artificial groups (morphogenera) composed of potentially unrelated taxa that share a suite of general characters (Read and Hickey, 1972). Common palm morphogenera include *Sabalites* G. Saporta (costapalmate leaves), *Phoenicites* A. Brongniart (pinnate leaves), and *Palmoxylon* (stems). Morphogenera are useful for documenting the presence and abundance of palms in a fossil flora and provide important information about the geographic distribution of the family through time, the composition of regional floras, and environmental conditions (e.g., Greenwood and West, 2017; Reichgelt et al., 2018). However, these fossils tell us relatively little about the deep evolutionary history of palms, including past taxonomic diversity, the tempo of diversification, and the biogeographic history of lineages. Extensive comparative work on modern stem anatomy has improved our understanding of the taxonomic affinities of *Palmoxylon* (Thomas and De Franceschi, 2012, 2013; Thomas and Boura, 2015; Nour-El-Deen et al., 2017), but is predominantly focused on the subfamily Coryphoideae and is yet to be widely applied to the numerous previously described occurrences of *Palmoxylon*. Consequently, much of our understanding of palm diversification is still based on studies of extant species, which may not accurately capture the true diversity and distributions of palms through time, particularly if extinction and major range shifts have played prominent roles in their history (Lieberman, 2005; Meseguer et al., 2015).

Fossils of reproductive structures such as flowers and fruits are rare in the geologic record but can possess informative morphological characters essential for systematic placement of palms below the family or subfamily level (Manchester et al., 2010, 2016; Allen, 2015; Matsunaga et al., 2019). Fossil fruits and flowers could therefore be key for understanding when major lineages of palms originated and their distributions through time. However, interpreting the fossil record of palm reproductive structures and describing new specimens presents significant challenges. Sometimes important morphological and anatomical characters are simply not preserved in the fossils, precluding taxonomic placement, or taxonomic determinations are based on limited character evidence and thus fossils are treated cautiously. Other barriers to identification include the large size of the family, substantial morphological diversity and convergence of traits, and a lack of accessible comparative data on reproductive morphology. The latter is particularly true of features that are not relevant to field identification but are useful for studying fossils, such as internal anatomy. These factors converge to make morphological comparisons unwieldy and the potential for taxonomic misidentification or imprecision high. Phylogenetic analyses can help resolve some of these issues but performing them requires thorough documentation of morpho-anatomical characters across the family.

To address this gap in our understanding of palm fruit structure and character distributions across Arecaceae, we performed a genus-level survey of modern fruit morphology using X-ray microcomputed tomography (μCT) and synthesized information from the literature, including the *Genera Palmarum* (Dransfield et al., 2008), morphological cladistic (Baker et al., 2009), and anatomical studies (Essig, 1977, 2002; Essig et al., 1999, 2001; Chapin et al., 2001; Essig and Hernandez, 2002; Romanov et al., 2011; Bobrov et al., 2012a, b). We summarize the data here to serve as a resource for describing new fossils, placing fossils in phylogenetic trees, and understanding morphological diversity in palms. To test the utility and limitations of fruit characters for understanding the systematic relationships of fossil palms, we analyzed the phylogenetic relationships of five previously described fossils and evaluated the accuracy of fossil placements using an approach we refer to as "phylogenetic pseudofossilization analysis." Diagnostic character suites of the family, subfamilies, and tribes are discussed, and patterns of morphospace occupation are examined. Finally, we provide recommendations for recognizing fossil palm fruits, assigning them to extant lineages, and using them to better understand the evolutionary history palms.

<h1>MATERIALS AND METHODS

<h2>X-ray μCT survey

Specimens of extant palm fruits were obtained on loan from herbarium collections at the Fairchild Tropical Botanic Garden (FTG), Royal Botanic Gardens, Kew (K), L. H. Bailey Hortorium Herbarium (BH), and the Florida Museum of Natural History (UF). Some specimens were collected from the living collection on the grounds of Fairchild Tropical Botanic Garden and subsequently dried prior to μCT scanning. Scans were performed at the University of Michigan Earth and Environmental Sciences μCT facility (UM CTEES) on a Nikon XTH 225ST industrial CT system (Nikon Metrology, Brighton, Michigan, USA). Whole fruits were scanned using 40–105 kV and 100–210 μA of X-ray power. We aimed to maximize resolution while keeping the entire specimen in the field of view, and thus, effective pixel size ranged from around 4 μm in the smallest specimen (*Hemithrinax ekmaniana* Burret.) to 122.5 μm in the largest (*Lodoicea maldivica* (J.F.Gmel.) Pers., which exceeded the field of view of a single scan). Exposure was set to 1.00–2.83 s, depending on the specimen, averaging two frames per projection. Scan parameters for each specimen are archived along with scan data on MorphoSource (project P776). Scans of 207 species were made, representing nearly all currently accepted genera (Appendix S1). Specimens of the genera *Tectiphiala* H.E.Moore, *Masoala* Jum., *Laccospadix* H.Wendl. & Drude., *Ammandra* O.F.Cook, *Guihaia* J.Dransf., S.K.Lee & F.N.Wei, *Deckenia* H. Wendl. ex Seem., and *Wendlandiella* Dammer could not be obtained and some characters were scored based on descriptions in the literature. Genera erected since the publication of *Genera Palmarum* (Dransfield et al., 2008*; Lanonia* A.J.Hend. & C.D.Bacon, *Saribus* Blume, *Sabinaria* R.Bernal & Galeano, *Jailoloa* Heatubun & W.J.Baker, *Manjekia* W.J.Baker & Heatubun, and *Wallaceodoxa* Heatubun & W.J.Baker) were also excluded, because neither specimens nor adequate fruit descriptions were obtained. Fossil seeds of *Sabal bigbendense* Manch., Wheeler, & Lehman (Manchester et al., 2010) were also μCT scanned to determine if they could be included in phylogenetic analyses, using similar settings as for extant palm fruits.

<h2>Morphological data set

The morphological data set of fruits was modified from the matrix of Baker et al. (2009). This allowed inclusion of fruit and gynoecial characters that we were unable to observe in some species, and left open the possibility of including other vegetative and reproductive characters from the original matrix in future analyses. When possible, we used the same species to document fruit morphology, but occasionally another species was substituted. Multiple species of some genera were studied, which revealed that many fruit characters scored did not vary between congeneric species. Characters were scored based on observations made from μCT scans, drawings and descriptions from *Genera Palmarum* (Dransfield et al., 2008), and anatomical descriptions and illustrations from the literature (Essig, 1977, 2002; Essig et al., 1999, 2001; Chapin et al., 2001; Essig and Hernandez, 2002; Baker et al., 2009; Romanov et al., 2011; Bobrov et al., 2012a, b; Manchester et al., 2016). Some of the original characters of Baker et al. (2009) were recoded or rescored to match our character definitions and hierarchies. The final matrix used in the phylogenetic and morphospace analyses contained 45 fruit and gynoecial characters, 13 of which originated from the Baker et al. (2009) matrix (Appendix S2).

<h2>Morphospace analysis

Morphospace analyses were performed to visualize similarity among major lineages of palms based on our data set of fruit characters. Due to significant amounts of missing data for some species, outgroup taxa (*Dasypogon* and *Kingia*) and several palms were removed from the morphological data set prior to use in morphospace analyses (*Saribus*, *Tectiphiala*, *Wendlandiella*, *Masoala*, *Laccospadix*, and *Clinosperma*). We first computed a distance matrix from the morphological character matrix, using the Maximum Observable Rescaled Distance (MORD; Lloyd, 2016) as the distance metric. The MORD distance metric scales distances from zero to one, with one as the maximum possible distance based on the observed characters. We also applied a correction for hierarchical characters, introduced by Hopkins and St. John (2018), which proportionally weights primary characters by their secondary characters using a tuning parameter (here set to 0.5), rather than treating inapplicable characters as missing data. This correction was applied to 20 characters but could not be used for all hierarchical characters in our data set because of missing data in primary (dependent) characters for some species, which the correction cannot accommodate. This distance matrix was then used in a principal coordinate ordination to visualize similarity in a two-dimensional space using the Cailliez procedure to

correct for negative eigenvalues (Cailliez, 1983), which applies a constant to all distances that are equal to the largest negative eigenvalue. This correction often reduces the amount of variation explained by the first two principal coordinate (PCo) axes (Nordén et al., 2018; Schaeffer et al., 2020), but was necessary for this data set because the largest absolute negative eigenvalue was greater than the largest positive value (Nordén et al., 2018). The results were plotted using the first two principal coordinate axes, which were the axes capturing the highest amount of variance in the data, with convex hulls delimiting subfamilies and tribes. All analyses were performed in R (v. 3.5.2) using the packages *claddis* (v. 0.3.0; Lloyd, 2016) and *ape* (v. 5.3; Paradis & Schliep, 2019) (Appendix S3).

<h2>Phylogenetic analysis of fossils

Fossil fruits were scored in the morphological matrix based on descriptions from the literature, direct examination of specimens, or both. The fossils that were selected met the following criteria: they (1) could confidently be assigned to palms, (2) represented Eocene or older occurrences in the fossil record, and (3) preserved sufficient scorable morphological characters. Eocene and older fossils were prioritized because of their greater potential for refining divergence-time estimates within Arecaceae, as node-dating analyses use the oldest occurrences of lineages. These fossils were *Hyphaeneocarpon indicum* Bande, Prakash, & Ambwani emend. Matsunaga, S.Y.Sm., Manch., Srivastava, & Kapgate (Matsunaga et al., 2019) and *Palmocarpon drypeteoides* (Mehrotra, Prakash & Bande) Manch., Bonde, Nipunage, Srivastava, Mehrotra & S.Y.Sm. (Manchester et al., 2016) from the Maastrichtian-Danian Deccan Intertrappean Beds of India (67–64 Ma), *Coryphoides poulseni* Koch from the Danian Agatdal Formation of Greenland (64–62 Ma; Koch, 1972), *Friedemannia messelensis* Collinson, Manch. & Wilde from the middle Eocene Messel oil shales of Germany (~47 Ma; Collinson, Manchester, & Wilde, 2012), and *Nypa burtini* Brongniart from the Eocene London Clay Formation (~47 Ma; Reid & Chandler, 1933). We made direct observations of *Hyphaeneocarpon indicum* and *Palmocarpon drypeteoides*, but *Friedemannia messelensis*, *Coryphoides poulseni*, and *Nypa burtini* were scored based on the original publications. *Hyphaeneocarpon indicum* was scored for 42 out of 45 characters (93%), *Palmocarpon drypeteoides* was scored for 36 characters (80%), *Coryphoides poulseni* and *Nypa burtini* were scored for 17 characters (38%), and *Friedemannia messelensis* was scored for 12 characters (27%).

The phylogenetic placement of these five fossil taxa was analyzed using maximum likelihood in the program RAxML (Stamatakis, 2014) with posterior non-parametric bootstrapping to evaluate clade support (using the RAxML rapid bootstrap algorithim; Stamatakis et al., 2008). The initial analyses included all five fossils and a matrix containing both molecular and morphological characters, hereafter referred to as "total-evidence." We used the extended majority-rule stopping criterion implemented in RAxML to determine convergence of bootstrap replicates (the "autoMRE" option; Pattengale et al., 2009), which stopped the run after 400 replicates. The molecular data set was assembled to match the taxon sampling in the morphological data and included 10 genes from GenBank: *18S*, *atpB*, *matK*, *ndhF*, *PRK*, *rbcL*, *RPB2*, *rps16*, *trnL-trnF* intergenic spacer, and *trnQ-rps16* intergenic spacer (Appendix S4). Each gene was initially aligned using MAFFT (Katoh and Standley, 2013) and refined with PRANK (Löytynoja, 2014). Alignments were visually inspected for obvious errors, but ultimately no manual adjustments were made except to trim alignment edges. All partitions were concatenated using SequenceMatrix (v1.8; Vaidya et al., 2011). The molecular data were separated into 10 partitions, one for each gene, with each analyzed using a general time-reversible model of rate substitution with gamma-distributed rate variation among sites (GTR+G model). The morphological characters were analyzed using the Markov k (Mk) model of morphological character evolution, corrected for ascertainment bias (i.e., Mkv).

The total-evidence analysis yielded a best-scoring tree topology consistent with that obtained using molecular data alone (without fossils), but with lower bootstrap values in many parts of the tree (Appendix S5). We attributed the relatively low node support of the total-evidence analysis to two possible sources of uncertainty. First, node support for otherwise well-supported relationships might be reduced if any of the fossils function as rogue taxa. Since fossils are represented by relatively few characters (~0.002% of columns), uncertainty in their position could also be exacerbated by bootstrap resampling of the full alignment. Second, genus-level relationships within several tribes, such as Areceae and Trachycarpeae, are poorly resolved in most molecular phylogenetic analyses (Baker et al., 2009; Bacon et al., 2012; Faurby et al., 2016), and our data set is no exception. Several of the fossils were placed within these groups, making it impossible to determine whether low support in the total-evidence analysis reflected uncertainty in the molecular data or the placement of fossils (i.e., the morphological data).

To disentangle these two sources of uncertainty, we first performed an analysis containing all fossil species using a backbone topological constraint, and then did separate constrained analyses for each fossil. The backbone constraint limits the tree search to topologies that conform to the supplied tree; because the fossils are not included as tips in the constraint tree, they can move freely and their placement is determined by the distribution of morphological characters. The topological constraint removed uncertainties associated with the molecular data, while analyzing each fossil independently enabled us to evaluate their placement in the tree without the influence of other fossils. The tree used as a constraint was constructed using the same molecular data set and parameters as the total-evidence analysis (described above), but with the morphological data excluded to avoid circularity (Appendix S6). The constrained searches using this tree were performed in RAxML under the Mkv model, using 100 bootstrap replicates each (Appendix S5). Affinities for each fossil were determined based on the shallowest subtending node for which support was very high in the constrained analysis (over 95% bootstrap support). In other words, starting from the fossil tip, we moved down the tree until we encountered a node for which bootstrap support was higher than 95%. Other factors were also considered and are discussed for each fossil below (see Discussion).

<h2>Phylogenetic pseudofossilization analyses

We performed a series of "phylogenetic pseudofossilization" analyses to test the limitations of the morphological data set for resolving the phylogenetic relationships of fossils, and consequently, the presumed accuracy of our fossil placements. Specifically, we tested whether the phylogenetic position of individual living taxa, as independently inferred using molecular data, could be recovered using the morphological data. Tips were selected arbitrarily from each tribe and removed from the backbone constraint tree, resulting in a set of trees with a single missing tip. Individual phylogenetic analyses were then performed using each backbone constraint tree and the morphological matrix (fossils removed; no molecular data), following the same procedure described above for the constrained analyses of fossils. We hereafter refer to the resulting tree as the "inferred tree," the unmodified constraint tree as the "reference tree," and the extant species removed from the constraint tree as "pseudofossils."

Taxa selected for pseudofossilization exhibit a wide range of data completeness ranging from 38–98% of characters analyzed (17–44 characters analyzed out of 45; Table 1). This range

reflects the fact that not all characters could be scored for many extant specimens due to preservation (most fruits were dried and thus some characters were ambiguous), or because specimens were not obtained and only limited information was available in the literature. Additionally, some characters are hierarchical and thus there are many inapplicable characters for some taxa. The fossil fruits we analyzed were scored for 27–93% of the total characters (12–43 characters), with *Friedemannia messelensis* scored for the fewest. To make the pseudofossilization analyses more comparable to the fossils in terms of character sampling, we degraded three pseudofossils down to the same 12 characters scored for *F. messelensis* (Table 1): *Areca*, *Ptychosperma*, and *Syagrus*. In total, 40 pseudofossilization analyses were performed on 37 taxa.

To summarize the results of these analyses, we evaluated the topological accuracy of the best-scoring tree and the uncertainty in pseudofossil placement across bootstrap replicates. To measure accuracy, we calculated Robinson-Foulds (RF) distances between the reference tree and each inferred tree. Here, RF values are equivalent to the number of bipartitions (splits) not shared by the two trees. To evaluate uncertainty, we considered two things: (1) the shallowest subtending node with bootstrap support over 95%, and the higher taxonomic group to which that node corresponds; (2) the distribution of node support throughout the tree. Exploring the latter is important because, in phylogenetic analyses of fossils with a backbone constraint, nodes with less than 100% support indicate placement of the fossil outside that group in a proportion of replicates, reflecting uncertainty in the data. To summarize the distribution of node support driven by variation in placement of pseudofossil tips in bootstrap replicates, we calculated pairwise RF distances between the reference tree and each bootstrap tree (100 for each pseudofossil tip) and took the mean and max values (Table 1). RF distances were computed using the R package *phangorn* v.2.5.3 in R (Schliep, 2011).

<h1>RESULTS

<h2>Morphological diversity of palm fruits

Palms exhibit tremendous diversity in fruit morphology (Figs. 1–7). Fruit size varies from a few millimeters (*Geonoma* Willd.) to ~50 cm in length (*Lodoicea maldivica*; Dransfield et al., 2008). Pericarp structure ranges from completely fleshy and parenchymatous (e.g., Fig. 3A–C), to highly sclerenchymatous and fibrous (e.g., Figs. 2, 3D–L); the epicarp can be smooth, bumpy,

prickly, corky (Figs. 4D, 5A, and 6F), or scaly (Fig. 1A). Palm fruits also occupy a broad spectrum of overall shape from spherical to fusiform (Fig. 6H) to highly irregular and deeply lobed (Fig. 3K, L). The endosperm can be homogeneous or ruminate (e.g., Fig. 3A), a character that is sometimes labile within species. Embryo position within seeds can be basal (Fig. 6J), variously lateral (Figs. 1B, 4A, B), or apical (Fig. 3D, F). The gynoecia from which fruits develop are similarly varied; they range from completely apocarpous to fully syncarpous and are composed of one (Fig. 6J–N), three (Fig. 3B, J), or more than three carpels (Fig. 5A). Moreover, many palms are pseudomonomerous with two carpels aborting early in floral development, but often retaining carpel traces as trilobed stigmas and vestigial locules (Fig. 6E, G; Dransfield et al., 2008). Consequently, although most palms produce single-seeded fruits, many species have more than one seed. Ovules can be anatropous, hemianatropous, campylotropous, or orthotropous with placentation apical, variously lateral, or basal (Dransfield et al., 2008).

Development also plays a role in fruit diversity. Endocarp, used here for the hard inner tissue of the pericarp that surrounds the seed, can develop from different regions of the pericarp: the locular epidermis, inner zone of the pericarp, or from the middle zone of the pericarp (Murray, 1973; Romanov et al., 2011; Bobrov et al., 2012a). We recognize that some of these structures do not represent "true" endocarps in the developmental sense (i.e., derived from the innermost pericarp layer), but opted for the functional definition here for expediency and because it is more practical for discussing fossils that lack developmental information. Furthermore, several characters appear related to where growth is concentrated within the gynoecium during fruit development, although the specific processes responsible for these features are unclear. These characters include (1) the position of stigmatic remains in mature fruits, (2) the position of ovule and seed attachment, and (3) whether fruits with more than one seed are deeply lobed. Stigmatic remains can be located almost anywhere on the fruit, and their position may depend on whether more than one carpel produces a mature seed. Nevertheless, the position of stigmatic remains is consistent within genera, except in the few of genera where seed number affects stigmatic position. In addition, the location of ovule placentation within the gynoecium often differs from the position of seed attachment in the mature fruit; this occurs in 61 genera in Arecoideae, five genera of Ceroxyloideae, and eight genera of Coryphoideae. For this reason, seed attachment cannot be used to infer ovule placentation, and vice versa. Finally, multiseeded fruits derived from syncarpous gynoecia can be either deeply lobed, resembling 2–3 smaller

fruits conjoined at the base (Fig. 3K, L), or simple with all seeds enclosed in a continuous pericarp (Fig. 3B, J), a condition we hereafter refer to as "multilocular" for expediency.

This variation makes it difficult to circumscribe characters by which all palm fruits can be universally recognized, making identification of fossil palm fruits especially challenging. Nevertheless, all palm fruits share the following traits: they develop from uniovulate carpels with a superior ovary, seeds are albuminous at maturity (contain endosperm), and embryos are small, conical to cylindrical, straight, and occupy a relatively small fraction of mature seed volume (Dransfield et al., 2008). Although there are almost certainly other characters shared among fruits of all palms, this survey revealed few common characters for which there are no major exceptions. It is more useful, therefore, to focus on the features that characterize major groups within Arecaceae. We highlight some of these below, with the caveat that there are often exceptions within these groups.

<h3>Calamoideae

Subfamily Calamoideae currently comprises 17 genera grouped into three tribes and nine subtribes (Baker and Dransfield, 2016), and is phylogenetically sister to all other palms (Baker et al., 2009; Couvreur et al., 2011; Faurby et al., 2016). Some genera are acaulescent or arborescent (e.g., *Raphia* P.Beauv., *Metroxylon* Rottb.) but most are lianas—the rattan palms. Flowers are unisexual in most species, but a few genera have bisexual flowers or both. Fruits of Calamoideae can be readily distinguished from those of other palms by their distinctive epicarp composed of basally oriented, imbricate scales (Fig. 1). These scales develop basipetally from outgrowths of the ovary surface, a process that begins early in gynoecial development (Bobrov et al., 2012b). At maturity, the scales contain tissues of both the epicarp and the mesocarp. Stigmatic remains are always apical (Fig. 1B, D).

In most genera, the mesocarp is not well differentiated to the inside of the scales and there is no endocarp surrounding the seed (Fig. 1B–F). The exception is *Eugeissona* Griff., which has a pericarp containing numerous longitudinal fibrovascular bundles and a prominent endocarp derived from the central zone of the pericarp (Bobrov et al., 2012b). Seed number ranges from one to three, and multiseeded fruits are multilocular. Embryos are either basal or lateral (Fig. 1B) and seeds are always attached basally. In many genera, the seed coat is either unevenly thickened on one side or has a thick, fleshy sarcotesta to attract seed dispersers (Fig.

1D, F). In many species, the seed coat forms a deep lateral intrusion into the endosperm (postament; Fig. 1C), similar to that of some members of Coryphoideae.

<h3>Nypoideae

Nypoideae is monotypic and contains only the extant species *Nypa fruticans* Wurmb, the mangrove palm. *Nypa* fruits are large and borne in dense globose heads, resulting in individual fruits that are roughly obovate, often laterally compressed, and angular in transverse section with longitudinal ridges (Fig. 2). Stigmatic remains form a prominent apical nub, referred to as an "umbo" (Fig. 2A, B). Fruit anatomy and development was described in detail by Bobrov et al. (2012a). The endocarp is thick (Fig. 2B, D), derived from the middle zone of the pericarp (Bobrov et al., 2012a), and has a round basal germination pore (Fig. 2B, E). It also has a thin longitudinal ridge that protrudes into the seed (Fig. 2D). The endosperm ranges from deeply ruminate to homogeneous (nonruminate). Fruits are water dispersed and the pericarp is dry at maturity, containing numerous longitudinal bundles (Fig. 2B, D).

<h3>Coryphoideae

Subfamily Coryphoideae includes 47 genera in eight tribes (Baker and Dransfield, 2016). All extant genera have palmate or costapalmate leaves, except for *Phoenix* L. (date palm), *Caryota* L., and *Arenga* Labill. ex DC., which have pinnate leaves. Floral morphology varies, but nearly all Coryphoideae have gynoecia with prominent styles elevating the stigma (Dransfield et al., 2008), a character that is otherwise uncommon. Coryphoideae have considerable variation in fruit morphology and can be separated into two major clades based in part on gynoecial structure: a syncarpous clade composed of tribes Borasseae, Caryoteae, Chuniophoeniceae, and Corypheae (*Corypha* L.), and a clade containing tribes Trachycarpeae, Cryosophileae, and Phoeniceae (*Phoenix*) and Sabaleae (*Sabal* Adans.). Members of this clade are either completely apocarpous or synstylous with free ovaries; the exception is *Sabal*, which is syncarpous. For convenience we refer to it hereafter as the "apocarpous clade," bearing in mind that *Sabal* is syncarpous.

Within the syncarpous clade, Corypheae, Chuniophoeniceae, and Caryoteae have fleshy fruits lacking prominent endocarp and longitudinal bundles within the pericarp (Fig. 3A–C). *Corypha* and Chuniophoeniceae (Fig. 3A) produce single-seeded fruits, whereas all the other

tribes have up to three seeds (Fig. 3B–L). Multiseeded fruits can either be multilocular (Caryoteae, Borasseae—subtribe Lataninae; Fig. 3B, J) or deeply lobed (Borasseae—subtribe Hyphaeninae; Fig. 3K, L). Tribe Borasseae has distinctive fruits that are typically large and fibrous (Fig. 3D, L). Most notably, *Lodoicea maldivica* produces the largest fruits among palms and the largest seeds of all extant plants (Tomlinson, 2006; Bellot et al., 2020). All genera have thick endocarps that originate from the middle zone of the pericarp (like *Nypa* and *Eugeissona*; Bobrov et al., 2012a), form pyrenes around each seed (e.g., Fig. 3J), and have apical germination pores consisting of holes or thin regions of the endocarp above the embryo. In *Satranala* J.Dransf. & Beentje (Fig. 3E, L), instead of the seedling germinating through the pore, the endocarp splits into two valves to release the seed, which is a germination mode unique to *Satranala* (Dransfield et al., 2008).

Most genera in the apocarpous clade have simple, fleshy, single-seeded fruits, and several members of tribe Cryosophileae are unicarpellate (Uhl and Moore, 1971; Fig. 4). Stigmatic remains are apical in all genera, except for *Sabal*. Many have a thin endocarp derived from the middle zone of the pericarp (Bobrov et al., 2012a), and lack germination pores. In most genera, the seed coat either forms a postament (Fig. 4A, D–F) or is irregularly thickened along one side (Fig. 4B, E). Embryos are usually lateral within the seed, although a few genera have embryos that are apically or basally attached (Fig. 4).

<h3>Ceroxyloideae

Ceroxyloideae includes eight genera in three tribes: Cyclospatheae (*Pseudophoenix* H.Wendl. ex Sarg.), Ceroxyleae (*Ceroxylon* Bonpl. ex DC., *Juania* Drude, *Oraniopsis* (Becc.) J.Dransf., and *Ravenea* H.Wendl. ex C.D.Bouché), and Phytelepheae (*Ammandra* O.F.Cook, *Aphandra* Barfod, and *Phytelephas* Ruiz & Pav.). Most genera have a prominent endocarp at maturity, which lacks a germination pore. Among species for which fruit anatomy has been studied in detail, the endocarp is composed of a single layer of palisade sclereids derived from the locular epidermis, and sometimes additional layers of sclerenchyma from the inner pericarp (Bobrov et al., 2012a). Seed number varies within the subfamily. *Pseudophoenix* fruits have up to three seeds and are deeply lobed when multi-seeded, Ceroxyleae produce a single seed, and Phytelepheae have multilocular fruits with up to 10 seeds (Fig. 5). The pericarp of Phytelepheae is dry, and composed of numerous pointed, corky protrusions formed by clusters of large radial

fiber bundles (Fig. 5A). In contrast, the pericarp of Ceroxyleae and *Pseudophoenix* is mostly fleshy, lacking significant fiber and fibrovascular bundles (Fig. 5B–D). In some genera, the endocarp is discontinuous at the point of seed attachment, forming a "hilar seam" (Fig. 5B). We documented this trait only in *Pseudophoenix* but it could be present in other genera with prominent endocarps, which we were not able to observe.

<h3>Arecoideae

The Arecoideae are the largest subfamily, comprising 108 genera, 14 tribes, and currently 10 genera unplaced at the tribal level; the largest tribe, Areceae, includes 11 subtribes. Fruits develop from syncarpous gynoecia composed of three carpels and many species are pseudomonomerous. Pericarp structure is variable within the subfamily, ranging from fleshy with no endocarp (6D) to fibrous with thick endocarp (Fig. 6A–N). However, many species have multiple layers of prominent longitudinal fiber and fibrovascular bundles in the pericarp (Fig. 6K). Like Ceroxyloideae, most genera with well-developed endocarps have an innermost layer of palisade sclereids derived from the locular epidermis (Essig, 1977, 2002; Essig and Young, 1979; Essig et al., 1999, 2001; Chapin et al., 2001; Essig and Hernandez, 2002; Fig. 6F–G, I–N), although this character has not been investigated in all groups. Endocarp opercula are common throughout Arecoideae and were not found in any other subfamilies (Fig. 6A, C, J, M). Germination pores lacking opercula were also observed in several genera.

Several tribes of Arecoideae occupy nonoverlapping regions of morphospace (Fig. 7). Some overlapping tribes are closely related (e.g., Areceae and Euterpeae) and the monotypic tribes are variously distributed throughout the morphospace, sometimes positioned distantly from sister lineages (e.g., *Sclerosperma* G.Mann & H.Wendl.). Despite segregation of many tribes in these analyses, it is difficult to circumscribe informative character suites for some groups. Below, we focus on clades that were most distinctive and for which there are documented fossil occurrences.

Tribe Cocoseae includes 17 genera in three subtribes: Attaleinae, Bactridinae, and Elaeidinae (Figs. 6A–C). All cocosoid palms have thick endocarps, derived from the locular epidermis and inner zones of the pericarp, with three circular germination pores (although some teratological specimens have more than three). In many genera, the germination pores contain opercula (Fig. 6A, C). These germination pores are diagnostic of Cocoseae and their position can

be informative for subtribe classification. Basal or subbasal germination pores are found only among Attaleinae (Fig. 6C), while subapical pores occur only in Bactridinae and Elaeidinae. Lateral germination pores are found in both subtribes (Fig. 6A). Subapical and subbasal pores are defined here as those occurring in the upper or lower thirds of the endocarp, respectively; lateral germination pores are positioned within the middle third of the endocarp, usually at the midline. In some cases, lateral pores are just above or below the midline, consistent with the occurrence of subapical or subbasal pores in their respective subtribes. Seed number is variable, with some genera consistently producing one seed (e.g., *Cocos* L.) and others up to three (e.g., *Butia* [Becc.] Becc., *Attalea* Kunth). Multiseeded fruits are multilocular with a continuous endocarp around all seeds (Fig. 6C). One-seeded fruits have aborted locules adjacent to the fertile locule, with corresponding germination pores. The pericarp usually has several layers of longitudinal fiber and fibrovascular bundles that are either relatively uniform in size or exhibit a subtle size gradient (Fig. 6A, B).

The Areceae (Fig. 6J–N) form the largest tribe among palms but have a number of distinctive characters that make them potentially recognizable in the fossil record. All members are pseudomonomerous and thus fruits are single seeded, usually lacking obvious traces of abortive carpels in mature fruits, although trifid stigmatic remains are sometimes present. Fruits tend to be relatively elliptical in shape, often retain remnants of the perianth at maturity, and some have apical stigmatic remains that, together with the pericarp, form a beak at the apical end of fruits above the seed. However, these characters are not universal in the tribe, and many do not have apical stigmatic remains. Longitudinal bundles of the pericarp exhibit either a narrow size gradient, or they consist of smaller bundles intermixed with larger ones with massive fibrous sheaths (Fig. 6K). The presence of these massive longitudinal bundles is restricted to Areceae and Geonomeae. Most genera have a prominent but relatively thin endocarp consisting of a single layer of palisade sclereids derived from the locular epidermis and sometimes the innermost cells of the pericarp (Fig. 6J–N). The endocarp often forms an operculum (Fig. 6J, M), as well as a hilar seam (Fig. 6N). A few genera have apically attached seeds—a character unique to Areceae.

<h2>Morphospace analyses

The morphospace plots (Fig. 7) are based on the first two PCo axes, which together capture 22.53% of the variation in extant fruit structure (14.49% and 8.04%, respectively). The low amount of total variation captured by the first two PCo axes is typical of morphospace analyses of phylogenetic morphological data sets, which tend to have substantial missing or inapplicable data, and for which negative eigenvalue corrections are often necessary (Nordén et al., 2018; Schaeffer et al., 2020). Moreover, phylogenetic morphological data sets are often generated such that characters are relatively uncorrelated (for instance, as opposed to landmark data), and thus more difficult to compress onto a few axes. Although such plots do not capture all of the variation in data, they nevertheless provide useful visualizations of character space and are consistent with our observations of specimens. As perhaps expected, we found that for the most part tribes within a subfamily are nonoverlapping, but as a whole there are some parts of the morphospace occupied by all subfamilies. The exceptions are tribes Areceae and Euterpeae, which occupy a unique region of morphospace along with the monogeneric tribes Roystoneae and Sclerospermeae.

<h2>Phylogenetic relationships of fossil palms

Placement of the fossil palms was the same across all analyses, except for some weakly supported differences in sister-taxon relationships for *Friedemannia messelensis* between total-evidence and constrained analyses. Results of total-evidence and constrained analyses are summarized in Fig. 8.

<h3>Nypa burtini

Pyritized fruits of *N. burtini* from the Eocene London Clay Formation preserve details of pericarp structure, including the basal germination pore and lateral internal ridge of the endocarp (Reid and Chandler, 1933). Unlike modern *Nypa*, the endocarp ridge does not extend the full length of the seed and is only present in the basal half. All analyses placed *N. burtini* sister to extant *Nypa fruticans* (Fig. 8), with 76% bootstrap support in the total-evidence analysis and 100% bootstrap support using a backbone constraint.

<h3>Hyphaeneocarpon indicum

Fruits of *H. indicum* were described from the Deccan Intertrappean Beds of India, and are late Maastrichtian–early Danian (~67–64 Ma) in age (Matsunaga et al., 2019). Results of the phylogenetic analyses were consistent with those of Matsunaga et al. (2019), who used a more limited morphological data set. In all analyses, bootstrap support for placement of *Hyphaeneocarpon* in the Hyphaeninae crown group, allied with *Bismarckia* and *Satranala*, was high (93% total-evidence, 100% constrained; Fig. 8). In contrast to the results of Matsunaga et al. (2019), in which *Hyphaeneocarpon* is sister to *Satranala* (posterior probability 0.53), *Hyphaeneocarpon* is here resolved as sister to *Bismarckia* with fairly high bootstrap support (87% total-evidence, 90% constrained).

<h3>Coryphoides poulseni

Fruits and seeds of *Coryphoides poulseni* originate from the Danian (64–62 Ma) Agatdal Formation of Nuussuaq, West Greenland (Koch, 1972). Koch (1972) documented several informative characters including some aspects of pericarp anatomy, the presence of a prominent intrusion of the seed coat into the endosperm (postament), elongate raphe, basal seed attachment, and lateral embryo position. The total-evidence analysis places *C. poulseni* in subtribe Trachycarpeae, with 90% bootstrap support for inclusion of *C. poulseni* in the Trachycarpeae total group (crown + stem group; Fig. 8). *Corphyoides poulseni* is positioned sister to *Licuala* Wurmb in subtribe Livistoninae, but node support for relationships within the tribe is generally low. Using a backbone constraint, support for its placement in crown Livistoninae is high (100%).

<h3>Palmocarpon drypeteoides

Fruits of *Palmocarpon drypeteoides* from the Deccan Intertrappean Beds of India are three-seeded and multilocular, with a thick endocarp bearing three subbasal germination pores, a layer of palisade sclereids lining each locule, and pericarp with longitudinal fibrovascular bundles (Manchester et al., 2016). The total-evidence analysis indicates moderate support for the placement of *P. drypeteoides* in the Attaleinae (tribe Cocoseae) total group (92%) and crown (88%). Placement within crown Attaleinae, allied with genera exhibiting subbasal germination pores, receives strong support when the molecular topology is constrained (100%; Fig. 8).

<h3>Friedemannia messelensis

Fossils of *F. messelensis* are compressed fruits and seeds from the early Eocene Messel oil shale of Germany (Collinson et al., 2012). Fruits are elliptical, with persistent perianth at the base and apical stigmatic remains. Despite their preservation as lignitized compressions, which precludes observation of many fruit and seed characters, Collinson et al. (2012) documented several important characters. These included multiple layers of longitudinal bundles in the pericarp and, by dissection of seeds from fruit specimens, apically attached seeds with an apical hilum and elongate raphe. Apical seed attachment is rare in palms and is only found in some members of tribe Areceae. The total-evidence analysis recovered strong support for inclusion of *F. messelensis* in the Areceae total group (97%; Fig. 8). Analyses employing a backbone constraint indicated high bootstrap support for inclusion of *F. messelensis* in crown group Areceae (100%), within the western Pacific clade (Ptychospermatinae, Archontophoenicinae, Basseliniinae, Carpoxylinae, *Dransfieldia* W.J.Baker & Zona, and *Heterospathe* Scheff.; 100% bootstrap support).

<h2>Phylogenetic pseudofossilization analyses

Results of the phylogenetic pseudofossilization analyses and summary metrics are listed in Table 1 and Fig. 8 (see Appendix S7 for all trees). Of the 37 taxa analyzed, most were placed accurately to subfamily, tribe, or subtribe level. For those that were not placed in the correct higher-level groups, the shallowest well-supported node nearly always reflected the true affinities, albeit sometimes with low precision. Two taxa were recovered in the wrong clade with high support: *Sclerosperma* and *Roystonea*, both of which were placed in Areceae (Appendix S7). Both of these represent extreme cases of morphological convergence of some fruit and gynoecial characters within Arecoideae and are discussed in detail below (see Discussion section). Overall, there is no apparent relationship between the number of characters scored and the accuracy of phylogenetic placement (Table 1). Some genera were placed accurately using relatively few characters (e.g., *Saribus*), while others exhibit poor accuracy using either many (e.g., *Orania*) or few characters (e.g., *Wendlandiella*). The variation in accuracy for taxa scored for few characters is probably related to whether the group has one or a few key characters uniting them, and the fact that the informativeness of individual characters is heterogeneous across the tree; in other words, characters that are crucial in some groups are not relevant in

others. For taxa scored for many characters, inaccurate placement likely results from taxon sampling issues related to removing the taxon from the comparative data set. Most of the taxa that were placed inaccurately in the best-scoring tree were genera forming monogeneric tribes; in fact, only one of these taxa was placed correctly (*Podococcus* G.Mann & H.Wendl.). This likely reflects the absence of morphologically similar close relatives to these taxa, because each genus was sampled once. Additionally, removing the taxon to analyze its phylogenetic relationships likely erodes informative character distributions throughout the tree, making it more difficult to accurately place taxa when morphologically similar close relatives are lacking (see Discussion). Our findings are overall consistent with previous studies demonstrating that, for recovering fossil relationships, extant taxon sampling of morphological characters is more important than the amount of missing data in fossils (Guillerme and Cooper, 2016).

<h1>DISCUSSION

<h2>Diversity of fruit structure and phylogenetic relationships of fossils

Despite the diversity in palm fruit structure and apparent convergent evolution of many traits, fruit characters carry strong taxonomic signal particularly below the subfamily level. Principal coordinate ordination of fruit characters helps to visualize this (Fig. 7). Convex hulls around subfamilies overlap, but within subfamilies most tribes occupy distinct regions of morphospace. Overlapping tribes are often closely related to one another, such as Euterpeae and Areceae (Arecoideae), or Trachycarpeae and Cryosophileae (Coryphoideae). These patterns of morphospace occupation are congruent with our observation that circumscribing character suites for each subfamily, to the exclusion of others, is more difficult than it is for tribes and other major clades within subfamilies.

Fruit characters observed in this survey are also informative for understanding systematic relationships of fossil species using phylogenetics (Fig. 8). *Nypa burtini* was resolved as sister to extant *Nypa fruticans* using 17 characters (including inapplicable states). *Hyphaeneocarpon indicum* (42 characters) is positioned as a crown member of subtribe Hyphaeninae in tribe Borasseae (Coryphoideae), with which it shares characters such as apical embryos and germination pores, basal stigmatic remains, and aborted carpels in mature fruits. *Coryphoides poulseni* (17 characters) was placed in subtribe Livistoninae of tribe Trachycarpeae (Coryphoideae). These affinities are reasonable because the combination of a basal postament

(deep intrusion of the seed coat) and lateral embryo are found only in subtribe Livistoninae. *Palmocarpon drypeteoides* (36 characters) resolves as a crown member of subtribe Attaleinae in tribe Cocoseae (Arecoideae), allied with other genera bearing subbasal germination pores and more than one seed. *Palmocarpon drypeteoides* was placed in subtribe Attaleinae by Manchester et al. (2016) based on those characters and other anatomical similarities with cocosoid palms.

Friedemannia messelensis was scored for 12 characters and was recovered within the Pacific clade of tribe Areceae (Fig. 8). Subtribe and genus-level relationships within Areceae have low support in the tree used as a backbone constraint and are similarly poorly resolved in other trees (Baker et al., 2009, 2011; Faurby et al., 2016), but a Pacific clade has nevertheless been recovered repeatedly in other studies. Despite support for existence of a Pacific clade, affinities with this clade to the exclusion of the rest of Areceae are not adequately justified by the morphology of the fossils. *Friedemannia messelensis* does resemble modern Areceae fruits, including members of the Pacific clade, in the overall structure of the fruits and has an apical hilum indicating apical attachment of the seed within the fruit. Apical seed attachment is found in six palm genera, all in Areceae. However, the character is not restricted to the Pacific clade. For these reasons, the most conservative affinities for *Friedemannia* are with the Areceae crown group. Moreover, until phylogenetic relationships among modern members of Areceae are better understood, and thus extant character distributions are better resolved, we will have little confidence in the taxonomic placement of any Areceae fossil below the tribe level.

<h2>Key fruit characters for fossil identification

Some characters frequently preserved in fossils are especially useful for determining systematic affinities and should be given particular attention when describing new specimens. They include pericarp structure, endocarp anatomy, embryo position, seed attachment, position of stigmatic remains, seed coat structure, and any other structures that indicate carpel number, such as vestigial locules. Key features of the pericarp include the organization and relative size of longitudinal bundles, and presence of other sclerenchyma within the pericarp such as radial fiber bundles. Endocarp morphology, anatomy, and developmental origin can also be informative, although the latter may be impossible to document without developmental stages preserved. Endocarp germination structures are found in several distantly related groups, but the

number, form, and position usually have systematic value. For example, germination pores with opercula are found only in Arecoideae and, in core arecoids, they are always basal. In Cocoseae, three germination pores are present, and their position can be diagnostic for subtribes. Apical germination pores are found in Borasseae, while basal germination pores occur in *Nypa* and *Eugeissona* (Calamoideae). The position of stigmatic remains and embryo attachment tend to be conserved within clades, while seed attachment is more variable but can be useful in combination with other traits. For *Friedemannia messelensis*, seed attachment proved to be an essential character because apical seed attachment is restricted to tribe Areceae.

The characters preserved in fossils will inevitably vary and few fossils will preserve all the characters that can be observed in living species, but notably all the key characters outlined above have been documented in fossils. Certain characters require exceptional preservation, such as embryos and abortive carpels, and are unlikely to be preserved in many fossils. Even in exceptionally preserved fossils, surveying multiple specimens may be necessary to document such characters. Other systematically informative characters pertain to lignified and relatively degradation-resistant tissues of the fruit such as endocarp, sclerenchyma of the pericarp, and the seed coat. The fact that these tissues have relatively high preservation potential may lead to taphonomic biases that affect our ability to detect certain clades in the fossil record. For instance, groups with fibrous pericarps and thick endocarps (e.g., Cocoseae, Borasseae) may be more readily preserved and recognized in the fossil record, whereas those that produce fleshy fruits lacking extensively lignified tissues (e.g., Chamaedoreae, Chuniophoeniceae) are less likely to be both preserved and identified when their fossils are recovered. A corollary of this is that some groups are more likely to be represented in the fossil record as seeds rather than whole fruits, such as many of the members of Coryphoideae that have fleshy pericarp and no endocarp at maturity. Most specimens of *Coryphoides poulseni* are seeds and there was only one complete specimen with intact pericarp. Similarly, numerous seeds resembling those of several genera of Coryphoideae have been described from the Eocene London Clay Formation (Reid and Chandler, 1933), but very few whole fruits are known. Unfortunately, with this data set seed characters by themselves are not very informative without knowing how they are oriented within the fruit. Although characters could be recoded to better accommodate these fossils, we found that seeds by themselves have few informative characters that are easily coded in a discrete

character framework. Methods that quantify shape and accommodate continuous characters could therefore be promising for identifying fossil palm seeds in the absence of fruit characters.

<h2>Phylogenetic pseudofossilization

Phylogenetic pseudofossilization analyses of extant taxa demonstrate that this data set can accurately place fruit specimens up to the subtribe level, and that employing a stringent node-support criterion is important for avoiding erroneous fossil placements. These results make us confident that the taxonomic affinities we propose for the fossils are accurate, insofar as we understand the morphology of these fossils. Additionally, the extant taxa for which pseudofossilization analyses struggled or failed to place also illustrates the importance of taxon sampling for establishing the character distributions necessary for accurate inference (Guillerme and Cooper, 2016) and may have implications for phylogenetic analyses of fossils with novel character combinations.

Taxa belonging to monogeneric tribes presented the greatest challenges to pseudofossilization analyses. Many of these taxa define boundaries of the morphospace occupied by the clades in which they are nested, and also possess characters that are both absent in close relatives and homoplasic throughout the broader group of interest. If the distribution of an informative character is defined by the taxon removed from the data set for pseudofossilization, the analyses unsurprisingly struggle to produce accurate placements. Because of these features, such taxa can also serve as an analog for fossils that exhibit novel and/or homoplasic character combinations that are discordant with character distributions among extant taxa.

Encouragingly, in our pseudofossilization analyses the true relationships of most of these taxa were reflected in the best-supported node and the overall distribution of node support throughout the tree (Table 1). Based on this result, we strongly recommend that a strict node-support criterion (higher than 95% bootstrap) be used when analyzing the relationships of fossils using this data set, and that the total distribution of node support be considered. We found that in some cases support for inaccurate relationships came close to the 95% threshold, and that in these cases uncertainty was distributed broadly across multiple clades (i.e., high RF values) rather than restricted to a single group. Therefore, additional caution should be used in interpreting the affinities of a fossil if broad distributions of support values are observed using

this data set and method. We note that for all the fossils analyzed here, node support was narrowly distributed within the clades in which fossils were placed.

Although our data set performed well in most pseudofossilization analyses, two taxa, *Roystonea* and *Sclerosperma*, represent "worst-case scenarios" for fossil placement. Both are members of monogeneric tribes and were placed in the wrong clade (Areceae) with very high support. This result is maybe unsurprising considering their position in the morphospace plots: both are positioned far from closest relatives, and in regions occupied exclusively by members of Areceae and Euterpeae (Fig. 7). We suspect this is driven by convergent evolution of several characters that are absent in close relatives of the two genera and otherwise found only in Areceae, including pseudomonomery and some aspects of pericarp anatomy. However, to our knowledge neither of these taxa possess fruit characters that are strictly apomorphic in Areceae (such as apical seed attachment as in *F. messelensis*), and therefore, we might be hesitant to accept crown placement in the tribe if these were real fossils. It is possible that their firm placement in Areceae could also be partly related to idiosyncrasies of the phylogenetic inference methods we used. We therefore stress the importance of seeking agreement between different sources of taxonomic information (in this case, the full spectrum of observed morphology vs. the phylogenetic results based on a limited sample of that spectrum) and considering which characters might be supporting inferred relationships. Although we hope this worst-case scenario is unlikely to be replicated in practice, because any real-world application of this data set would not omit any extant taxa, it illustrates that it is possible that attempts to accurately place a fossil using this data set could be confounded by extreme morphological convergence. This is arguably a risk for any phylogenetic analysis of fossils, especially those of single structures rather than whole plants, and for these reasons we emphasize here the precept that phylogenetic results are best treated as hypotheses.

<h2>Palm fruit fossil record

Numerous fossil palm fruits and seeds have been documented from Cretaceous and Cenozoic localities worldwide. Many of these occurrences have been summarized by Harley (2006). Here we discuss a few of the oldest records, many of which need careful re-evaluation before being used in macroevolutionary studies of palms, as well as those used as calibration fossils. The oldest putative palm fruits are those of *Hyphaeneocarpon aegyptiacum* Vaudois-

Miéja & Lejal-Nicol from the Aptian (113–125 Ma) of Egypt (Vaudois-Miéja and Lejal-Nicol, 1987), but the age and morphology of the specimens make them questionable as early palms. The fossils are large, pyriform, and have a structure interpreted as the endocarp with a round germination pore. Unfortunately, no anatomy is preserved in the specimens. Although large, roughly pyriform fruits occur in some species of modern *Hyphaene*, few other characters are preserved that suggest affinities with palms. Two other fossil fruits that should be reinvestigated to confirm age and affinities with palms are *Cocoopsis* sp. Fliche and *Astrocaryopsis* sp. Fliche, described from Cenomanian strata of France (Fliche, 1894). In particular, *Cocoopsis* was described as having the characteristic endocarp pores of Cocoseae, but no figures of the specimens accompany the description.

Seeds of *Sabal bigbendense* and *Sabal bracknellense* from the Campanian Aguja Formation in Texas were placed in the modern genus based on morphological similarities (Manchester et al., 2010). The μCT scans of *S. bigbendense* performed by us revealed a thickened region of the seed coat along the hilum and a well-preserved embryo positioned laterally–subapically within the seed relative to the hilum (Fig. 4H, I). The thickening of the seed coat near the hilum and lateral–subapical embryo position support relationships with *Sabal* and the "apocarpous clade" more generally. Unfortunately, the lack of whole fruits precluded inclusion of *S. bigbendense* and other seed fossils in our phylogenetic analyses. Nevertheless, our observations of the original specimens do not refute the placement of *S. bigbendense* in *Sabal*. Its age and strong resemblance to modern *Sabal* make it a potentially important calibration fossil, and we think several different approaches are defensible given what we know about the morphology of *S. bigbendense*. Calibration of the *Sabal* stem node (e.g., Bellot et al., 2020) is reasonable, given the strong resemblance with extant *Sabal* seeds, but the lack of apomorphic characters for *Sabal* make it inappropriate as a crown calibration for the genus (Sauquet, 2013). A more conservative view would be to employ it as a calibration for the crown node of the "apocarpous clade" of Coryphoideae, because the characters of the seed coat and embryo are found throughout the clade and are not unique to *Sabal*. We note this more conservative approach does not consider the overall shape of the seeds, which might be informative in a comparative framework.

Fossils of *Tripylocarpa aestuaria* Gandolfo & Futey originate from the Danian Salamanca Formation (63.3–61.9 Ma) of Argentina (Futey et al., 2012), and were placed in

subtribe Attaleinae of tribe Cocoseae based on morphological comparisons and phylogenetic analysis. It has since been used as a calibration fossil for Attaleinae in a variety of studies (e.g., Meerow et al., 2015; Freitas et al., 2016; Barrett et al., 2019). However, we think affinities with Cocoseae are equivocal because *T. aestuaria* has a single apical germination pore (Futey et al., 2012) rather than three pores, as is characteristic of tribe Cocoseae. Germination pores in Attaleinae are either lateral or subbasal but never apical; subapical pores are found in Bactridinae. Among palms, fruits with a single apical germination pore occur only in Borasseae and consist of large holes or very thin portions of the endocarp. In contrast, the structure interpreted as a germination pore in *T. aestuaria* is a narrow channel in the thick apical zone of the endocarp. While *T. aestuaria* is intriguing, its relationships with Cocoseae or another group of palms should be treated as uncertain until further anatomical and morphological details are documented. We note that our assessment of *T. aestuaria* does not necessarily cast doubt on analyses that employed it as a calibration because *Palmocarpon drypeteoides* is of a similar age and has well-supported affinities with Attaleinae.

Compressions of palm fruits were recovered from the Paleocene Cerrejón Formation of Colombia (~60–58 Ma; Gomez-Navarro et al., 2009) and identified as *Nypa* sp. and c.f. *Cocos*, the latter of which has served as a calibration for Cocoseae (Couvreur et al., 2011; Faurby et al., 2016). These interpretations are reasonable given their overall morphology and co-occurrence with palm leaf fossils. However, we note that the fossils lack diagnostic characters of these genera, and the *Cocos* and *Nypa*-like features (large size, ovate shape, apical stigmatic remains, and longitudinal fibers or striations) are found also in Borasseae. In this case it seems unlikely these are borassoid palms given the presence of pinnate leaves at the locality and geographic distributions of Borasseae, both modern and fossil. We mention it only to make the point that these features alone do not characterize *Cocos* to the exclusion of other taxa, and this can be said of other *Cocos*-like fossils described from compressions or casts (e.g., Shukla et al., 2012; Srivastava and Srivastava, 2014; Singh et al., 2016). Such considerations are especially important for fossils used as calibrations. This is not to criticize other workers, as we recognize that identifying new fossils and vetting calibrations is difficult and infrequently straightforward. These examples simply help to illustrate some of the challenges of identifying palm fruits, especially compression fossils, and their bearing on using such fossils in downstream analyses.

<h2>Implications for palm macroevolution

The age and phylogenetic relationships of the fossils analyzed here provide new information on the history of several palm lineages and indicate earlier origins for these clades than previously estimated (Table 2). Three of these fossils are suitable as new node calibrations in divergence-time analyses: *Coryphoides poulseni*, *Palmocarpon drypeteoides*, and *Friedemannia messelensis*. *Hyphaeneocarpon indicum* has recently been used to calibrate the crown node of subtribe Hyphaeninae of tribe Borasseae, pulling the origin of the syncarpous clade into the Cretaceous (Bellot et al., 2020). The fossils investigated here reveal that this Cretaceous diversification is not restricted to the syncarpous clade, and that the Late Cretaceous and early Paleogene diversification of palms likely established many of the groups we recognize today.

Resolving the phylogenetic relationships of fossils not only provides important information on the age of clades, but also their historical distributions. Considering fossil distributions in studies of historical biogeography is essential for developing accurate hypotheses on the origin of groups and the role of processes like long-distance dispersal, climate, and tectonic changes in the assembly of modern floras. This is especially true where extinction rates are sufficient to erase evidence of historical distributions from modern species ranges (Meseguer et al., 2015). The historical distributions of groups implied by some of the fossils we analyzed would probably be difficult to predict based on the ranges of modern species alone. This is demonstrated by incongruities between the fossils discussed in this paper and some ancestral range reconstructions (ARR) for palms. At the broadest spatial scales (seven regions: North America, South America, Africa-Arabia, Indian Ocean, India, Eurasia [to Wallace's Line], and the Pacific) ARRs at the family level do, for the most part, predict the historical distributions of the fossils analyzed here (Baker and Couvreur, 2013), although do they not capture the extensive geographic range of fossil Nypoideae (Gee, 2001). However, at finer scales, ARRs for Trachycarpeae/Livistoninae (Bacon et al., 2012) and Borasseae (Bellot et al., 2020) are unable to reconstruct the historical presence of these groups in Greenland (*Coryphoides poulseni*) and India (*Hyphaeneocarpon indicum*), respectively, perhaps because they were simply not parameterized to accommodate these possibilities. For these reasons we contend that either fossils should be incorporated into analyses of historical biogeography (e.g., Wood et al., 2013; Meseguer et al., 2015; Landis et al., 2020), or conclusions drawn from such analyses should be

reconciled with empirical data from the fossil record. The phylogenetic analyses we present in this paper, which remove some uncertainty from the calculus, will hopefully enable greater consideration of paleontological information in future biogeographic studies.

<h1>CONCLUSIONS AND PERSPECTIVES

The fossil record presents numerous challenges to scientists who study the evolution of plants and their interactions with Earth systems through time, but we hope this study illustrates that some of these challenges can be overcome to unlock the wealth of information in paleontological samples. Answering many fundamental questions in evolution requires taxonomic or phylogenetic information on fossils, something that is impossible to obtain without thorough study of extant diversity—a particularly daunting and time-consuming task for large and diverse clades. Lifetimes of expertise are therefore often critical for making seemingly small advances in understanding the fossil record. Moreover, the nature of paleobotanical work is such that we are often tasked with identifying fossils for which we have relatively limited knowledge, and thus, works that assemble information scattered in the literature can be especially useful.

In this study we synthesized several large bodies of research on the morphology of palm fruits and augmented this with 3D morphological data from µCT scans. We found that the diversity in fruit morphology and lack of straightforward synapomorphies for many clades can make identifying fossil palms below the family level very difficult using the types of taxonomic comparisons on which paleobotanists historically rely, an exception being Cocoseae. Phylogenetic analyses are thus crucial for resolving affinities of many fossil fruits. We demonstrated that relationships can be reconstructed with reasonable accuracy using fruit characters, but that thorough taxon sampling is essential and node support must be considered to avoid erroneous taxonomic placement. Our data set can serve as a useful starting point for identifying new fossils and will hopefully be improved by future studies. The phylogenetic relationships suggested by our analyses of previously published fruit fossils highlight the fact that there is still a great deal to be discovered about the early diversification of palms, and that the fossil record has much to contribute toward elucidating patterns and processes in palm evolution.

We propose several ways forward for palm researchers, whether they study fossils, living species, or both. First, we note that the ambiguous relationships of fossils to modern taxa is a

barrier to applying paleobotanical data to studies of macroevolution and biogeography. Conservative application of fossil data in such studies is necessary because information from fossils can strongly influence the outcome of analyses (Ho and Phillips, 2009; Warnock et al., 2012). This is one reason we advocate that careful phylogenetic analyses be performed by paleobotanists whenever possible, although we concede that this is not always feasible owing to mode or quality of preservation. Nevertheless, permineralizations are abundant in the fossil record, particularly those of palm stems and roots. While the phylogenetic utility of vegetative anatomy is yet to be fully explored, this is a promising area of future study. The extensive research that has been done on stem anatomy (Thomas and De Franceschi, 2013; Thomas and Boura, 2015), root anatomy (Seubert, 1996a, b, 1997, 1998a, b), and leaf anatomy (Tomlinson et al., 2011) means that there is a wealth of existing comparative data that can be tapped for these purposes.

Characterizing the fossil record of Arecaceae is important because it will allow us to better understand macroevolutionary and macroecological processes in the deep past. One outstanding question that spans both of these areas involves the relationship between palm evolution and environmental changes during the Cenozoic, particularly the origin and expansion of modern tropical rainforests. Today ~90% of palms are restricted to tropical rainforest biomes (Couvreur et al., 2011). Palms are thus an important group of organisms for understanding the origins of modern tropical rainforests and the evolutionary consequences of the expansion of these environments during the Paleogene (Burnham and Johnson, 2004; Morley, 2011; Couvreur and Baker, 2013; Jaramillo, 2019). Untangling these relationships will require further research on modern palm biology, continued study of the fossil record, and integration of perspectives from paleontological and neontological fields. This includes developing comparative data sets of palm morphology and anatomy, especially traits related to growth in different environments. In light of new fossil data, estimates of evolutionary tempo in palms should be refined using the wide variety of available methods, including those that can accommodate stratigraphic data and incorporate fossils as tips in the phylogeny (Heath et al., 2014; Zhang et al., 2016; Barido-Sottani et al., 2019). Finally, compiling occurrence data from the fossil record will be essential for independently testing hypotheses derived from extant taxa, exploring the relationships between palm distributions and climate, and understanding the role of extinction in shaping modern diversity and distributions. Living species and fossils provide imperfect but complementary data

on the history of life, and palms are among the few angiosperm families with such a long and dense fossil record. As such, we have a rare and precious opportunity to obtain a more comprehensive understanding of palm evolution. Progress in the areas described above, as well as the creative application of new data and methods to old questions, will undoubtedly lead to greater knowledge of the biology and long history of these glorious plants.

<h1>ACKNOWLEDGMENTS

We are indebted to the countless individuals who collect the specimens that make studies like ours possible. We also acknowledge Drs. John Dransfield, Natalie Uhl, and Madeline Harley who constructed the morphological matrix of Baker et al. (2009), which formed the basis of our study. In addition, we are grateful to Associate Editor Dr. Bruce Tiffney, Dr. William Baker, and one anonymous reviewer, whose constructive and informative comments significantly improved the quality of the manuscript. For access to collections and assistance with specimen loans we thank Harry Smith and Dr. William Baker of the Royal Botanic Gardens, Kew, Dr. Brett Jestrow of Fairchild Tropical Botanic Garden, Drs. Anna Stalter and Kevin Nixon of L.H. Bailey Hortorium Herbarium, Drs. Steven Manchester and Hongshan Wang of Florida Museum of Natural History. Undergraduate research assistants Ashley Hamersma, Mireille Farjo, Serena Safawi, and Jasimine Ash were instrumental in processing, analyzing, and managing μCT data. We also thank James Saulsbury for helpful conversations and thoughtful feedback. This study includes data produced in the CTEES facility at University of Michigan, supported by the Department of Earth & Environmental Sciences and College of Literature, Science, and the Arts. This research was generously funded by a National Science Foundation Doctoral Dissertation Improvement Grant to K.K.S.M. and S.Y.S. (DEB 1701645).

<h1>AUTHOR CONTRIBUTIONS

K.K.S.M. led conception and design of project with S.Y.S. S.Y.S. provided resources, and S.Y.S. and K.K.S.M. contributed funding. K.K.S.M. gathered, analyzed, and visualized the data. K.K.S.M. wrote the initial manuscript draft; K.K.S.M. and S.Y.S. reviewed and edited further drafts.

<h1>DATA AVAILABILITY

X-ray μCT data of palm fruits generated by this study are archived on MorphoSource under project number P776: "3D morphology of palm fruits (Arecaceae)" (https://www.morphosource.org).

The morphological data matrix is available in the online supporting information, as well as on MorphoBank (https://morphobank.org).

<h1>SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. List of palm species μCT scanned and their specimen numbers. Note that specimens labeled FTG (H) are accessioned in the FTG herbarium and those labeled FTG (G) were collected on the grounds by K.K.S.M.

Appendix S2. Morphological character matrix files. The matrix is provided in a nexus file formatted for Mesquite, with an embedded tree file (the molecular backbone tree used in this study) for visualizing character distributions. Also included is a list of character definitions and explanations of how characters were scored.

Appendix S3. Source files and R code for morphospace analyses.

Appendix S4. List of GenBank accession numbers for the DNA sequence data.

Appendix S5. Input and tree files for all phylogenetic analyses of fossils.

Appendix S6. Input and tree files for the molecular tree used as a backbone constraint.

Appendix S7. Reference tree and inferred trees for all phylogenetic pseudofossilization analyses.

<h1>LITERATURE CITED

Allen, S. E. 2015. Fossil Palm Flowers from the Eocene of the Rocky Mountain Region with Affinities to *Phoenix* L. (Arecaceae: Coryphoideae). *International Journal of Plant Sciences* 176: 586–596.

Bacon, C. D., W. J. Baker, and M. P. Simmons. 2012. Miocene dispersal drives island radiations in the palm tribe trachycarpeae (Arecaceae). *Systematic Biology* 61: 426–442.

Baker, W. J., and T. L. P. Couvreur. 2013. Global biogeography and diversification of palms

sheds light on the evolution of tropical lineages. I. Historical biogeography. *Journal of Biogeography* 40: 274–285.

Baker, W. J., and J. Dransfield. 2016. Beyond *Genera Palmarum*: progress and prospects in palm systematics. *Botanical Journal of the Linnean Society* 182: 207–233.

Baker, W. J., M. V. Norup, J. J. Clarkson, T. L. P. Couvreur, J. L. Dowe, C. E. Lewis, J. C. Pintaud, et al. 2011. Phylogenetic relationships among arecoid palms (Arecaceae: Arecoideae). *Annals of Botany* 108: 1417–1432.

Baker, W. J., V. Savolainen, C. B. Asmussen-Lange, M. W. Chase, J. Dransfield, F. Forest, M. M. Harley, et al. 2009. Complete generic-level phylogenetic analyses of palms (Arecaceae) with comparisons of supertree and supermatrix approaches. *Systematic Biology* 58: 240–256.

Barido-Sottani, J., G. Aguirre-Fernández, M. J. Hopkins, T. Stadler, and R. Warnock. 2019. Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence times under the fossilized birth-death process. *Proceedings of the Royal Society B: Biological Sciences* 286: 20190685.

Barrett, C. F., B. T. Sinn, L. T. King, J. C. Medina, C. D. Bacon, S. C. Lahmeyer, and D. R. Hodel. 2019. Phylogenomics, biogeography and evolution in the American genus *Brahea* (Arecaceae). *Botanical Journal of the Linnean Society* 190: 242–259.

Bellot, S., R. P. Bayton, T. L. P. Couvreur, S. Dodsworth, W. L. Eiserhardt, H. W. Pritchard, L. Roberts, and P. E. Toorop. 2020. On the origin of giant seeds: the macroevolution of the double coconut (*Lodoicea maldivica*) and its relatives (Borasseae, Arecaceae). *New Phytologist* 228: 1134-1148.

Berry, E. W. 1914. The Upper Cretaceous and Eocene floras of South Carolina and Georgia. *US Geological Survey, Professional Paper* 84: 1–200.

Berry, E. W. 1916. A petrified palm from the Cretaceous of New Jersey. *American Journal of Science* 41: 193–197.

Bobrov, A. V. F. C., J. Dransfield, M. S. Romanov, and E. S. Romanova. 2012. Gynoecium and

fruit histology and development in *Eugeissona* (Calamoideae: Arecaceae). *Botanical Journal of the Linnean Society* 168: 377–394.

Bobrov, A. V. F. C., D. H. Lorence, M. S. Romanov, and E. S. Romanova. 2012. Fruit development and pericarp structure in *Nypa fruticans* Wurmb (Arecaceae): A comparison with other palms. *International Journal of Plant Sciences* 173: 751–766.

Burnham, R. J., and K. R. Johnson. 2004. South American palaeobotany and the origins of neotropical rainforests. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 1595–1610.

Cailliez, F. 1983. The analytical solution of the additive constant problem. *Psychometrika* 48: 305–308.

Chapin, M. H., F. B. Essig, and J.-C. Pintaud. 2001. The morphology and histology of the fruits of Pelagodoxa (Arecaceae): taxonomic and biogeographical implications. *Systematic Botany* 26: 779–785.

Collinson, M. E., S. R. Manchester, and V. Wilde. 2012. Fossil fruits and seeds of the Middle Eocene Messel biota, Germany. *Abh. Senckenberg Ges. Naturforsch.* 570: 1–251.

Couvreur, T. L. P., and W. J. Baker. 2013. Tropical rain forest evolution: Palms as a model group. *BMC Biology* 11: 2–5.

Couvreur, T. L. P., F. Forest, and W. J. Baker. 2011. Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. *BMC Biology* 9: 44.

Dransfield, J., N. W. Uhl, C. B. Asmussen, W. J. Baker, M. M. Harley, and C. E. Lewis. 2008. Genera Palmarum: the evolution and classification of palms. Kew Publishing, Richmond, Surrey, UK.

Eldrett, J. S., D. R. Greenwood, I. C. Harding, and M. Huber. 2009. Increased seasonality through the Eocene to Oligocene transition in northern high latitudes. *Nature* 459: 969–973.

Essig, F. B. 1977. A systematic histological study of palm fruits. I. The Ptychosperma alliance.

Systematic Botany 2: 151–168.

Essig, F. B. 2002. A systematic histological study of palm fruits. VI. Subtribe Linospadicinae (Arecaceae). *Brittonia* 54: 196–201.

Essig, F. B., L. Bussard, and N. Hernandez. 2001. A systematic histological study of palm fruits. IV. Subtribe Oncospermatinae (Arecaceae). *Brittonia* 53: 466–471.

Essig, F. B., and N. Hernandez. 2002. A systematic histological study of palm fruits. V. Subtribe Archontophoenicinae (Arecaceae). *Brittonia* 54: 65–71.

Essig, F. B., T. J. Manka, and L. Bussard. 1999. A systematic histological study of palm fruits. III. Subtribe Iguanurinae (Arecaceae). *Brittonia* 51: 307–325.

Essig, F. B., and B. E. Young. 1979. A Systematic Histological Study of Palm Fruits. I. The Ptychosperma Alliance. *Systematic Botany* 4: 16–28.

Faurby, S., W. L. Eiserhardt, W. J. Baker, and J.-C. Svenning. 2016. An all-evidence species-level supertree for the palms (Arecaceae). *Molecular Phylogenetics and Evolution* 100: 57–69.

Fliche, P. 1894. Sur les fruits de palmiers trouvés dans le cénomanien environs de Sainte-Menehould. *Comptes rendus Hebdomadaires des Séances de l'Academie des Sciences* 118: 889–890.

Freitas, C., A. W. Meerow, J. C. Pintaud, A. Henderson, L. Noblick, F. R. C. Costa, C. E. Barbosa, and D. Barrington. 2016. Phylogenetic analysis of *Attalea* (Arecaceae): insights into the historical biogeography of a recently diversified neotropical plant group. *Botanical Journal of the Linnean Society* 182: 287–302.

Futey, M. K., M. A. Gandolfo, M. C. Zamaloa, R. Cúneo, and G. Cladera. 2012. Arecaceae Fossil Fruits from the Paleocene of Patagonia, Argentina. *Botanical Review* 78: 205–234.

Gee, C. T. 2001. The mangrove palm *Nypa* in the geologic past of the new world. *Wetlands Ecology and Management* 9: 181–194.

Gomez-Navarro, C., C. Jaramillo, F. Herrera, S. L. Wing, and R. Callejas. 2009. Palms (Arecaceae) from a Paleocene rainforest of northern Colombia. *American Journal of Botany* 96: 1300–1312.

Greenwood, D. R., and C. K. West. 2017. A fossil coryphoid palm from the Paleocene of western Canada. *Review of Palaeobotany and Palynology* 239: 55–65.

Guillerme, T., and N. Cooper. 2016. Effects of missing data on topological inference using a Total Evidence approach. *Molecular Phylogenetics and Evolution* 94: 146–158.

Harley, M. M. 2006. A summary of fossil records for Arecaceae. *Botanical Journal of the Linnean Society* 151: 39–67.

Harley, M. M., and W. J. Baker. 2001. Pollen aperture morphology in Arecaceae: Application within phylogenetic analyses, and a summary of the fossil record of palm-like pollen. *Grana* 40: 45–77.

Heath, T. A., J. P. Huelsenbeck, and T. Stadler. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences* 111: E2957–E2966.

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

Hopkins, M. J., and K. St. John. 2018. A new family of dissimilarity metrics for discrete character matrices that include inapplicable characters and its importance for disparity studies. *Proceedings of the Royal Society B: Biological Sciences* 285: 20181784.

Jaramillo, C. 2019. 140 Million Years of Tropical Biome Evolution. *The Geology of Colombia* 2: 1–28.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Koch, B. E. 1972. Coryphoid fruits and seeds from the Danian of Nûgssuaq, West Greenland. *The Geological Survey of Greenland* 99: 1–37.

Landis, M. J., D. A. R. Eaton, W. L. Clement, B. Park, E. L. Spriggs, P. W. Sweeney, E. J. Edwards, and M. J. Donoghue. 2020. Joint phylogenetic estimation of geographic movements and biome shifts during the global diversification of *Viburnum*. *Systematic Biology* 70: 67-85.

Lloyd, G. T. 2016. Estimating morphological diversity and tempo with discrete character-taxon matrices: Implementation, challenges, progress, and future directions. *Biological Journal of the Linnean Society* 118: 131–151.

Löytynoja, A. 2014. Phylogeny-aware alignment with PRANK BT - Multiple Sequence Alignment Methods. *In* D. J. Russell [ed.], 155–170. Humana Press, Totowa, NJ, USA.

Manchester, S. R., S. D. Bonde, D. S. Nipunage, R. Srivatava, R. C. Mehrotra, and S. Y. Smith. 2016. Trilocular Palm Fruits from the Deccan Intertrappean Beds of India. *International Journal of Plant Sciences* 177: 633–641.

Manchester, S. R., T. M. Lehman, and E. A. Wheeler. 2010. Fossil Palms (Arecaceae, Coryphoideae) Associated with Juvenile Herbivorous Dinosaurs in the Upper Cretaceous Aguja Formation, Big Bend National Park, Texas. *International Journal of Plant Sciences* 171: 679–689.

Matsunaga, K. K. S., S. R. Manchester, R. Srivastava, D. K. Kapgate, S. Y. Smith, and F. Museum. 2019. Fossil palm fruits from India indicate a Cretaceous origin of Arecaceae tribe Borasseae. *Botanical Journal of the Linnean Society* 190: 260–280.

Meerow, A. W., L. Noblick, D. E. Salas-Leiva, V. Sanchez, J. Francisco-Ortega, B. Jestrow, and K. Nakamura. 2015. Phylogeny and historical biogeography of the cocosoid palms (Arecaceae, Arecoideae, Cocoseae) inferred from sequences of six WRKY gene family loci. *Cladistics* 31: 509–534.

Morley, R. J. 2011. Cretaceous and tertiary climate change and the past distribution of megathermal rainforests. *In:* Tropical Rainforest Responses to Climatic Change. Springer Praxis Books. Springer, Berlin, Heidelberg.

Murray, S. G. 1973. The formation of endocarp in palm fruits. *Principes* 17: 91–102.

Nour-El-Deen, S., R. Thomas, and W. El-Saadawi. 2017. First record of fossil Trachycarpeae in Africa: three new species of *Palmoxylon* from the Oligocene (Rupelian) Gebel Qatrani Formation, Fayum, Egypt. *Journal of Systematic Palaeontology* 16: 741–766.

Nordén, K. K., T. L. Stubbs, A. Prieto-Márquez, and M. J. Benton. 2018. Multifaceted disparity approach reveals dinosaur herbivory flourished before the end-Cretaceous mass extinction. *Paleobiology* 44: 620–637.

Pan, A. D., A. D. Pan, B. F. Jacobs, B. F. Jacobs, J. Dransfield, J. Dransfield, W. J. Baker, and W. J. Baker. 2006. The fossil history of palms (Arecaceae) in Africa and new records from the Late Oligocene (28–27 Mya) of north-western Ethiopia. *Botanical Journal of the Linnean Society* 151: 69–81.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.

Pattengale, N. D., M. Alipour, O. R. P. P. Bininda-Emonds, B. M. E. E. Moret, and A. Stamatakis. 2009. How many bootstrap replicates are necessary? *In* S. Batzoglou [ed.], Research in Computational Molecular Biology, 184–200. Springer Berlin Heidelberg, Berlin, Heidelberg, Germany.

Pross, J., L. Contreras, P. K. Bijl, D. R. Greenwood, S. M. Bohaty, S. Schouten, J. A. Bendle, et al. 2012. Persistent near-tropical warmth on the Antarctic continent during the early Eocene epoch. *Nature* 488: 73–77.

Read, R. W., and L. J. Hickey. 1972. A revised classification of fossil palm and palm-like leaves. *International Association for Plant Taxonomy* 21: 129–137.

Reichgelt, T., C. K. West, and D. R. Greenwood. 2018. The relation between global palm distribution and climate. *Scientific Reports* 8: 2–12.

Reid, E. M., and M. E. J. Chandler. 1933. The London Clay Flora. British Museum (Natural History), London, United Kingdom.

Romanov, M. S., A. V. F. C. Bobrov, D. S. A. Wijesundara, and E. S. Romanova. 2011. Pericarp development and fruit structure in borassoid palms (Arecaceae-Coryphoideae-Borasseae). *Annals of Botany* 108: 1489–1502.

Sauquet, H. 2013. A practical guide to molecular dating. *Comptes Rendus Palevol* 12: 355–367.

Schaeffer, J., M. J. Benton, E. J. Rayfield, and T. L. Stubbs. 2020. Morphological disparity in theropod jaws: comparing discrete characters and geometric morphometrics. *Palaeontology* 63: 283–299.

Schliep, K. P. 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics* 27: 592–593.

Seubert, E. 1996a. Root anatomy of palms II. Calamoideae. *Feddes Repertorium* 107: 43–59.

Seubert, E. 1996b. Root anatomy of palms III. Ceroxyloideae, Nypoideae, Phytelephantoideae. *Feddes Repertorium* 107: 597–619.

Seubert, E. 1997. Root anatomy of palms: I. Coryphoideae. *Flora* 192: 81–103.

Seubert, E. 1998a. Root anatomy of palms IV. Arecoideae, part 1, general remarks and descriptions on the roots. *Feddes Repertorium* 109: 89–127.

Seubert, E. 1998b. Root anatomy of palms IV. Arecoideae, part 2 systematic implications. *Flora* 109: 231–247.

Shukla, A., R. C. Mehrotra, and J. S. Guleria. 2012. *Cocos sahnii* Kaul: A *Cocos nucifera* L.-like fruit from the Early Eocene rainforest of Rajasthan, western India. *Journal of Biosciences* 37: 769–776.

Singh, H., A. Shukla, and R. C. Mehrotra. 2016. A fossil coconut fruit from the early Eocene of

Gujarat. *Journal of the Geological Society of India* 87: 268–270.

Srivastava, R., and G. Srivastava. 2014. Fossil fruit of *Cocos* L. (Arecaceae) from Maastrichtian-Danian sediments of central India and its phytogeographical significance. *Acta Palaeobotanica* 54: 67–75.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

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

Thomas, R., and A. Boura. 2015. Palm stem anatomy: phylogenetic or climatic signal? *Botanical Journal of the Linnean Society* 178: 467–488.

Thomas, R., and D. De Franceschi. 2012. First evidence of fossil Cryosophileae (Arecaceae) outside the Americas (early Oligocene and late Miocene of France): Anatomy, palaeobiogeography and evolutionary implications. *Review of Palaeobotany and Palynology* 171: 27–39.

Thomas, R., and D. De Franceschi. 2013. Palm stem anatomy and computer-aided identification: The Coryphoideae (Arecaceae). *American Journal of Botany* 100: 289–313.

Tomlinson, P. B. 2006. The uniqueness of palms. *Botanical Journal of the Linnean Society* 151: 5–14.

Tomlinson, P. B., J. W. Horn, and J. B. Fisher. 2011. The Anatomy of Palms: Arecaceae - Palmae. Oxford University Press.

Uhl, N. W., and H. E. Moore. 1971. The palm gynoecium. *American Journal of Botany* 58: 945–992.

Vaidya, G., D. J. Lohman, and R. Meier. 2011. SeqenceMatrix: Cladistics multi-gene data sets with character set and codon information. *Cladistics* 27: 171–180.

Vajda, V., and A. Bercovici. 2014. The global vegetation pattern across the Cretaceous-Paleogene mass extinction interval: A template for other extinction events. *Global and Planetary Change* 122: 29–49.

Vaudois-Miéja, N., and A. Lejal-Nicol. 1987. Paléocarpologie africaine: apparition dés l'Aptien en Égypte d'un palmier (*Hyphaeneocarpon aegyptiacum* n. sp.). *Comptes Rendus de l'Académie des Sciences Paris Série II* 304: 233–238.

Warnock, R. C. M., Z. Yang, and P. C. J. Donoghue. 2012. Exploring uncertainty in the calibration of the molecular clock. *Biology Letters* 8: 156–159.

Wood, H. M., N. J. Matzke, R. G. Gillespie, and C. E. Griswold. 2013. Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. *Systematic Biology* 62: 264–284.

Zhang, C., T. Stadler, S. Klopfstein, T. A. Heath, and F. Ronquist. 2016. Total-Evidence Dating under the Fossilized Birth-Death Process. *Systematic Biology* 65: 228–249.

Table 1. Results of phylogenetic pseudofossilization. Values in column labeled "Chars." refer to the number of characters scored (out of 45) for each taxon and does not include inapplicable states. Reference affinities for each taxon are based on the reference tree, while inferred affinities indicate placement in the best-scoring maximum likelihood tree from pseudofossilization analyses, regardless of bootstrap support. Robinson-Foulds (RF) distances are reported for each inferred tree (the best-scoring tree) relative to the reference tree (lower values are better). The RF mean and RF max values summarize the pairwise RF distances between the reference tree and each bootstrap tree. Best-supported affinities are based on the shallowest subtending node with bootstrap support above 95%. Here the modifier "crown" in parentheses indicates that the best-supported node is one distal to the crown node for the group. If no modifier is indicated, the crown node is the best-supported node. See the supplementary materials for all trees. All affinities shown in bold are those that are incorrect based on the reference tree. Note that while the affinities from the best-scoring tree are incorrect for several taxa, the best-supported affinities are accurate for all but *Sclerosperma* and *Roystonea* (see Discussion). Abbreviation "sis." = "sister to". * = taxa forming monogeneric tribes; ** = taxa artificially reduced to the same 12 characters as *Friedemannia messelensis*.

Genus	Chars.	Reference affinities	Inferred affinities from best-scoring tree	RF	RF mean	RF max	Best-supported affinities
Ptychosperma	41	Arecoideae, Areceae, Ptychospermatinae	**Arecoideae, Areceae, Basselininae**	1 2	10. 5	 16	Areceae, Pacific clade
*Ptychosperma***	12	Arecoideae, Areceae, Ptychospermatinae	**Arecoideae, Areceae, Archontophoeniceae**	1 2	 6	 6	Areceae, Pacific clade

Areca	36	Arecoideae, Areceae, Arecinae	Arecoideae, Areceae, Arecinae	2	1.92	20	Areceae (crown)
*Areca***	12	Arecoideae, Areceae, Arecinae	**Arecoideae, Rheinhardtieae**	8	24	24	Arecoideae (crown)
Chamaedorea	32	Arecoideae, Chamaedoreae	Arecoideae, Chamaedoreae	0	16.1	18	Arecoideae (crown)
Wendlandiella	17	Arecoideae, Chamaedoreae	**Ceroxyloideae, Ceroxyleae**	6	11.1	26	Arecoideae+Ceroxyloideae
Attalea	43	Arecoideae, Cocoseae, Attaleinae	Arecoideae, Cocoseae, Attaleinae	6	6.0	6	Attaleinae (crown)
Syagrus	44	Arecoideae, Cocoseae, Attaleinae	Arecoideae, Cocoseae, Attaleinae	6	6.2	8	Attaleinae (crown)
*Syagrus***	12	Arecoideae, Cocoseae, Attaleinae	Arecoideae, Cocoseae, Attaleinae	8	8.7	22	Attaleinae
Euterpe	38	Arecoideae, Euterpeae	Arecoideae, Euterpeae	4	3.3	4	Euterpeae (crown)
Oenocarpus	29	Arecoideae, Euterpeae	Arecoideae, Euterpeae	2	11.1	22	Areceae+Euterpeae
Prestoea	37	Arecoideae, Euterpeae	Arecoideae, Euterpeae	4	3.6	18	Euterpeae (crown)
Calyptrogyne	45	Arecoideae, Geonomeae	Arecoideae, Geonomeae	4	3.2	4	Geonomeae (crown)
Geonoma	38	Arecoideae, Geonomeae	**Arecoideae, Euterpeae**	2	13.9	22	Core arecoid clade

Iriartella	29	Arecoideae, Iriarteae	Arecoideae, Iriarteae	4	4.0	18	Arecoideae+Ceroxyloideae
Socratea	32	Arecoideae, Iriarteae	**Coryphoideae, Caryoteae**	1	15.0	24	Coryphoideae+Ceroxyloideae+Arecoideae
Leopoldinia		Arecoideae, sis.		1	10.		
*a**	28	Pelagodoxeae	**Arecoideae, sis. Roystoneae**	2	8	28	Arecoideae (crown)
Manicaria		Arecoideae, core arecoid		1	11.		
*	38	clade	**Ceroxyloideae, Phytelepheae**	2	3	14	Arecoideae+Ceroxyloideae
*Orania**	40	Arecoideae, POS clade	**Arecoideae, sis. Manicarieae**	8	8.2	12	Arecoideae (crown)
*Pelagodoxa**	34	Arecoideae, sis. Leopoldinieae	**Arecoideae, sis. Leopoldinieae**	0	2.0	22	Arecoideae
*Podococcus**	30	Arecoideae, POS clade	Arecoideae, POS clade	2	2.4	12	POS+RRC clade
*Sclerosperma**	38	Arecoideae, POS clade	**Arecoideae, Areceae**	3	30.5	32	**Areceae**
*Reinhardtia**	38	Arecoideae, RRC clade	**Arecoideae, Geonomeae**	1	14.6	18	Arecoideae (crown)
*Roystonea**	41	Arecoideae, RRC clade	**Arecoideae, Areceae**	1	20.8	30	**Areceae**
Calamus	22	Calamoideae, Calameae	Calamoideae, Calameae	4	4.2	8	Calameae (crown)
*Eugeissona**	40	Calamoideae, sis. Calameae	**Nypoideae**	8	6.2	10	Calamoideae+Nypoideae

Mauritia	23	Calamoideae, Lepidocaryeae, Mauritiinae	Calamoideae, Lepidocaryeae, Mauritiinae	0	6.8	8	Calamoideae
Juania	30	Ceroxyloideae, Ceroxyleae	Ceroxyloideae, Ceroxyleae	4	4.0	14	Arecoideae+Ceroxyloideae
*Pseudophoenix**	35	Ceroxyloideae, sis Phytelepheae+Ceroxyleae	**Arecoideae, sis. Manicarieae**	1 0	9.8	16	Arecoideae+Ceroxyloideae
Ammandra	28	Ceroxyloideae, Phytelepheae	Ceroxyloideae, Phytelepheae	0	0.6	2	Phytelepheae (crown)
Bismarckia	42	Coryphoideae, Borasseae, Hyphaeninae	Coryphoideae, Borasseae, Hyphaeninae	0	1.6	8	Borasseae
Caryota	25	Coryphoideae, Caryoteae	Coryphoideae, Caryoteae	0	30.0	34	Coryphoideae+Ceroxyloideae+Arecoideae
Chuniophoenix	23	Coryphoideae, Chuniophoeniceae	Coryphoideae, Chuniophoeniceae	2	2.2	10	Coryphoideae
*Corypha**	31	Coryphoideae, sis. Borasseae	**Coryphoideae, sis. Chuniophoeniceae**	4	11.8	30	Coryphoideae+Ceroxyloideae+Arecoideae
Thrinax	29	Coryphoideae, Cryosophileae	Coryphoideae, Cryosophileae	6	5.8	16	Cryosophileae (crown)
*Phoenix**	30	Coryphoideae, sis. Trachycarpeae	Coryphoideae, Trachycarpeae	6	9.3	16	Coryphoideae, Apocarpous clade
*Sabal**	34	Coryphoideae, sis. Cryosophileae	**Coryphoideae, Syncarpous clade**	8	7.4	20	Coryphoideae

		Coryphoideae,	**Coryphoideae, sis.**	1			
Livistona	33	Trachycarpeae, Livistoninae	**Phoeniceae+Trachycarpeae**	0	8.1	18	Coryphoideae
		Coryphoideae,	Coryphoideae, Trachycarpeae,				
Saribus	17	Trachycarpeae, Livistoninae	Livistoninae	2	5.2	30	Coryphoideae

Table 2. Palm fruit fossils suitable as new node calibrations. Groups with which relationships are strongly supported in phylogenetic analyses are indicated, as well as the estimated ages of those clades from Baker and Couvreur (2013) and the age of the fossils. Some key characters of the fossils that support their phylogenetic relationships are listed. Note that these relationships are based on multiple characters, which are discussed in the text.

Species	**Group**	**Key characters**	**Fossil age (million years)**	**Estimated node age (95% HPD Baker & Couvreur, 2013)**
Coryphoides poulseni	Crown Trachycarpeae	Lateral embryo, basal postament	64–62	47.15–22.98
Friedemannia messelensis	Crown Areceae	Apical hilum & seed attachment	47	42.42–25.95
Palmocarpon drypeteoides	Crown Attaleinae	Three subbasal germination pores	67–64	49.78–23.29

			Calamoideae, sis.				
*Nypa**	40	Nypoideae	**Eugeissoneae**	8	8.5	18	Arecaceae (crown)

Figure 1. Calamoideae. Images B–F from µCT scans. (A) External view of fruits showing basally oriented pericarp scales. From left to right: *Metroxylon salmonense* (20 mm; K000754987), *Raphia farinifera* (10 mm; FTG76039), *Lepidocaryum tenue* (5 mm; FTG136527), and *Pigafetta filaris* (1 mm; FTG88176). (B) Longitudinal section (LS) of *Lepidocaryum tenue* (tribe Lepidocaryeae) showing apical stigmatic remains (arrowhead), lateral embryo (asterisk), and uniformly thin seed coat (arrow). Note lack of endocarp or fiber bundles in pericarp. Scale = 5 mm. FTG136527. (C) LS of *Oncocalamus mannii* (tribe Calameae). Seed with lateral postament (arrow). Scale = 5 mm. BH000104592. (D) LS of *Mauritiella armata* (tribe Lepidocaryeae). Seed is shrunken but shows remnants of thick, fleshy sarcotesta of seed coat (arrow). Note apical stigmatic remains (arrowhead). Scale = 2 mm. FTG117555. (E) Transverse section (TS) of *Pigafetta filaris* (tribe Calameae) with thickened sarcotesta (arrow). Scale = 1 mm. FTG88176. (F) TS of *Plectocomia mulleri* (tribe Calameae). Seed with thickened sarcotesta (arrow). Scale = 5 mm. BH000154523.

Figure 2. *Nypa fruticans* (Nypoideae). Images B–E from µCT scans. (A) External view of fruit showing obovate shape and deep longitudinal grooves. Note apical stigmatic remains forming structure referred to as an "umbo" (arrow). Scale = 20 mm. (B) Longitudinal section of fruit shown in (A). Pericarp with numerous longitudinal fiber and fibrovascular bundles to outside of thick endocarp ("e"). Note basal germination pore of endocarp (asterisk). Scale = 10 mm. (C) 3D model of endocarp seen laterally, segmented from µCT scan shown in (B). Scale = 10 mm. (D) Transverse section of fruit from A–C. Endocarp ("e") forms longitudinal ridge intruding into seed (arrow). Note numerous fiber and fibrovascular bundles of pericarp in transverse section (white dots). Scale = 10 mm. (E) Basal view of endocarp model from C, showing circular germination pore of endocarp. Scale = 10 mm. FTG84164.

Figure 3. Coryphoideae, syncarpous clade. Images A–J from µCT scans. (A) Transverse section (TS) of *Tahina spectabilis* (tribe Chuniophoeniceae). Note thin endocarp ("e") and deeply ruminate endosperm forming radial furrows in seed (arrow). Scale = 5 mm. K000525955 (holotype). (B) TS of *Arenga engleri* (tribe Caryoteae). Fruit is trilocular with three seeds. Note lack of prominent endocarp and fibrovascular bundles in pericarp. Scale = 2 mm. FTG10076. (C) Longitudinal section (LS) of *Caryota mitis* (tribe Caryoteae). Note lack of prominent endocarp and remnants of fleshy pericarp, shrunken around seed. Scale = 2 mm. FTG89-34 A. (D) LS of *Medemia argun* (tribe Borasseae). Fruit with prominent endocarp ("e") and apical germination

pore consisting of gap in endocarp above the embryo (below asterisk). Note deeply ruminate endosperm and two abortive carpels basally (arrows). Scale = 5 mm. K000208672. (E) LS of *Satranala decussilvae* (tribe Borasseae). Endocarp thick, externally sculptured with prominent ridges (arrowhead). Apical ridge (arrow) functions as germination valve. Seed is shrunken, forming pockets around endosperm ruminations. Scale = 10 mm. K000525955. (F) LS of *Hyphaene thebaica* (tribe Borasseae). Note apical germination pore (below asterisk), consisting of thinner zone of endocarp with sparser fiber bundles. Scale = 10 mm. FTG136617. (G) Off-median LS of *Bismarckia nobilis* (tribe Borasseae), showing two abortive carpels forming basal bulges (arrows). Endocarp ("e") is thick, composed of interwoven fiber bundles. Scale = 5 mm. FTG76031. (H) TS of immature fruit of *Borassus madagascariensis* (tribe Borasseae). Specimen fresh collected and scanned prior to drying. Section taken near base, passing through three empty locules. Note perianth remnants surrounding fruit (arrow). (I) TS of *Borassus flabellifer* pyrene (seed + endocarp). Endocarp ("e") is thick, forming a ridge that intrudes laterally into seed (arrow). Scale = 10 mm. FTG10156. (J) TS of *Borassodendron machadonis* (tribe Borasseae). Fruit is trilocular, with three pyrenes. Endocarp forms multiple longitudinal ridges that intrude into seed (arrow). Note embryo in transverse section (arrowhead). Scale = 10 mm. FTG68387B. (K&L) One- and two-seeded fruits of *Medemia argun* (K) and *Satranala decussilvae* (L). Fruits are deeply lobed when more than one carpel matures, appearing as two fruits fused at base. Scale = 20 mm. K000208672 (K), K000525955 (L).

Figure 4. Coryphoideae, apocarpous clade. All images from µCT scans. (A) Longitudinal section (LS) of *Livistona benthamii* (tribe Livistoninae). Seed with prominent lateral postament (seed coat intrusion; arrow), lateral embryo (asterisk in embryo cavity), and thin endocarp ("e"). Scale = 2.5 mm. FTG2001-0637B. (B) LS of *Rhapidophyllum hystrix* (tribe Trachycarpeae). Fruit formed from two out of three unfused carpels, connected at base near perianth remnants. Note that pericarp is not fused, and seed coat thickened on one side (arrows), opposite embryo. Scale = 5 mm. FTG16959. (C) Transverse section (TS) of *Schippia concolor* (tribe Cryosophileae). Note embryo in seed (arrow), lack of endocarp, and fleshy pericarp shrunken into thin layer (arrowhead). Scale = 10 mm. FTG2002-0575B. (D) LS of *Johannesteijsmannia altifrons* (tribe Livistoninae). Endocarp ("e") prominent, thickened basally. Seed with basal postament (arrow). Note corky pericarp with irregular, warty protrusions. Scale = 5 mm. K000933830. (E) LS of *Acoelorraphe wrightii* (tribe Livistoninae). Note prominent endocarp

("e"), thickened region of seed coat (arrow), and longitudinal fibrovascular bundles adjacent endocarp, seen in grazing section (arrowhead). Scale = 1 mm. FTG10066. (F) LS of *Leucothrinax morrissii* (tribe Cryosophileae). Seed has prominent lateral postament (arrow) and lateral embryo (asterisk). Note that seed is loose within dried fruit and has rotated from original position. Apical stigmatic remains at arrowhead. Scale = 1 mm. FTG10528. (G) LS of *Sabal palmetto* (tribe Sabaleae) seed for comparison with *Sabal bigbendense* fossil (H&I). Note thickened seed coat basally (arrow) and lateral embryo (arrowhead). Scale = 2 mm. UF1158. (H) LS of *Sabal bigbendense* fossil seed. Note darker area in seed (arrow), which is the thickened zone of seed coat, and lateral embryo (arrowhead). Scale = 3 mm. UF402-53789. (I) Translucent volume rendering of specimen in (H), with embryo indicated by arrowhead. Scale = 3 mm. UF402-53789.

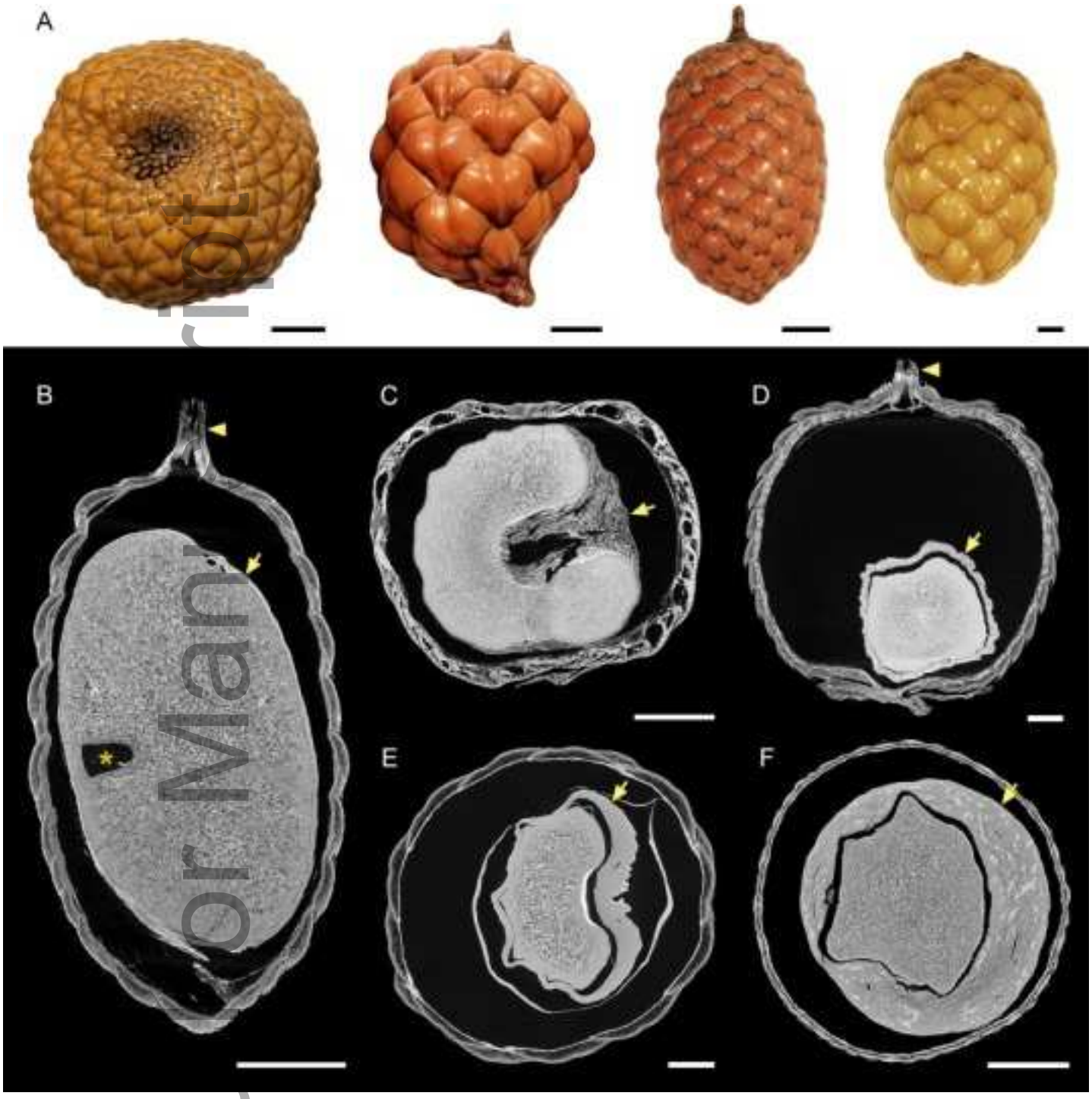
Figure 5. Ceroxyloideae. All images from µCT scans. (A) Transverse section (TS) of *Ammandra decasperma* (tribe Phytelepheae). Fruit is immature, all eight locules lacking seeds. Pericarp is composed of corky warts formed by numerous radial fiber bundles like in *Pelagodoxa henryana* (see Fig. 6F). Scale = 1 cm. FTG60393. (B) Longitudinal section (LS) of *Pseudophoenix vinifera* (tribe Cyclospatheae). Note two abortive carpels basally (arrows) and thin endocarp ("e"), which is discontinuous at the point of seed attachment, forming hilar seam ("h"). FTG814015. (C) TS of *Oraniopsis appendiculata* (tribe Ceroxyleae). Note thin, shrunken pericarp (arrowhead) to the outside of the seed coat (arrow), and absence of prominent endocarp. Scale = 5 mm. BH000154548. (D) LS of *Juania australis* (tribe Ceroxyleae). There is at least one large fibrovascular bundle in the pericarp (arrow), which was otherwise mostly fleshy. Note lack of prominent endocarp and embryo cavity in endosperm (asterisk). Scale = 2 mm. Moore 9368 (Kew).

Figure 6. Arecoideae. All images from µCT scans. (A–C) Tribe Cocoseae. (A) Median transverse section (TS) of *Jubaeopsis caffra*, passing through lateral germination pore of endocarp. Note small scattered vascular bundles in pericarp, embryo to the inside of the germination pore (left of asterisk), hollow cavity in seed endosperm, and thin operculum in germination pore (arrow). Scale = 5 mm. K001083912. (B) TS of *Cocos nucifera*. Endosperm is hollow, with the seedling haustorium in the center (asterisk). Note numerous small longitudinal fiber and fibrovascular bundles in the pericarp to the outside of the endocarp. Scale = 2 cm. BH000199147. (C) Longitudinal section (LS) of *Butia capitata* with two locules. Note subbasal

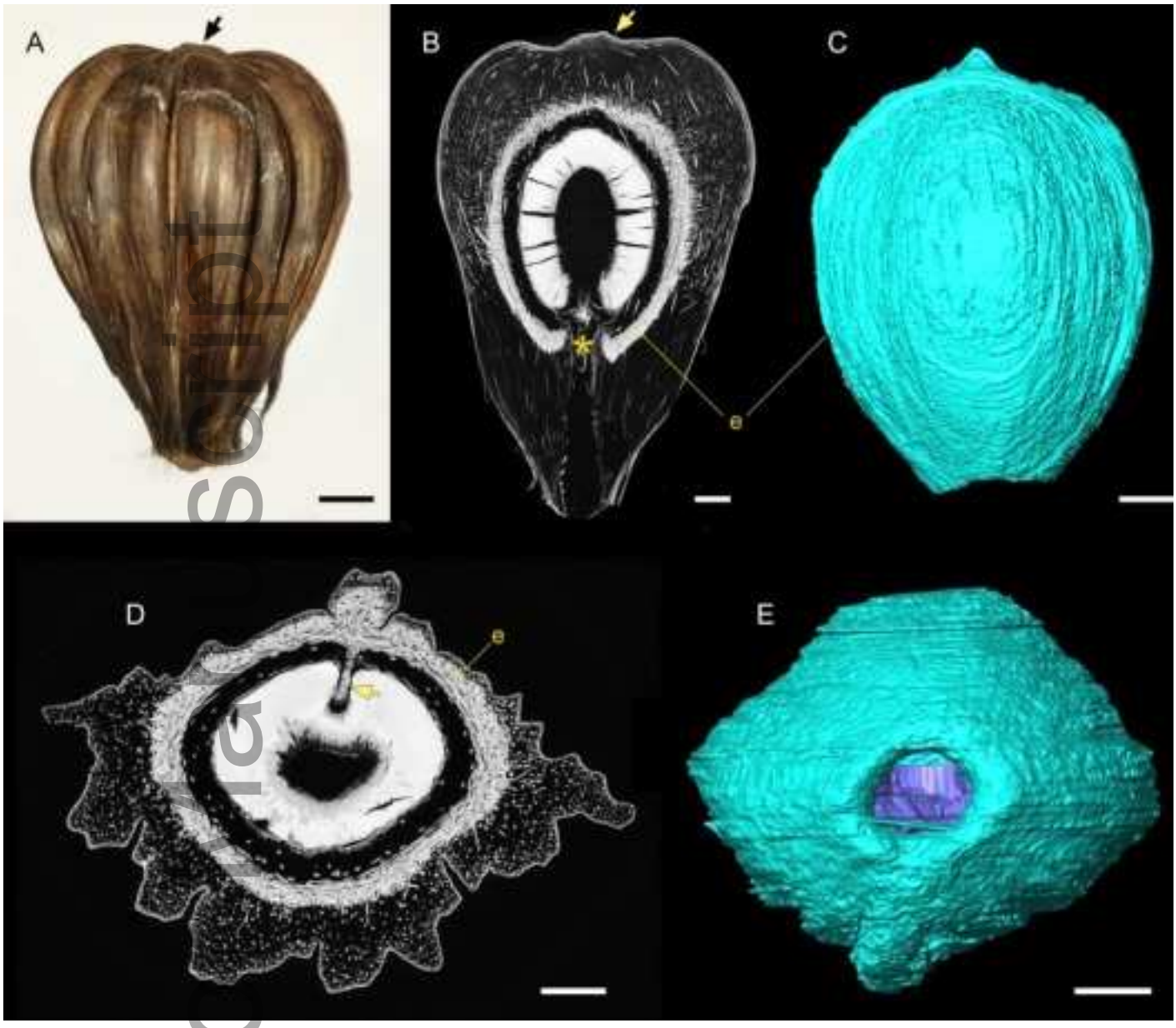
germination pore with operculum (arrow), and apical stigmatic remains (arrowhead). Scale = 2 mm. FTG76645. (D) TS through *Iriartella setigera* (tribe Iriarteae). Note absence of prominent endocarp and longitudinal fibrovascular bundles. Scale = 2 mm. K0001244565. (E) LS of *Euterpe oleracea* (tribe Euterpeae) with two abortive carpels basally (arrows). Note endosperm ruminations, corresponding to the deep radial cracks in the seed. Scale = 2 mm. FTG72880. (F) TS of *Pelagodoxa henryana* (tribe Pelagodoxeae). Note pericarp of numerous large corky warts composed of radial fiber bundles and thin but prominent endocarp. BH000154524 (G) LS of *Orania lauterbauchiana* (tribe Oranieae). Fruit has a very thin endocarp and a thickened region of the seed coat at the hilum (asterisk). Note abortive carpel traces basally (arrow). Scale = 5 mm. K000114185. (H) LS of *Podococcus barteri* (tribe Podococceae). Note slender, elongate shape and small lateral embryo (arrow). Scale = 5 mm. K000114526. (I) LS of *Sclerosperma profiziana* (tribe Sclerospermeae). Note basal embryo (arrow) and thin endocarp. Scale = 5 mm. Profizi 841 (Kew). (J) LS of *Cyphokentia* (*Moratia*) *cerifera* (tribe Areceae). Endocarp is thick with a prominent basal operculum beneath embryo (asterisk), which is shrunken. Note large flattened fibrovascular bundle seen apically (arrow). Scale = 2 mm. BH000154527. (K) TS of *Wodyetia bifurcata* (tribe Areceae). Note prominent endocarp to the inside of thick zone of compacted longitudinal fiber and fibrovascular bundles with massive sheaths (arrow). Scale = 5 mm. FTG140799. (L) TS of *Ptychococcus paradoxus* (tribe Areceae). Endocarp comprised of locular epidermis (thin white layer at arrow) and sclerenchymatous inner zone of pericarp in which large longitudinal fibrovascular bundles are embedded (arrowhead). Scale = 5 mm. FTG82784. (M) LS of *Brongniartikentia lanuginosa* (tribe Areceae). Note basal embryo (arrow) and prominent operculum in endocarp. Scale = 2 mm. BH000154515. (N) TS of *Acanthophoenix rubra* (tribe Areceae). Endocarp is discontinuous at region of seed attachment, forming a hilar seam ("h"). Note thickening of seed coat at hilum (arrow). Scale = 2 mm. Vaughan 851 (Kew). Labels: e = endocarp, o = operculum.

Figure 7. Morphospace plots based on fruit characters from morphological matrix. First (*x*-axis, 14.49%) and second (*y*-axis, 8.04%) principal coordinate axes are shown. Convex hulls delimit subfamilies and tribes. For tribes, only those with more than three genera are shown with convex hulls. Note that for Arecoideae, POS corresponds to the clade formed by the monogeneric tribes Podococceae, Oranieae, and Sclerospermeae.

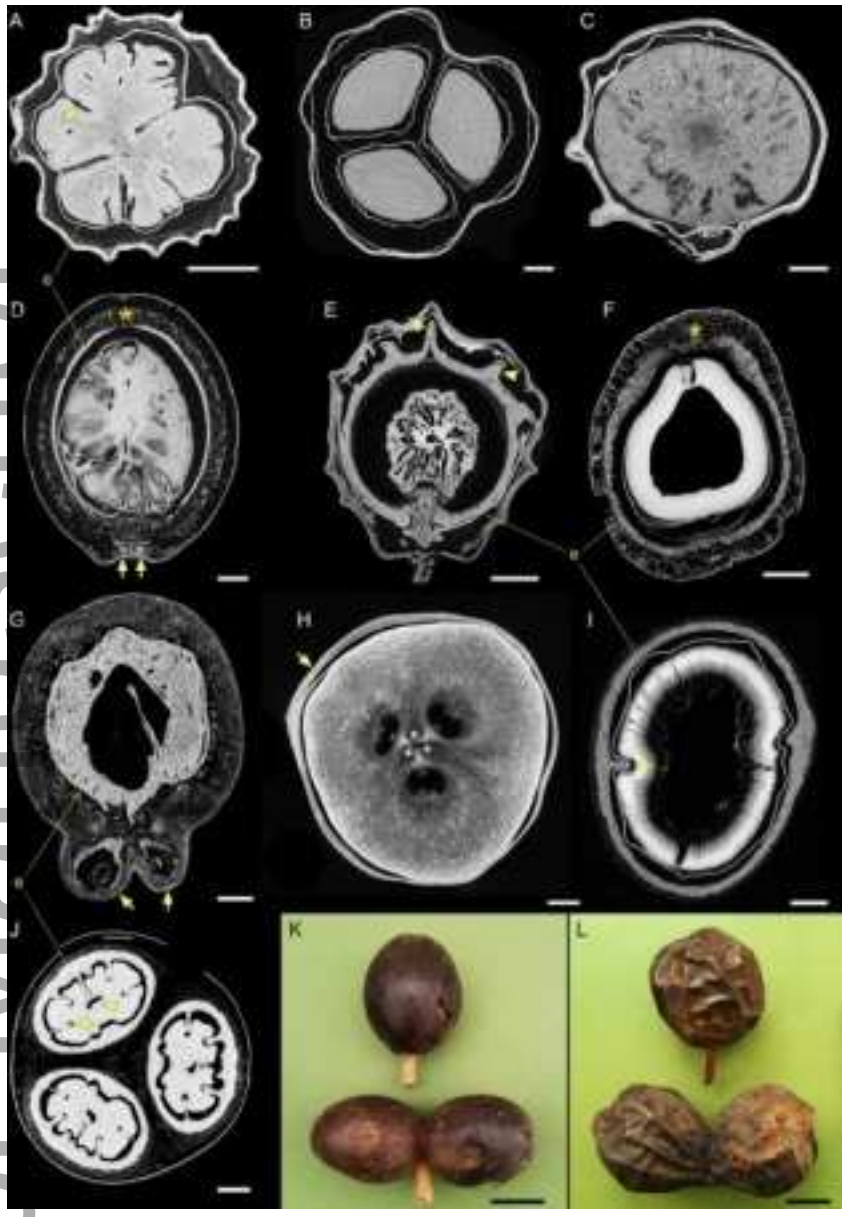
Figure 8. Results of phylogenetic analyses of the fossils (constrained and total-evidence), and pseudofossilization analyses, summarized on the total-evidence tree. Tree is drawn with uniform branch lengths for clarity. Branches with thicker lines subtend nodes with bootstrap support of 90% or higher in the total-evidence analysis. Tribes and subtribes without fossils are shown collapsed (black triangles), with labels indicating the number of collapsed tips in parentheses. Fossil species are indicated with bold text and dagger symbol. Numbers above branches indicate bootstrap support for the shallowest well-supported clade for fossil placement, based on backbone constraint analyses. Numbers in gray circles correspond to tips excluded from the backbone constraint trees and the resulting shallowest well-supported nodes in phylogenetic pseudofossilization analyses; for clarity only pseudofossil placements for members of clades containing fossils are shown (see Table 1 for full results). Relevant tribes and other major clades indicated with bubbles: Ar = tribe Areceae, Co = tribe Cocoseae, Tr = tribe Trachycarpeae, Bo = tribe Borasseae, RRC = Roystoneae, Reihnhardtieae, and Cocoseae, POS = Podococceae, Oranieae, and Sclerospermeae. See Appendices S5 and S7 for all trees.



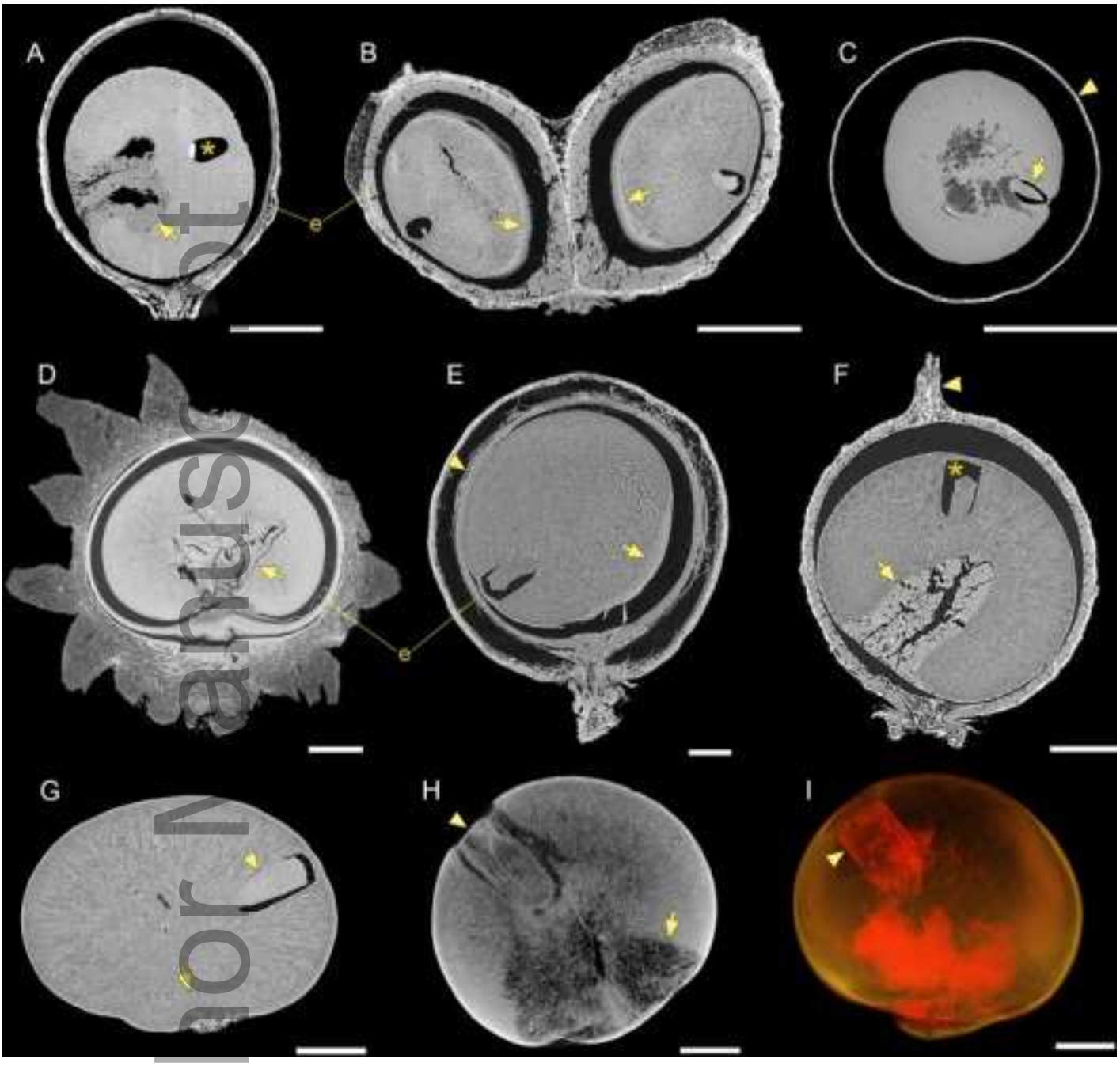
ajb2_1616_f1.jpg



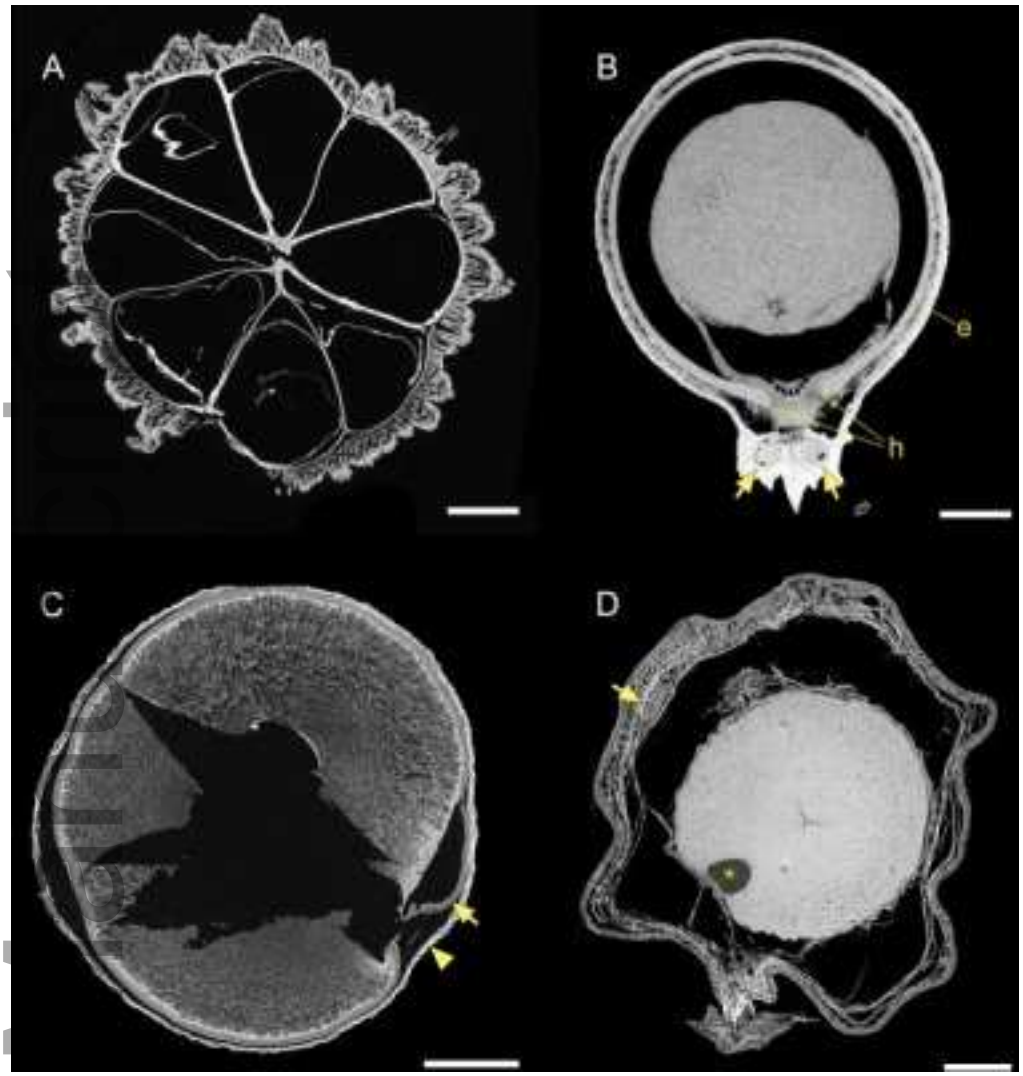
ajb2_1616_f2.jpg



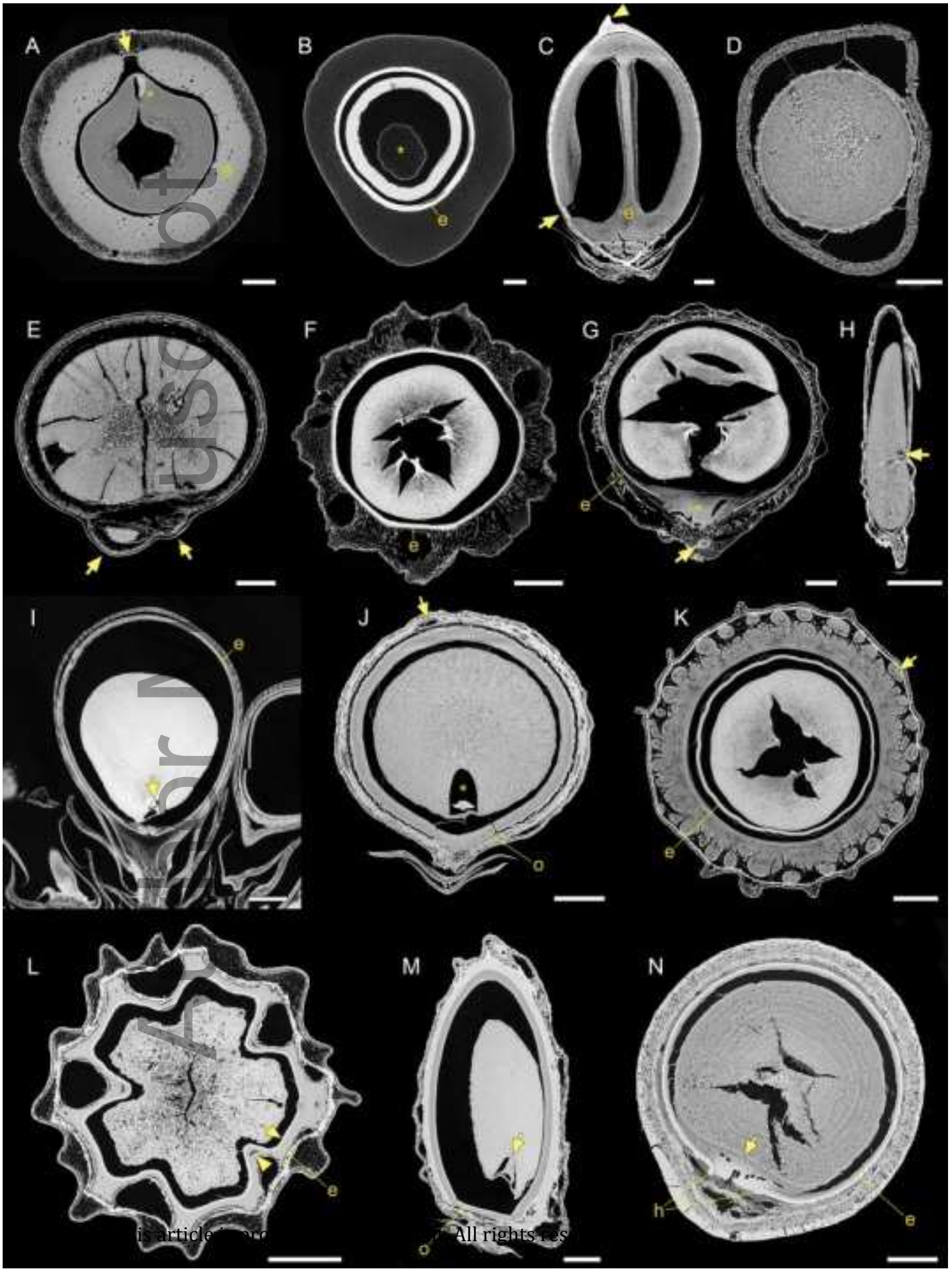
ajb2_1616_f3.jpg



ajb2_1616_f4.jpg



ajb2_1616_f5.jpg



ajb2_1616_f6.jpg