MDPI

Article

UNRAVELLING THE PHYLOGENOMIC RELATIONSHIPS OF THE MOST DIVERSE AFRICAN PALM GENUS *Raphia* (CALAMOIDEAE, ARECACEAE)

Andrew J. Helmstetter [1,†], Suzanne Mogue Kamga [2], Kevin Bethune [3], Thea Lautenschläger [4], Alexander Zizka [5], Christine D. Bacon [6], Jan J. Wieringa [7], Fred Stauffer [8], Alexandre Antonelli [9], Bonaventure Sonké [10] and Thomas L.P. Couvreur [11,*]

1 IRD, DIADE, Univ Montpellier, Montpellier, France; andrew.j.helmstetter@gmail.com
2 Université de Yaoundé I, Ecole Normale Supérieure, Département des Sciences Biologiques, Laboratoire de Botanique systématique et d'Ecologie, B.P. 047, Yaoundé Cameroon ; mogueblue@yahoo.com
3 IRD, DIADE, Univ Montpellier, Montpellier, France; kevin.bethune@hotmail.fr
4 Department of Biology, Institute of Botany, Faculty of Science, Technische Universität Dresden, 01062 Dresden, Germany ; thea.lautenschlaeger@tu-dresden.de
5 German Center for Integrative Biodiversity Research (iDiv) Halle -Leipzig-Jena, Leipzig, Germany
6 Department of Biological and Environmental Sciences and Gothenburg Global Biodiversity Centre, University of Gothenburg, Gothenburg, Sweden; christine.bacon@bioenv.gu.se
7 Naturalis Biodiversity Center, Darwinweg 2, 2333 CR, Leiden, the Netherlands; jan.wieringa@naturalis.nl
8 Department of Botany and Plant Biology, Conservatory and Botanical Garden of the City of Geneva, University of Geneva, Geneva, Switzerland; Fred.Stauffer@ville-ge.ch
9 Department of Biological and Environmental Sciences and Gothenburg Global Biodiversity Centre, University of Gothenburg, Gothenburg, Sweden; and Royal Botanic Gardens, Kew, Richmond, UK; a.antonelli@kew.org
10 Université de Yaoundé I, Ecole Normale Supérieure, Département des Sciences Biologiques, Laboratoire de Botanique systématique et d'Ecologie, B.P. 047, Yaoundé Cameroon; bsonke_1999@yahoo.com
11 IRD, DIADE, Univ Montpellier, Montpellier, France; Thomas.Couvreur@ird.fr
* Correspondence: Thomas.Couvreur@ird.fr
† Current address: FRB-Cesab, Institut Bouisson Bertrand, 5, rue de l'Ecole de médecine - 34000 Montpellier, France

Version March 11, 2020 submitted to Plants

Abstract: Palms are conspicuous floristic elements across the tropics. In continental Africa, even though there are less than 70 documented species, they are omnipresent across the tropical landscape. The genus *Raphia* has 20 accepted species in Africa and one species endemic to the Neotropics. It is the most economically important genus of African palms with most of its species producing food and construction material. *Raphia* is divided into five sections based on inflorescence morphology. Nevertheless, the taxonomy of *Raphia* is problematic with no intra-generic phylogenetic study available. We present a phylogenetic study of the genus using a targeted exon capture approach sequencing of 56 individuals representing 18 out of the 21 species. Our results recovered five well supported clades within the genus. These reflect to a certain extent the sections as defined based on inflorescence morphology. Overall, morphological based identifications agreed well with our phylogenetic analyses, with 12 species recovered as monophyletic based on our sampling. Species delimitation analyses recovered 17 or 23 species depending on the confidence level used. Species delimitation is especially problematic in the Raphiate and Temulentae sections. In addition, our clustering analysis using SNP data suggested that individual clusters matched geographic distribution. The Neotropical species *R. taedigera* is supported as a distinct species, rejecting the hypothesis of a recent introduction into South America. Our analyses support the hypothesis that the *Raphia* individuals from Madagascar are potentially a distinct species different from the widely

distributed *R. farinifera*. In conclusion, our results support the infra generic classification of *Raphia* based on inflorescence morphology, which is shown to be phylogenetically useful. Classification and species delimitation within sections remains problematic even with our phylogenomic approach. Certain widely distributed species could potentially contain cryptic species. More in-depth studies should be undertaken using morphometrics, increased sampling and more variable markers. Our study provides a robust phylogenomic framework that enables further investigation on the biogeographic history, morphological evolution and other eco-evolutionary aspects of this charismatic, socially and economically important palm genus.

Keywords: Africa, exons, Madagascar, rain forests, phylogenomics, *Raphia*, sequence capture

1. Introduction

Palms are iconic floristic elements across the tropics both in terms of diversity and the natural resources they provide, playing important roles for the welfare of rural and urban people at equatorial latitudes. Worldwide, there are an estimated 2500 palm species [1], mainly occurring in tropical rain forests. Africa, however, harbours less than 70 species (excluding Madagascar) [2,3], a pattern that contrasts strongly with the Neotropics or South East Asia, which contain 800 and 1200 species respectively [1,4,5]. Despite this low diversity, palms are omnipresent across the African landscape, particularly in the tropical rain forests of the continent [2,6].

Among African palms, the genus *Raphia* (subfamily Calamoideae, tribe Raphiaeae) is the most species rich, with 21 species described to date [2,7]. Of these, one, *R. taedigera*, is endemic to the Neotropics, with a disjunct distribution in Brazil and central America. The presence of this species in the Neotropics was suggested as either pre-Colombian and natural (biogeographic long distance dispersal/vicariance [8,9]) or as recently naturalized by Africans during the slave trade some 400 years ago [6,10,11]. *Raphia* species mainly occur in tropical rain forests, most often in swampy or periodically inundated areas where they can dominate the vegetation, producing dense monospecific stands (known as "Raphiales" in French). A few species, however, have adapted to drier conditions restricted to river systems in the Sahel or southern Africa.

Raphia is the most economically important genus of African palms across tropical African communities. One recent study documented over 100 different uses across the genus, with the most important ones being extraction of palm wine, grubs and construction material [12,13]. Exploitation of its species in the wild also represents an important source of income for populations across tropical Africa, especially for low-income households [12,14,15]. In addition, *Raphia* species play vital ecological roles in wet land ecosystems [16] where they dominate the landscape, such as in peatlands of the Congo Basin where they are highly abundant [17]. *Raphia* dominated swamps are also important ecosystems for the protection for critically endangered animals such as the lowland gorillas because hard to access or cultivate (e.g. [18]).

Raphia species are massive palms with very long pinnate leaves. One species holds the record for the longest measured leaf in angiosperms, reaching up to 25 meters (*R. regalis*). The trunk is generally above-ground and is solitary or clustered, while two species (*R. regalis* and *R. vinifera*) have very short or subterranean (acaulescent) trunks. When present, the trunk can be covered by old leaves or a dense network of fibres, which can be curly or straight, an important character to identify species (e.g. [19,20]). *Raphia* species are monoecious, with male and female flowers on the same individual and are hapaxanthic, meaning that individual stems die after a single flowering event [1]. The inflorescences structure is relatively simple and branched to two orders [1]. The first and second order branches, or rachillae, are referred to as the "partial inflorescence" [21]. The shape and overall morphology of these partial inflorescences are one of the most important taxonomic characters for species identification and to define the different sections of the genus [20,21].

Despite its importance, *Raphia* remains one of the least understood palm genera in terms of taxonomy and phylogenetic relationships [1,20]. This is mainly due to their massive size, making them difficult to collect for non-specialists, which leads to few herbarium specimens or specimens that are incomplete or fragmentary. Several attempts have been undertaken to tackle the taxonomy of the genus, beginning in the early 1900s with the first complete monograph of the genus [22]. This was followed by more regional attempts through the last century [23,24]. The last major revision of the genus was undertaken by Otedoh [21], who placed species into five different sections based on the structure of the partial inflorescence: Moniliformes (including the subsection Erectae), Temulentae, Raphiate, Flabellatae and Obclavatae.

The six species within the Moniliformes section are characterized by thin and easily breakable rachillae when fresh (1 B). Otedoh [21] also created a subsection, Erectae, where he placed two species in which the inflorescences are defined as erect (*R. autralis, R. regalis*) (1 G, O). The Temulentae section has robust and tightly appressed rachillea. The partial inflorescences are racquet-shaped with the apical second order rachillae shorter than the basal ones (1 E). This section contains three (possibly four) species, including one of the most widespread and important species *R. hookeri*. With seven species, the Raphiate section is the most complex group of the genus. Several species are only known from a few collections or just the type. This section is characterized by species having second order rachillae that are robust (thick) but loosely disposed between them (1 D). The inflorescence within this section can be semi-erect or drooping (1 I). The Flabellatae section contains two species with very characteristic partial inflorescence structures. The second order rachillae are tightly packed in a single plane being racket-shaped in appearance (1 F). The inflorescence also has very conspicuous bracts that cover completely or partially the partial inflorescences (1 O). Finally, the Obclavatae section contains one species (*R. sudanica*) with distinct club-shaped and compact partial inflorescences with large bracts covering too (1 C).

To date, no in depth morphological or molecular phylogenetic study of *Raphia* has been undertaken. The current phylogenetic analysis of the Calamoideae subfamily only included a single species, namely *R. farinifera* [25]. The main objective of this study is to generate a densely sampled phylogenetic tree of the genus and test the validity of the taxonomic sections of Otedoh [21]. In particular, we test if the partial inflorescence structure has a phylogenetic signal and is useful for *Raphia* species classification. In addition, by sampling several individuals per morphologically identified species, we also tested species limits and monophyly. In order to achieve these objectives we sequenced more than 150 palm specific nuclear markers across 56 *Raphia* accessions. We used a species delimitation approach to define species limits and generated SNP data to study at fine-scale genetic relationships in identified species complexes.

2. Results

2.1. DNA sequencing

We sequenced 56 individuals representing 18 species or 87.5% of the species diversity within the genus. A total of 15.4 million reads were generated and mapped to the reference exons belonging to 176 genes of the Heyduk et al. [26] bait kit. Across all *Raphia* and outgroup individuals the average coverage depth was 139.6x. We identified 102 genes for which 75% of the exon length was recovered in at least 25% of individuals. 20 loci were flagged by Hybpiper as paralogs because multiple assembled contigs matched a single reference locus. Those that occurred in the 75/25 set were removed, resulting in a final dataset of 85 supercontigs equalling 162kb of sequence data. Our SNP calling approach applied filters on mapping quality (>40%) depth (>25), quality by depth (>2), minimum depth across individuals (>10) minor allele frequency (>0.01) and we excluded monomorphic site. This ultimately yielded 915 and 1,627 high-quality, biallelic SNPs for the *R. hookeri* and *R. zamiana* species complexes, respectively (see below).

2.2. Evolutionary history of Raphia

We generated two phylogenetic hypotheses for *Raphia* using two distinct methods. The first analysis was conducted based on a gene-tree coalescent approach using ASTRAL while the second inferred phylogenetic relationships based on a concatenated approach using IQ-TREE.

Support varied throughout the *Raphia* ASTRAL tree - about 50% of branches had a local posterior probability (LPP) above 75% (see Figure 1 in the main text). Major clades were well supported (LPP > 80%) while relationships towards the tips of the tree generally had lower support. The final normalized quartet score, the proportion of quartet trees that agree with the species tree, was 65%, indicating that there is gene tree conflict in the genus.

The IQ-TREE concatenated approach (see Figure A1 in the supplementary materials) had increased bootstrap support compared to ASTRAL. More than 88% of branches had bootstrap support greater than 75%. The best partitioning scheme put the 85 loci into 20 different partitions. Major clades were again well-supported in this tree (bootstrap > 80%).

Our phylogenetic analyses recovered five well supported clades. Overall, these clades corresponded with the sections as defined by Otedoh [21]. *Raphia regalis* was always inferred with strong support as sister to the rest of the genus independent of the inference method (Figures 1, A1). When comparing the two phylogenetic approaches we identify a topological difference in the phylogenetic placement of the section Temulentae, the species *R. matombe* and the Moniliformes and Flabellatae sections 2. In the IQ-TREE we find weak support for the Temulentae to be sister to all *Raphia* (except *R. regalis*) (Figure 2a) yet the ASTRAL tree indicates with higher support that Temulentae is sister to a clade containing *R. matombe*, Moniliformes & Flabellatae (Figure 2b).

The relationships between species in the Raphiate section are weakly to moderately supported in both analyses (Figures 1,A1). Nevertheless, we do recover monophyletic groups in some species consistent with prior morphological identifications. This is the case for individuals of *R. laurentii* and *R. monbuttorum*, which despite low support are monophyletic. Furthermore, both these species are recovered as sister, with moderate to high support. However, our species delimitation analysis suggests that individuals identified under both species are conspecific (Figure 1 A). Support is generally higher in the ASTRAL tree, even when taking into account different gene histories, so we suggest that the ASTRAL tree represent a more accurate reconstruction of the phylogeny of *Raphia* (Fig. 2b) so we will principally refer to the relationships in this tree from now on.

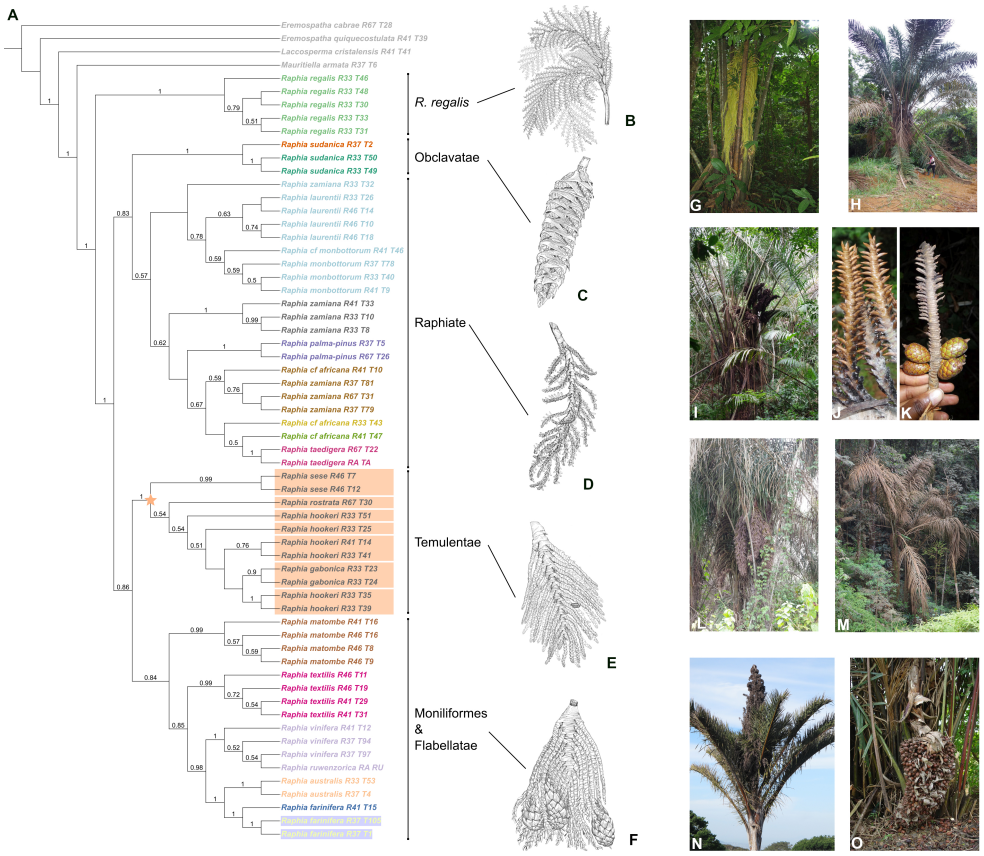

Figure 1. A: Cladogram of the genus *Raphia* inferred using 85 gene trees and ASTRAL. Values of local posterior probabilities are shown above the branches. Branch lengths are represented in Figure A2. Individuals are color coded based on the hypothesis of species delimitation inferred using SODA with $\alpha = 0.01$. A single clade, the Temulentae section marked with a star and referred to as "hookeri complex" in the main text, varied between our two values of α. The orange boxes represent the species limits using SODA with a more stringent value of $\alpha = 0.005$. Tip names contain the species name as well as the sequencing ID. B: *R. regalis* partial inflorescence representing the Moniliformes section, but see discussion. C: *R. sudanica* inflorescence, representing the Obclavatae section. D: *R. palma-pinus* inflorescence, representing the Raphiate section. E: *R. hookeri* inflorescence, representing the Temulentae section. F: *R. farinifera* inflorescence, representing the Flabellatae section. G: *R. regalis*, note the inflorescence subtended by the leaves (Couvreur 398, Cameroon). H: *R. zamiana* (Mogue Kamga 17, Gabon). I: *R. monbuttorum* (Couvreur 1212, Cameroon). J: detail of *R. monbuttorum* rachillae (Couvreur 1212, Cameroon). K: detail of *R. laurentii* rachillae (Mogue Kamga 39, Democratic Republic of Congo). L: *R. hookeri* (no voucher, Cameroon). M: *R. gabonica* (Mogue Kamga 22, Gabon). N: *R. australis* (no voucher, South Africa, Kirstenbosch Botanic Garden). O: Inflorescence of *R. vinifera* (Couvreur 638, Cameroon). B-F: Drawings reproduced from [23]; Photos G-J, L-O: T.L.P. Couvreur; Photo K: S. Mogue Kamga.

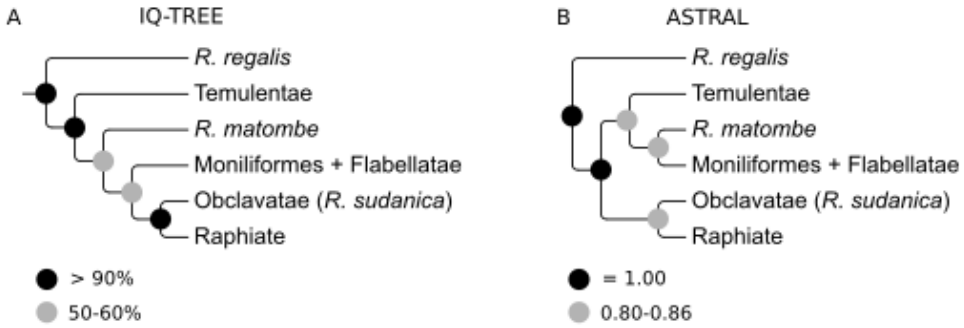

Figure 2. Major incongruences between the (**A**) concatenation (IQ-TREE) and (**B**) gene tree (ASTRAL) phylogenetic approaches. Both trees have been modified to show the relationships among major *Raphia* clades. Support values are indicated on the nodes as either (**a**) bootstrap or (**b**) local posterior probability.

2.3. Species delimitation

Our species delimitation approach yielded between 17 (α = 0.005) and 23 (α = 0.01) species in *Raphia* genus (Figure 1). Higher values of α split a clade of closely related individuals (marked with a star in Figure 1), predominantly belonging to *R. hookeri*, into seven different species. Generally, our species delimitation results corresponded to in field morphological classification of *Raphia* species using available floras (e.g. [19,20]). In some cases we found that SODA split individuals belonging *a priori* to a single species into multiple species, for example *R. farinifer a* and *R. sudanica* (Fig. 1). Conversely, individuals assigned to different species such as *R. laurentii* and *R. monbottorum* were classified as the same species after SODA delimitation independent of α values. In general, the support among different species as delimited by SODA was high (Figure 1).

2.4. Fine scale structure in two species-complexes

To further explore genetic structure among our two main species complex, namely the "zamiana complex" and the "hookeri complex" (marked with a triangle and star in Figure 1), we used SNPs extracted from the sequence data to look at the variation among individuals. The "hookeri" complex showed little evidence of clustering, with most individuals evenly spread out on the first two principal component (PC) axes (Figure 3a). We observed two major groups of >8 individuals in the "zamiana" complex along PC1 (Figure 3b), separating all of the *R. laurentii* and *R. monbuttorum* from the rest of the individuals. The first two PCs in both analyses explained 7-10% of the variance in the dataset. In general, our SNP data supports SODA species delimitation as the assigned species grouped together along one or both of the first two PCs in most cases (Figure 3a, b). Finally, our SNP data revealed that individuals within the "hookeri" complex clustered into four major groups (Figure 3a, c): the single individual from Togo; individuals from western Cameroon; individuals from East Cameroon and individuals from Gabon.

The plotting of these complexes on maps of the sampling region reveals that the delimited species cluster geographically (Figure 3c, d). In the hookeri complex, the *R. sese* individuals were sampled at a great distance from each other and *R. gabonica* falls in the middle of the *R. hookeri* distribution range. Likewise, in the zamiana complex *R. laurentii* and *R. monbottorum* are widespread, overlapping with other taxa. Many of the delimited species co-occur or are adjacent to one another in Cameroon.

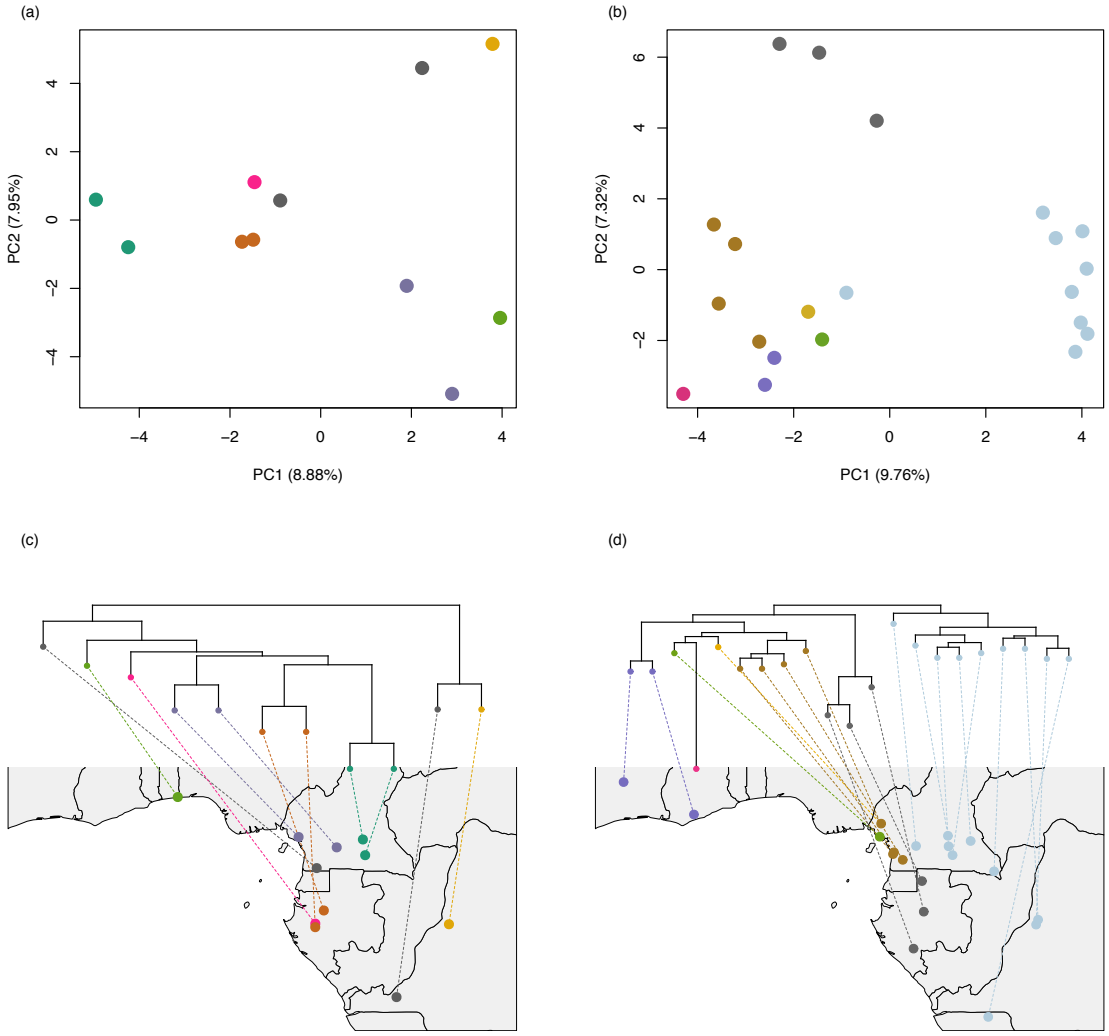

Figure 3. (**a**) Scatterplot of *R. hookeri* complex based on 915 SNPs. (**b**) Scatterplot of *R. zamiana* complex based on 1627 SNPs. Clades representing the (**c**) *R. hookeri* and (**d**) *R. zamiana* complexes were extracted from the ASTRAL (Figure 1) tree and linked to their locations on a map of central Africa. Individuals are coloured by the colours corresponding to SODA species delimitation for $\alpha = 0.01$ in (**a**) & (**c**) and $\alpha = 0.005$ in (**b**) & (**d**). An individual belonging to *R. taedigera* (RA_TA) is not shown in panels (**a**) & (**c**) due to missing data.

3. Discussion

3.1. Synthesizing morphology and molecules: the sections of Otedoh reevaluated.

Our phylogenomic analyses of *Raphia* provide a novel and overall well supported phylogenetic framework for this important African genus (Figure 1). Although some of the morphology based sections of Otedoh [21] were recovered, we also recovered some topological differences (Figure 1).

The Moniliformes and Flabellatae are not recovered as monophyletic. The Moniliformes are split into two clades (Figure 1), while the two Flabellatae species (*Raphia regalis*, *R. australis*) are not recovered as sister. In all analyses, the acaulescent central African species *Raphia regalis* is recovered with strong support as sister to the rest of the genus (Figures 1, A1). This species, together with *R. australis*, were placed within the subsection "erectae" [20,21] because the inflorescences were suggested to be "erect", in contrast to the rest of the *Raphia* species whose inflorescences are hanging or semi-erect. Our results do not support this classification, as *R. australis* is recovered as sister to *R. farinifera* (of the Flabelattae section, Figure 1) and phylogenetically divergent from *R. regalis*. A closer observation in the field showed that only the inflorescences of *R. australis* are truly erect (Figure 1 N). In contrast, the inflorescences of *R. regalis* appear erect but are in fact "supported" by the large leaves and not truly erect (Figure 1 G).

The phylogenetic placement of the Moniliformes species *R. matombe* from the Democratic Republic of the Congo and Angola is different between the two types of analyses. The close relationships between these two sections is not surprising. The inflorescences, although different in some aspects such as the clearly racket-shaped partial inflorescences in the section Flabellatae, show certain similarities not encountered in other *Raphia* species. Both have thin rachillae and the partial inflorescences are subtended by large showy bracts at least in the younger stages of development. These morphological similarities thus support the close phylogenetic relationships recovered here between these two sections.

The Obclavate section, composed of the sole species *R. sudanica*, is recovered with strong or moderate support as sister to the Raphia section. This species presents a unique inflorescence structure within the genus that is reduced and compressed into a cylindrical shape (Figure 1 C), with large bracts covering the inflorescences almost completely [20,21,27]. In addition, and in contrast to most species, *R. sudanica* thrives within the drier regions of the Sahel. These distinctive characters and its phylogenetic position support it being placed in its own section, confirming the classification of Otedoh [21].

Finally, the two remaining sections, Raphiate and Temulentae, are recovered as monophyletic, although with varying levels of support from strong to moderate (Figures 1,A1). This also confirms the classification of Otedoh [21] and the usefulness of partial inflorescence shapes in the classification of *Raphia* species.

Our results suggest that certain sections erected by Otedoh [21] are not monophyletic and need to be re-evaluated. Differences in phylogenetic relationships between the concatenated and coalescent approaches have been increasingly reported in the genomic era [28]. Our results were similar to those in Couvreur et al. [19] where higher bootstrap support were obtained when using the concatenation approach, despite the coalescent approach highlighting considerable gene tree conflict. Here, we favour the phylogenetic hypothesis recovered when using the coalescent approach (Figure 1) because these methods allow gene history to be taken into account [Some theoretical paper + [29]] and provide an arguably more realistic reconstruction of phylogenetic relationships when using a large number of independently evolving nuclear markers as used here. Our analyses suggest that we can retain five sections, only slightly different than those initially defined by Otedoh [21]. Three sections have been reconstructed in the phylogeny: Obclavatae (with its only species *R. sudanica*), Raphiate and Temulentae. The latter two sections are internally complicated, and more discussion about the phylogenetic relationship within sections is provided below. The main problem thus comes from the Moniliformes and Flabellatae sections, which are not monophyletic. *Raphia regalis* should be placed in a section of its own, linked to its unique morphology being an acaualescent species with large

inflorescences subtended in between large leaves (Figure 1 G). Finally, the last section should regroup all the other species from both the Moniliformes and Flabellatae. In both cases, we shall refrain here from erecting a new section because it is out of the scope of this paper.

3.2. *Species delimitation and species complexes*

Phylogenetic relationships between species are well to weakly resolved depending on the section, as discussed below.

3.2.1. The Moniliformes/Flabellatae section

Within the Moniliformes/Flabellatae section species relationships are generally strongly supported (Figures 1, A1) and several species are recovered as monophyletic (*R. australis*, *R. farinifera*, *R. matombe*) while species limits in others are less clear (*R. textilis*, *R. vinifera*).

Once again, there is a conflict between the concatenated and coalescent analyses. *Raphia textilis* is recovered as monophyletic with strong support (Figure 1). Nevertheless, there is little doubt that these samples represent the same species as they are morphologically similar. This is also confirmed by the species delimitation analysis at both levels of α (Figure 1,A2).

Another result recovered is the close relationship of the two montane species of *Raphia*: *R. ruwenzorica* been included within *R. vinifera*. Both species occupy a similar ecological and altitudinal range, despite being geographically separated by ca. 2,500 km. *Raphia vinifera*, which has long been mis-identified with *R. mambillensis* (now a synonym of *R. vinifera* [30]), is very common in the Cameroon Volcanic Line (CVL) in Cameroon and Nigeria, where it grows between 1,200 and 2,000m in grassland/open vegetation and is very abundant along streams and rivers [10,30]. *Raphia ruwenzorica* occurs between 800 and 1,500 m in the Albertine rift region in eastern Democratic Republic of the Congo and Burundi and has been suggested to grow in "savanna country" along valleys [20,21,31,32]. In addition, both species present similar partial inflorescence that is flat and racket shaped. However, both species differ markedly in their port with *R. ruwenzorica* reported to have a distinct tall truck reaching up to 15 m [10,21,31] whereas as *R. vinifera* is acaulescent or with a short trunk (less than 1 m; [30]). This, in addition to the 2,500+ km separating these species, suggests that they could be recognized as distinct, despite our results. Interestingly, an intraspecific CVL / Albertine rift disjunction has been documented in different taxa such as *Isolona congolana* (Annonaceae, [33]) and *Prunus africana* [34].

Raphia farinifera is the most widespread species of *Raphia*, occurring from West Africa to East Africa and Madagascar, and has also been reported from the Republic of Congo and Angola [23,32,35–40]. Our limited (3) but widespread sampling (West Africa and Madagascar) of individuals clustered together with maximum support (Figures 1,A1). However, our species delimitation analysis suggests that the Malagasy individual (R41_T15) is a different species (Figure 1). *Raphia* individuals from Madagascar were initially described as a different species (*R. ruffia*) [22] and the name subsequently synonymized with *R. farinifera* [32]. In Madagascar, *Raphia* is widely used (one of the most useful palms) and, today at least, not found in natural forests across the island [41]. This has led to the hypothesis that *Raphia* was introduced 1,500 years ago during the first wave of human colonization of the island [41]. However, Beccari [22] (p. 53) writes that the Malagasy species "prefers the vicinity of the sea where it forms whole forests in swampy places especially on the East coast" suggesting that it did at one point in time occur naturally and abundantly. Our sampling is not extensive enough to answer this question conclusively, but our results suggest that Malagasy individuals might indeed belong to a different species (*R. ruffia*) as concluded by Beccari [22] (p. 53). Finally, *R. farinifera* is recovered as sister to *R. australis*, a relationship already suggested based on morphology [32].

3.2.2. The Raphiate section

One of the most complex and least understood sections is the Raphiate section, which contains seven to eight species. Some of these species are poorly known and rarely collected, sometimes only known from a single poor quality specimen (*R. mannii*; R. *longiflora*; *R. africana*). In our study we were

not able to sample *R. mannii* and *R. longiflora*, thus our results for this section are still incomplete. Indeed, these species are morphologically similar [19], having clustering stems covered with straight fibres in addition to having semi-erect inflorescences when young (Figure 1 I), a unique character for the genus. Nevertheless, it is hard morphologically to consider these two species as conspecific. Indeed, the shape of the rachillae is quite different between these species (Figure 1 J,K). *Raphia laurentii* is characterized by rather thick rachillae covered by numerous tightly packed rachis bracts leading to an overall digitate aspect of the rachillae (Figure 1 J). In *R. monbuttorum* the rachillae are thinner and the rachis' bracts are less tightly packed around the rachillae (Figure 1 K). These differences appear to be consistent and provide useful identification characters [19].

Raphia zamiana was recently described [7]. Our broad sampling of this species, however, recovers *R. zamiana* as polyphyletic, with individuals grouping into two main clades, flagged as two different species by our species delimitation analysis (Figures 1, A2). Interestingly, these two species are geographically distinct, with one clade sampled across Gabon and one across Cameroon (Figure 3b, d), the latter containing the type of *R. zamiana*. The Gabon cluster is particularly well supported in both analyses. At this point, however, it is hard to pin point clear morphological characters differentiating these two clusters, as extensive field observations have yet to distinguish them properly.

We sampled two individuals of the Neotropical species *R. taedigera*, both from Brazil. As expected from the morphology of the partial inflorescence [21], this species grouped within the Raphiate section (Figure 1). Both individuals clustered together with strong support, and, in turn, were recovered as sister to either *R. africana* (Figure 1) or *R. palma-pinus* (Figure A1, in both cases with weak support values. Otedoh [21], following certain authors [23,42] suggested that *R. taedigera* was very close morphologically to a species identified as *R. vinifera*. However, early on the taxonomic concept of *R. vinifera* has been confusing, erroneously mistaking this species for a Raphiate type species [23,42]. Mogue Kamga et al. [30] clarified the situation showing that the name *R. vinifera* refers to a Flabelattae species mainly occurring in the CVL. To date, it remains unclear to what species Otedoh and others [21,23,42] were referring to when invoking *R. vinifera*.

Despite these taxonomic confusions, our results provide some results as to the origin of *R. taedigera*. It has been hypothesised that this species originated as a result of vicariance during the breakup of Gondwana [4]. If this were the case we would have expected that *R. taedigera* to be sister to the African species. The deeply nested position of *R. taedigera* within the genus does not support this hypothesis. Instead, our results lend some support to the conclusion of Otedoh [10] who suggested that *R. taedigera* did not show any "primitive" characters within the genus. Otedoh went further to suggest that *R. taedigera* was the result of a recent introduction in South and Central America during the slave trade some 400 years ago [10,11]. Our species delimitation results suggest, however, that *R. taedigera* is a valid species (Figure 1), at least based on the individuals sampled from Brazil. Finally, Otedoh also suggested the presence of *R. taedigera* in coastal west-central Africa [10,11]. However, to date we have not been able to locate this species in African collections and this hypothesis remains doubtful [1]. Our phylogenetic analyses suggest that *R. taedigera* is genetically quite different from other *Raphia* species (Figure A2) supporting the hypothesis that it must have dispersed to the Neotropics more than 400 years ago. This would fit with paleoecological data from Nicaragua documenting *R. taedigera* pollen over the last 2,500 years [9]. A more detailed sampling of *R. taedigera* from the Neotropics together with a dated molecular phylogeny approach will provide a better understanding of the biogeographic history of this interesting trans-Atlantic disjunction.

3.2.3. The Temulentae section

This section contains the species referred to as the "wine" palms [20,21] with three species previously included in this section (*R. hookeri*, *R. rostrata*, *R. sese*), all of which are sampled here. In addition, our results show that the newly described species, *R. gabonica* [7], is also part of the Temulentae section. This was not clear at the time of the publication as the partial inflorescence suggested a possible relationship with the Moniliformes section [7]. Overall, species identified based

on morphology clustered together (e.g. *R. gabonica*, *R. sese*) with strong or low support. Nevertheless, all four species show a very close phylogenetic proximity, suggesting that this section could be regarded as one large species complex. Indeed, depending on the level of stringency, our species delimitation analysis recovered either seven distinct species or one single species (Figures 1, A2). It is important to note that changing levels of α did not impact species delimitation in the other sections. Morphologically, however, these species are different and can easily be identified in the field, which is partly supported by our phylogenetic analysis. For example, *R. gabonica* resembles *R. hookeri* in the clearly visible single stem covered with characteristic curly fibers, but differs markedly by being a *terra firma* low-density species with thin (Moniliformes-like) and densely packed rachillae. In contrast, *R. hookeri* is a swampy species, growing in large, monodominant stands with robust and more evenly-spaced rachillae [7,19]. In the same way, *R. rostrata* is characterized by a small but clustering stem with curly mixed with straight hanging fibers and occurs along rivers with strong currents [19].

Raphia hookeri is recovered here as polyphyletic, possibly including four different cryptic species. This is one of the most important, abundant and widespread *Raphia* species and its overall morphology is rather constant across its range. However, individuals appear to be geographically structured like in *R. zamiana* (see above). Interestingly, this mirrors patterns of genetic structure recovered across a wide range of central African plant species [43,44], including *R. zamiana*.

4. Conclusions

Our results provide a new step forward in understanding the phylogenetic relationships and taxonomy within this major African palm genus. We show that the morphological sections based on partial inflorescence shape defined by Otedoh [21] are relatively robust overall, even though two sections will need to be grouped and redefined morphologically. Our results also uncover important species delimitation problems defined here as species complexes (*R. hookeri, R. zamiana*) that must be solved if we are to have a thorough understanding of *Raphia* systematics. Given the economic and ecological importance of *R. hookeri*, clarifying its species delimitation will be important in the future. Different approaches could rely on more in-depth population level studies using more variable markers (e.g. microsatellites) combined with detailed morphometric measurements as has been done in other African tree species [45]. We show here that the Heyduk et al. bait kit [26] is useful for understanding relationships within the *Raphia* genus and between species as in other groups [**?**], although it appears to be limited for untangling species complexes. Resolving relationships within *Raphia* will thus rely on more data, including increased intra-species sampling, detailed morphological studies in certain species and larger baiting kits e.g. [46].

5. Materials and Methods

5.1. Species sampling, library preparation and DNA sequencing

We sampled a total of 56 individuals (see A1 for details) representing 18 out of the 21 species accepted to date [7,19] and representing all sections described by Otedoh [21]. In order to collect proper material for sequencing, several field trips were undertaken across several African countries including Ivory Coast, Ghana, Gabon, Cameroon, Angola and the Demographic Republic of the Congo between 2012 and 2017. We were not able to access material from three accepted species: *R. gentiliana*, *R. mannii* and *R. longiflora*. We sampled two to seven individuals per species in order to test for monophyly. However, only a single specimen was available for *R. ruwenzorica*. Finally, we sampled four species within Calamoideae as outgroups: *Eremospatha cabrae*, *Eremospatha quiquecostulata*, *Laccosperma cristalensis* and *Mauritiella armata* following [25,47]. We extracted DNA from leaves dried in silicagel, except for one individual of *R. taedigera* and the only individual of *R. ruwenzorica* for which DNA was extracted from herbarium dried material.

Methods for DNA extraction, preparation of sequencing libraries, hybridization, Illumina MiSeq DNA sequencing and read cleaning followed [19]. In brief, barcoded Illumina libraries were

constructed based on a modified protocol of Rohland and Reich [48]. We hybridized DNA to defined exons using the palm-specific nuclear baiting kit of Heyduk et al. [26]. This kit allows to sequence exons from 176 nuclear genes across the palm family.

5.2. Contig assembly and multi-sequence alignment

We used HybPiper (v1.2) [49] to process our cleaned reads (following [19]) to obtain sequences corresponding to the target exons plus associated intronic sequence data (referred to as supercontigs). We aligned each set of supercontigs using MAFFT (v7.305) [50] with the –auto option and cleaned these alignments with GBLOCKS (v0.91b) [51] using the default parameters and all allowed gap positions.

To identify a suitable set of loci for phylogenetic inference we selected only those supercontigs that had 75% of their exon length reconstructed in at least 25% of individuals (referred to as 75/25). We used only those loci in which at least 75% of the exon length was recovered because the use of fragmented sequences is known to increase gene tree error, whereas the number of individuals has little effect as long as the gene tree is accurate [52].

5.2.1. Paralog identification

HybPiper flags potential paralogs when multiple contigs are discovered mapping well to a single reference sequence. We ran hybpiper on the 837 exons that made up the baiting kit [26], identified flagged loci and constructed exon trees using RAxML (v8.2.9) [53]. We examined each tree to determine whether putative paralogs formed a species clade. When sequences concerning more than three individuals were flagged for a locus, we examined whether the 'main' and alternative sequences formed separate clades. If so this locus was classified as a paralog and discarded from the dataset. For each gene, we then calculated at the proportion of exons that we confirmed as paralogs after inspection. If this proportion was <%50 we removed the entire gene from our analyses.

5.3. Coalescent phylogenetic inference

Individual gene trees were constructed with 100 bootstraps and the GTRGAMMA model using RAxML (v8.2.9) [53] (option "-f a"). If after inference, branches had bootsrap support values >10 they were collapsed using the program nw_ed [54] because this approach has been shown to improve the accuracy of ASTRAL [29]. We used the selected 75/25 gene trees as our input to run ASTRAL-III (v5.5.11) [29] using the default options.

5.4. Species delimitation

After constructing our ASTRAL tree we used the associated approach SODA [55]. Simulations using this approach have shown it to be of similar accuracy or more accurate [55] than other popular species delimitation methods such as BPP [56] at a fraction of the computational cost. SODA uses frequencies of quartet topologies to determine if each branch in a guide tree inferred from gene trees (i.e. the ASTRAL tree from above) is likely to have a positive length. This identifies where in the tree coalescence is random, and where it is non-random. It then uses the results to infer a new, extended species tree that defines boundaries among species. We used two cut-off values of α (confidence level): 0.01 and 0.005.

5.5. Maximum-likelihood phylogenetic inference

After suitable loci were identified we filled any missing individuals in each alignment with an empty sequence. We then concatenated all aligned loci using the *pxcat* function in the program phyx [57]. We used IQ-TREE (v1.6.8; [58])to infer a maximum likelihood tree of all individuals. We partitioned our dataset so that each supercontig had a separate substitution model and used the following options when running the program: "-m MFP+MERGE -rcluster 10 -bb 1000 -alrt 1000". We selected the optimal partitioning scheme using ModelFinder [59], choosing the best model based on

Bayesian Information Criterion (BIC) score and merging genes until model fit stopped increasing. We also used rcluster [60] to decrease computational load. We made use of the ultrafast bootstrapping ([61]; 1000 replicates) and the SH-like approximate likelihood ratio test ([62]; 1000 replicates) to assess branch support in the tree.

5.6. SNP calling

To call SNPs we first used SeCaPr (v1.1.4; [63]) to build a psuedoreference. After filtering out low coverage and paralogous loci, consensus sequences are built and combined to form a reference file that is closer to the study group than the original, and will recover more data. We mapped our cleaned, paired reads to this new, dataset-specific reference using BWA (v0.7.12; [64]). Duplicates were removed and we called SNPs using the program HaplotypeCaller in GATK (v4.0; [65]). We applied thresholds to mapping quality (>40%) depth (>25), quality by depth (>2), minimum quality across all individuals (>10) and minor allele frequency (>0.01) to filter SNPs using bcftools (v1.8; [66]). We kept only biallelic SNPs and excluded monomorphic sites.

5.7. Genetic clustering

We performed Discriminant Analysis of Principal Components (DAPC) [67] to identify genetic clusters in two species complexes of *Raphia*. We used the function *find.clusters* in the R package 'adegenet' [68] to infer the number of clusters using successive K-means with 100,000 iterations per value of k up to k = 20. We used BIC to identify the best-fitting number of clusters. We then used the function *dapc* [67] to define the diversity among the clusters identified. We chose the optimum number of axes to use with the function *optim.a.score*.

Author Contributions: Conceptualization, AJH, TLPC and BS; methodology KB and AH; validation AJH, KB and MKS; formal analysis AJH; resources LT, FS, CDB, AZ, AA, MKS, TLPC; data curation KB, AJH and MKS; writing–original draft preparation AJH and TLPC; writing–review and editing, TO COMPLETE; supervision, TLPC and BS; project administration, TLPC; funding acquisition, TLPC. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by Agropolis Fondation (RAPHIA project) under the reference ID 1403-026 through the «Investissements d'avenir» program (ANR-10-LABX-0001-01; ISITE MUSE: ANR-16-IDEX-0006) and Agence Nationale de la Recherche (ANR-15- CE02-0002-01), both to TLPC. This work was also part of the PALM-A-GOV (project number: 1702-007) project to TLPC, publicly funded through the ANR (the French National Research Agency) under the "Investissements d'avenir" program with the reference ANR-10-LABX001-01 Labex Agro and coordinated by Agropolis Fondation within the I-SITE MUSE (ANR-16-IDEX-0006). AZ is funded by iDiv via the German Research Foundation (DFG FZT 118), specifically through sDiv, the Synthesis Centre of iDiv. C.D.B. is funded by the Swedish Research Council (2017-04980). A.A. is funded by the Swedish Research Council, the Swedish Foundation for Strategic Research, the Knut and Alice Wallenberg Foundation and the Royal Botanic Gardens, Kew. ADD funding others here

Acknowledgments: The authors acknowledge the IRD itrop HPC (South Green Platform) at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper. We are grateful to the Centre National de la Recherche Scientifique et Technique (CENAREST) in Gabon and the Agence National des Parques Nationaux (ANPN) for delivering the necessary research permits for our field work in Gabon in 2015 and 2016 (permits AR0020/16; AR0036/15 (CENAREST) and AE16014 (ANPN)). Prof. Paul Henri Bourobou Bourobou is thanked for his assistance in obtaining these research permits. Fieldwork in Cameroon was undertaken under the "accord cadre de cooperation" between the IRD and Ministère de la Recherche Scientifique et Technique (MINRESI). TU Dresden maintains a Memorandum of Understanding with the Instituto Nacional da Biodiversidade e Áreas de Conservação (INBAC) of the Ministério do Ambiente da República de Angola that includes collection and export permits.. FOR THOSE HOW HAVE PERMITS PLEASE ADD OR SEND TO THOMAS

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CVL	Cameroon Volcanic Line
SNP	Single Nucleotide Polymorphism
PC	Principal Components Analysis

Appendix A Supplementary Tables

Genus	Species epithet	ID	Collector (herbarium)	Country	Latitude	Longitude	Run	Tag	No. mapp	Mean dep	Stdev dep
Eremospatha	*cabrae*	R227	Couvreur 1165 (WAG)	Cameroon	-1.03044	10.51881	RUN67	TAG-28	441790	315.48	469.205
Eremospatha	*quiquecostulata*	R162	Couvreur 1079 (WAG)	Gabon	-5.02068	15.1545	RUN41	TAG39	91982	46.1243	68.317
Laccosperma	*cristalensis*	R164	Couvreur 1142 (WAG)	Gabon	-5.73485	14.2162	RUN41	TAG41	105250	54.4592	86.7057
Mauritiella	*armata*	R135	Couvreur 257 (NY)	Bolivia	-1.45585	12.5863	RUN37	TAG6	406142	213.159	1244.61
Raphia	*africana (cf)*	R072	Couvreur 971 (WAG)	Cameroon	5.48034	10.05056	RUN41	TAG10	121972	63.533	69.0331
Raphia	*africana (cf)*	R174	No voucher, close to Couvreur 971 (WAG)	Cameroon	-0.8852	18.1337	RUN41	TAG47	50272	26.0156	28.656
Raphia	*africana (cf)*	R77	No voucher, close to Couvreur 971 (WAG)	Cameroon	-2.37125	11.16443	RUN33	TAG43	315915	159.931	188.084
Raphia	*australis*	R130	MBC 99 874D (SANBI)	Cultivated, Kirstenbosch National Botanical Garden, South Africa	-33.9878	18.43262	RUN37	TAG4	913167	473.13	627.945
Raphia	*australis*	R92	MBC 99 874D (SANBI)	Cultivated, Kirstenbosch National Botanical Garden, South Africa	-33.9878	18.43262	RUN33	TAG53	255329	129.65	156.525
Raphia	*farinifera*	R127	Baker 1357 (K)	Cultivated, Royal Botanic Gardens, Kew, U.K.	N/A	N/A	RUN37	TAG105	486723	250.42	309.47
Raphia	*farinifera*	R128	Baker 1410 (K)	Burkina Faso	5.39625	-1.38278	RUN37	TAG1	476130	245.7	292.787
Raphia	*farinifera*	R132	Dransfield 7516 (K)	Madagascar	-1.84988	13.85903	RUN41	TAG15	133441	68.7643	85.4757
Raphia	*gabonica*	R34	Kamga Mogue 22 (WAG)	Gabon	2.49409	10.34844	RUN33	TAG23	284243	143.822	174.031
Raphia	*gabonica*	R36	Kamga Mogue 23 (WAG)	Gabon	3.4825	13.59469	RUN33	TAG24	212852	108.979	136.141
Raphia	*hookeri*	R124	Couvreur 984 (WAG)	Cameroon	3,52108	11,74376	RUN41	TAG14	78554	40.8646	44.4911
Raphia	*hookeri*	R37	Kamga Mogue 25 (WAG)	Gabon	3.59972	11.2877	RUN33	TAG25	278678	142.219	175.028
Raphia	*hookeri*	R63	Kamga Mogue 1 (WAG)	Cameroon	1.59848	11.62294	RUN33	TAG35	319096	162.68	210.004
Raphia	*hookeri*	R69	Kamga Mogue 12 (WAG)	Cameroon	4.11224	9.56915	RUN33	TAG39	197508	100.49	122.815
Raphia	*hookeri*	R71	Kamga Mogue 26 (WAG)	Cameroon	4.87064	9.26582	RUN33	TAG41	292201	149.509	181.385
Raphia	*hookeri*	R89	Michon 01 (G)	Togo	-2.25428	11.14284	RUN33	TAG51	201880	103.103	124.193
Raphia	*laurentii*	R186	Kamga Mogue 39 (WAG)	Democratic Repubic of Congo	-7.94817	15.83894	RUN46	TAG10	87838	47.1055	51.2042
Raphia	*laurentii*	R198	Kamga Mogue 42 (WAG)	Democratic Repubic of Congo			RUN46	TAG14	86786	46.5246	51.1853
Raphia	*laurentii*	R208	Lautenschläger 806 (JACQ)	Angola	3.137672	9.971397	RUN46	TAG18	64379	34.3542	37.7478
Raphia	*laurentii*	R40	Ayole 01 (YA)	Cameroon	3.19962	10.51772	RUN33	TAG26	272509	138.397	168.066
Raphia	*matombe*	R134	19392103 (BR)	Cultivated, Meide Botanical Garden, Belgium	-0.1473	11.726	RUN41	TAG16	76250	39.5975	44.2225
Raphia	*matombe*	R181	Kamga Mogue 37 (WAG)	Democratic Repubic of Congo	-0.8642	18.1458	RUN46	TAG8	55883	29.9284	31.68
Raphia	*matombe*	R183	Kamga Mogue 38 (WAG)	Democratic Repubic of Congo	-0.60673	18.2468	RUN46	TAG9	66676	35.6465	38.4587
Raphia	*matombe*	R206	Lautenschläger 1095 (JACQ)	Angola	2.339	10.6025	RUN46	TAG16	59006	31.6789	34.7279
Raphia	*monbottorum*	R173	Kamga Mogue 31 (WAG)	Cameroon	-5.65308	14.3181	RUN41	TAG46	98860	50.8202	56.6257
Raphia	*monbottorum*	R059	Kandem 211 (WAG)	Cameroon	6.366477	10.8946	RUN41	TAG9	104568	54.2312	58.009
Raphia	*monbottorum*	R66	Kamga Mogue 04 (WAG)	Cameroon	3.58237	13.14197	RUN37	TAG78	407023	208.475	245.939
Raphia	*monbottorum*	R70	Kamga Mogue 13 (WAG)	Cameroon	4.87036	9.26579	RUN33	TAG40	352731	180.143	231.782
Raphia	*palma-pinus*	R133	Ouatara & Stauffer 14 (G)	Ghana	-1.40695	12.5712	RUN37	TAG5	792856	411.261	507.288
Raphia	*palma-pinus*	TC-S1328	Stauffer 857 (G)	Ivory Coast	7.24598	-5.39625	RUN67	TAG-26	459813	318.875	309.014
Raphia	*regalis*	R55	Couvreur 685 (WAG)	Cameroon	3.885056	14.39931	RUN33	TAG30	184893	94.4991	128.908
Raphia	*regalis*	R56	Couvreur 753 (WAG)	Cameroon	3.07485	13.3663	RUN33	TAG31	210853	107.898	137.475
Raphia	*regalis*	R58	Couvreur 398 (WAG)	Cameroon	3.97037	13.2367	RUN33	TAG33	388670	196.877	269.794
Raphia	*regalis*	R81	Weiringa 8539 (WAG)	Gabon	-2.3714	11.16399	RUN33	TAG46	161079	81.8365	100.151
Raphia	*regalis*	R83	Weiringa 8333 (WAG)	Gabon	7.55826	2.19247	RUN33	TAG48	150544	76.7195	98.4529
Raphia	*rostrata*	R229	Kamga Mogue 43 (WAG)	Cameroon	-0.82955	10.52294	RUN67	TAG-30	476709	329.286	313.366
Raphia	*ruwenzorica*	RA_RU	Robyns 4039 (BR) - herbarium sample	Democratic Repubic of Congo			N/A	N/A	N/A	N/A	N/A
Raphia	*sese*	R179	Kamga Mogue 36 (WAG)	Democratic Repubic of Congo	-0.88513	18.1337	RUN46	TAG7	62972	33.7432	37.0717
Raphia	*sese*	R195	Kamga Mogue 41 (WAG)	Democratic Repubic of Congo	-6.74019	16.20117	RUN46	TAG12	88659	47.5337	50.4421
Raphia	*sudanica*	R129	Bayton 70 (K)	Burkina Faso	10.59388889	5.30694444	RUN37	TAG2	471997	246.119	299.37
Raphia	*sudanica*	R86	Michon 09 (G)	Benin	8.98317	1.49297	RUN33	TAG49	379755	193.881	231.648
Raphia	*sudanica*	R88	Michon 56 (G)	Togo	6.39183	2.67703	RUN33	TAG50	355099	181.913	220.289
Raphia	*taedigera*	Env0563	MBC 94803 A (MBC)	Cultivated, Montgomery Botanical Garden	N/A	N/A	RUN67	TAG-22	666580	446.883	465.14
Raphia	*taedigera*	RA_TA	Noblick 5015 (K) - (herbarium sample)	Brazil	-1.5666	-48.73333	N/A	N/A	N/A	N/A	N/A
Raphia	*textilis*	R149	Couvreur 743 (WAG)	Gabon	0.6059	10.4118	RUN41	TAG29	132014	68.4236	80.0799
Raphia	*textilis*	R151	Couvreur 1075 (WAG)	Gabon	4.18332	13.10538	RUN41	TAG31	80524	41.722	45.5424
Raphia	*textilis*	R192	Kamga Mogue 40 (WAG)	Democratic Repubic of Congo	-6.14997	15.40333	RUN46	TAG11	78738	42.0848	46.2868
Raphia	*textilis*	R209	Lautenschläger 1086 (JACQ)	Angola	-0.07916	11.00836	RUN46	TAG19	50163	26.882	30.0834
Raphia	*vinifera*	R105	Couvreur 638 (WAG)	Cameroon	10.59389	-5.30694	RUN41	TAG12	87308	45.4189	49.1511
Raphia	*vinifera*	R113	No voucher, but close to Couvreur 638 (WAG)	Cameroon	10.59389	-5.30694	RUN37	TAG94	286322	145.81	177.287
Raphia	*vinifera*	R116	No voucher, but close to Couvreur 638 (WAG)	Cameroon	10.59389	-5.30694	RUN37	TAG97	367120	189.172	220.664
Raphia	*zamiana*	R07	Kamga Mogue 17 (WAG)	Gabon	6.27413	10.51091	RUN33	TAG8	217692	110.68	136.444
Raphia	*zamiana*	R09	Kamga Mogue 17 (WAG)	Gabon	3.52108	11.74376	RUN33	TAG10	228828	116.278	144.853
Raphia	*zamiana*	R154	Couvreur 1122 (WAG)	Gabon	4.12977	9.21399	RUN41	TAG33	56712	29.1983	31.1171
Raphia	*zamiana*	R230	Kamga Mogue 45 (WAG)	Cameroon	2.15329	15.7367	RUN67	TAG-31	468825	318.682	376.272
Raphia	*zamiana*	R57	Couvreur 427 (WAG)	Cameroon	3.07684	13.36761	RUN33	TAG32	411897	209.629	268.899
Raphia	*zamiana*	R93	Ayole 20 (YA)	Cameroon	3.23685	10.02514	RUN37	TAG79	300001	153.527	177.289
Raphia	*zamiana*	R95	Ayole 32 (YA)	Cameroon	2.80897	10.52734	RUN37	TAG81	420757	215.428	250.534

Table A1. Table 1: Herbarium specimen collector and number and herbarium deposited, coordinates of collection and country for all samples included in our study. All samples were extracted from silicagel dried leaves, expect when stated otherwise. The last five columns refer to the sequencing identification (TAG and INDEX used) and different sequencing statistics.

Appendix B Supplementary Figures

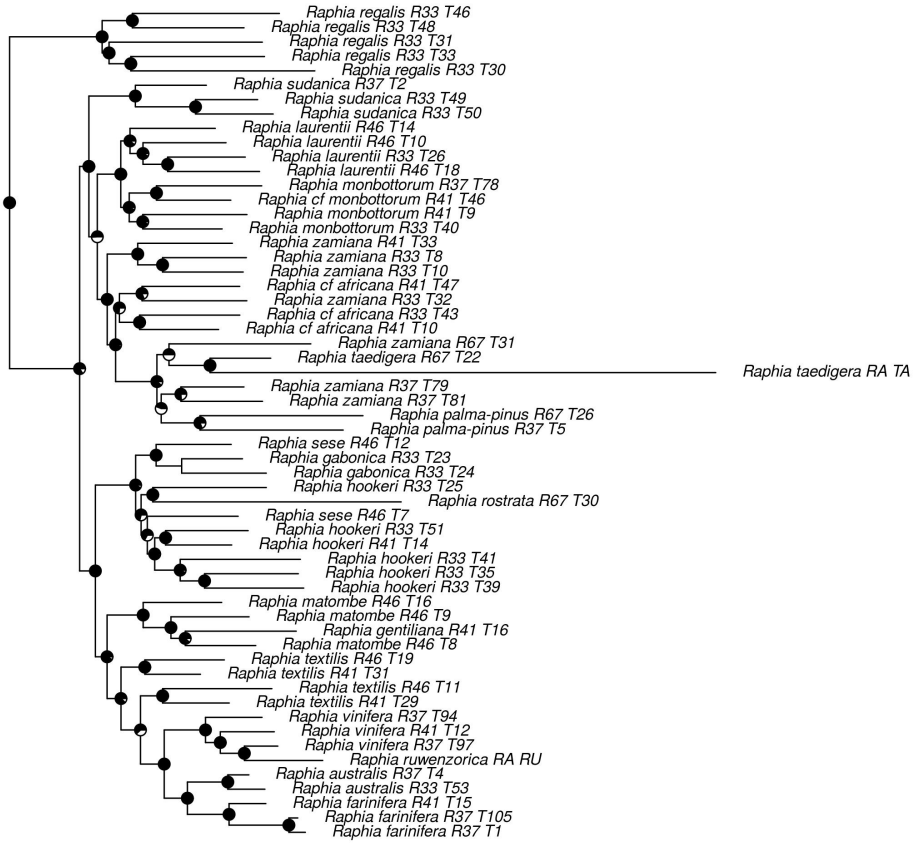

Figure A1. IQTREE *Raphia* inferred using 162kb of sequence data. Values for ultrafast bootstrap support are depicted on nodes.

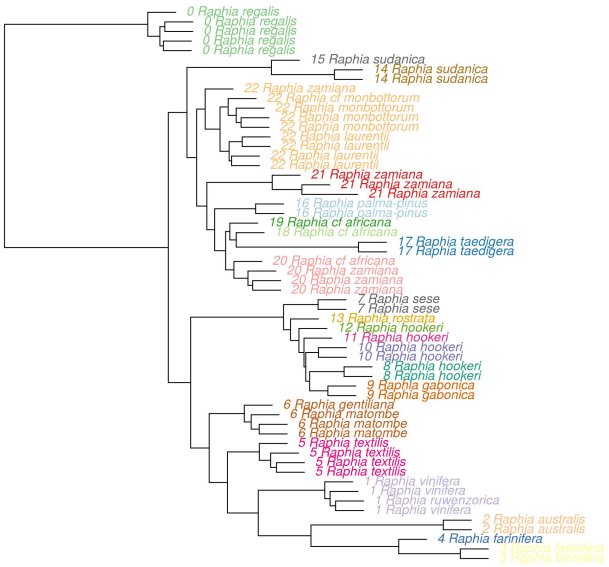

Figure A2. ASTRAL tree of *Raphia* including inferred branch lengths (except terminal branch lengths) and tip labels coloured with species delimitation as inferred with SODA ($\alpha = 0.01$).

References

1. Dransfield, J.; Uhl, N.; Asmussen, C.B.; Baker, W.; Harley, M.M.; Lewis, C.E. *Genera Palmarum: The evolution and classification of palms*; Kew Publishing: Kew, 2008.
2. Stauffer, F.W.; Ouattara, D.N.; Roguet, D.; da Giau, S.; Michon, L.; Bakayoko, A.; Ekpe, P. An update to the African palms (Arecaceae) floristic and taxonomic knowledge, with emphasis on the West African region. *Webbia* **2017**, *72*, 1–14. doi:10.1080/00837792.2017.1313381.
3. Cosiaux, A.; Gardiner, L.M.; Stauffer, F.W.; Bachman, S.P.; Sonké, B.; Baker, W.J.; Couvreur, T.L.P. Low extinction risk for an important plant resource: Conservation assessments of continental African palms (Arecaceae/Palmae). *Biological Conservation* **2018**, *221*, 323–333. doi:10.1016/j.biocon.2018.02.025.
4. Moore, H. Palms in the tropical forest ecosystems of Africa and South America. In *Tropical forest ecosystems of Africa and South America: a comparative review*; Meggers, B.; Ayensu, E.; Duckworth, W., Eds.; Smithsonian Institution Press: Washintong DC, 1973; pp. 63–88.
5. Couvreur, T. Odd man out: Why are there fewer plant species in African rain forests? *Plant Systematics and Evolution* **2015**, *301*, 1299–1313. doi:10.1007/s00606-014-1180-z.
6. Dransfield, J. The palms of Africa and their relationships. In *Modern systematic studies in African botany*; Goldblatt, P.; Lowry, P.P., Eds.; Missouri Botanical Garden Press, 1988; pp. 95–103.
7. Mogue Kamga, S.; Niangadouma, R.; Stauffer, F.W.; Sonké, B.; Couvreur, T.L.P. Two new species of Raphia (Palmae/Arecaceae) from Cameroon and Gabon. *PhytoKeys* **2018**, *111*, 17–30. doi:10.3897/phytokeys.111.27175.
8. Urquhart, G.R. Long-term Persistence of Raphia taedigera Mart. Swamps in Nicaragua1. *Biotropica* **1999**, *31*, 565–569. doi:10.1111/j.1744-7429.1999.tb00403.x.
9. Urquhart, G.R. Paleoecological evidence of Raphia in the Pre-Columbian Neotropics. *Journal of Tropical Ecology* **1997**, *13*, 783–792. doi:10.1017/S0266467400010993.
10. Otedoh, M. Systematic Studies in Raphia Palms. Doctoral Thesis, University of Reading, UK., 1976.
11. Otedoh, M.O. The African origin of Raphia taedigera – Palmae. *J. Niger. Inst. Oil Palm Res* **1977**, *42*, 11–16.
12. Kamga, M.; Brokamp, G.; Cosiaux, A.; Awono, A.; Fürniss, S.; Barfod, A.S.; Muafor, F.J.; Le Gall, P.; Sonké, B.; Couvreur, T. Use and cultural significance of Raphia palms. *Economic Botany*.

13. Lautenschläger, T.; Neinhuis, C., Eds. *Riquezas naturais de Uíge – uma breve introdução sobre o estado atual, a utilização, a ameaça e a preservação da biodiversidade*, technische universität dresden ed.; 2014.
14. Obahiagbon, F.I. A review of the origin, morphology, cultivation, economic products, health and physiological implications of raphia palm. *African Journal of Food Science* **2009**, *3*, 447–453. 00011.
15. Couvreur, T.L.P.; Fumtim, J. A l'Echelle du Raphia / On Raphia and Man, 2017.
16. Profizi, J.P. Swampy Area Transformations by Exploitation of Raphia hookeri (Arecaceae) in Southern Benin (West Africa). *Human Ecology* **1988**, *16*, 87–94.
17. Dargie, G.C.; Lewis, S.L.; Lawson, I.T.; Mitchard, E.T.A.; Page, S.E.; Bocko, Y.E.; Ifo, S.A. Age, extent and carbon storage of the central Congo Basin peatland complex. *Nature* **2017**, *542*, 86–90. doi:10.1038/nature21048.
18. Rainey, H.J.; Iyenguet, F.C.; Malanda, G.A.F.; Madzoké, B.; Santos, D.D.; Stokes, E.J.; Maisels, F.; Strindberg, S. Survey of Raphia swamp forest, Republic of Congo, indicates high densities of Critically Endangered western lowland gorillas Gorilla gorilla gorilla. *Oryx* **2010**, *44*, 124–132. doi:10.1017/S003060530999010X.
19. Couvreur, T.L.P.; Helmstetter, A.J.; Koenen, E.J.M.; Bethune, K.; Brandão, R.D.; Little, S.A.; Sauquet, H.; Erkens, R.H.J. Phylogenomics of the Major Tropical Plant Family Annonaceae Using Targeted Enrichment of Nuclear Genes. *Frontiers in Plant Science* **2019**, *9*, 1941. doi:10.3389/fpls.2018.01941.
20. Tuley, P. *The palms of Africa*; The Tendrine Press, Trendrine Press: Zennor, 1995.
21. Otedoh, M.O. A revision of the genus Raphia Beauv. (Palmae). *J. Niger. Inst. Oil Palm Res* **1982**, *6*, 145–189. 00011.
22. Beccari, O. Studio monografico del genere "Raphia". *Webbia* **1910**, *3*, 37–130. doi:10.1080/00837792.1910.10669558.
23. Russell, T.A. The Raphia Palms of West Africa. *Kew Bulletin* **1965**, *19*, 173–196. doi:10.2307/4108027.
24. Tuley, P.; Russell, T.A. The Raphia palms reviewed. *Nigerian Field* **1966**, *31*, 54–65.
25. Baker, W.J.; Dransfield, J.; Hedderson, T.A. Phylogeny, character evolution, and a new classification of the calamoid palms. *Systematic Botany* **2000**, *25*, 297–322.
26. Heyduk, K.; Trapnell, D.W.; Barrett, C.F.; Leebens-Mack, J. Phylogenomic analyses of species relationships in the genus Sabal (Arecaceae) using targeted sequence capture. *Biological Journal of the Linnean Society* **2016**, *117*, 106–120. doi:10.1111/bij.12551.
27. Ouattara, D.N.; Stauffer, F.W.; Bakayoko, A. Lectotypification de Raphia sudanica A. Chev.(Arecaceae, Calamoideae), et commentaires sur la biologie et la conservation de l'espèce. *Adansonia* **2014**, *36*, 53–61. 00002.
28. Liu, X.; Fu, Y.X. Exploring population size changes using SNP frequency spectra. *Nature Genetics* **2015**, *47*, 555–559. doi:10.1038/ng.3254.
29. Zhang, C.; Sayyari, E.; Mirarab, S. ASTRAL-III: Increased scalability and impacts of contracting low support branches. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017, Vol. 10562 LNBI, pp. 53–75. doi:10.1007/978-3-319-67979-2_4.
30. Mogue Kamga, S.; Sonké, B.; Couvreur, T.L.P. Raphia vinifera (Arecaceae; Calamoideae): Misidentified for far too long. *Biodiversity Data Journal* **2019**, *7*, e37757. doi:10.3897/BDJ.7.e37757.
31. Robyns, W.; Tournay, R. Monocotylées nouvelles ou critiques de la région du Parc National Albert (Congo belge). *Bulletin du Jardin botanique de l'État a Bruxelles* **1955**, *25*, 239–260. doi:10.2307/3667057.
32. Stauffer, F.W.; Ouattara, D.; Stork, A.L. Palmae. In *Tropical African Flowering Plants: Monocotyledons 2*; Lebrun, J.P.; Stork, A.L., Eds.; Conservatoire et Jardin botaniques de la Ville de Genève, Switzerland.: Genève, 2014; Vol. 8, pp. 326–354.
33. Couvreur, T. Monograph of the syncarpous African genera *Isolona* and *Monodora* (Annonaceae). *Systematic Botany Monographs* **2009**, *87*, 1–150.
34. Kadu, C.a.C.; Schueler, S.; Konrad, H.; Muluvi, G.M.M.; Eyog-Matig, O.; Muchugi, A.; Williams, V.L.; Ramamonjisoa, L.; Kapinga, C.; Foahom, B.; Katsvanga, C.; Hafashimana, D.; Obama, C.; Geburek, T. Phylogeography of the Afromontane Prunus africana reveals a former migration corridor between East and West African highlands. *Molecular Ecology* **2011**, *20*, 165–178. doi:10.1111/j.1365-294X.2010.04931.x.
35. Dale, I.R. Palms of Kenya. *The Journal of the East Africa and Uganda Natural History Society* **1938**, *23*, 183–186.
36. Russell, T.A. Palmae. In *Flora of West Tropical Africa*, 2 ed.; Hutchinson, J.; Dalziel, J.M.; Hepper, F., Eds.; Crown Agents for Oversea Governments and Administrations: London, 1968; Vol. 3, pp. 159–169.

37. Dransfield, J. Palmae. In *Flora of Tropical East Africa*; Polhill, R., Ed.; A.A. Balkema: Rotterdam, 1986; pp. 1–58.
38. Dransfield, J.; Beentje, H.J. *The palms of Madagascar*; Royal Botanic Gardens and International Palm Society: Kew, 1995. 00134.
39. Baker, W.; Smith, G.F. Palmae. In *Plants of Angola*; Figueiredo, E., Ed.; Number 23 in Strelitzia, 2008; pp. 172–173.
40. Faye, A.; Moutsamboté, J.; Couvreur, T. Palms in southern Republic of Congo. *PALMS* **2016**, *59*, 181–190.
41. Dransfield, J.; Rakotoarinivo, M. The biogeography of Madagascar palms. In *The Biology of Island Floras*, Cambridge University Press ed.; Bramwell, D.; Caujapé-Castells, J., Eds.; 2011; pp. 179–196.
42. Mann, G.; Wendland, H. On the Palms of Western Tropical Africa. *Transactions of the Linnean Society of London* **1864**, *24*, 421–440. doi:10.1111/j.1096-3642.1863.tb00165.x.
43. Hardy, O.J.; Born, C.; Budde, K.; Daïnou, K.; Dauby, G.; Duminil, J.Ô.; Ewédjé, E.E.B.; Gomez, C.; Heuertz, M.; Koffi, G.K.; Lowe, A.J.; Micheneau, C.; Ndiade-Bourobou, D.; Piñeiro, R.; Poncet, V. Comparative phylogeography of African rain forest trees: A review of genetic signatures of vegetation history in the Guineo-Congolian region. *Comptes Rendus - Geoscience* **2013**, *345*, 284–296. doi:10.1016/j.crte.2013.05.001.
44. Helmstetter, A.J.; Amoussou, B.E.N.; Bethune, K.; Kandem, N.G.; Kakaï, R.G.; Sonké, B.; Couvreur, T.L.P. Phylogenomic data reveal how a climatic inversion and glacial refugia shape patterns of diversity in an African rain forest tree species. *bioRxiv* **2019**, p. 807727. doi:10.1101/807727.
45. Lissambou, B.J.; Couvreur, T.L.P.; Atteke, C.; Stévart, T.; Piñeiro, R.; Dauby, G.; Monthe, F.K.; Ikabanga, D.U.; Sonké, B.; M'batchi, B.; Hardy, O.J. Species delimitation in the genus Greenwayodendron based on morphological and genetic markers reveals new species. *TAXON* **2019**, *68*, 442–454. doi:10.1002/tax.12064.
46. Loiseau, O.; Olivares, I.; Paris, M.; de La Harpe, M.; Weigand, A.; Koubínová, D.; Rolland, J.; Bacon, C.D.; Balslev, H.; Borchsenius, F.; Cano, A.; Couvreur, T.L.P.; Delnatte, C.; Fardin, F.; Gayot, M.; Mejía, F.; Mota-Machado, T.; Perret, M.; Roncal, J.; Sanin, M.J.; Stauffer, F.; Lexer, C.; Kessler, M.; Salamin, N. Targeted Capture of Hundreds of Nuclear Genes Unravels Phylogenetic Relationships of the Diverse Neotropical Palm Tribe Geonomateae. *Frontiers in Plant Science* **2019**, *10*. doi:10.3389/fpls.2019.00864.
47. Baker, W.J.; Savolainen, V.; Asmussen-Lange, C.B.; Chase, M.W.; Dransfield, J.; Forest, F.; Harley, M.M.; Uhl, N.W.; Wilkinson, M. Complete generic-level phylogenetic analyses of palms (Arecaceae) with comparisons of supertree and supermatrix approaches. *Systematic Biology* **2009**, *58*, 240–256. doi:10.1093/sysbio/syp021.
48. Rohland, N.; Reich, D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research* **2012**, *22*, 939–946. doi:10.1101/gr.128124.111.
49. Johnson, M.G.; Gardner, E.M.; Liu, Y.; Medina, R.; Goffinet, B.; Shaw, A.J.; Zerega, N.J.C.; Wickett, N.J. HybPiper: Extracting Coding Sequence and Introns for Phylogenetics from High-Throughput Sequencing Reads Using Target Enrichment. *Applications in Plant Sciences* **2016**, *4*, 1600016. doi:10.3732/apps.1600016.
50. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **2013**, *30*, 772–780. doi:10.1093/molbev/mst010.
51. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **2000**, *17*, 540–552. doi:10.1093/oxfordjournals.molbev.a026334.
52. Sayyari, E.; Mirarab, S. Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Molecular biology and evolution* **2016**, *33*, 1654–1668, [1601.07019]. doi:10.1093/molbev/msw079.
53. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313, [arXiv:10.1093/bioinformatics/btu033]. doi:10.1093/bioinformatics/btu033.
54. Junier, T.; Zdobnov, E.M. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* **2010**, *26*, 1669–1670. doi:10.1093/bioinformatics/btq243.
55. Rabiee, M.; Mirarab, S. SODA: Multi-locus species delimitation using quartet frequencies. *bioRxiv* **2019**, p. 869396. doi:10.1101/869396.
56. Yang, Z. The BPP program for species tree estimation and species delimitation. *Current Zoology* **2015**, *61*, 854–865. doi:10.1093/czoolo/61.5.854.
57. Brown, J.W.; Walker, J.F.; Smith, S.A. Phyx: Phylogenetic tools for unix. *Bioinformatics* **2017**, *33*, 1886–1888. doi:10.1093/bioinformatics/btx063.

589 58. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic
590 algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **2015**,
591 *32*, 268–274. doi:10.1093/molbev/msu300.
592 59. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.; Von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast model
593 selection for accurate phylogenetic estimates. *Nature Methods* **2017**, *14*, 587–589. doi:10.1038/nmeth.4285.
594 60. Lanfear, R.; Calcott, B.; Kainer, D.; Mayer, C.; Stamatakis, A. Selecting optimal partitioning schemes for
595 phylogenomic datasets. *BMC Evolutionary Biology* **2014**, *14*, 82. doi:10.1186/1471-2148-14-82.
596 61. Hoang, D.T.; Chernomor, O.; Von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast
597 bootstrap approximation. *Molecular Biology and Evolution* **2018**, *35*, 518–522. doi:10.1093/molbev/msx281.
598 62. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms
599 and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0.
600 *Systematic Biology* **2010**, *59*, 307–321. doi:10.1093/sysbio/syq010.
601 63. Andermann, T.; Cano, Á.; Zizka, A.; Bacon, C.; Antonelli, A. SECAPR—a bioinformatics pipeline for the
602 rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments.
603 *PeerJ* **2018**, *6*, e5175. doi:10.7717/peerj.5175.
604 64. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*
605 **2009**, *25*, 1754–1760, [1303.3997]. doi:10.1093/bioinformatics/btp324.
606 65. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.;
607 Altshuler, D.; Gabriel, S.; Daly, M.; DePristo, M.A. The Genome Analysis Toolkit: a MapReduce
608 framework for analyzing next-generation DNA sequencing data. *Genome research* **2010**, *20*, 1297–303.
609 doi:10.1101/gr.107524.110.
610 66. Li, H.; Barrett, J. A statistical framework for SNP calling, mutation discovery, association mapping and
611 population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993.
612 doi:10.1093/bioinformatics/btr509.
613 67. Jombart, T.; Devillard, S.; Balloux, F. Discriminant analysis of principal components: A new method for the
614 analysis of genetically structured populations. *BMC Genetics* **2010**, *11*, 94. doi:10.1186/1471-2156-11-94.
615 68. Jombart, T. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**,
616 *24*, 1403–1405. doi:10.1093/bioinformatics/btn129.