Phylogenomics of the Major Tropical Plant Family Annonaceae Using Targeted Enrichment of Nuclear Genes

Thomas L. P. Couvreur[1*†], Andrew J. Helmstetter[1†], Erik J. M. Koenen[2], Kevin Bethune[1], Rita D. Brandão[3], Stefan A. Little[4], Hervé Sauquet[4,5] and Roy H. J. Erkens[3]

[1] IRD, UMR DIADE, Univ. Montpellier, Montpellier, France, [2] Institute of Systematic Botany, University of Zurich, Zurich, Switzerland, [3] Maastricht Science Programme, Maastricht University, Maastricht, Netherlands, [4] Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université-Paris Saclay, Orsay, France, [5] National Herbarium of New South Wales (NSW), Royal Botanic Gardens and Domain Trust, Sydney, NSW, Australia

Targeted enrichment and sequencing of hundreds of nuclear loci for phylogenetic reconstruction is becoming an important tool for plant systematics and evolution. Annonaceae is a major pantropical plant family with 110 genera and ca. 2,450 species, occurring across all major and minor tropical forests of the world. Baits were designed by sequencing the transcriptomes of five species from two of the largest Annonaceae subfamilies. Orthologous loci were identified. The resulting baiting kit was used to reconstruct phylogenetic relationships at two different levels using concatenated and gene tree approaches: a family wide Annonaceae analysis sampling 65 genera and a species level analysis of tribe Piptostigmateae sampling 29 species with multiple individuals per species. DNA extraction was undertaken mainly on silicagel dried leaves, with two samples from herbarium dried leaves. Our kit targets 469 exons (364,653 bp of sequence data), successfully capturing sequences from across Annonaceae. Silicagel dried and herbarium DNA worked equally well. We present for the first time a nuclear gene-based phylogenetic tree at the generic level based on 317 supercontigs. Results mainly confirm previous chloroplast based studies. However, several new relationships are found and discussed. We show significant differences in branch lengths between the two large subfamilies Annonoideae and Malmeoideae. A new tribe, Annickieae, is erected containing a single African genus *Annickia*. We also reconstructed a well-resolved species-level phylogenetic tree of the Piptostigmteae tribe. Our baiting kit is useful for reconstructing well-supported phylogenetic relationships within Annonaceae at different taxonomic levels. The nuclear genome is mainly concordant with plastome information with a few exceptions. Moreover, we find that substitution rate heterogeneity between the two subfamilies is also found within the nuclear compartment, and not just plastomes and ribosomal DNA as previously shown. Our results have implications for understanding the biogeography, molecular dating and evolution of Annonaceae.

Keywords: Annonaceae, rain forests, systematics, transcriptomes, Piptostigmateae, herbarium

1. INTRODUCTION

Targeted enrichment followed by high throughput sequencing of hundreds or even thousands of loci for phylogenetic reconstruction is becoming a golden standard in plant evolutionary biology (Cronn et al., 2012; Xi et al., 2013; Mariac et al., 2014; Barrett et al., 2016). Targeted capture is a genome reduction approach whereby selected regions of the genome are "captured" or hybridized in solution using site-specific baits (also called probes). Once the baits hybridize to the targeted regions, the rest of the genome is discarded, and only the regions of interest are sequenced (Grover et al., 2012) increasing the read depth of the regions of interest. This is in contrast to earlier approaches where the focus lay on recovering complete genomes (e.g., Staats et al., 2013). Multiplexing approaches based on individual barcodes enable fast and cost effective sequencing of multilocus sequence data for phylogenetic or phylogeographic inference (McCormack et al., 2013). Targeted capture of sequence data, either nuclear or plastid, has been used to reconstruct phylogenetic relationships at several taxonomical scales from within angiosperms (Stull et al., 2013; Johnson et al., 2018) to within families (Mandel et al., 2014; Barrett et al., 2016), or to increase phylogenetic resolution in species-rich clades (Nicholls et al., 2015) or at infra-specific levels (Zellmer et al., 2012; Faye et al., 2016). The difficult step concerns the identification of the targeted regions and the design of the baits. This relies on genome-wide information such as transcriptomes or full genomes to identify low copy orthologous genes useful for phylogenetic inference. These data are not always available for non-model groups or clades, especially for tropical lineages. However, increased availability of transcriptomes across angiosperms (Wickett et al., 2014) and lowered sequencing costs allow access to more genomic resources useful in non-model groups.

Annonaceae is a pantropical family of trees, shrubs and lianas and is among the most species-rich lineages of tropical rain forest plants (Chatrou et al., 2012a). To date Annonaceae contain 110 genera and around 2,430 species (Chatrou et al., 2012b; Guo et al., 2017b; Chaowasku et al., 2018; Xue et al., 2018). Phylogenetic analyses of the family started in the 1990's with a morphology-based phylogenetic tree published by Doyle and Le Thomas (1994) providing a first cladistic understanding of relationships within the family. In the molecular era, chloroplast sequence data have been the main markers for Annonaceae family and species level phylogenetic reconstructions (Doyle et al., 2000; Mols et al., 2004; Richardson et al., 2004; Pirie et al., 2006; Su et al., 2008; Couvreur et al., 2011a; Chatrou et al., 2012b; Guo et al., 2017a,b; Thomas et al., 2017). Most of these studies, however, used a relatively small number of markers (but see Hoekstra et al. (2017); Guo et al. (2018); Lopes et al. (2018) for notable exceptions). For instance, the latest classification update of Annonaceae is based on eight chloroplast markers (Chatrou et al., 2012b). Although this led to an overall well-supported generic phylogenetic tree, several relationships remained weakly supported. In addition, species level phylogenies of different Annonaceae genera are, in general, moderately supported based on plastid markers alone (e.g., Pirie et al., 2006; Erkens et al., 2007;

Couvreur et al., 2011b; Chaowasku et al., 2014; Thomas et al., 2017).

Annonaceae is subdivided into four subfamilies and 15 tribes (Chatrou et al., 2012b; Guo et al., 2017b). Subfamilies Annonoideae and Malmeoideae contain over 90% of all species in the family (1,515 and 783, respectively; Guo et al., 2017b) when compared to the other two smaller subfamilies Anaxagoreoideae and Ambavioideae. Based on plastid phylogenetic analyses, Annonoideae and Malmeoideae showed contrasting overall branch lengths, initially leading to them to be known as "long branch" and "short branch" clades, respectively (Richardson et al., 2004). Recently, it has been shown that these phylogenetic branch differences are linked to different substitution rates within the chloroplast and ribosome (Hoekstra et al., 2017).

To date, no nuclear phylogenetic Annonaceae tree has been generated, mainly because of difficulties in amplifying nuclear genes using standard Sanger approaches (Pirie et al., 2005). In this study, we aimed to generate for the first time a nuclear phylogenetic tree of Annonaceae at both generic and species level using target sequence capture and next generation sequencing. We designed and tested a baiting kit useful at several taxonomic levels with the ultimate aim of reconstructing the Annonaceae Tree of Life, including all species in a well-resolved phylogenetic tree based on a wide coverage of the nuclear genome. Specifically, we wanted to answer several questions: Can we generate a baiting kit useful for Annonaceae wide phylogenetic analyses? Are plastid and nuclear phylogenetic relationships concordant in Annonaceae? Do we detect differences in branch lengths based on nuclear data between Annonoideae and Malmeoideae?

2. MATERIALS AND METHODS

2.1. Nuclear Bait Design

2.1.1. Sampling
Nuclear baits were designed based on the analysis of transcriptomes sequenced from five species of Annonaceae, two from Malmeoideae and three from Annonoideae (**Table 1**). For four species, leaf material was collected at the Utrecht Botanical Garden (The Netherlands) and stored in RNAlater storage solution (Sigma-Aldrich, St. Louis, USA). Flowers of *Marsypopetalum littorale* (DIV030) were collected with forceps at the time of anthesis on July 11 2013 at the Botanic Gardens, Vienna (Austria). Flowers were placed in Falcon tubes and immediately submerged in liquid nitrogen, then stored at $-80°$ C.

2.1.2. RNA Extraction
Each leaf piece was homogenized using liquid nitrogen. RNA isolation was conducted using a CTAB approach and a two-step chloroform:isoamylalcohol (24:1) extraction. RNA was finally eluted in RNase-free water and treated with DNase I (Qiagen, Hilden, Germany). RNA extraction from floral material was based on 100 mg of material and followed modified protocols (Kolosova et al., 2004; Dal Cin et al., 2005). RNA pellets were suspended in RNAse free water.

TABLE 1 | Species used to sequence transcriptomes and design the nuclear bating kit.

Subfamily	Tribe	Species	Accession number	Country	Locality	Material
Malmeoideae	Miliuseae	*Marsypopetalum littorale*	ANN120494 / ANN 128	Austria	Botanischer Garten der Universität Wien	Flower
Annonoideae	Bocageae	*Mkilua fragrans*	91GR01649	The Netherlands	Utrecht Botanical Garden	Leaf
Annonoideae	Monodoreae	*Monodora myristica*	E84GR00389	The Netherlands	Utrecht Botanical Garden	Leaf
Annonoideae	Uvarieae	*Monanthotaxis whytei*	84GR00388	The Netherlands	Utrecht Botanical Garden	Leaf
Malmeoideae	Miliuseae	*Sapranthus microcarpus*	1998GR00992	The Netherlands	Utrecht Botanical Garden	Leaf

2.1.3. RNA Quantification, Library Preparation, and Sequencing

RNA was quantified in a Qubit 2.0 instrument (Invitrogen, Carlsbad CA, USA) using the Quant-iT RNA assay kit (Invitrogen, Carlsbad CA, USA) according to the manufacturer's protocol. A total of 4 ug RNA was used for library preparation using the TruSeq® Stranded mRNA Library prep kit (Illumina, San Diego CA, USA) following the manufacturer's protocol for low throughput samples (Part 15031047 Rev. E). An alternate fragmentation protocol was followed to obtain insert sizes of 200 bp as described in the appendix of the manufacturer's protocol. The obtained libraries were quantified using a Qubit 2.0 instrument (Invitrogen, Carlsbad CA, USA) and Qubit dsDNA HS assay kit (Invitrogen, Carlsbad CA, USA). Samples were pooled to a final concentration of 4 nM and sequenced on an Illumina HiSeq 2000 sequencer (Illumina, San Diego CA, USA) using 2 × 100 bp chemistry.

2.1.4. Bait Design

Reads generated from the transcriptome of each species were assembled *de novo* using Trinity (v2.2.0) (Grabherr et al., 2011) using default values. Using the resulting contigs, orthologous loci were selected following these steps. Candidate Open Reading Frames (ORF) were extracted and translated using TransDecoder (https://github.com/TransDecoder) resulting in a number of full protein sequences (start and stop codons present). These were then clustered together based on protein similarity using CD-HIT (v4.6.6) (Li and Godzik, 2006; Fu et al., 2012), to reduce redundancy due to allelic diversity or alternatively spliced isoforms. This step was repeated four times, with percentage identity cutoffs of 0.99, 0.97, 0.95, and 0.9, leading to four sets of protein sequences with different levels of reduced redundancy. We then blasted the resulting sets of proteins for each species against itself ("SelfBLAST") and discarded proteins with multiple hits in the same species, to reduce the prevalence of gene families and keep only low-copy genes. In order to identify genes shared between 3, 4, and 5 species, we applied a 5-way reciprocal best hit (RBH) search using BLASTp (protein to protein BLAST, default values, word size = 3), implemented with a custom python script, and repeated for each CD-HIT percentage identity cutoff. This step was undertaken twice: once with *Monodora myristica* Dunal (Annonoideae) and once with *Marsypopetalum littorale* (Blume) B.Xue & R.M.K.Saunders (Malmeoideae) as focal taxa against which other taxa were compared. This allowed us to identify sets of proteins common within each subfamily and amongst Annonaceae genera. Shared proteins were then aligned using

MAFFT (v7.305) (Katoh and Standley, 2013) with the G-INS-i method and resulting alignments were trimmed using BMGE (v1.12) (Criscuolo and Gribaldo, 2010) to remove sites with more than 50% gaps. Finally, phylogenetic relationships between the five species for each protein alignment were reconstructed using RAxML (v8.2.9) (Stamatakis, 2014). We only selected proteins for which phylogenetic relationships were congruent with known Annonaceae relationships between these five taxa. Following the phylogenetic tree of Annonaceae (Chatrou et al., 2012b; Guo et al., 2017b) we expect *[(Mkilua,(Monodora, Monanthotaxis)),(Marsypopetalum, Sapranthus)]*. Indeed, our sampling of transcriptomes comes from phylogenically distant Annonaceae species in two major subfamilies. We assume that these relationships based on plastid markers will not be different at the nuclear level.

We further filtered these identified low copy orthologous sequences by selecting genes whose protein sequences are (i) more divergent between species than the percentage identity cutoff used in CD-HIT step, (ii) had <5% length variation among species and (iii) were longer than 300 bp. Then, we excluded homologous sequences where the percentage identity between different pairs was different from what we expect from known evolutionary relationships of those taxa (see above for relationships). Genes with a percentage identity of two phylogenetically closest species >0.75 and two furthest species >0.7 were selected. These threshold values allowed selection of a reasonable number of variable genes. For sequences longer than 1,500 bp we were more stringent (identity of closest species >80). Finally, we excluded genes with a high variability of identity percentage (variance >130).

One particularity is that we kept redundant exon sequences between species when possible. Thus if a same exon region was present in four different species, baits were generated for all four "variants" of that exon. This was done to increase capture efficiency across the whole family for that exon. This is similar in concept to a degenerated primer pair for example.

Final baiting sets were synthesized at Arbor Biosciences (MYbaits http://www.arborbiosci.com/) to 120 bp long, with 3X tiling and with a 40 bp spacing between the start of neighboring baits.

2.2. Phylogenetic Analyses of Annonaceae
2.2.1. Taxon Sampling

In order to validate our nuclear-baiting kit for family-wide Annonaceae phylogenetics, we undertook a two-level analysis. First, we aimed to reconstruct, for the first time, a generic

level nuclear phylogenetic tree of Annonaceae (referred to as "Annonaceae" analyses). We sampled 65 genera (out of 110) from all four subfamilies (see **Supplementary Data Sheet**) representing a total of 11 tribes out of the 15 currently recognized tribes (Chatrou et al., 2012b; Guo et al., 2017b). From one to 12 species were sampled per tribe. All genera were represented by one species, except for the pantropical genus *Xylopia* (two species included: one from South America, one from Africa) and the paraphyletic genus *Friesodielsia* (Guo et al., 2017a) for which we sampled one African species and one South-East Asian species.

Second, to assess its usefulness for reconstructing species-level relationships, we undertook a species-level sampling of the African tribe Piptostigmateae (referred to as "Piptostigmateae" analyses). This tribe comprises seven genera (Ghogue et al., 2017) most of them recently revised (Versteegh and Sosef, 2007; Couvreur et al., 2009, 2015; Marshall et al., 2016; Ghogue et al., 2017). We sampled a total of 29 out of 39 species covering all seven genera (see **Supplementary Data Sheet**). One to five individuals were sampled per species, plus a number of outgroup taxa, leading to a total of 83 individuals included in our analyses. All DNA was extracted from silicagel dried leaves, except for two samples, where DNA was extracted from air dried herbarium samples collected in 2000.

2.2.2. DNA Extraction, Library Preparation, In-solution Hybridization, and Sequencing

DNA was extracted from silica gel dried leaves following the MATAB and chloroform separation methods of Mariac et al. (2014) and Scarcelli et al. (2006). The full protocol is provided in **Supplementary Data Sheet Protocol 1**. Illumina libraries were constructed following a modified protocol of Rohland and Reich (2012) using 6-bp barcodes and Illumina indexes to allow for multiplexing at different levels. Extra steps were added to the Rohland and Reich (2012) protocol to allow for amplification and in-solution hybridization as follows: total DNA for each individual was sheared using a Bioruptor Pico (Diagenode, Liége, Belgium) to a mean target size of 500 bp. DNA was then repaired, ligated and nick filled-in before an 8–11 cycle prehybridization PCR was performed. After clean-up and quantification, libraries were bulked, mixed with biotin-labeled baits and hybridized to the targeted regions using the bait kit designed above. The hybridized biotin-labeled baits were then immobilized using streptavidin-coated magnetic beads. A magnetic field was applied and supernatant containing unbounded DNA was discarded. Enriched DNA fragments were then eluted from the beads and amplified in a 14–16 cycle real-time PCR to complete adapters and generate final libraries. Libraries were sequenced on an Illumina HiSeq v3 platform pair end and length of 150 bp (Illumina, SAn Diego CA, USA) at CIRAD facilities (Montpellier, France) with around 18 pmol of the capture-amplified DNA libraries deposited on the flowcell. The full step-by-step protocol is provided in **Supplementary Data Sheet Protocol 2**.

2.2.3. Bioinformatics

Demultiplexing with a 0-mistmatch threshold was undertaken using the demultadapt script (https://github.com/Maillol/demultadapt). Adapters were removed using cutadapt (v1.2.1)

(Martin, 2011) with the default parameters. Reads were filtered according to their length (>35 bp) and quality mean values (Q > 30) using a custom script (https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl). Forward and reverse sequences were paired according to their name in the fastq files using a comparison script, adapted from TOGGLe (Tranchant-Dubreuil et al., 2018). A terminal trimming of 6 bp was performed on reverse sequences to ensure removal of barcodes in case of sequences shorter than 150 bp using the fastx trimmer script which is part of the fastx toolkit (https://github.com/agordon/fastx_toolkit). A final sorting by sequence identifier was done on the stq file before further analyses.

2.2.4. Contig Assembly and Multi-Sequence Alignment

We used the pipeline HybPiper (v1.2) (Johnson et al., 2016) under the default settings to process our cleaned data. Briefly, reads were mapped to targets using BWA (v0.7.12) (Li and Durbin, 2009) and those reads that were successfully mapped were assembled into contigs using SPAdes (v3.11.1) (Bankevich et al., 2012). Exonerate (Slater and Birney, 2005) was then used to align the assembled contigs to their associated target sequence. Additionally, if contigs were slightly overlapping (Johnson et al., 2016) they were combined into "supercontigs" which contained both target and off-target sequence data. Exonerate was run a second time so that introns could be more accurately identified. We aligned each set of supercontigs using MAFFT (v7.305) (Katoh and Standley, 2013) with the "–auto" option and cleaned these alignments with GBLOCKS (v0.91b) (Castresana, 2000) using the default parameters and all allowed gap positions ("-b5=a"). The number of parsimony informative sites were calculated for each supercontig alignment using the "pis" function in the ape R package (Paradis et al., 2004).

2.2.5. Paralog Identification

HybPiper flags potential paralogs when multiple contigs are discovered mapping well to a single reference sequence. The program uses coverage and identity to a reference to choose a "main" sequence and denotes the others as potential paralogs. We took flagged loci and constructed gene trees using RAxML (v8.2.9) (Stamatakis, 2014). We examined each tree to determine whether putative paralogs formed a species clade, in which case these were unlikely true paralogs and we could continue with the main sequence selected by HybPiper. If the "main" and alternative sequences formed separate clades they were likely true paralogs so we removed the entire locus from the downstream analyses.

2.2.6. Phylogenetic Inference

We inferred trees using a generic-level dataset of Annonaceae and a species-level dataset of the tribe Piptostigmateae (see **Supplementary Data Sheet**). For each dataset we identified those exons that had 75% of their length reconstructed in 75% of individuals. 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 (Sayyari et al., 2017). We chose loci that were found in 75% of taxa to limit

the amount of missing data but still recover loci that were not shared among distantly related clades. We then took the corresponding supercontigs (i.e., targeted exon regions plus surrounding off-target captured sequence data) for phylogenetic inference. We initially tried tree inference using exons only but support was generally lower than trees based on supercontigs and the topology was similar so we decided to proceed using supercontigs (results not shown). All trees were reconstructed unrooted (as is mandatory in ASTRAL, see below) and then rooted post-inference. The Annonaceae tree was rooted on *Anaxagorea crassipetala* based on previous phylogenetic trees inferred with chloroplast DNA (Chatrou et al., 2012b; Guo et al., 2017b) showing it is the earliest diverging genus of the family. The Piptostigmateae tree was rooted using the outgroup taxon *Annona glabra* from the Annonoideae sub family (Chatrou et al., 2012b).

2.2.6.1. Coalescent approach

Individual gene trees were constructed using RAxML (v8.2.9) (Stamatakis, 2014) with the "-f a" option and 100 bootstrap replicates. We used a GTRGAMMA model which estimates the rates in the general time-reversible (GTR) matrix, the alpha shape parameter of the Gamma model of rate heterogeneity and the empirical base frequencies for the locus. Branches with boostrap support >10 were collapsed using Newick Utilities program nw_ed (Junier and Zdobnov, 2010). This has been shown to improve the accuracy of inferred species trees (Zhang et al., 2017). We then ran ASTRAL-III (v5.5.11) (Zhang et al., 2017) using the gene trees that corresponded to our selected loci, first using the default parameters, then using "-t 1" to output quartet support values.

2.2.6.2. Concatenation approach

After alignment, we filled missing individuals at each locus with an empty sequence and concatenated aligned loci using the pxcat function in the program phyx (Brown et al., 2017). A GTRGAMMA model was defined using the partition file output by pxcat leading to a different model for each alignment (Brown et al., 2017). A rapid bootstrap analysis in RAxML (v8.2.9) (Stamatakis et al., 2008; Stamatakis, 2014) using the "-f a" option with 100 replicates was performed followed by a thorough ML search on the original alignment.

3. RESULTS

3.1. Baiting Kit

The trinity analyses recovered a total of 297,193 raw contigs for *Marsypopetalum littorale*, 342,592 for *Sapranthus microcarpus*, 102,025 for *M. myristica*, 194,497 for *Monanthotaxis whytei* and 164 635 *Mkilua fragrans*. After the orthologous low copy selection process we ended up with 80,317, 68,039, 68,039, 22,326, 33,794 contigs, respectively. Targeted exon regions varied from 300 to 6,072 bp. A total of 813 exonic regions were finally selected after our filtering process. However, a number of these regions were present several times as different variants of the same exon: 236 were present twice, 31 were present three times, and 15 present four times. Thus, the baiting kit effectively

targets 469 unique exonic regions with a total capture footprint of 364,653 bp (available as **Supplementary Table 1**). The final MyBaits Annonaceae baiting kit was made of 11,583 baits 120 bp long. This baiting kit is available upon request at Arbor Biosciences (http://www.arborbiosci.com). The fastq (R1 and R2) sequences for all individuals are available in Genbank SRA under Bioproject number PRJNA508895 (http://www.ncbi.nlm.nih.gov/bioproject/508895).

3.2. Number of Loci Retrieved and Variability

Overall, we recovered good average depth for all individuals sequenced (275; min: 45; max: 726) and on average we recovered over 90% of our target with a coverage greater than 10x (**Supplementary Data Sheet**).

For the Annonaceae analyses, we recovered 98 to 468 loci depending on the percentage of reconstructed targeted exon length for the same percentage of individuals (**Table 2**). We recovered at least a fraction of all but one of the targeted loci (468 of 469). A total of 331 loci were reconstructed for over 75% of their length and for over 75% of individuals and this "75/75"

TABLE 2 | Variability of number of captured of loci using the Annonaceae bait kit.

	> 0%	≥ 25%	≥ 50%	≥ 75%	95%
Annonaceae	468	466	449	331	98
Piptostigmateae	469	466	453	379	170

The values indicate the percentage of individuals for which we reconstructed the same percentage of the targeted exon region. Phylogenetic analyses were carried out using the 75/75 datasets.
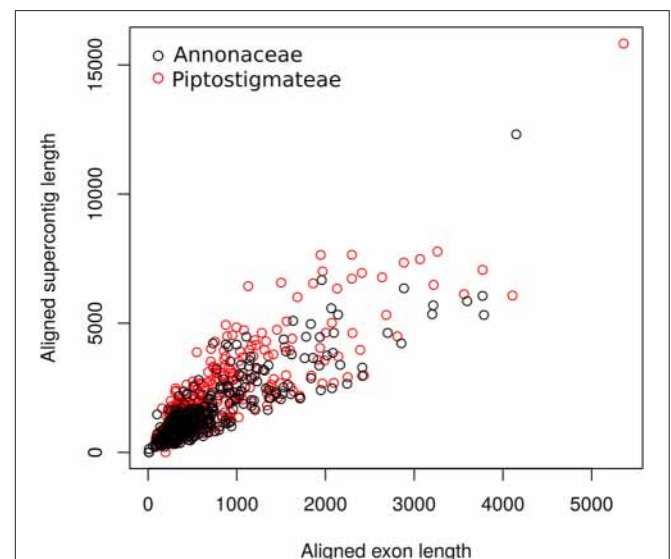

FIGURE 1 | Scatterplot of targeted exon length against supercontig length (targeted exon plus off target data) at each locus. Lengths were calculated post-alignment. Black open circles: Annonaceae alignments; Red open circles: Piptostigmateae alignments.

subset was used for tree inference. The total length of these 331 loci was 545 610 bp.

For the Piptostigmateae analyses, we captured between 170 and 469 loci (**Table 2**). We captured at least some of all the targeted exons in this dataset. A total of 379 loci were retrieved under the 75/75 rule and used for the phylogenetic inference. The total length of these 379 loci was 766,373 bp.

We recovered substantial amounts of non-target sequence data per locus. The typical aligned length of supercontigs was more than twice that of exon alignments (**Figure 1**). On average we recovered 828 bp of off-target sequence data in our Annonaceae dataset and 1 241 bp in our Piptostigmateae dataset to be used for downstream phylogenetic inference.

We calculated 75/75 loci for each Annonaceae subfamily and identified 179 loci common across each subfamily locus set. We found loci that were shared between all but three subfamily combinations, and loci that were specific to each subfamily except Ambavoideae (**Figure 2A**). The set of 75/75 loci used in the Piptostigmateae tree inference contained almost all of those used in the Annonaceae tree inference as well as 51 additional loci (**Figure 2B**).

At the Annonaceae level, post alignment length of the 331 supercontigs varied from 407 to 12,314 bp long with a mean of 1,648 bp (**Table 3**). For the Piptostigmateae matrix, the supercontigs varied from 492 to 15,824 bp long with a mean of 2,022 bp (**Table 3**). In both cases, the longer the supercontigs the more Parsimony Informative Sites were retrieved (**Figure 3C**).
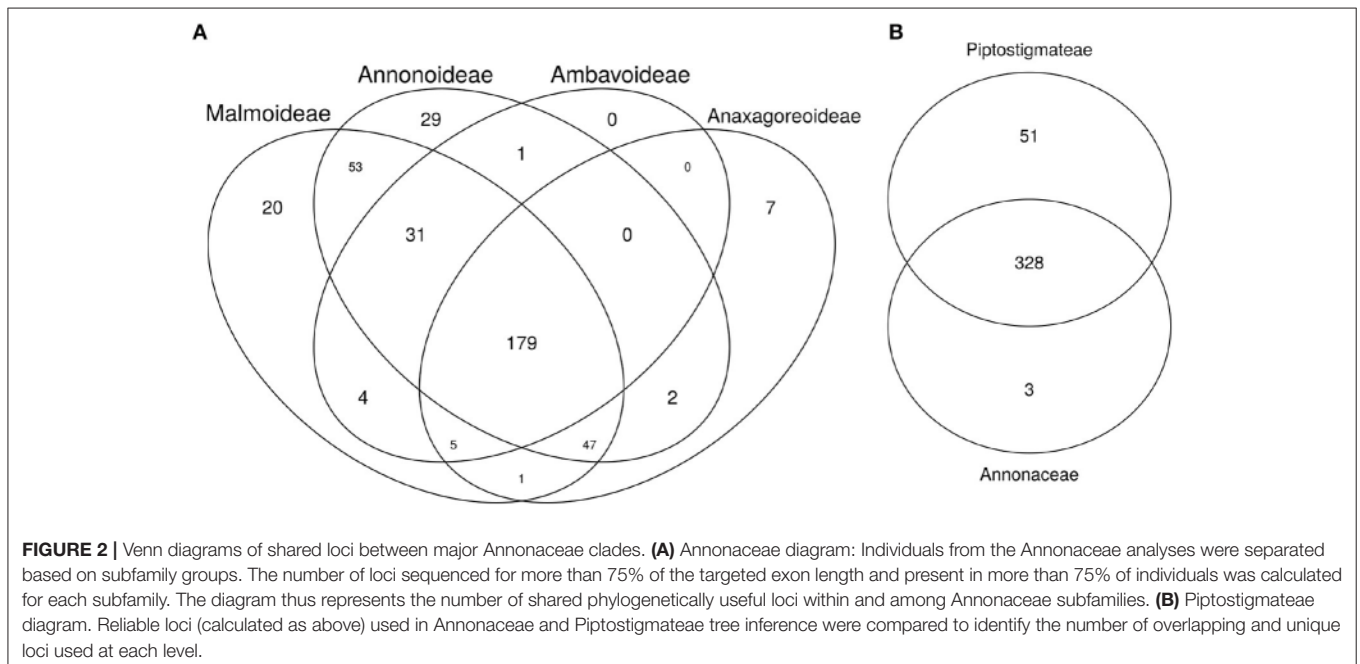
3.3. Family-Level Analyses

We visually inspected loci flagged as putative paralogs by hybpiper to verify whether there was a clear phylogenetic distinction between main and alternative paralogs. This lead to 14 loci where paralogy seemed probable and these were removed from our 75/75 locus subset. After removal, the post alignment
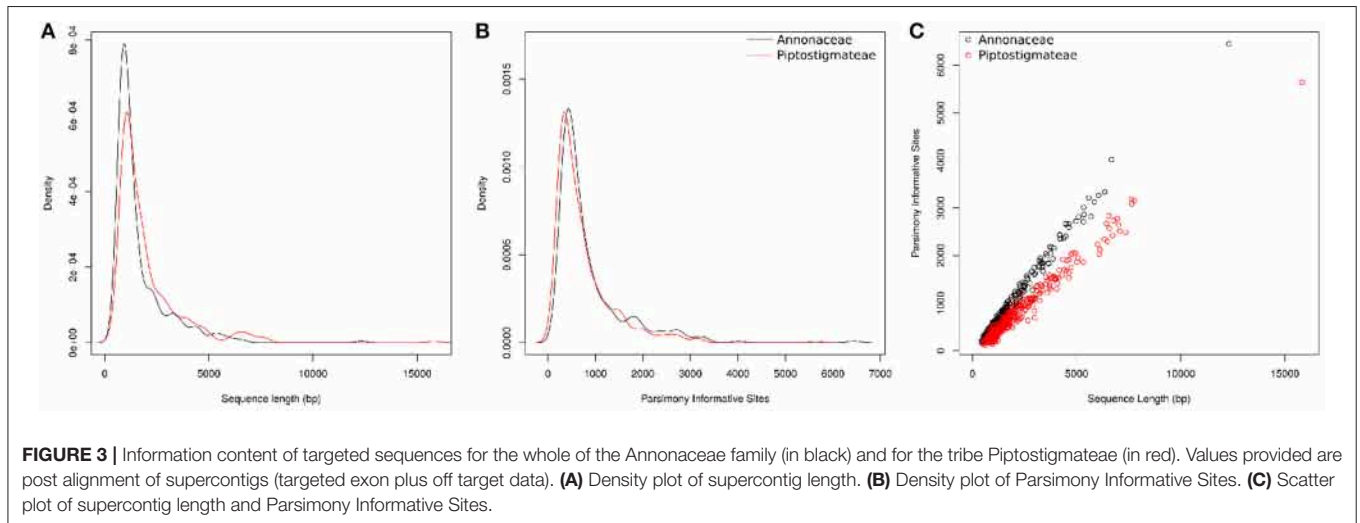
concatenated length of the remaining 317 loci was 533,322 bp. Mean supercontig length was 1,648 bp long, with on average 861 parsimony informative sites and the longest supercontig was 12,314 bp long (**Table 3**, **Figure 3A**).

Support was generally high throughout the Annonaceae ASTRAL tree, with approximately 86% of branches possessing local posterior probabilities (LPP) of more than 90% (**Figure 4A**). An estimated 95% of input quartet trees were in the species tree indicating a generally low level of gene tree conflict in the family. However, assessing quartet support at nodes (**Figure 4B**) revealed gene tree conflict was high in the Miliuseae. The RAxML tree (**Figure 5**) was inferred with very high overall levels of bootstrap support—88% of branches had 100 bootstrap support. The topologies produced by the ASTRAL and RAxML approaches were very similar. RAxML consistently conferred higher support in instances of gene tree conflict such as the placement of Miliuseae genera. One of the most conspicuous differences between approaches is the placement of the genus *Annickia*. This genus is inferred to be sister to the Malmeoideae in the ASTRAL tree while it is sister to the rest of the Piptostigmateae in the

TABLE 3 | Length and parsimony informative sites statistics based on the aligned Annonaceae and Piptostigmateae matrices.

		Supercontig	PIS
Annonaceae	Mean	1 648.4	861.7
	SD	1 325.7	746.6
	Min	407	169
	Max	12 314	6 450
Pipotstigmateae	Mean	2 022.1	743.4
	SD	1 621.7	637.9
	Min	492	122
	Max	15,824	5,639

FIGURE 2 | Venn diagrams of shared loci between major Annonaceae clades. **(A)** Annonaceae diagram: Individuals from the Annonaceae analyses were separated based on subfamily groups. The number of loci sequenced for more than 75% of the targeted exon length and present in more than 75% of individuals was calculated for each subfamily. The diagram thus represents the number of shared phylogenetically useful loci within and among Annonaceae subfamilies. **(B)** Piptostigmateae diagram. Reliable loci (calculated as above) used in Annonaceae and Piptostigmateae tree inference were compared to identify the number of overlapping and unique loci used at each level.

FIGURE 3 | Information content of targeted sequences for the whole of the Annonaceae family (in black) and for the tribe Piptostigmateae (in red). Values provided are post alignment of supercontigs (targeted exon plus off target data). **(A)** Density plot of supercontig length. **(B)** Density plot of Parsimony Informative Sites. **(C)** Scatter plot of supercontig length and Parsimony Informative Sites.

RAxML tree, albeit with low support. In addition, several genera are recovered in different positions than previously inferred using plastid markers, such as *Artabotrys* and *Sanrafaelia*.

3.4. Piptostigmateae Analyses

After visual verification we identified 24 probable paralogs, 23 of which were found in our 75/75 set of loci and removed from downstream analyses. The remaining 356 loci contained a total of 743,325 bp of sequence data. Mean supercontig length was 2,022 bp long, with on average 743 parsimony informative sites and the longest supercontig was 15,824 bp long (**Table 3**, **Figure 3B**).

We found similar high levels of node support in our Piptostigmateae trees when compared to the Annonaceae analyses, despite the reduced evolutionary distance between taxa (**Figure 6A**). 85% of branches had LPP values >0.9. Gene tree conflict was again low (91% of gene tree quartets were represented in the species tree) but slightly higher than in the Annonaceae tree. This conflict was principally found within species (**Figure 6B**), for example *Piptostigma goslineanum* formed a well-supported clade but branches within possessed low quartet support and LPP. A small number of nodes found deeper in the genera *Greenwayodendron* and *Piptostigmata* were also poorly supported (**Figure 6A**). As in the Annonaceae analysis, we found increased support in all areas of the Piptostigmateae RAxML tree (**Figure 7**) and the topologies generated by the two inference methods were similar. Likewise, RAxML tended to give higher levels of support in those areas of the tree where ASTRAL estimated low LPP (**Figure 7**). Like in the Annonaceae tree the placement of *Annickia* changed between approaches either sister to (RAxML) or diverging before (ASTRAL) the rest of the Piptostigmateae split from the rest of the Malmeoideae genera sampled.

4. DISCUSSION

4.1. An Annonaceae-Wide Nuclear Baiting Kit

Generating "baiting kits" for targeted enrichment of hundreds of nuclear loci using Next Generation Sequencing is becoming

a major component of plant phylogenetic studies (Cronn et al., 2012; Zellmer et al., 2012; Nicholls et al., 2015; Heyduk et al., 2016; Kadlec et al., 2017; Johnson et al., 2018). Here, we present a baiting kit that can be used to undertake phylogenetic studies across the large pantropical Annonaceae family at different taxonomic levels (e.g., **Figures 5**, **7**). The Annonaceae kit enables potential capture of 469 unique exonic regions (**Table 2**). Although the kit was designed using comparative transcriptomic analyses of five species from two of the most diverse Annonaceae subfamilies (Annonoideae and Malmeoideae), we successfully captured shared loci within the two other smaller subfamilies (Ambavioideae and Anaxagoreoideae). Under our 75%/75% rule (see methods), 179 loci were common across all major subfamilies (**Figure 2A**). However, few loci were recovered unique to Anaxagoroideae (7) and none for Ambavoideae (**Figure 2A**). These results indicate that this baiting kit has a wide taxonomic breadth of utility, and will be useful to reconstruct phylogenetic relationships across the family. We show that our baiting kit is effective at recovering a similar number of loci when we focus on a specific tribe with a species level sampling. For the Piptostigmateae tribe, and before removal of potential paralogs, we recovered 379 loci under the 75/75 rule (**Table 2**). Thus, at a lower taxonomic level we recovered slightly more loci than at the family level (331 loci at 75/75). In addition, 87% of loci recovered for Piptostigmateae were common to the family wide Annonaceae analysis (**Figure 2B**) highlighting that the baiting kit can be used to simultaneously capture ingroup and outgroup taxa when targeting specific Annonaceae clades. This cross-family flexibility comes from the way we designed our baits. First, we used a top-down approach: baits were designed and optimized based on five species sampled from across the family. We thus identified orthologous exons that were common across our sample. Such approaches were already successful within other tropical plant families such as palms (Heyduk et al., 2016), where the baiting kit has been successfully used to sequence species in different clades of the family (Comer et al., 2016). In contrast, phylogenetic relationships within the large Neotropical genus *Inga* (Nicholls et al., 2015) or the mega diverse genus *Erica* (Kadlec et al., 2017) were reconstructed using baits designed from transcriptomes of species mostly within or closely
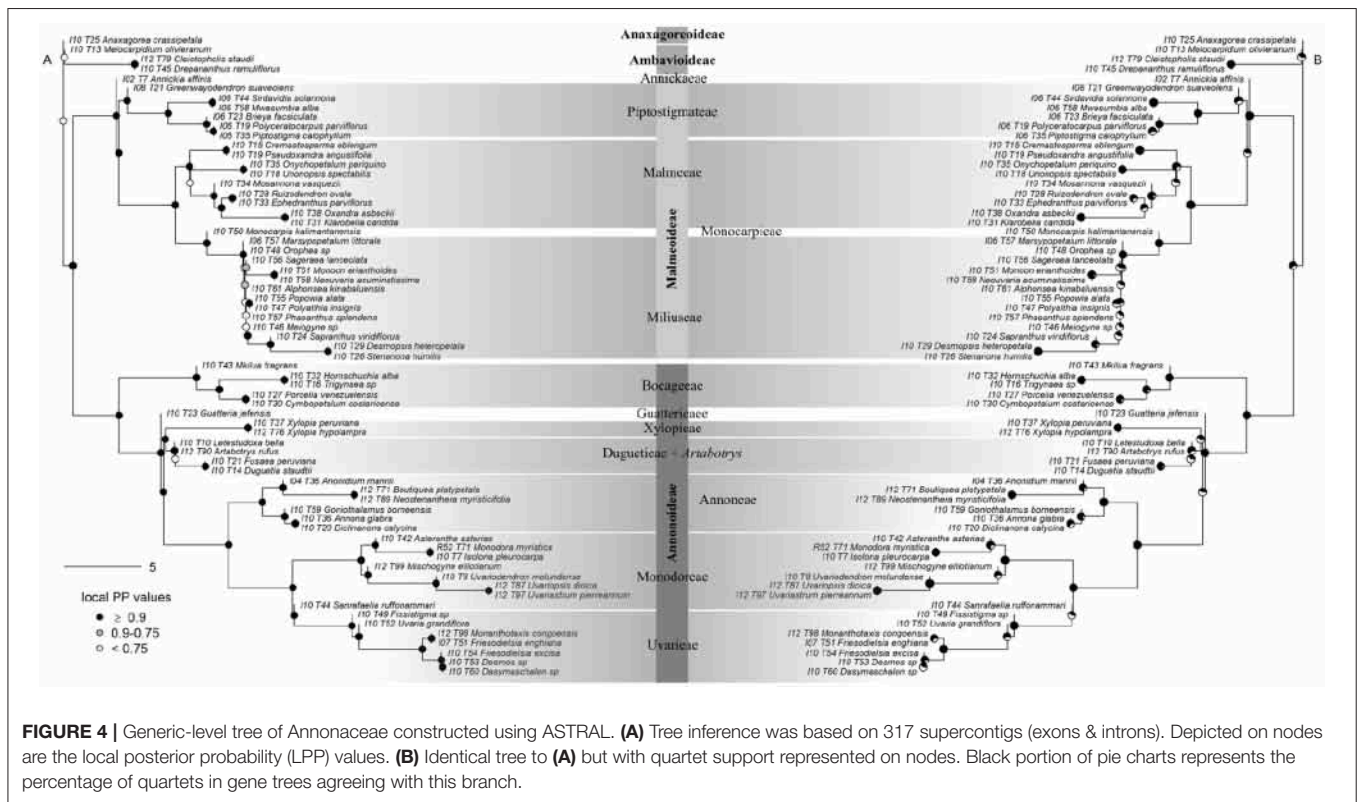
FIGURE 4 | Generic-level tree of Annonaceae constructed using ASTRAL. **(A)** Tree inference was based on 317 supercontigs (exons & introns). Depicted on nodes are the local posterior probability (LPP) values. **(B)** Identical tree to **(A)** but with quartet support represented on nodes. Black portion of pie charts represents the percentage of quartets in gene trees agreeing with this branch.

related to target genera or study clade. In the latter studies, the baiting kit was shown to successfully cpature exons across the Ericaceae family too (Kadlec et al., 2017). Second, we also used a "degenerated" kit, meaning that for 282 (out of 469) identified orthologous exons we included different species-specific exon variants increasing capture capacity across the family. However, we did not undertake explicit analyses or resequencing to test if our degenerate kit improved capture across Annonaceae or not.

We recovered around 500–1,000 bp of off-target sequence data (i.e., sequence data flanking our targeted regions). Off-target sequences are mostly composed of intronic data, but it has been shown that non-baited exon regions can also be captured in the process (Kadlec et al., 2017). Annonaceae do not have a sequenced genome yet, so it is hard to know precisely what we are recovering in off-target sequences. Recovery of off-target sequence data is an important property that allows the same baiting kit to be informative at several taxonomic levels, which is clearly shown by this and other studies (Kadlec et al., 2017). In a preliminary study we also used the same baiting kit to successfully capture 354 loci at infraspecific levels within *Greenwayodendron suaveolens* (Engl. & Diels) Verdc. providing good phylogeographic resolution (unpublished).

Finally, for two samples, DNA was extracted directly from herbarium specimens (INDEX6_TAG49 *Greenwayodendron usambaricum* (Verdc.) Lissambou, Hardy & Couvreur and INDEX6_TAG50 *Annickia kummerae* (Engl. & Diels) Setten & Maas) leading to good overall coverage and sequencing depth (**Supplementary Data Sheet**). This confirms the utility of targeted enrichment approaches to sequence hundreds of nuclear
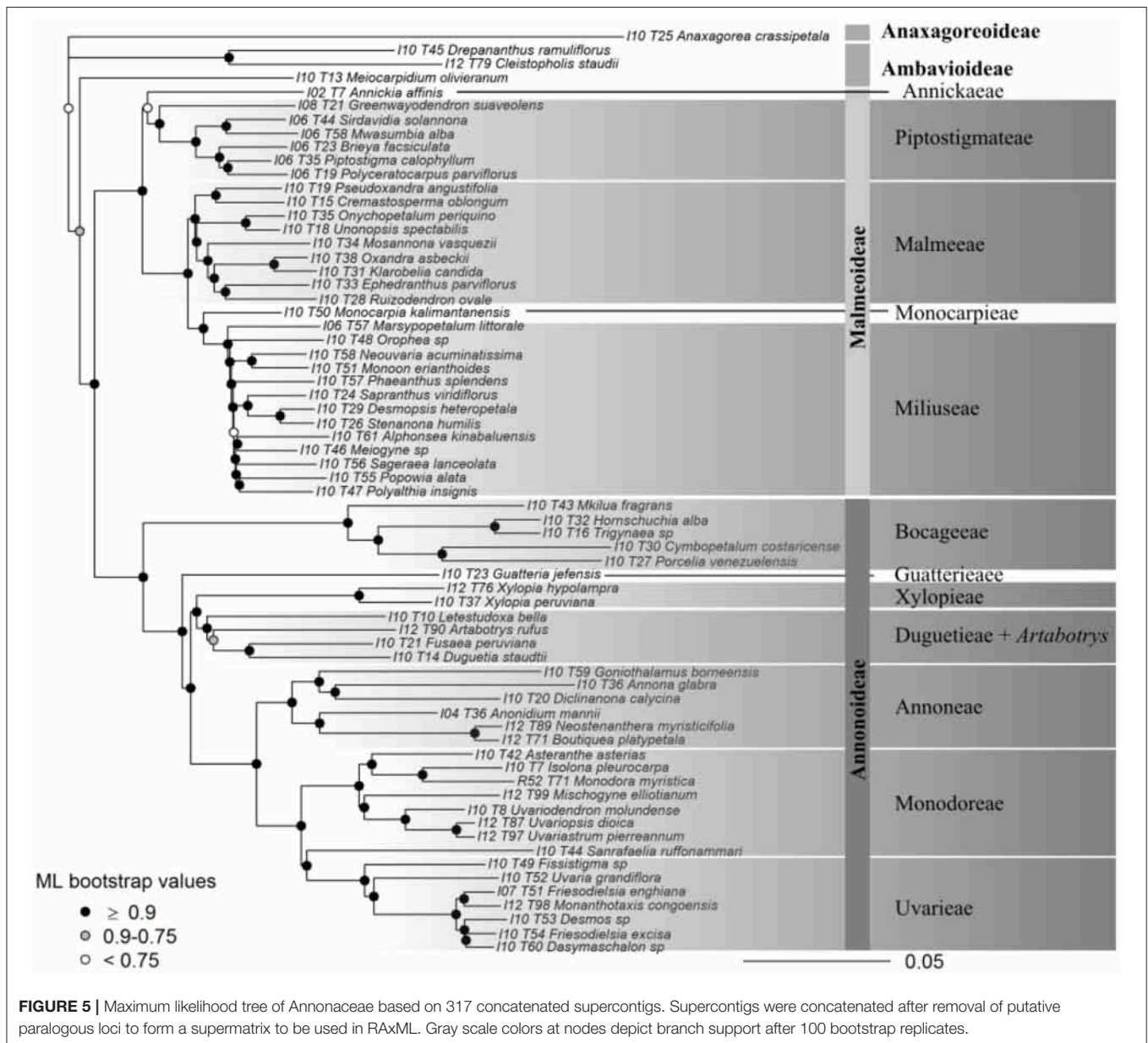
genes form herbarium preserved material (Hart et al., 2016). DNA extracted from Annonaceae herbarium specimens have been successfully and routinely sequenced in the past even from 100+ year-old specimens (Erkens et al., 2008). Thus our results are encouraging and suggest that our kit will be useful to tap into the "genomic treasure trove" (Staats et al., 2013) of Annonaceae herbarium specimens.

4.2. Substitution Rates

Our Maximum Likelihood concatenated phylogenetic tree (**Figure 5**) based only on nuclear loci shows a clear difference in branch lengths between the two large Annonaceae subfamilies Annonoideae and Malmeoideae. Although no formal tests were undertaken here, our results suggest that this increased substitution rate within Annonoideae when compared to Malmeoideae affects the entire genome as already suggested by Hoekstra et al. (2017). This genome-wide among-lineage rate heterogeneity can have important impacts when estimating divergence times (e.g., Bellot and Renner, 2014; Hoekstra et al., 2017; Testo et al., 2018). This will need to be taken into account in further biogeographical analyses of the family (Hoekstra et al., 2017), for example using Random Local Clock approaches (Drummond and Suchard, 2010; Bellot and Renner, 2014; Testo et al., 2018).

4.3. Next Generation Annonaceae Phylogenetics

We present the first phylogenomic analyses of Annonaceae based on 317 (excluding paralogs) nuclear loci sampled for 65

FIGURE 5 | Maximum likelihood tree of Annonaceae based on 317 concatenated supercontigs. Supercontigs were concatenated after removal of putative paralogous loci to form a supermatrix to be used in RAxML. Gray scale colors at nodes depict branch support after 100 bootstrap replicates.

out of 110 (60%) currently recognized genera (**Figures 4**, **5**). Overall our results mostly agree with previous Annonaceae wide phylogenetic analyses solely based on plastid regions (Richardson et al., 2004; Su et al., 2008; Couvreur et al., 2011a; Chatrou et al., 2012b; Guo et al., 2017a,b). We rooted the Annonaceae tree using *Anaxagorea crassipetala* based on the results of these studies. In the future it would be useful to use a outgroup Magnoliales species to confirm this relationship and resolve the position of the early-diverging *Meiocarpidium*. However, the age of the oldest extant Annonaceae clades (Couvreur et al., 2011a; Pirie and Doyle, 2012) means that incongruence between plastid and nuclear markers is unlikely. Deep level relationships are congruent between the concatenated and gene tree approaches and all four sub families are recovered with maximum support (**Figures 4**, **5**). The clade most affected by our methods is that

of the genera of the Miliuseae tribe. Indeed, some of these intergeneric relationships are well supported in the concatenated approach, but not in the coalescent approach. Our results, however, do provide some new insights into the phylogeny of Annonaceae, which we describe here.

4.3.1. Ambavioideae
4.3.1.1. Meiocarpidium
The position of the monotypic genus *Meiocarpidium* is moderately supported as sister to Annonaceae excluding *Anaxagorea* and Ambavoideae in the concatenated maximum likelihood tree (**Figure 5**), and weakly supported as sister to the rest of the Ambavoideae in the gene tree (**Figure 4**). About 1/3 of loci support the quartet relationship of *Meiocarpidium* as sister to the rest of the Ambavioideae (**Figure 4B**) leading however
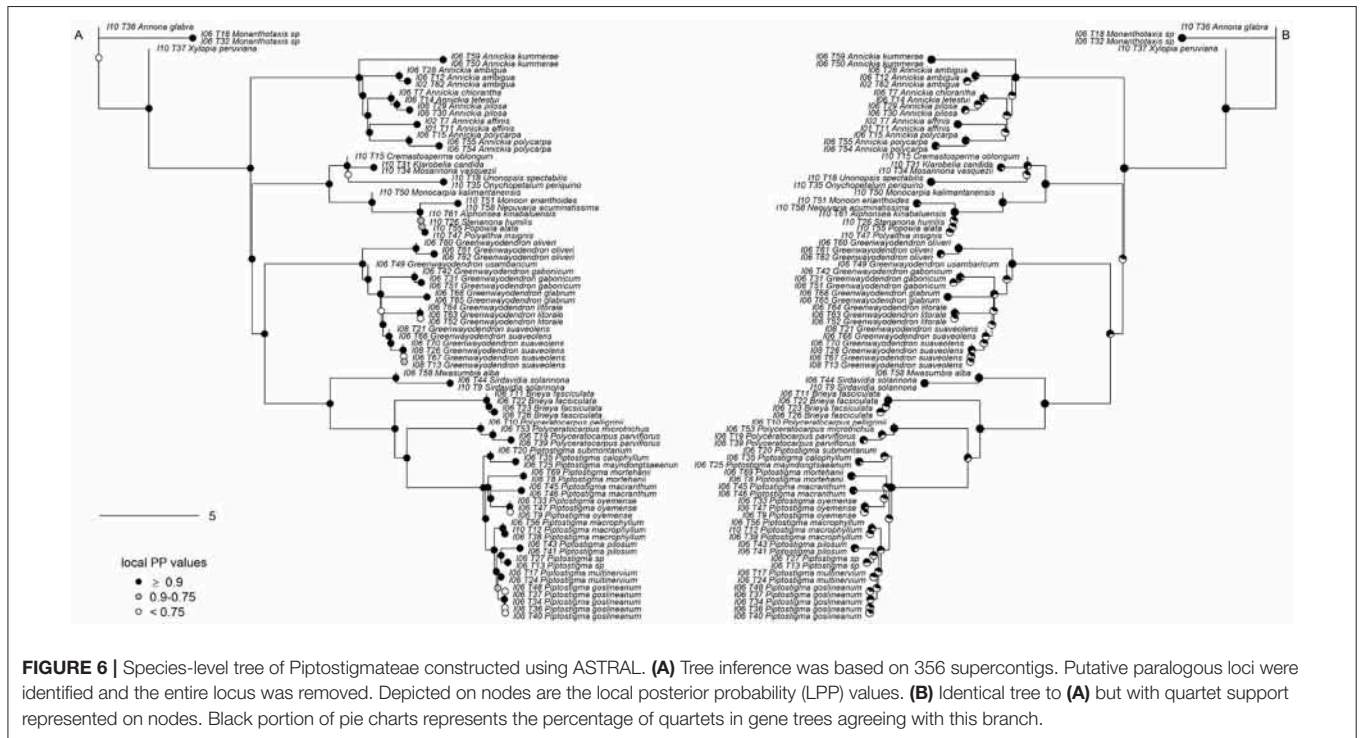
FIGURE 6 | Species-level tree of Piptostigmateae constructed using ASTRAL. **(A)** Tree inference was based on 356 supercontigs. Putative paralogous loci were identified and the entire locus was removed. Depicted on nodes are the local posterior probability (LPP) values. **(B)** Identical tree to **(A)** but with quartet support represented on nodes. Black portion of pie charts represents the percentage of quartets in gene trees agreeing with this branch.

to low LPP support. *Meiocarpidium* was already suggested to be early diverging within Annonaceae especially based on floral and pollen characters (Le Thomas, 1980, 1981) and its placement was problematic in past phylogenetic analyses. Indeed, this genus was recovered as sister to the rest of Ambavioideae based on eight plastid loci (Chatrou et al., 2012b; Guo et al., 2017a) with moderate bootstrap support. Here, despite substantial increase in sequence data the position of *Meiocarpidium* remains uncertain.

4.3.1.2. Guatteria

Our phylogenomic analysis sheds light on one of the major, unresolved phylogenetic questions within the Annonoideae: the position of the tribe Guatterieae (Erkens et al., 2009). This tribe consists solely of the species-rich Neotropical genus *Guaterria* (Maas et al., 2015). So far family-wide chloroplast-based analyses were not able to determine the position of this tribe with certainty (Erkens et al., 2009; Couvreur et al., 2011a; Chatrou et al., 2012b; Guo et al., 2017b). Here, we show with high support that Guatterieae is sister to all Annonoideae except the tribe Bocageeae (**Figures 4**, **5**). This result, however, is only based on a single *Guaterria* species, and thus needs to be confirmed based on a larger sample of this genus.

4.3.1.3. Artabotrys

The unexpected position of the paleotropical genus *Artabotrys* renders the tribe Xylopieae polyphyletic. The relationship between tribes Xylopieae and Duguetieae are strongly supported in both our gene tree and concatenated analyses (**Figures 4**, **5**). However, *Artabotrys* is always recovered as nested with high support within the Duguetieae clade (sampled genera *Letestudoxa*, *Fusaea*, *Duguetia*). These results contrast with what was inferred based on plastid data where *Artabotrys*

was systematically recovered as sister with strong support to the large pantropical genus *Xylopia* (Richardson et al., 2004; Chatrou et al., 2012a; Thomas et al., 2015; Guo et al., 2017b). This relationship, however, was not recovered by cladistic analyses of morphological data (Doyle and Le Thomas, 1994) or *rbcL* sequence alone (Doyle et al., 2000). The tribe Xylopieae was erected following the phylogenetic analyses of eight plastid markers, but it was stressed that clear morphological synapomorphies between the two constituents genera were lacking (Chatrou et al., 2012b). Indeed, *Artabotrys* is a genus of lianas that has evolved special hook-shaped terminal inflorescences whereas *Xylopia* contains large tree species with non-hook shaped axillary inflorescences. It is beyond the scope of this paper to detail differences and similarities, but our results suggest that the phylogenetic placement of *Artabotrys* should be more thoroughly investigated, and the taxonomic implications reviewed. Increased sampling of both genera (5 out of 7 genera sampled here) and species could potentially help resolve this question.

4.3.1.4. Sanrafaelia

Our results suggest that the Tanzanian endemic genus *Sanrafaelia* is excluded from tribe Monodoreae, and is in fact sister with strong to the mainly climbing tribe Uvarieae (**Figures 4**, **5**). Based on plastid markers *Sanrafaelia* together with its East African sister genus *Ophrypetalum* were recovered with moderate support as sister to a clade of African genera (Couvreur et al., 2008; Chatrou et al., 2012b; Guo et al., 2017b). Even though one synapomorphy was identified to support the erection of the Monodoreae tribe (sessile monocarps; Couvreur et al., 2008; Chatrou et al., 2012b) it appears that this tribe might have to be
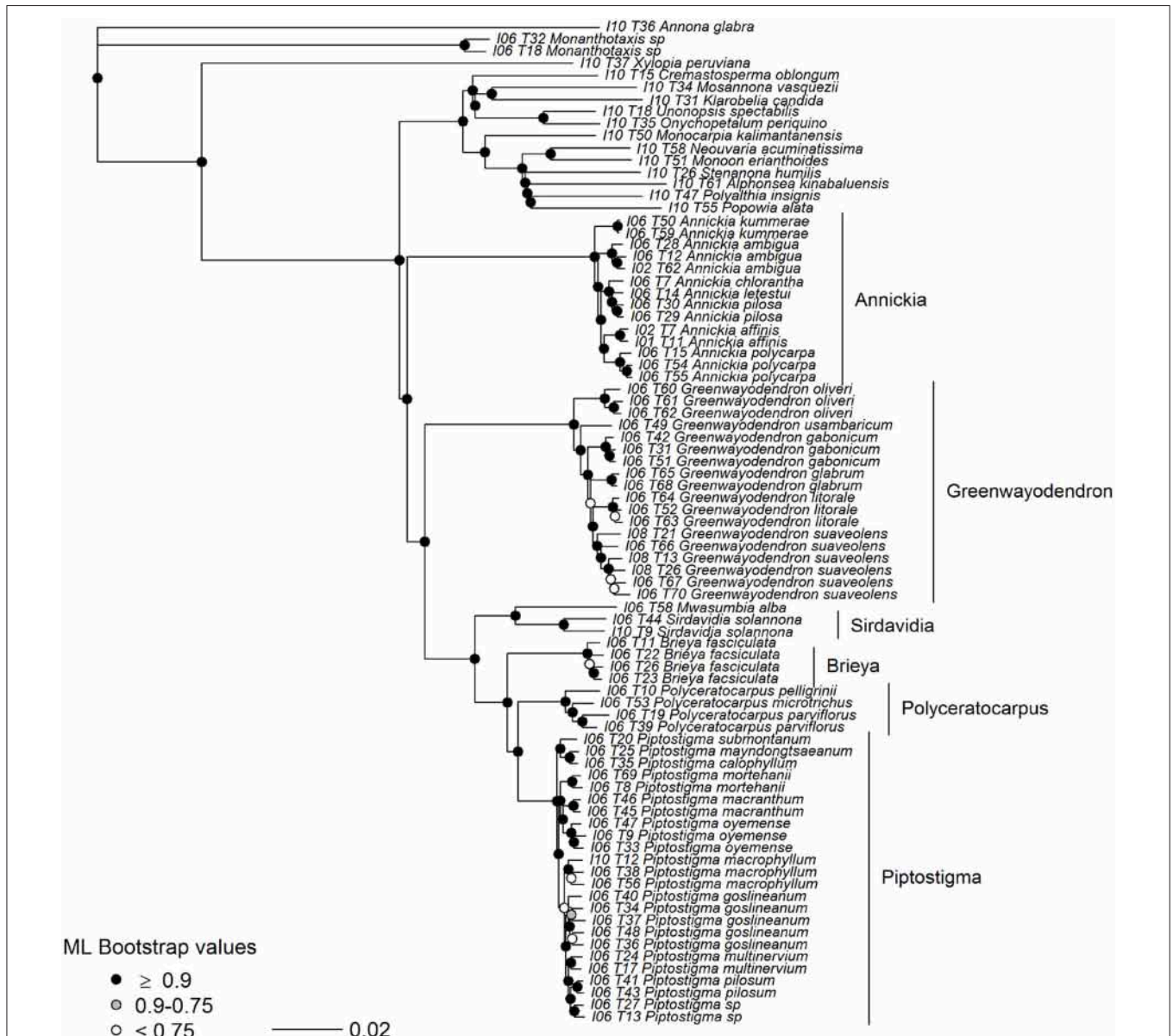
FIGURE 7 | Maximum likelihood tree of Piptostigmateae based on 356 concatenated supercontigs. Supercontigs were concatenated after removal of individuals with putative paralogous sequences from alignments to form a supermatrix to be used in RAxML. Node colors show support based on 100 bootstrap replicates.

revisited in the light of nuclear data, and especially the inclusion of the genus *Ophrypetalum*.

4.3.2. Malmeoideae
4.3.2.1. Annickia
Our data suggest that the African tribe Piptostigmateae is potentially not monophyletic. Indeed, *Annickia* is resolved as sister with high support to the rest of the Malmeoideae in the gene tree analysis (**Figure 4**), but weakly supported as sister to the rest of the Piptostigmateae genera in the concatenated analysis (**Figure 5**). These results were also confirmed when we increased sampling in the tribe (**Figures 6**, **7**) highlighting discordance between the gene tree and the concatenated analyses. In the

gene tree analysis, not all loci recovered the *Annickia* sister to Malmeoideae relationship (41% in the Annonaceae analysis and 64% in the Piptostigmateae analysis). This clearly underlines some conflict in the proper placement of the genus. Nevertheless, this still translated into high LPP support for this branch because the alternative quartet topologies are recovered by significantly fewer loci. Substantial gene tree conflict cannot be accounted for by the concatenation approach (Liu et al., 2015) and this could be the reason why we recovered different relationships in the RAxML trees (**Figures 5**, **7**). Discordance in phylogenetic relationships between concatenated and gene tree coalescent approaches based on phylogenomic data have been reported (Xi et al., 2014). We thus suggest that the gene tree analysis using

ASTRAL provides a better inference of true relationships than the concatenated approach (Liu et al., 2015).

The placement of *Annickia* has been problematic in previous analyses as well. It has either been weakly inferred as sister to the rest of Malmeoideae (Doyle et al., 2000) or as sister to the rest of the Piptostigmateae genera based on several plastid loci (Pirie et al., 2006; Chatrou et al., 2012b; Guo et al., 2017b). Morphological and pollen character-based cladistic analyses in contrast underlined affinities of *Annickia* with other members of the Malmeoideae such as *Unonopsis* (Doyle and Le Thomas, 1994). *Annickia* was included in the tribe Piptostigmateae (Chatrou et al., 2012b) because it was resolved (but not supported) as sister to that tribe and probably because *Annickia* is, as all other Piptostigmateae genera, strictly African. *Annickia* presents several morphological differences with other Piptostigmateae genera such as in pollen (columelar vs. granular infratectum) (Doyle and Le Thomas, 2012) and floral characters (one vs. several ovules, more than 20 carpels vs. less than 20) (Doyle and Le Thomas, 1994; Versteegh and Sosef, 2007; Couvreur et al., 2009, 2015). Based on the present phylogenomic analyses taking into account gene trees, both at generic and species level sampling, and morphology we suggest that the genus *Annickia* does not belong in the Piptostigmateae, and propose the description of a new tribe, Annickieae.

Annickieae Couvreur, tribus nov.—TYPE GENUS: *Annickia* (van Setten and Maas, 1990). Trees, indumentum with single, bifid, trifid, fasciculate or stellate hairs; flowers bisexual, solitary, terminal; perianth actinomorphic, 6 free tepals in two opposite whorls of 3, differentiated in an outer whorl of 3 sepals and an inner whorl of 3 valvate petals. Stamens 110–175, filaments short, anthers extrorse; carpels 35–70, with a single ovule per carpel; monocarps free, stipitate. Comprising a single genus, *Annickia*, with eight species endemic to continental Africa (Versteegh and Sosef, 2007). The sister relationship of Annickieae with the rest of the Malmeoideae based on our gene tree analysis will have implications in understanding the evolution and biogeography of Annonaceae, especially within the Malmeoideae subfamily (e.g., Couvreur et al., 2011a; Doyle and Le Thomas, 2012; Thomas et al., 2017).

4.3.2.2. Miliuseae

Our results show that resolving the backbone of the Miliuseae tribe will be challenging even when using a phylogenomic approach. As when using plastid sequence data, branches leading up to the different sampled genera are short. However, most of the nodes are strongly supported in the concatenated approach (**Figure 5**) but the gene tree analysis shows high levels of discordance between loci (**Figure 4**). Nevertheless, our nuclear phylogenetic tree agrees with previous plastid-based analyses (Mols et al., 2004; Chatrou et al., 2012b; Xue et al., 2012; Chaowasku et al., 2014; Guo et al., 2017b). For example, *Sapranthus*, *Desmopsis* and *Stenanona* form a well-supported clade which is also the case when using plastid markers (Chatrou et al., 2012b; Ortiz-Rodriguez et al., 2016; Guo et al., 2017b). The coalescent analysis highlights that most backbone relationships are supported by just one third of loci with low LPP support (**Figure 4**). Our results support the idea that the Miliuseae

underwent a period of rapid diversification leading to widespread incomplete lineage sorting. However, it must be noted that our sampling of this tribe is far from complete, and more genera and species will need to be added to gain a true understanding of this clade.

4.4. Piptostigmateae Species-Level Phylogenetic Tree

Besides the position of *Annickia* (see above), our phylogenomic analysis of tribe Piptostigmateae (**Figures 6**, **7**) recovers maximally supported relationships between genera, confirming previous plastid phylogenetic trees (Couvreur et al., 2009, 2015; Ghogue et al., 2017; Guo et al., 2017b). Our results confirm the generic status of *Brieya* as a phylogenetically separate genus from *Piptostigma* (Ghogue et al., 2017). The coalescent analysis (**Figure 6**) shows that the monophyly of all genera is supported by a high percentage of loci (>94%).

Relationships between species within genera are overall well-supported in the concatenated analysis and with high LPP in the coalescent approach. Our results support the latest taxonomic revision of the genus *Greenwayodendron* with five phylogenetically different species (Lissambou et al., 2018) instead of the long held view of two species (Le Thomas, 1969). This includes the recognition of the East African taxon *Greenwayodendron usambaricum* (Verdc.) Lissambou, Hardy & Couvreur as a distinct species from *Greenwayodendron suaveolens* (Engl. & Diels) Verdc.. In addition, our results are also concordant with the latest revision of the genus *Piptostigma* (Ghogue et al., 2017). However, we did identify two individuals which appear to belong to a new undescribed species (labeled as sp. in **Figures 6**, **7**). Finally, using our baiting kit, all species were recovered as monophyletic. This suggests that overall, species circumscription in this tribe is well-established.

5. CONCLUSION

We show here that even within a tropical plant family with overall no available genomic resources, it is possible to generate nuclear markers useful for family wide and species level phylogenetic analyses. Thus, for other tropical plant lineages, having access to fresh material of a few non related species covering the diversity of the studied group is potentially enough to generate a phylogenetically useful nuclear baiting kit. In addition, these family specific kits could be used in tandem with the angiosperm universal baiting kit recently developed (Johnson et al., 2018). In the era of phylogenomics, it will be important not only to generate clade specific baits, but also try and include baits used across angiosperms for enabling easier sharing and compilation of data around the plant Tree of Life (Eiserhardt et al., 2018).

The nuclear baiting kit designed here enables the potential capture of up to 469 loci across Annonaceae. We show that more than 300 loci can be captured across the whole family with good coverage. The kit is thus a good tool to generate a well-resolved species-level phylogenetic tree of all Annonaceae species. This in turn, can be used to address

numerous taxonomical or evolutionary questions within the family. It remains however to be tested if we can capture exons from closely related plant families such as Eupomatiaceae or other Magnoliales families, especially for molecular dating. This will be important as the most relevant fossils for dating Annonaceae are placed within other magnoliid families (Massoni et al., 2015a,b).

The phylogenetic relations within the Miliusieae tribe are hard to resolve even when using a large number of nuclear loci. This remains a problem for Annonaceae phylogenetics. Potentially, the addition of full chloroplast sequences or plastomes might provide added resolution. It remains to be seen how much resolution and support we recover when reconstructing the phylogenetic relationships of species-rich genera such as *Guatteria* (Erkens et al., 2007; Maas et al., 2015) or *Goniothalamus* (Tang et al., 2015). We hope that our newly developed baiting kit and inferred phylogenetic trees will enable and encourage studies of the evolutionary history of Annonaceae at a range of taxonomic levels.

AUTHOR CONTRIBUTIONS

TC and AH conceived the study. RE and HS contributed transcriptome data. RB, SL, and KB generated the data. TC, AH, EK, KB, RB, and SL analyzed the data. TC and AH wrote the article. All authors have edited and contributed to the writing and reviewed the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01941/full#supplementary-material

REFERENCES

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Barrett, C. F., Bacon, C. D., Antonelli, A., Cano, A., and Hofmann, T. (2016). An introduction to plant phylogenomics with a focus on palms. *Bot. J. Linnean Soc.* 182, 234–255. doi: 10.1111/boj.12399

Bellot, S., and Renner, S. S. (2014). Exploring new dating approaches for parasites: the worldwide Apodanthaceae (Cucurbitales) as an example. *Mol. Phylogenet. Evol.* 80, 1–10. doi: 10.1016/j.ympev.2014.07.005

Brown, J. W., Walker, J. F., and Smith, S. A. (2017). Phyx: phylogenetic tools for unix. *Bioinformatics* 33, 1886–1888. doi: 10.1093/bioinformatics/btx063

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Chaowasku, T., Damthongdee, A., Jongsook, H., Ngo, D. T., Le, H. T., Tran, D. M., et al. (2018). Enlarging the monotypic Monocarpieae (Annonaceae, Malmeoideae): recognition of a second genus from Vietnam informed by morphology and molecular phylogenetics. *Candollea* 73, 261–275. doi: 10.15553/c2018v732a11

Chaowasku, T., Thomas, D. C., Van der Hammen, R. W., Smets, E. F., Mols, J. B., and Chatrou, L. W. (2014). A plastid DNA phylogeny of tribe Miliuseae: Insights into relationships and character evolution in one of the most recalcitrant major clades of Annonaceae. *Am. J. Bot.* 101, 691–709. doi: 10.3732/ajb.1300403

Chatrou, L. W., Erkens, R. H. J., Richardson, J. E., Saunders, R. M. K., and Fay, M. F. (2012a). The natural history of Annonaceae. *Bot. J. Linnean Soc.* 169, 1–4. doi: 10.1111/j.1095-8339.2012.01242.x

Chatrou, L. W., Pirie, M., Erkens, R. H., Couvreur, T. L. P., Neubig, K. M., Abbott, J., et al. (2012b). A new subfamilial and tribal classification of the pantropical flowering plant family Annonaceae informed by molecular phylogenetics. *Bot. J. Linnean Soc.* 169, 5–40. 00126. doi: 10.1111/j.1095-8339.2012.01235.x

Comer, J. R., Zomlefer, W. B., Barrett, C. F., Stevenson, D. W., Heyduk, K., and Leebens-Mack, J. H. (2016). Nuclear phylogenomics of the palm subfamily Arecoideae (Arecaceae). *Mol. Phylogenet. Evol.* 97, 32–42. doi: 10.1016/j.ympev.2015.12.015

Couvreur, T. L. P., Helmstetter, A. J., Koenen, E. E., Bethume, K., Brandao, R. D., Little, S., et al. (2018). Phylogenomics of the major tropical plant family Annonaceae using targeted enrichment of nuclear genes. *bioRxiv [Preprint].* 440925. doi: 10.1101/440925

Couvreur, T. L. P., Pirie, M., Chatrou, L. W., Saunders, R. M. K., Su, Y., Richardson, J. E., et al. (2011a). Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. *J. Biogeogr.* 38, 664–680. doi: 10.1111/j.1365-2699.2010.02434.x

Couvreur, T. L. P., Porter-Morgan, H., Wieringa, J. J., and Chatrou, L. W. (2011b). Little ecological divergence associated with speciation in two African rain forest tree genera. *BMC Evol. Biol.* 11:296. doi: 10.1186/1471-2148-11-296

Couvreur, T. L. P., Richardson, J. E., Sosef, M. S. M., Erkens, R. H. J., and Chatrou, L. W. (2008). Evolution of syncarpy and other morphological characters in African Annonaceae: a posterior mapping approach. *Mol. Phylogenet. Evol.* 47, 302–318. doi: 10.1016/j.ympev.2008.01.018

Couvreur, T. L. P., Sonke, B., Niangadouma, R., and Sauquet, H. (2015). Sirdavidia, an extraordinary new genus of Annonaceae from Gabon. *PhytoKeys* 46, 1–19. doi: 10.3897/phytokeys.46.8937

Couvreur, T. L. P., Van der Ham, R., Mbele, Y., Mbago, F., and Johnson, D. (2009). Molecular and morphological characterization of a new monotypic genus of Annonaceae, *Mwasumbia*, from Tanzania. *Syst. Bot.* 34, 266–276. doi: 10.1600/036364409788606398

Criscuolo, A., and Gribaldo, S. (2010). BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10:210. doi: 10.1186/1471-2148-10-210

Cronn, R., Knaus, B. J., Liston, A., Maughan, P. J., Parks, M., Syring, J. V., et al. (2012). Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.* 99, 291–311. doi: 10.3732/ajb.1100356

Dal Cin, V., Danesin, M., Rizzini, F. M., and Ramina, A. (2005). RNA extraction from plant tissues: the use of calcium to precipitate contaminating pectic sugars. *Mol. Biotechnol.* 31, 113–119. doi: 10.1385/MB:31:2:113

Doyle, J., and Le Thomas, A. (1994). Cladistic analysis and pollen evolution in Annonaceae. *Acta Bot. Gall.* 141, 149–170. doi: 10.1080/12538078.1994.10515148

Doyle, J. A., Bygrave, P., and Le Thomas, A. (2000). "Implications of molecular data for pollen evolution in Annonaceae," in *Pollen & Spores: Morphology and Biology*, eds M. Harley, C. Morton, and S. Blackmore (Kew: Royal Botanic Gardens), 259–284.

Doyle, J. A., and Le Thomas, A. (2012). Evolution and phylogenetic significance of pollen in Annonaceae. *Bot. J. Linnean Soc.* 169, 190–221. doi: 10.1111/j.1095-8339.2012.01241.x

Drummond, A. J., and Suchard, M. A. (2010). Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 8:114. doi: 10.1186/1741-7007-8-114

Eiserhardt, W. L., Antonelli, A., Bennett, D. J., Botigué, L. R., Burleigh, J. G., Dodsworth, S., et al. (2018). A roadmap for global synthesis of the plant tree of life. *Am. J. Bot.* 105, 614–622. doi: 10.1002/ajb2.1041

Erkens, R. H., Chatrou, L. W., Maas, J. W., van der Niet, T., and Savolainen, V. (2007). A rapid diversification of rainforest trees (Guatteria; Annonaceae) following dispersal from Central into South America. *Mol. Phylogenet. Evol.* 44, 399–411. doi: 10.1016/j.ympev.2007.02.017

Erkens, R. H., Cross, H., Maas, J. W., Hoenselaar, K., and Chatrou, L. W. (2008). Assessment of age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. *Blumea Biodivers. Evol. Biogeogr. Plants* 53, 407–428. doi: 10.3767/000651908X608052

Erkens, R. H. J., Maas, J. W., and Couvreur, T. L. P. (2009). From Africa via Europe to South America: migrational route of a species-rich genus of Neotropical lowland rain forest trees (Guatteria, Annonaceae). *J. Biogeogr.* 36, 2338–2352. doi: 10.1111/j.1365-2699.2009.02162.x

Faye, A., Deblauwe, V., Mariac, C., Richard, D., Sonké, B., Vigouroux, Y., and Couvreur, T. L. P. (2016). Phylogeography of the genus Podococcus (Palmae/Arecaceae) in Central African rain forests: climate stability predicts unique genetic diversity. *Mol. Phylogenet. Evolut.* 105, 126–138. doi: 10.1016/j.ympev.2016.08.005

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Ghogue, J.-P., Sonké, B., and Couvreur, T. L. P. (2017). Taxonomic revision of the African genera *Brieya* and *Piptostigma* (Annonaceae). *Plant Ecol. Evol.* 150, 173–216. doi: 10.5091/plecevo.2017.1137

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Grover, C. E., Salmon, A., and Wendel, J. F. (2012). Targeted sequence capture as a powerful tool for evolutionary analysis. *Am. J. Bot.* 99, 312–319. doi: 10.3732/ajb.1100323

Guo, X., Hoekstra, P. H., Tang, C. C., Thomas, D. C., Wieringa, J. J., Chatrou, L. W., et al. (2017a). Cutting up the climbers: Evidence for extensive polyphyly in Friesodielsia (Annonaceae) necessitates generic realignment across the tribe Uvarieae. *Taxon* 66, 3–19. doi: 10.12705/661.1

Guo, X., Tang, C. C., Thomas, D. C., Couvreur, T. L. P., and Saunders, R. M. K. (2017b). A mega-phylogeny of the Annonaceae: taxonomic placement of five enigmatic genera and support for a new tribe, Phoenicantheae. *Sci. Rep.* 7:7323. doi: 10.1038/s41598-017-07252-2

Guo, X., Thomas, D. C., and Saunders, R. M. K. (2018). Gene tree discordance and coalescent methods support ancient intergeneric hybridisation between Dasymaschalon and Friesodielsia (Annonaceae). *Mol. Phylogenet. Evol.* 127, 14–29. doi: 10.1016/j.ympev.2018.04.009

Hart, M. L., Forrest, L. L., Nicholls, J. A., and Kidner, C. A. (2016). Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65, 1081–1092. doi: 10.12705/655.9

Heyduk, K., Trapnell, D. W., Barrett, C. F., and Leebens-Mack, J. (2016). Phylogenomic analyses of species relationships in the genus Sabal (Arecaceae) using targeted sequence capture. *Biol. J. Linnean Soc.* 117, 106–120. doi: 10.1111/bij.12551

Hoekstra, P. H., Wieringa, J. J., Smets, E., Brandão, R. D., Lopes, J. C., Erkens, R. H. J., et al. (2017). Correlated evolutionary rates across genomic compartments in Annonaceae. *Mol. Phylogenet. Evol.* 114, 63–72. doi: 10.1016/j.ympev.2017.05.026

Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., et al. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl. Plant Sci.* 4:1600016. doi: 10.3732/apps.1600016

Johnson, M. G., Pokorny, L., Dodsworth, S., Botigué, L. R., Cowan, R. S., Devault, A., et al. (2018). A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Syst. Biol.* syy086. doi: 10.1093/sysbio/syy086

Junier, T., and Zdobnov, E. M. (2010). The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. *Bioinformatics* 26, 1669–1670. doi: 10.1093/bioinformatics/btq243

Kadlec, M., Bellstedt, D. U., Maitre, N. C. L., and Pirie, M. D. (2017). Targeted NGS for species level phylogenomics: made to measure or one size fits all? *PeerJ* 5:e3569. doi: 10.7717/peerj.3569

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kolosova, N., Miller, B., Ralph, S., Ellis, B. E., Douglas, C., Ritland, K., et al. (2004). Isolation of high-quality RNA from gymnosperm and angiosperm trees. *Biotechniques* 36, 821–824. doi: 10.2144/04365ST06

Le Thomas, A. (1969). "Annonacées," in *Flore du Gabon*, Vol. 16, ed A. Aubréville (Paris: Museum National d'Histoire Naturelle), 1–371.

Le Thomas, A. (1980). Ultrastructural characters of the pollen grains of African Annonaceae and their significance for the phylogeny of primitive angiosperms (first part). *Pollen Spores* 22, 267–342.

Le Thomas, A. (1981). Ultrastructural characters of the pollen grains of African Annonaceae and their significance for the phylogeny of primitive angiosperms (second part). *Pollen Spores* 23, 1–36.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

Lissambou, B.-J., Hardy, O. J., Atteke, C., Stevart, T., Dauby, G., Mbatchi, B., et al. (2018). Taxonomic revision of the African genus Greenwayodendron (Annonaceae). *Phytokeys* 114, 55–93. doi: 10.3897/phytokeys.114.27395

Liu, L., Xi, Z., Wu, S., Davis, C. C., and Edwards, S. V. (2015). Estimating phylogenetic trees from genome-scale data. *Ann. N. Y. Acad. Sci.* 1360, 36–53. doi: 10.1111/nyas.12747

Lopes, J. C., Chatrou, L. W., Mello-Silva, R., Rudall, P. J., and Sajo, M. G. (2018). Phylogenomics and evolution of floral traits in the Neotropical tribe Malmeeae (Annonaceae). *Mol. Phylogenet. Evol.* 118, 379–391. doi: 10.1016/j.ympev.2017.10.020

Maas, P. J. M., Westra, L., Guerrero, S. A., Lobão, A. Q., Scharf, U., Zamora, N., et al. (2015). Confronting a morphological nightmare: revision of the Neotropical genus Guatteria (Annonaceae). *Blumea Biodivers. Evol. Biogeogr. Plants* 60, 1–219. doi: 10.3767/000651915X690341

Mandel, J. R., Dikow, R. B., Funk, V. A., Masalia, R. R., Staton, S. E., Kozik, A., et al. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: an example from the compositae. *Appl. Plant Sci.* 2:1300085. doi: 10.3732/apps.1300085

Mariac, C., Scarcelli, N., Pouzadou, J., Barnaud, A., Billot, C., Faye, A., et al. (2014). Cost effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Mol. Ecol. Resour.* 14, 1103–1113. doi: 10.1111/1755-0998.12258

Marshall, A., Couvreur, T. L. P., Summers, A., Deere, N., Luke, W., Ndangalasi, H. J., et al. (2016). A new species in the tree genus Polyceratocarpus (Annonaceae) from the Udzungwa Mountains of Tanzania. *PhytoKeys* 63, 63–76. doi: 10.3897/phytokeys.63.6262

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Massoni, J., Couvreur, T. L., and Sauquet, H. (2015a). Five major shifts of diversification through the long evolutionary history of Magnoliidae (angiosperms). *BMC Evol. Biol.* 15:49. doi: 10.1186/s12862-015-0320-6

Massoni, J., Doyle, J., and Sauquet, H. (2015b). Fossil calibration of Magnoliidae, an ancient lineage of angiosperms. *Palaeontol. Electron.* 18, 1–25. doi: 10.26879/435

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., and Brumfield, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol. Phylogenet. Evol.* 66, 526–538. doi: 10.1016/j.ympev.2011.12.007

Mols, J. B., Gravendeel, B., Chatrou, L. W., Pirie, M., Bygrave, P., Chase, M., and Kessler, P. (2004). Identifying clades in Asian Annonaceae: monophyletic genera in the polyphyletic Miliuseae. *Am. J. Bot.* 91, 590–600. doi: 10.3732/ajb.91.4.590

Nicholls, J. A., Pennington, R. T., Koenen, E. J., Hughes, C. E., Hearn, J., Bunnefeld, L., et al. (2015). Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus Inga (Leguminosae: Mimosoideae). *Front. Plant Sci.* 6:710. doi: 10.3389/fpls.2015.00710

Ortiz-Rodriguez, A. E., Ruiz-Sanchez, E., and Ornelas, J. F. (2016). Phylogenetic relationships among members of the Neotropical clade of Miliuseae (Annonaceae): generic non-monophyly of Desmopsis and Stenanona. *Syst. Bot.* 41, 815–822. doi: 10.1600/036364416X693928

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412

Pirie, M. D., Chatrou, L. W., Erkens, R. H. J., Maas, J. W., Van der Niet, T., Mols, J., et al. (2005). "Phylogeny reconstruction and molecular dating in four Neotropical genera of Annonaceae: the effect of taxon sampling in age estimation," in *Plant Species-Level Systematics: New Perspectives on Pattern and Process*, eds F. T. Bakker, L. W. Chatrou, B. Gravendeel, and P. Pelser (Ruggell: Gantner Verlag), 149–174.

Pirie, M. D., Chatrou, L. W., Mols, J. B., Erkens, R. H. J., and Oosterhof, J. (2006). 'Andean-centred' genera in the short-branch clade of Annonaceae: testing biogeographical hypotheses using phylogeny reconstruction and molecular dating. *J. Biogeogr.* 33, 31–46. doi: 10.1111/j.1365-2699.2005. 01388.x

Pirie, M. D., and Doyle, J. A. (2012). Dating clades with fossils and molecules: the case of Annonaceae. *Bot. J. Linnean Soc.* 169, 84–116. doi: 10.1111/j.1095-8339.2012.01234.x

Richardson, J. E., Chatrou, L. W., Mols, J. B., Erkens, R. H., and Pirie, M. D. (2004). Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philos. Trans. R. Soc. Lond. B Ser.* 359, 1495–1508. doi: 10.1098/rstb.2004.1537

Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22, 939–946. doi: 10.1101/gr.128124.111

Sayyari, E., Whitfield, J. B., and Mirarab, S. (2017). Fragmentary gene sequences negatively impact gene tree and species tree reconstruction. *Mol. Biol. Evol.* 34, 3279–3291. doi: 10.1093/molbev/msx261

Scarcelli, N., Tostain, S., Vigouroux, Y., Agbangla, C., Daïnou, O., and Pham, J. L. (2006). Farmers' use of wild relative and sexual reproduction in a vegetatively propagated crop. The case of yam in Benin. *Mol. Ecol.* 15, 2421–2431. doi: 10.1111/j.1365-294X.2006.02958.x

Slater, G. S., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinform.* 6:31. doi: 10.1186/1471-2105-6-31

Staats, M., Erkens, R. H., Vossenberg, B. v. d., Wieringa, J. J., Kraaijeveld, K., Stielow, B., et al. (2013). Genomic treasure troves: complete genome sequencing of herbarium and Insect Museum Specimens. *PLOS ONE*, 8(7):e69189. doi: 10.1371/journal.pone.0069189

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web-servers. *Syst. Biol.* 57, 758–771. doi: 10.1080/10635150802429642

Stull, G. W., Moore, M. J., Mandala, V. S., Douglas, N. A., Kates, H. R., Qi, X., et al. (2013). A targeted enrichment strategy for massively parallel sequencing of angiosperm plastid genomes. *Appl. Plant Sci.* 1:1200497. doi: 10.3732/apps.1200497

Su, Y., Smith, G., and Saunders, R. M. (2008). Phylogeny of the basal angiosperm genus Pseuduvaria (Annonaceae) inferred from five chloroplast DNA regions, with interpretation of morphological character evolution. *Mol. Phylogenet. Evol.* 48, 188–206. doi: 10.1016/j.ympev.2008.03.028

Tang, C. C., Thomas, D. C., and Saunders, R. M. (2015). Molecular phylogenetics of the species-rich angiosperm genus Goniothalamus (Annonaceae) inferred from nine chloroplast DNA regions: synapomorphies and putative correlated evolutionary changes in fruit and seed morphology. *Mol. Phylogenet. Evol.* 92, 124–139. doi: 10.1016/j.ympev.2015.06.016

Testo, W., Field, A., and Barrington, D. (2018). Overcoming among-lineage rate heterogeneity to infer the divergence times and biogeography of the clubmoss family Lycopodiaceae. *J. Biogeogr.* 45, 1929–1941. doi: 10.1111/jbi.13373

Thomas, D. C., Chatrou, L. W., Stull, G. W., Johnson, D. M., Harris, D. J., Thongpairoj, U.-S., et al. (2015). The historical origins of palaeotropical intercontinental disjunctions in the pantropical flowering plant family Annonaceae. *Perspect. Plant Ecol. Evol. Syst.* 17, 1–16. doi: 10.1016/j.ppees.2014.11.001

Thomas, D. C., Tang, C. C., and Saunders, R. M. K. (2017). Historical biogeography of Goniothalamus and Annonaceae tribe Annoneae: dispersal-vicariance patterns in tropical Asia and intercontinental tropical disjunctions revisited. *J. Biogeogr.* 44, 2862–2876. doi: 10.1111/jbi.13086

Tranchant-Dubreuil, C., Ravel, S., Monat, C., Sarah, G., Diallo, A., Helou, L., et al. (2018). TOGGLe, a flexible framework for easily building complex workflows and performing robust large-scale NGS analyses. *bioRxiv [Preprint]*. 245480. doi: 10.1101/245480

van Setten, A. K., and Maas, P. J. M. (1990). Studies in Annonaceae. XIV. Index to generic names of Annonaceae. *Taxon* 39, 675–690. doi: 10.2307/1223393

Versteegh, P. C., and Sosef, M. S. M. (2007). Revision of the African genus Annickia (Annonaceae). *Syst. Geogr. Plants* 77, 91–118.

Wickett, N. J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., et al. (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. U.S.A.* 111, E4859–E4868. doi: 10.1073/pnas.1323926111

Xi, Z., Liu, L., Rest, J. S., and Davis, C. C. (2014). Coalescent versus concatenation methods and the placement of Amborella as sister to water lilies. *Syst. Biol.* 63, 919–932. doi: 10.1093/sysbio/syu055

Xi, Z., Rest, J. S., and Davis, C. C. (2013). Phylogenomics and coalescent analyses resolve extant seed plant relationships. *PLoS ONE* 8:e80870. doi: 10.1371/journal.pone.0080870

Xue, B., Su, Y. C. F., Thomas, D. C., and Saunders, R. M. K. (2012). Pruning the polyphyletic genus Polyalthia (Annonaceae) and resurrecting the genus Monoon. *Taxon* 61, 1021–1039.

Xue, B., Tan, Y.-H., Thomas, D. C., Chaowasku, T., Hou, X.-L., and Saunders, R. M. K. (2018). A new Annonaceae genus, Wuodendron, provides support for a post-boreotropical origin of the Asian-Neotropical disjunction in the tribe Miliuseae. *Taxon* 67, 250–266. doi: 10.12705/672.2

Zellmer, A. J., Hanes, M. M., Hird, S. M., and Carstens, B. C. (2012). Deep phylogeographic structure and environmental differentiation in the carnivorous plant *Sarracenia alata*. *Syst. Biol.* 61, 763–777. doi: 10.1093/sysbio/sys048

Zhang, C., Sayyari, E., and Mirarab, S. (2017). "ASTRAL-III: increased scalability and impacts of contracting low support branches," in *Comparative Genomics, Lecture Notes in Computer Science*, eds A. McLysaght and D. H. Huson (Cham: Springer), 53–75. doi: 10.1007/978-3-319-67979-2_4