Organelle Phylogenomics and Extensive Conflicting Phylogenetic Signals in the Monocot Order Poales

Hong Wu[1,2], Jun-Bo Yang[1], Jing-Xia Liu[1], De-Zhu Li[1*] and Peng-Fei Ma[1*]

[1] Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China,
[2] University of Chinese Academy of Sciences, Beijing, China

The Poales is one of the largest orders of flowering plants with significant economic and ecological values. Reconstructing the phylogeny of the Poales is important for understanding its evolutionary history that forms the basis for biological studies. However, due to sparse taxon sampling and limited molecular data, previous studies have resulted in a variety of contradictory topologies. In particular, there are three nodes surrounded by incongruence: the phylogenetic ambiguity near the root of the Poales tree, the sister family of Poaceae, and the delimitation of the xyrid clade. We conducted a comprehensive sampling and reconstructed the phylogenetic tree using plastid and mitochondrial genomic data from 91 to 66 taxa, respectively, representing all the 16 families of Poales. Our analyses support the finding of Bromeliaceae and Typhaceae as the earliest diverging groups within the Poales while having phylogenetic relationships with the polytomy. The clade of Ecdeiocoleaceae and Joinvilleaceae is recovered as the sister group of Poaceae. The three families, Mayacaceae, Eriocaulaceae, and Xyridaceae, of the xyrid assembly diverged successively along the backbone of the Poales phylogeny, and thus this assembly is paraphyletic. Surprisingly, we find substantial phylogenetic conflicts within the plastid genomes of the Poales, as well as among the plastid, mitochondrial, and nuclear data. These conflicts suggest that the Poales could have a complicated evolutionary history, such as rapid radiation and polyploidy, particularly allopolyploidy through hybridization. In sum, our study presents a new perspicacity into the complex phylogenetic relationships and the underlying phylogenetic conflicts within the Poales.

Keywords: Poales, phylogenomic conflict, plastome, mitochondrial, nuclear

INTRODUCTION

The order, Poales is a large group of flowering plants in the monocotyledons and belongs to the Commelinid clade, which includes the other three orders of Arecales, Commelinales, and Zingiberales (Angiosperm Phylogeny Group IV [APG IV] et al., 2016). With more than 20,000 species, Poales accounts for about 7% of the angiosperm and 33% of the monocot diversity, respectively (Givnish et al., 2010; Bouchenak-Khelladi et al., 2014; Alves et al., 2015). The species diversity of Poales is extremely uneven among these families. The largest family is Poaceae having about 12,000 species and the smallest one is Ecdeiocoleaceae with only three species

(Christenhusz and Byng, 2016; Hochbach et al., 2018). These species are widely distributed around the world, from the equator to the pole, from floating aquatic plants to the most water-deficient deserts, and most soil types (Stevens, 2001 onward). Moreover, they are generally becoming the dominant species in their ecological communities, such as the grasses (Poaceae) in the savanna and grassland and sedges (Cyperaceae) in the wetland (Linder and Rudall, 2005). Many species of Poales also have significant economic values with Poaceae as the most economically important family in the plant kingdom (Vallée et al., 2016). This family includes many food crops, e.g., rice (*Oryza sativa* L.), wheat (*Triticum aestivum* L.), and maize (*Zea mays* L.), as well as a variety of bamboos that have multiple applications (Saarela et al., 2018). The pineapple [*Ananas comosus* (L.) Merr.] in Bromeliaceae is a famous tropical fruit and an ornamental plant (Chen et al., 2019). *Typha orientalis* C. Presl and *T. angustifolia* L. from Typhaceae are widely used in weaving and paper industries (Sun, 1992).

Based on a phylogeny using the plastid *rbc*L gene, Duvall et al. (1993) first proposed a composition of 16 families of Poales. Since then, the delineation of Poales has gradually been transformed based on the combined morphological and plastid DNA evidence (Kellogg and Linder, 1995; Angiosperm Phylogeny Group [APG], 1998; Chase et al., 2000; Bremer, 2002). In APG IV, two families, Anarthriaceae and Centrolepidaceae were merged into Restionaceae (Angiosperm Phylogeny Group IV [APG IV] et al., 2016). However, the phylogenetic relationships among them are disputed and the delimitation of the Restionaceae is still problematic (Linder and Rudall, 1993; Bremer, 2002; Michelangeli et al., 2003; Briggs et al., 2014; Hochbach et al., 2018).

The 16 families of Poales can be generally divided into five clades or grades (Linder and Rudall, 2005). The Bromeliaceae, Rapateaceae, and Typhaceae comprise the early diverging grade. The remaining four clades are called "core Poales," which include the cyperid, xyrid, restiid, and graminid clades. The cyperid clade is strongly supported, including Cyperaceae, Juncaceae, and Thurniaceae. As the probable sister group of the cyperid clade, the xyrid clade consists of Eriocaulaceae, Mayacaceae, and Xyridaceae. However, the phylogenetic position and the relationship of the xyrid clade are still ambiguous (Givnish et al., 2018; Hochbach et al., 2018). The Restionaceae, Centrolepidaceae, and Anarthriaceae form the restiid clade and sister to the graminid clade, which encompasses the remaining four families, Ecdeiocoleaceae, Flagellariaceae, Joinvilleaceae, and Poaceae (Hochbach et al., 2018).

The phylogeny of Commelinid has always been a hot topic in the tree of life of monocots and the position of Poales in it has been determined (Barrett et al., 2016). However, a few studies focused on the Poales despite their high ecological and economical significance. The first study focusing on the Poales with a large-scale dataset was provided by Givnish et al. (2010) who sequenced 81 plastid genes of 34 representative species from 15 families. Although the backbone phylogeny of Poales has been reconstructed in this study, there are still many uncertainties about its phylogenetic relationships. First, in contrast to the earliest divergence of Bromeliaceae suggested by Chase et al.

(2006), Givnish et al. (2010), Barrett et al. (2016), Givnish et al. (2018), and Hochbach et al. (2018) used the *mat*K to reveal the Bromeliaceae and Typhaceae in an early diverging polytomy and this was supported by a large number of plastid-, mitochondrial- and nuclear-based studies in which Bromeliaceae + Typhaceae were resolved as the early diverging group (Christin et al., 2008; Soltis et al., 2011; Bouchenak-Khelladi et al., 2014; Hertweck et al., 2015; Baker et al., 2021). Moreover, McKain et al. (2016) used the transcriptome data to show that Typhaceae was the first lineage to diverge within the Poales followed by Bromeliaceae. Interestingly, Darshetkar et al. (2019) analyzed the 81 plastid genes dataset like Givnish et al. (2010) with different software and models and obtained the conflicting result with Typhaceae being the first diverging lineage followed by Bromeliaceae. Second, many studies supported Ecdeiocoleaceae as sister to Poaceae (Givnish et al., 2010, 2018; Barrett et al., 2016; Darshetkar et al., 2019; Li et al., 2019). However, the Ecdeiocoleaceae + Joinvilleaceae clade was revealed to be a sister to Poaceae with increased taxon sampling (Bouchenak-Khelladi et al., 2014; McKain et al., 2016; Hochbach et al., 2018; Baker et al., 2021). Third, the relationship involving the Mayacaceae and the xyrid clade is enigmatic. Earlier studies suggested that the Mayacaceae was either within or closely related to the xyrid clade (Michelangeli et al., 2003; Linder and Rudall, 2005; Givnish et al., 2010, 2018; Darshetkar et al., 2019), but other studies suggested that it was within the cyperid clade (Chase et al., 2000, 2006; Janssen and Bremer, 2004) or between the xyrid clade and the cyperid clade (Davis et al., 2004; Hertweck et al., 2015; McKain et al., 2016). Recently, Baker et al. (2021) used 353 nuclear genes to construct the tree of life of angiosperms and found that the Mayacaceae was resolved as an early diverging lineage of Poales, only after the divergence of Bromeliaceae + Typhaceae. In short, these conflicting phylogenetic relationships may be due to the different molecular markers and/or the sparse taxon sampling, as well as various phylogenetic reconstruction methods used in previous studies.

Phylogenomics is an effective way to reconstruct the tree of life based on the genome-scale data (Delsuc et al., 2005; Yu and Zhang, 2006; Zou and Ge, 2008; Zeng et al., 2014; Wen et al., 2015). The data sources of phylogenomics can be either from organelle genomes or from nuclear genomes. In plants, organelle genomes include the plastid genome and mitochondrial genome, which mostly follow a uniparental inheritance (Zeng et al., 2014). In general, the nuclear genome has a faster evolutionary rate and the mitochondrial genome has a slower evolutionary rate while the plastid genome has a moderate rate (Tian and Li, 2002; Wei et al., 2005). The plastid genome is widely used for plant phylogenomic studies (Wei et al., 2005; Leseberg and Duvall, 2009; Gao et al., 2010; Davis et al., 2014; Zeng et al., 2014; Zhao et al., 2021). Plastid phylogenomics has successfully solved a number of plant phylogenetic problems involving taxonomic categories from high to low, e.g., the tree of life of angiosperms at the ordinal level (Li et al., 2019), phylogenetic relationships at the familial level of Malpighiales (Xi et al., 2012), or within the families of Rosaceae and Leguminosae (Zhang et al., 2017, Zhang R. et al., 2020), and the establishment of the phylogenetic framework of the bamboo

tribe, Arundinarieae (Ma et al., 2014, 2017). Compared with the plastid genome, the plant mitochondrial genome has a slower evolutionary rate but a higher genomic rearrangement rate and thus poses challenges for assembly (Wei et al., 2005). These features restrict its application in the study of plant phylogeny. However, some studies have shown that mitochondrial genes could provide additional evolutionary information and are useful in reconstructing the plant phylogeny (Norman and Gray, 2001; Perrotta et al., 2002; Guo and Ge, 2004; Qiu et al., 2010; Bock et al., 2014; Sun et al., 2015; Folk et al., 2016).

In phylogenomics, hundreds to thousands of DNA loci are used, and the phylogenetic discordance or conflicting gene trees appear frequently. The reasons could be the stochastic error and the systematical error, and more often, biological factors including horizontal gene transfer, hybridization, introgression, gene duplication and loss, incomplete lineage sorting, and non-allelic gene conversion (Zou and Ge, 2008; Degnan and Rosenberg, 2009; Sun et al., 2015; Harpak et al., 2017; Kapli et al., 2020). A lot of studies have revealed inconsistencies among plastid, mitochondria, and nuclear phylogenies in plants (Wendel et al., 1995; Sun et al., 2015; Vargas et al., 2017; Hochbach et al., 2018; Jost et al., 2021). In addition, several recent studies suggested a phylogenetic discordance within the plastome at varied evolutionary scales (Gonçalves et al., 2019; Walker et al., 2019; Zhang R. et al., 2020; Yang et al., 2021), questioning the traditional concept of plastomes as a single inherited unit, or at least treating them uncritically in phylogenetic analyses (Gonçalves et al., 2019). More empirical studies are demanded to describe and understand the source, scope, and consequences of conflicting phylogenetic signals in the plastid genome (Yang et al., 2021).

Here, we adopted the classification of 16 families of Poales in this study as the basis for analyses. We reconstructed the phylogeny of Poales with plastid and mitochondrial genomes with at least two taxa or samples for each of the 16 families. The aims of this study are to (1) resolve phylogenetic relationships of key nodes for Poales involving the early diverging grade, xyrid clade, and the sister group of Poaceae; (2) explore the potential conflict within the plastid genome in the Poales phylogeny; and (3) compare the phylogenies built from different plant genomes. With the broadest taxon sampling and the genomic-scale level data, our study resolved ambiguous phylogenetic relationships within the Poales and provided new insights into the conflicting signals among different genomes.

MATERIALS AND METHODS

Plant Materials, Sequencing, Assembly, and Annotation

We selected and sampled representative species in considering the total species of each family of the Poales. All 16 families had a sampling of at least two species. We increased samplings for large families accordingly and 19 species for both Poaceae and Cyperaceae, eight species for Bromeliaceae, four to six species for five families with a total number of 50–1,000 species, and two to four species for eight families less than 50 species. Our sampling could represent the species diversity and phylogenetic

diversity of families as far as possible. For the two main systematic problems concerned, we increased the sampling of Bromeliaceae, Typhaceae, Ecodeiocoleaceae, and Joinvilleaceae by two to three species (or individuals).

Illumina sequencing of genomic DNA was undertaken with about 2–10 GB of raw data with 150 bp paired-end reads generated for each sample. Plastomes and mitochondria gene sequences were assembled *de novo* using the GetOrangelle pipeline (Jin et al., 2020). For mitochondria gene sequence, we selected a reference mitochondria genome (*Oryza sativa* L. Indica Group, NC_007886) to blast and retrieve the output contigs by GetOrangelle, many of which were derived from the mitochondrial genome, and further assembled them in Geneious v9.1.4 (Kearse et al., 2012). We also downloaded 33 published plastomes of Poales for analyses, totaling 99 Poales accessions representing 91 species from 50 genera and 16 families. Meanwhile, we selected seven species of Commelinales and Zingiberales as outgroup taxa based on previous studies (Barrett et al., 2016). The corresponding species voucher and GenBank accession numbers are listed in **Supplementary Table 1**.

For the mitochondrial dataset, we obtained sequences of 48 taxa (50 accessions) and downloaded 15 additional complete mitochondrial genomes. As the combination of multi-locus data of representative taxa with single loci from multiple species can generate reliable higher-level phylogenies (Talavera et al., 2021), the *cob* gene sequences of seven species were also included, resulting in a total of 66 species (72 accessions) representing 16 families and 43 genera of Poales. Since there are no published mitochondrial genomes available for the Commelinales and Zingiberales, we selected two species of Arecales [*Phoenix dactylifera* L. (NC_016740) and *Cocos nucifera* L. (NC_031696)] and one species of Asparagales [(*Allium cepa* L. (NC_030100)] as the outgroup. The corresponding species voucher and GenBank accession numbers are listed in **Supplementary Table 2**.

The assembled plastomes were annotated with the PGA software (Qu et al., 2019), followed by manual examination and adjustment in Geneious v9.1.4 (Kearse et al., 2012). Mitochondrial genes were annotated in Geneious v9.1.4 (Kearse et al., 2012) according to the gene annotation information of three grass mitochondrial genomes [*Oryza sativa* Indica Group, NC_007886, *Sorghum bicolor* (L.) Moench, NC_008360 and *Zea mays* L. NB, NC_007982], and mitochondrial genes with more than 80% similarity with reference sequence genome were selected for annotation and tree construction.

Phylogenomic Analysis

We extracted the coding regions in PhyloSuite (Zhang D. et al., 2020), including 80 protein-coding, 4 rRNA and 30 tRNA genes of plastomes, and 28 mitochondrial genes, respectively. We obtained two plastid matrices, 114 genes (114PG), 80 protein-coding genes (80PG), and one mitochondrial matrix of 28 genes (28MG). The nucleotides were first translated into amino acid sequences and aligned with MAFFT v.5 (Katoh et al., 2005) software, and then we used PAL2NAL (Suyama et al., 2006) to obtain the corresponding nucleotide alignment. The ambiguously aligned regions were deleted by Gblocks (Castresana, 2000), and the parameters of allowed gap positions included all, with half, and none for the above three matrices.

TABLE 1 | Characteristics of plastid and mitochondrial matrices used for phylogenetic analyses of Poales.

Matrix	No. of species	No. of genes	Length (bp)	No. of parsimony-informative sites	No. of variable sites	Missing data
114PG	106	114	83,266	31,490	8,233	12.70%
114PG-all	106	114	74,643	30,499	7,231	9.90%
114PG-half	106	114	71,445	29,436	6,832	9.30%
114PG-no	106	105	34,176	13,395	2,887	12.70%
114PG-12	106	114	58,068	22,695	5,878	0.00%
80PG	106	80	75,593	30,287	7,495	13.80%
80PG-all	106	80	67,728	29,349	6,548	10.80%
80PG-half	106	80	64,968	28,407	6,220	10.20%
80PG-no	106	72	31,115	12,908	2,494	1.90%
80PG-12	106	80	50,396	21,492	5,140	13.80%
80PG-3	106	80	25,198	8,795	2,355	13.80%
28MG	75	28	23,709	6,608	2,571	26.40%
28MG-all	75	28	22,383	6,324	2,452	25.40%
28MG-half	75	28	21,704	6,186	2,374	24.90%
28MG-no	75	28	14,529	3,624	1,478	25.30%
80PG_OR_EJ12	106	79	66,179	27,894	6,049	9.70%
80PG-all_OR_EJ12	106	79	60,600	27,135	5,443	7.50%
80PG-half_OR_EJ12	106	79	64,794	28,401	6,219	10.00%
80PG-no_OR_EJ12	106	71	29,775	12,168	2,382	9.70%
80PG_OR_BT123	106	78	74,600	29,986	7,439	14.00%
80PG-all_OR_BT123	106	78	66,735	29,048	6,492	11.00%
80PG-half_OR_BT123	106	78	63,984	28,107	6,164	10.30%
80PG-no_OR_BT123	106	71	30,921	12,857	2,483	2.00%
80PG_OR_EMX123	106	78	58,013	24,020	4,888	2.60%
80PG-all_OR_EMX123	106	78	54,738	23,602	4,647	2.20%
80PG-half_OR_EMX123	106	79	59,706	25,120	5,502	6.10%
80PG-no_OR_EMX123	106	70	27,972	11,895	2,286	2.00%

PG, plastid gene; MG, mitochondrial gene; all, using Gblocks that allow gap positions with all; half, using Gblocks that allow gap positions with a half; no, using Gblocks that allow gap positions with none; −12, 1st + 2nd codon positions of the matrix; −3, 3rd codon; OR, outlier removed.

Matrices of the 1st + 2nd codon positions of 114PG, and 1st + 2nd of 80PG matrix, as well as the 3rd codon positions, were, respectively, obtained by SEAVIEW (Gouy et al., 2009). In total, we acquired 11 plastid matrices and four mitochondrial matrices for concatenate and coalescent analyses (**Table 1**).

For concatenate method, Bayesian inference (BI), maximum likelihood (ML), and maximum parsimony (MP) were employed. ML analyses were conducted in IQ-TREE v.1.6.10 (Nguyen et al., 2015) and RAxML v.8.2.12 (Stamatakis, 2015), respectively. IQ-TREE was performed with Ultrafast bootstrap with 1,000 replicates and the best model (**Supplementary Table 3**), and other default parameters. RAxML was implemented with 1,000 replicates using the GTRGAMMA model and other default parameters. MP analyses were conducted by PAUP*4.0b10 (Cummings, 2004). A heuristic search was executed with 1,000 replicates, random addition, and the tree bisection-reconnection (TBR) branch swapping with the MULTrees option. The bootstrap (BS) method with a heuristic search was performed with 1,000 replicates. BI analyses were implemented using the Mrbayes 3.6.2 (Ronquist et al., 2012) plugin in PhyloSuite (Zhang D. et al., 2020) and the best model with Corrected Akaike information criterion (AICc) was selected by ModelFinder (Kalyaanamoorthy et al., 2017). The Markov chain Monte Carlo (MCMC) algorithm was performed with 2,000,000 (plastid

matrices) and 6,000,000 (mitochondrial matrix) generations. Every 1,000 generations sampled one tree with the first 25% of generations abandoned as burnt-in. The trees that remained after reaching a stationary state with the average standard deviation of the split frequencies less than 0.01 were recognized as the consensus trees.

For coalescent analysis, we used ASTRAL-III (Zhang et al., 2018) to infer the species tree with the 80 (80PG matrix) and 72 (80PG-no matrix) plastid gene trees estimated from RAxML with 100 replicates. The branches with BS less than 10% in the gene tree were collapsed using the "nw_ed" code of utilities tool (Junier and Zdobnov, 2010). FigTree v. 1.4.4 was used to visualize the phylogenetic trees (Rambaut, 2012). In general, we defined full support as the posterior probability (PP) = 1.00, BS values = 100%, and local posterior probability (LPP) = 1.00; strong support as PP ≥ 0.99, BS ≥ 85%, LPP ≥ 0.9; moderate support as 0.9 ≤ PP < 0.99, 70% ≤ BS < 85%, and 0.85 ≤ LPP < 0.9; and weak support as PP < 0.9, BS < 70%, and LPP < 0.85.

Quantification Branch Support Values

To further quantify branch support values, we used the Quartet Sampling (QS) method with 1,000 replicates (Pease et al., 2018). This method can also distinguish branches with low

information from those with multiple highly supported but mutually exclusive phylogenetic relationships. Three QS scores were used: (1) Quartet concordance (QC) is the frequency of the concordant quartet inferred over both discordant quartets. (2) Quartet differential (QD) indicates that whether one alternative relationship is sampled more often than the other. (3) Quartet informativeness (QI) is the proportion of replicates that were informative (Pease et al., 2018). The quartet sampling outputs were visualized by the R-script QS_visualization.[1]

Quantification of Phylogenetic Signal for Alternative Tree Topologies

We assessed the phylogenetic signals for three sets of conflicting topologies based on previous studies (Shen et al., 2017, 2021; Zhang R. et al., 2020; Yang et al., 2021). Briefly, we computed the site-wise log-likelihood (SLS) using IQ-TREE and the differences in gene-wise log-likelihood scores (ΔGLS) between conflicting topologies using the Perl scripts of Shen et al. (2021). These analyses allowed us to quantify the phylogenetic signal distribution of alternative topologies at the site and gene levels, and to visualize the proportion of genes that support each topology. The three sets examined were as follows: (a) EJ for the sister relationship of Poaceae, EJ1 of [Poaceae, (Ecdeiocoleaceae, Joinvilleaceae)] vs. EJ2 of [Joinvilleaceae, (Ecdeiocoleaceae, Poaceae)]; (b) BT for the relationship of the early diverging group of Poales, BT1 of [(Typhaceae, Bromeliaceae), (Rapateaceae, core Poales)] vs. BT2 of Typhaceae, [Bromeliaceae, (Rapateaceae, core Poales)] vs. BT3 of Bromeliaceae [Typhaceae, (Rapateaceae, core Poales)]; and (c) EMX for the relationship of Mayacaceae, Eriocaulaceae, and Xyridaceae, EMX1 of EMX1 of Mayacaceae [Eriocaulaceae, (Xyridaceae (restiids, graminids))] vs. EMX2 of (Mayacaceae, Eriocaulaceae), [Cyperids, (Xyridaceae (restiids, graminids))] vs. EMX3 of (Mayacaceae, Eriocaulaceae), [(restiids, graminids), Xyridaceae].

In the analysis of supermatrix, just one or two outlier genes could have a significant effect on the phylogenetic topology (Brown and Thomson, 2016; Shen et al., 2017; Walker et al., 2019; Yang et al., 2021). In order to test this potential effect, we re-built phylogenies after removing the outlier genes in these matrices. The outlier genes were defined as those with phylogenetic signals deviating from a Gaussian-like distribution (Zhang R. et al., 2020; Yang et al., 2021). After removing these genes (**Supplementary Table 5**), we obtained 12 matrices (**Table 1**), i.e., 80PG_outlier_removed_EJ12 (ycf2 removed), 80PG-all_outlier_removed_EJ12 (ycf2), 80PG-half_outlier_removed_EJ12 (ycf2), 80PG-no_outlier_removed_EJ12 (ndhF), 80PG_outlier_removed_BT123 (ycf3 and petD), 80PG-all_outlier_removed_BT123 (ycf3 and petD), 80PG-half_outlier_removed_BT123 (ycf3 and petD), 80PG-no_outlier_removed_BT123 (ycf3), 80PG_outlier_removed_EMX123 (ycf1 and ycf2), 80PG-all_outlier_removed_EMX123 (ycf1 and ycf2), 80PG-half_outlier_removed_EMX123 (ycf1), and 80PG-no_outlier_removed_EMX123 (ndhA and psaB), for analyses.

Test of Topological Concordance

We further tested the conflict and the concordance of gene trees and species trees of 80PG and 80PG-no matrices using PhyParts (Smith et al., 2015). We used the Phyx v1.01 (Brown, 2019) to re-root the gene tree and species tree. Meanwhile, the BS support value that is higher than 50% was retained while those lower than 50% did not provide conflict or concordance information. We then separately mapped the 80 and 72 gene trees onto the corresponding species trees by PhyPart. The output of PhyPart was visualized by the script, phypartspiecharts.py (Johnson, 2020).

RESULTS

Taxon Sampling, Plastome, and Mitochondrial Matrix Characteristics

We newly assembled plastomes for 59 species and obtained 53 complete plastomes and 11 plastomes with gaps. The majority of these genomes were annotated to have 68–80 protein-coding genes, 4 rRNA genes, and 16–30 tRNA genes. The two main plastid gene matrices included 114 genes (114PG) and 80 genes (80PG), and the aligned sequences were 83,266 bp and 75,593 bp, respectively. The proportion of missing data ranged from 0.00 to 13.8% for the 11 matrices.
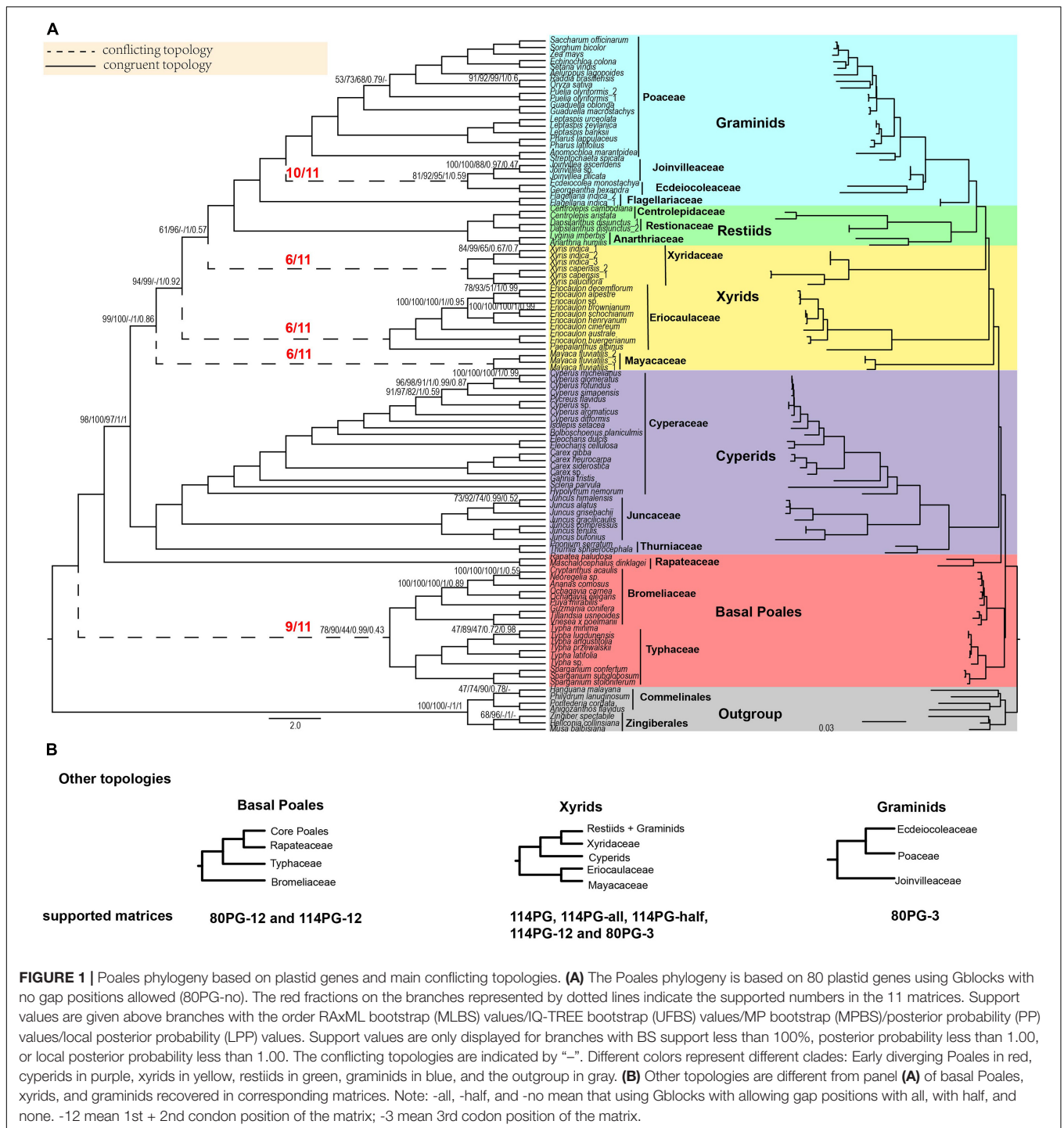
Three complete mitochondrial genomes and 46 at the scaffold level were extracted for 11–28 protein-coding genes for phylogenetic analysis. The alignment of the 28MG matrix was 23,709 bp in length. The proportion of missing data of four matrices ranged from 24.90 to 26.40%. The detailed information of all 15 matrices can be found in **Table 1**.

Plastid Phylogenetic Tree

For ML analyses, phylogenetic trees of Poales constructed using 11 plastid matrices shared a largely consistent topology, except that the relationships among the three families of Mayacaceae, Eriocaulaceae, and Xyridaceae were different (**Figure 1** and **Supplementary Figures 1–10**). The support values of different matrices varied narrowly with generally high support (**Figure 1** and **Supplementary Figures 1–10**). Therefore, the ML topology of the 80 PG-no matrix was selected for discussion, and the BS of RAxML (MLBS) and IQ-TREE (UFBS) were provided on branches.

In the phylogenetic trees of nine plastid matrices, the grouping of Typhaceae and Bromeliaceae was the first lineage diverging within the Poales (**Figure 1** and **Supplementary Figures 1–10**). This node received weak to strong support (BS: 58–90%), and then sister to Rapateaceae and the grouping of the other families with full support. For the remaining two matrices of the 1st + 2nd codon positions of 114PG-12 and 80PG-12, the Bromeliaceae was the earliest diverging lineage with weak support (BS 48–70%).

The cyperid clade received full support in all 11 matrices. In addition, the phylogenetic relationships within it were identical in all the matrices with the Cyperaceae sister to Juncaceae and then the sister to Thurniaceae (**Figure 1** and **Supplementary Figures 1–10**).

FIGURE 1 | Poales phylogeny based on plastid genes and main conflicting topologies. **(A)** The Poales phylogeny is based on 80 plastid genes using Gblocks with no gap positions allowed (80PG-no). The red fractions on the branches represented by dotted lines indicate the supported numbers in the 11 matrices. Support values are given above branches with the order RAxML bootstrap (MLBS) values/IQ-TREE bootstrap (UFBS) values/MP bootstrap (MPBS)/posterior probability (PP) values/local posterior probability (LPP) values. Support values are only displayed for branches with BS support less than 100%, posterior probability less than 1.00, or local posterior probability less than 1.00. The conflicting topologies are indicated by "–". Different colors represent different clades: Early diverging Poales in red, cyperids in purple, xyrids in yellow, restiids in green, graminids in blue, and the outgroup in gray. **(B)** Other topologies are different from panel **(A)** of basal Poales, xyrids, and graminids recovered in corresponding matrices. Note: -all, -half, and -no mean that using Gblocks with allowing gap positions with all, with half, and none. -12 mean 1st + 2nd condon position of the matrix; -3 mean 3rd codon position of the matrix.

There were two different topologies found for the xyrid clade (**Figure 1** and **Supplementary Figures 1–10**). For the first one, the Eriocaulaceae was the sister to Mayacaceae with a close relationship to the cyperid clade with weak to strong support (BS = 58–99%), and the Xyridaceae became an independent lineage that was close to the restiid clade in five matrices (114PG, 114PG-all, 114PG-half, 114PG-12, and 80PG-3) with full support. For the second one, the xyrid clade

was collapsed with Mayacaceae, Eriocaulaceae, and Xyridaceae diverging sequentially along the backbone phylogeny of Poales and the support was BS = 72–99, 61–100, and 100%, respectively.

In the 11 matrices, the restiid clade was all fully supported. The Restionaceae was sister to Centrolepidaceae, and then sister to Anarthriaceae with full support (**Figure 1** and **Supplementary Figures 1–10**). The restiid clade was sister to the graminid clade, which was fully supported in all the 11 matrices. The

Flagellariaceae was sister to the other three families in all matrices with full support. The Ecdeiocoleaceae was sister to Joinvilleaceae in 10 matrices with weak to strong support (BS = 57–92%) and then sister to Poaceae with full support. The 80PG-3 matrix found the grouping of Ecdeiocoleaceae and Poaceae with moderate support (BS = 70–87%) and then sister to Joinvilleaceae with full support (**Figure 1** and **Supplementary Figures 1–10**).

We selected two matrices (80PG and 80PG-no) to run ASTRAL, MP, and BI analyses. For ASTRAL, the early diverging topology of [(Bromeliaceae, Typhaceae) (Rapateacaea, core Poales)] received weak support in 80PG (LPP = 0.61) and 80PG-no (LPP = 0.43). The sister group of Ecdeiocoleaceae and Joinvilleaceae was found in 80PG-no with weak support (LPP = 0.59), while Joinvilleaceae diverged first followed by Ecdeiocoleaceae + Poaceae with moderate support (LPP = 0.86) in 80PG. The other relationships were similar to 80PG-no by ML analyses (**Figure 1** and **Supplementary Figures 11**, **12**).

For MP analyses, the early diverging group of 80PG and 80PG-no was Bromeliaceae + Typhaceae with weak support of maximum parsimony BS (MPBS) = 58 and 44%, respectively, and Ecdeiocoleaceae + Joinvilleaceae was sister to Poaceae with moderate to strong support (MPBS = 74 and 94%). The 80PG matrix found that Eriocaulaceae was sister to Xyridaceae with weak support (MPBS = 53%) and then sister to Mayacaceae with weak support (MPBS = 54%). In the 80PG-no matrix, we obtained a clade of Eriocaulaceae + Mayacaceae with weak support (MPBS = 51%) which was sister to Xyridaceae with strong support (MPBS = 98%). The topology of the other clades was similar to the ML analyses (**Figure 1** and **Supplementary Figures 13**, **14**).

For BI analyses, 80PG and 80PG-no generated the same topology as the ML analyses (114PG-no, 80PG, 80PG-all, 80PG-half, and 80PG-no). The Bromeliaceae + Typhaceae clade was the early diverging group with strong support [posterior probability (PP) = 0.99] in two matrices. The Ecdeiocoleaceae + Joinvilleaceae clade was sister to Poaceae with full support (PP = 1.00) in 80PG-no and weak support (PP = 0.52) in 80PG. The internal relationship of the xyrid clade was Eriocaulaceae, Mayacaceae, and Xyridaceae diverging in sequence. The topology of the other branches was similar to the ML analyses (**Figure 1** and **Supplementary Figures 15**, **16**).

Mitochondrial Phylogenetic Tree

We used four matrices of mitochondrial genes to reconstruct the phylogenetic tree of Poales. In ML analyses, the topologies remained the same except for the phylogenetic relationships involving the three families of Eriocaulaceae, Mayacaceae, and Xyridaceae. In all four matrices, the early diverging group was Typhaceae followed by Bromeliaceae with weak to strong support BS = 72–99% (**Figure 2** and **Supplementary Figures 17–22**). The xyrid clade was revealed to be paraphyletic. In the IQ-tree, the Mayacaceae diverged first and was followed by Eriocaulaceae + Xyridaceae in the 28MG and 28MG-all matrices, and the support for these two nodes was weak BS (38–69%) (**Figure 2** and **Supplementary Figure 18**). For the 28MG-half and 28MG-no matrices, the Mayacaceae also diverged first while the relationship between Eriocaulaceae and

Xyridaceae was unresolved (**Supplementary Figures 20**, **22**). In RAxML, the Mayacaceae diverged first, and Eriocaulaceae and Xyridaceae successively diverged in only 28MG-no matrix with weak support (MLBS = 26–71%) (**Supplementary Figure 21**) and they were sisters to each other with weak support of MLBS = 34–42% for the other three matrices. The restiid clade also became paraphyletic and the Anarthriaceae (*Anarthria humilis* Nees) was embedded in the Restionaceae with weak to strong support BS = 68–93%. The Restionaceae was sister to the cyperid clade with strong support (BS = 85–97%). Surprisingly, the Centrolepidaceae was also embedded in the Cyperid clade and was sister to Thurniaceae with strong support (BS = 85–97%). The sister relationship between Thurniaceae + Centrolepidaceae and Cyperaceae + Juncaceae received full support. The topology of the graminid clade in all the matrices is consistent with the Flagellariaceae in the basal position and Ecdeiocoleaceae + Joinvilleaceae sister to Poaceae with strong support (BS = 89–100%) except for the 28MG-no matrix with weak support (MLBS = 60%) in RAxML (**Figure 2** and **Supplementary Figures 17–22**).
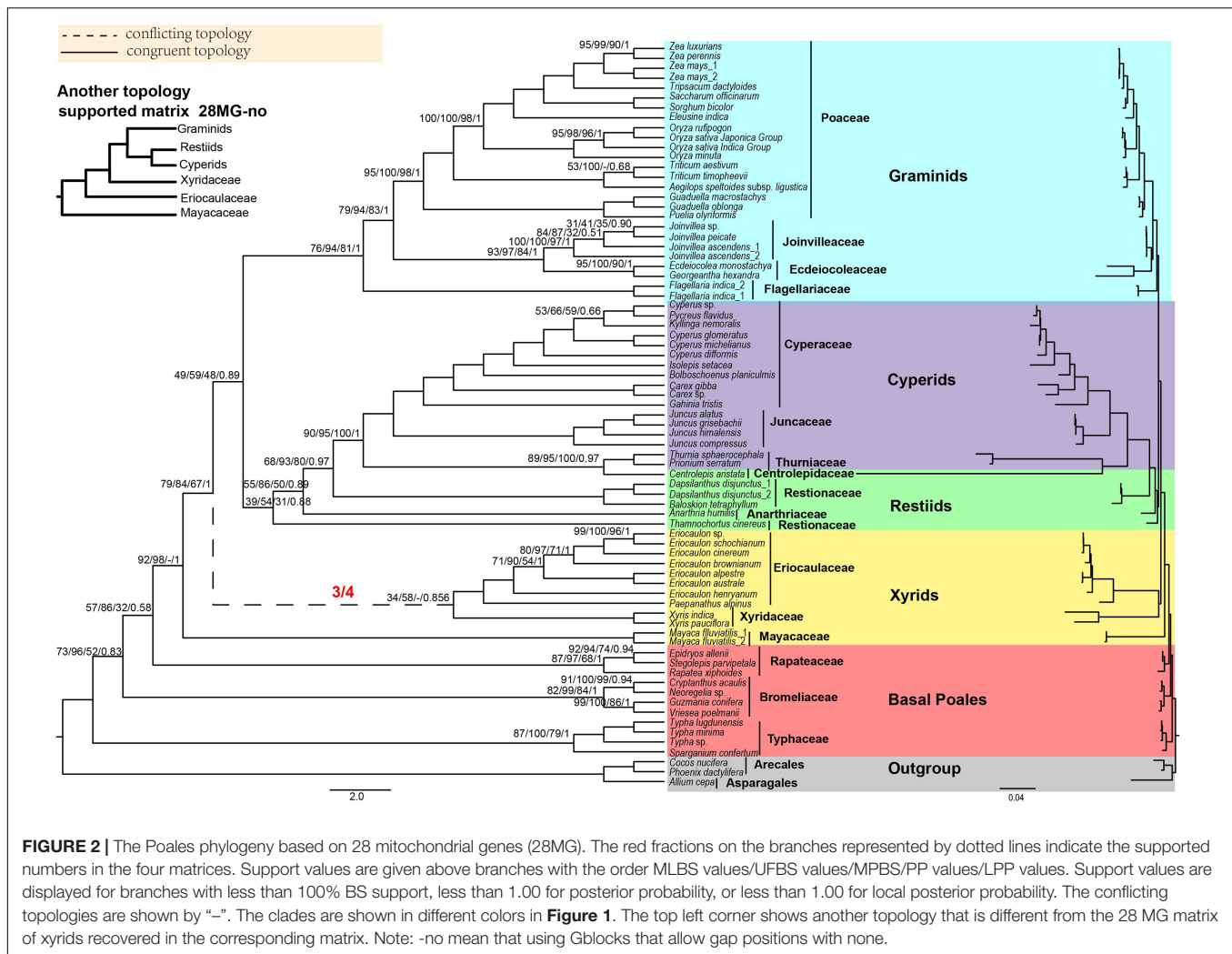
For MP analyses, we only analyzed the 28MG matrix. The early diverging family was the Typhaceae and then Bromeliaceae followed by Rapateaceae with MPBS of 52, 32, and 92% (**Supplementary Figure 23**), respectively. Afterward, the Mayacaceae was sister to Xyridaceae with weak support (MPBS = 68%) and the Eriocaulaceae formed a single clade. The Ecdeiocoleaceae and Joinvilleaceae were sisters with moderate support (MPBS = 84%) and as sisters to Poaceae with strong support (MPBS = 98%). The topologies of other branches were concordant with the RAxML tree of 28MG (**Figure 2** and **Supplementary Figure 23**).

For BI analyses, the topology of 28MG was consistent with the RAxML tree of this matrix (**Figure 2** and **Supplementary Figure 24**). The early diverging families of Typhaceae, Bromeliaceae, and Rapateaceae were separated in turn with PP = 0.83, 0.58, and 1.00. The Ecdeiocoleaceae was sister to Joinvilleaceae and this grouping was then sister to Poaceae and also with full support (**Supplementary Figure 24**).

Quantification of Branch Support Values

We chose two matrices (80PG-no and 28MG) to quantify branch support values. The QC score of ≥0.5 was considered to be strong to manifest support among quartets (Pease et al., 2018; Larson et al., 2020). The full support (QC = 1) was obtained for the monophyly of each family (**Supplementary Figure 25**).

In the 80PG-no matrix, we found no support (QC = -0.048) for Bromeliaceae + Typhaceae (**Supplementary Figure 25**). The monophyly of cyperid and restiid was in full support (QC = 1). The sister relationship between Thurniaceae and Cyperaceae + Juncaceae was also in full support (QC = 1). The Mayacaceae, Eriocaulaceae, and Xyridaceae diverged in sequence with no support (QC = -0.34 and -0.2) and weak support (QC = 0.12) for the node placing Xyridaceae sister to restiid + graminid. Within the graminid clade, the Ecdeiocoleaceae + Joinvilleaceae received moderate support (QC = 0.38) while receiving full support (QC = 1) for this grouping as sister to Poaceae. In contrast, the QD = 0 denoted
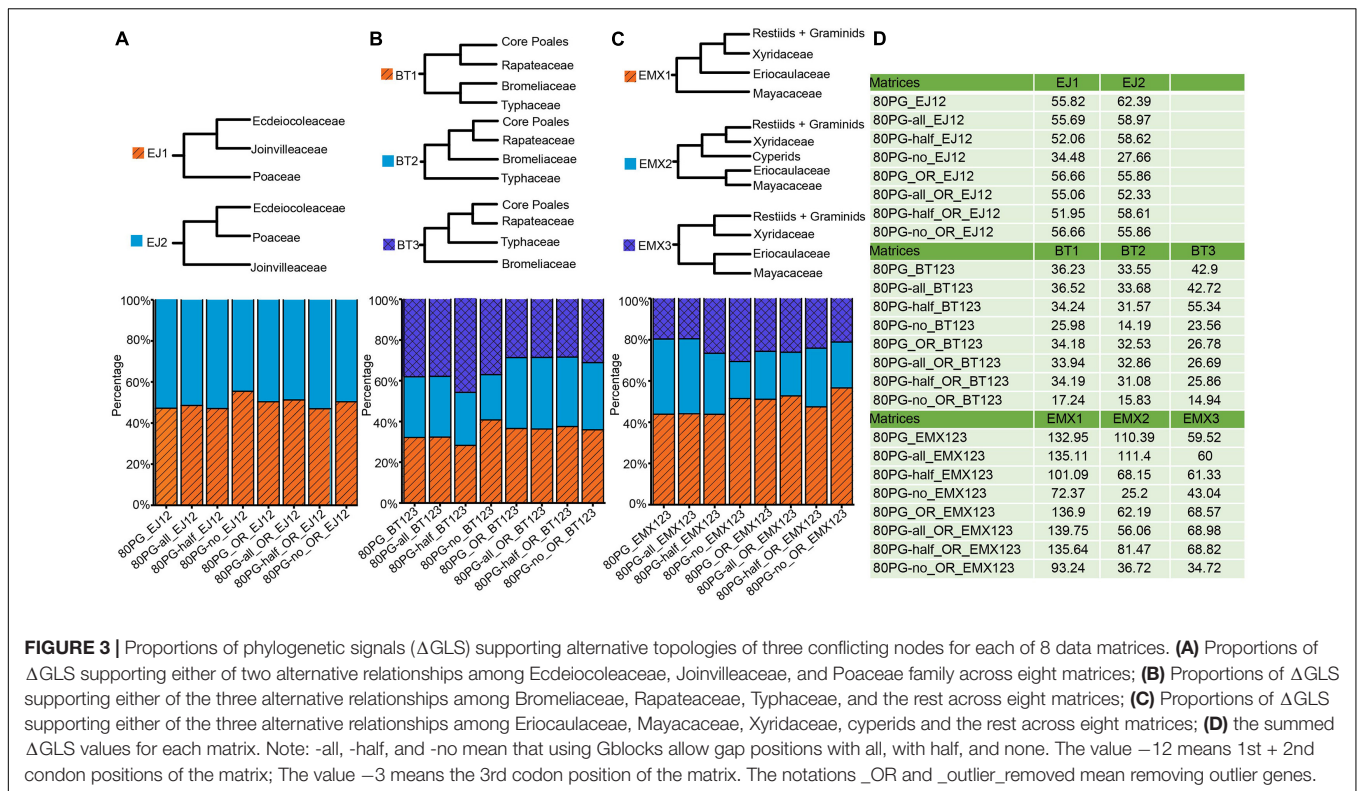
FIGURE 2 | The Poales phylogeny based on 28 mitochondrial genes (28MG). The red fractions on the branches represented by dotted lines indicate the supported numbers in the four matrices. Support values are given above branches with the order MLBS values/UFBS values/MPBS/PP values/LPP values. Support values are displayed for branches with less than 100% BS support, less than 1.00 for posterior probability, or less than 1.00 for local posterior probability. The conflicting topologies are shown by "–". The clades are shown in different colors in **Figure 1**. The top left corner shows another topology that is different from the 28 MG matrix of xyrids recovered in the corresponding matrix. Note: -no mean that using Gblocks that allow gap positions with none.

strong alternative relationships about the nodes within the graminid clade, and in general, the scores of less than 0.3 could be considered as that discordant quartets tend to be heavily skewed toward the conflicting topology (Pease et al., 2018; Larson et al., 2020). The sister groups of Ecdeiocoleaceae and Joinvilleaceae had a low QD score of 0.017 indicating the skew in discordance meaning the possible presence of a supported secondary evolutionary history. Similarly, the nodes connecting Mayacaceae, Eriocaulaceae, and Xyridaceae also received low QD scores (0.27 and 0.044) and thus strong support for alternative evolutionary history. The QD score for the Bromeliaceae + Typhaceae was 0.61, which indicated that inconsistent topologies occurred with relatively equal frequency. The QI scores for all the nodes and the relationships between families were all above 0.76, meaning enough phylogenetic information for these nodes.

For the mitochondrial matrix 28MG, we also obtained no-support (QC = −0.039) for the node placing the Bromeliaceae as the first diverging lineage within the Poales, as well as for the Eriocaulaceae + Xyridaceae (QC = −0.63) (**Supplementary Figure 26**). Moreover, the support was weak for the phylogenetic

relationships within the graminid clade with QC scores from 0.13 to 0.37. The QD scores for the nodes connecting Bromeliaceae and Typhaceae and Eriocaulaceae + Xyridaceae were the same as 0.16, indicating a majority of quartets supporting one of the alternative discordant quartet arrangements. The sister relationship between Cyperaceae + Juncaceae and Thurniaceae + Centrolepidaceae had QD = 0 and all the discordant trees sampled were only one of the two alternative topologies. The QI score for the node leading to Typhaceae and Rapateaceae was 0.4, while the other interfamilial relationships were all supported by QI > 0.68.

Quantification of Phylogenetic Signals for Alternative Topologies

We examined phylogenetic signals for three major conflicting topologies for the Poales phylogeny. Phylogenetic signals for the alternative resolutions of each conflicting topology are shown in **Figure 3** and **Supplementary Table 4**. For the conflict involving Ecdeiocoleaceae and Joinvilleaceae, we examined ΔGLS values between EJ1 and EJ2. The proportions of phylogenetical signals

FIGURE 3 | Proportions of phylogenetic signals (ΔGLS) supporting alternative topologies of three conflicting nodes for each of 8 data matrices. **(A)** Proportions of ΔGLS supporting either of two alternative relationships among Ecdeiocoleaceae, Joinvilleaceae, and Poaceae family across eight matrices; **(B)** Proportions of ΔGLS supporting either of the three alternative relationships among Bromeliaceae, Rapateaceae, Typhaceae, and the rest across eight matrices; **(C)** Proportions of ΔGLS supporting either of the three alternative relationships among Eriocaulaceae, Mayacaceae, Xyridaceae, cyperids and the rest across eight matrices; **(D)** the summed ΔGLS values for each matrix. Note: -all, -half, and -no mean that using Gblocks allow gap positions with all, with half, and none. The value −12 means 1st + 2nd condon positions of the matrix; The value −3 means the 3rd codon position of the matrix. The notations _OR and _outlier_removed mean removing outlier genes.

for EJ1 and EJ2 were basically the same (EJ1: 47.04–55.49% vs EJ2: 44.51–52.96%). We found nearly identical proportions of signals of 46.99–51.27% and 49.64–53.01% in the four new matrices for EJ1 and EJ2 after removing the outlier genes (**Supplementary Table 5**), respectively.

For the Bromeliaceae and Typhaceae, computation of ΔGLS values showed ambiguous proportions of sites supporting anyone of the three different topologies of BT1, BT2, and BT3 (**Figure 3**). The proportions of phylogenetical signal for BT1, BT2, and BT3 ranged from 28.26 to 40.77%, from 22.27 to 29.83%, and from 36.97 to 45.68%, respectively. The 80PG-half showed a higher proportion of phylogenetic signals supporting BT3 (45.68%). After removing the outlier genes, we found that the proportions of signals in the four new matrices of BT1, BT2, and BT3 were still low from 35.91 to 37.52%, from 32.97 to 35.15%, and from 28.38 to 31.12%, respectively.

For the relationships among the three families of Eriocaulaceae, Mayacaceae, and Xyridaceae, computation of ΔGLS values also showed ambiguous proportions of sites (**Figure 3**). The higher proportions of phylogenetic signal from 43.84 to 51.47% were shown for EMX1 while lower values from 17.92 to 36.45% and 19.58 to 30.61% were obtained for EMX2 and EMX3, respectively. We observed increased proportions of signals for EMX1 while decreased values for EMX 2 and EMX3 after removing the outlier genes.

Test of Topological Concordance

The topology of ASTRAL trees of the 80PG and 80PG-no matrix was a bit different (**Supplementary Figure 27**). The one was about the relationship among Eriocaulaceae, Mayacaceae,
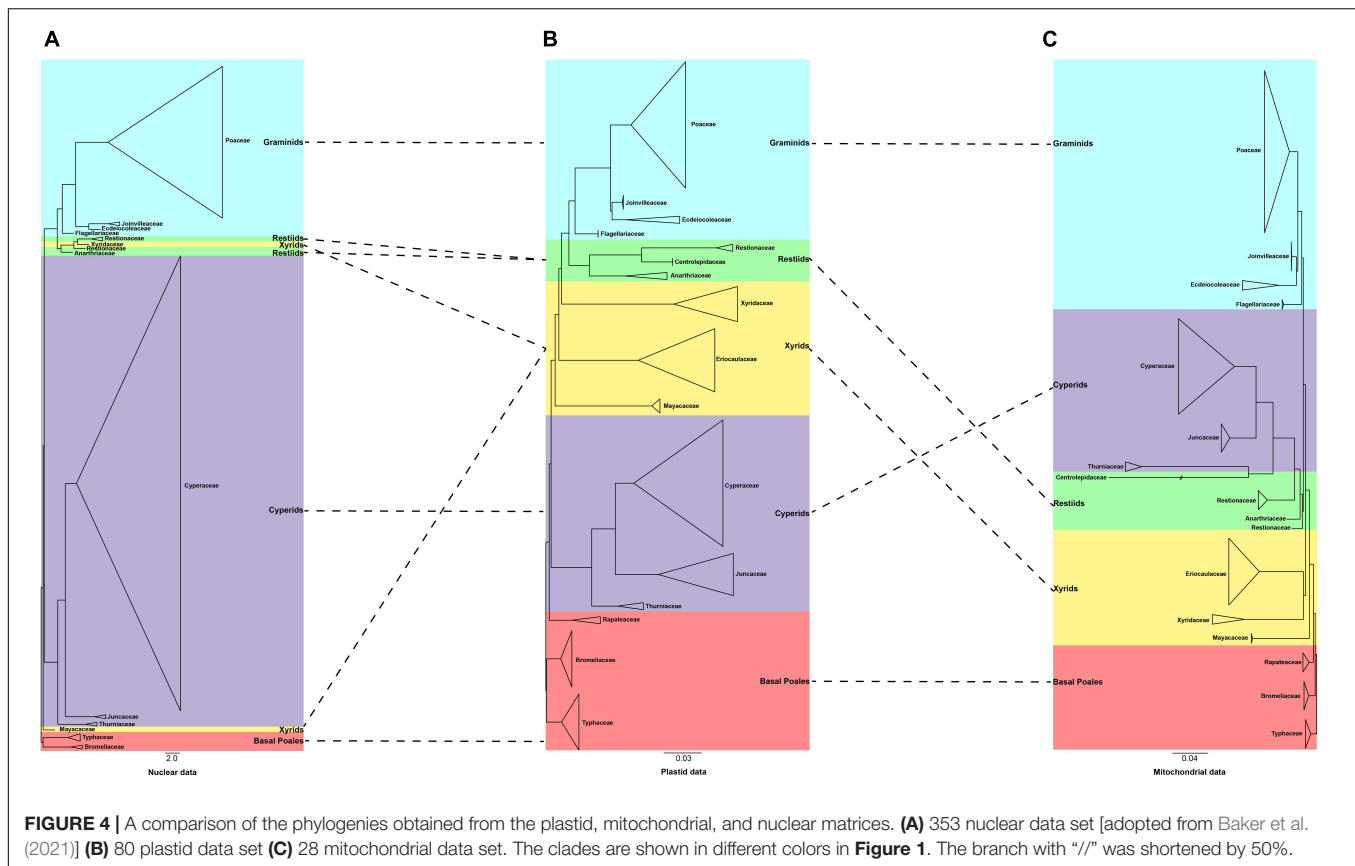
and Xyridaceae, and another involved Ecdeiocoleaceae and Joinvilleaceae. In 80PG, the Mayacaceae diverged first and Eriocaulaceae and Xyridaceae formed a sister group, while the three families diverged sequentially in 80PG-no. The other was about whether the Joinvilleaceae diverged first (80PG) or grouped with the Ecdeiocoleaceae (80PG-no).

Within the 80 genes of the 80PG matrix, only two genes supported the Bromeliaceae + Typhaceae, but 10 genes against this topology and the remaining genes were uninformative on these relationships. In comparison, none supported and seven genes rejected this topology out of the 72 genes in the 80PG-no matrix. The cyperid clade had 41 genes supported and 15 genes rejected from the 80 genes and a high number of 51 genes supported from the 72 genes, indicating that this clade is stable.

The two nodes connecting Eriocaulaceae, Mayacaceae, and Xyridaceae conflicted with a few supported genes of three and four, respectively, and both had 17 rejected genes in 80PG. In 80PG-no, the Eriocaulaceae + Xyridaceae also only had seven supported genes but 23 rejected genes. In contrast, the restiid clade was stable with high support of 38 and 52 genes in 80PG and 80PG-no, respectively. There were more rejected genes than supported genes for the Ecdeiocoleaceae + Joinvilleaceae (20 vs. 12) in 80PG-no and the first divergence of the Joinvilleaceae (26 vs. 14) in 80PG.

Comparison of Organelle and Nuclear Phylogenies

We further compared our plastid and mitochondrial phylogenies with the recently published nuclear tree of Poales

FIGURE 4 | A comparison of the phylogenies obtained from the plastid, mitochondrial, and nuclear matrices. **(A)** 353 nuclear data set [adopted from Baker et al. (2021)] **(B)** 80 plastid data set **(C)** 28 mitochondrial data set. The clades are shown in different colors in **Figure 1**. The branch with "//" was shortened by 50%.

(Baker et al., 2021; **Figure 4**). We found that the position of the graminid remained unchanged. The conflict concentrated on the early diverging group, the cyperid and restiid clades, and the xyrid assembly. Taking the plastid tree as a reference, we explored the conflict in detail. In the plastid phylogeny, the phylogenetic placement of the families of the xyrid assembly is between the cyperid clade and the restiid clade, while in the mitochondrial phylogeny, the position changes and is located between the early diverging taxa and the cyperid clade. For the nuclear data, the placement of the Mayacaceae is the same as that of mitochondria and the Xyridaceae is clustered into the Restionaceae (Baker et al., 2021). The cyperid clade has a close relationship with the early diverging group in the plastid data, while it is located between the xyrid assembly, Mayacaceae and restiid clade in the nuclear data (Baker et al., 2021), and between the graminid clade and the restiid clade in the mitochondrial data. The restiid clade is sister to the graminid clade in the plastid and nuclear data while sister to the cyperid clade in the mitochondrial data.

DISCUSSION

The Early Diverging Poales

The early diverging Poales include three families: Bromeliaceae, Rapateaceae, and Typhaceae. However, the relationships among them are variable in recent studies. The first divergence of Bromeliaceae was supported by the analyses of 75, 77, or 81

plastid genes (Givnish et al., 2010, 2018; Barrett et al., 2016) and the combination of one mitochondrial gene, two rDNA genes, and four plastid genes (Chase et al., 2006). In contrast, Typhaceae was estimated as the first diverging lineage followed by Bromeliaceae based on the analyses of 81 plastid trees in other studies (Darshetkar et al., 2019; Li et al., 2019) or nuclear genes (McKain et al., 2016). Moreover, the sister relationship between Bromeliaceae and Typhaceae was also revealed in some studies, such as using the *rbc*L and *ndh*F genes (Christin et al., 2008; Bouchenak-Khelladi et al., 2014), the low-copy nuclear gene *PHYC* (Hertweck et al., 2015), and 353 nuclear genes (Baker et al., 2021). Analyses of the concatenated plastid genes all uncovered the sister relationship of Bromeliaceae and Typhaceae with weak to strong support (**Figure 1** and **Supplementary Figures 1–10**). Similarly, the coalescent ASTRAL analyses also supported Bromeliaceae as sister to Typhaceae despite weak support (**Supplementary Figures 11**, **12**).

Nevertheless, the mitochondrial trees revealed the Typhaceae as the earliest diverging group with weak to strong support (**Figure 2** and **Supplementary Figures 17–24**), and this result is in conflict with the plastid phylogeny. The conflict between plastid and mitochondrial phylogenies may be due to the evolutionary histories of these two subcellular compartments being unlinked (Sun et al., 2015) and/or incomplete lineage sorting (Lee et al., 2018). The branches leading to the Bromeliaceae and Typhaceae are very short, indicating that they may endure extreme changes in the rates of molecular evolution

(Givnish et al., 2018). Moreover, the short internal branches and conflicting topology of Typhaceae and Bromeliaceae in trees could also be due to rapid radiation as suggested before (hence lack of sufficient phylogenetic signals) (Barrett et al., 2013; Hertweck et al., 2015; Hochbach et al., 2018). Taken together, we consider the phylogenetic relationship between Bromeliaceae and Typhaceae as polytomy Hochbach et al. (2018) and as the earliest diverging lineages of Poales followed by Rapateaceae.

The Non-monophyly of the Xyrid Assembly

The phylogenetic relationships among Eriocaulaceae, Mayacaceae, and Xyridaceae within the xyrid clade remain unresolved in previous studies. The Eriocaulaceae was placed as sister to Xyridaceae (Christin et al., 2008) or Mayacaceae (Givnish et al., 2010) in analyses using plastid genes, while the recent plastome study recognized the topology of [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (restiids, graminids)))] (Li et al., 2019). Using 353 nuclear genes, Baker et al. (2021) suggested that the Mayacaceae was an early diverging lineage within the Poales following the divergence of Bromeliaceae and Typhaceae, while the Xyridaceae was embedded in the restiid clade.

Five of our concatenated plastid datasets (114PG, 114PG-all, 114PG-half, 114PG-12, and 80PG-3) revealed Eriocaulaceae as sister to Mayacaceae with a close relationship to the cyperid clade, and the Xyridaceae became an independent clade that was close to the restiid clade (**Supplementary Figures 1**–**3**, **5**, and **10**). The other six matrices supported the topology [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (restiids, graminids)))] (**Figure 1** and **Supplementary Figures 4**, **6**–**9**). The two conflicting topologies are obtained probably, as a result, the tRNA and rRNA genes have different evolutionary histories from the protein-coding genes in the plastid genome, e.g., faster substitution rate. The third codon positions can mutate more frequently than the first and second positions and thus may experience mutation saturation leading to the phylogenetic artifact (Zeng et al., 2017). In comparison, the ASTRAL analyses revealed the topology [Mayacaceae, ((Eriocaulaceae, Xyridaceae), (restiids, graminids))] and [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (restiids, graminids)))] with weak and strong support in 80PG and 80PG-no, respectively (**Supplementary Figures 11**, **12**). For the mitochondrial data, two different topologies emerged, [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (restiids, graminids)))] and [Mayacaceae, ((Eriocaulaceae, Xyridaceae), (restiids, graminids))]. This result is likely due to the insufficient informative sites in the mitochondrial genes. Whatever be the final phylogenetic resolution of the xyrid clade, it appears that the monophyly of this clade would not be achieved and we can consider it paraphyletic. However, for the convenience of communication, we suggest tentatively keeping the name and calling it the paraphyletic xyrid assembly.

The Restiid Clade

The internal relationship of the restiid clade is still doubtful (Briggs et al., 2014; Hochbach et al., 2018). Previous studies based on the plastid, nuclear, or combined plastid, nuclear,

and mitochondrial data (Chase et al., 2006; Bouchenak-Khelladi et al., 2014; Hochbach et al., 2018; Baker et al., 2021) supported a monophyletic Anarthriaceae, which was sister to the Restionaceae. However, the relationship between Centrolepidaceae and Restionaceae had two different topologies. The sister relationship between Centrolepidaceae and Restionaceae was supported by a majority of previous studies with plastid, nuclear, and mitochondrial data (e.g., Chase et al., 2006; Givnish et al., 2010; Bouchenak-Khelladi et al., 2014; McKain et al., 2016; Darshetkar et al., 2019; Li et al., 2019; Baker et al., 2021). In other studies, the Centrolepidaceae was embedded in Restionaceae based on *rbc*L and *atp*B (Bremer, 2002) and the combined data (*mat*K, *Phy*B, and *Topo*6) (Hochbach et al., 2018). In our study, the topology of [Anarthriaceae, (Restionaceae, Centrolepidaceae)] is revealed by plastid data with strong support in both the concatenate and coalescent methods.

However, in our mitochondrial data, the Anarthriaceae is embedded in the Restionaceae. This result may be due to the fact that only one mitochondrial gene of *cob* was available for *Thamnochortus cinereus* H.P. Linder in analyses and this gene is short of phylogenetic information. Moreover, the Centrolepidaceae is placed in the cyperid clade and as a sister to the Thurniaceae with strong support. The delimitation of the Restionaceae and the placement of the Centrolepidaceae has thus not been fully resolved, and more molecular data and taxon samplings are needed for further analysis.

The Graminid Clade

The monophyly of the graminid clade with the Flagellariaceae as the first diverging group is supported by all previous studies (e.g., Chase et al., 2006; Christin et al., 2008; Soltis et al., 2011; Bouchenak-Khelladi et al., 2014; Hertweck et al., 2015; McKain et al., 2016; Hochbach et al., 2018; Darshetkar et al., 2019; Li et al., 2019; Baker et al., 2021). However, the relationship within the remaining families of Ecdeiocoleaceae, Joinvilleaceae, and Poaceae is still blurry. Both the topologies of [Joinvilleaceae, (Ecdeiocoleaceae, Poaceae)] and [(Joinvilleaceae, Ecdeiocoleaceae), Poaceae] were revealed in previous studies (Givnish et al., 2010; Bouchenak-Khelladi et al., 2014; Barrett et al., 2016; McKain et al., 2016; Hochbach et al., 2018; Darshetkar et al., 2019; Li et al., 2019; Baker et al., 2021). Our study supports the latter topology using the concatenate method, as well as the coalescent analysis of the 80PG. In contrast, the coalescent analysis of the 80PG-no generates the topology [Joinvilleaceae, (Ecdeiocoleaceae, Poaceae)] with moderate support. The conflict between coalescent and concatenate methods could be caused by the limitations of ASTRAL when largely uninformative loci exist as in the 80PG-no (Yang et al., 2021). The mitochondrial data also support the topology of [(Joinvilleaceae, Ecdeiocoleaceae), Poaceae] with high support values and we advocate a topology of [Flagellariaceae, ((Joinvilleaceae, Ecdeiocoleaceae), Poaceae)] for the graminid clade.

Conflicting Signals in Plastid Phylogeny

Based on branch support value and phylogenetic signal analyses, we observed the extensive presence of conflicts among plastid loci involving several long-questioned nodes

(e.g., the relationship of Bromeliaceae and Typhaceae, Joinvilleaceae and Ecdeiocoleaceae, and the xyrid assembly). The no-support means a majority of quartets favor one of the alternative discordant quartet arrangement history (Pease et al., 2018). We found no support for the topologies [(Bromeliaceae, Typhaceae), (Rapateaceae, core Poales)] and [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (graminids, restiids)))]. The ΔGLS values show almost the same proportions of sites supporting three kinds of topologies referring to the Bromeliaceae and Typhaceae, illustrating the conflict with three different topologies, and their relationships may be better treated as polytomy at present. In addition, the topology of [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (graminids, restiids)))] and [(Joinvilleaceae, Ecdeiocoleaceae), Poaceae] both have low QD scores, suggesting that there is a skew between two inconsistent topologies. This result points to a given branch having a biased biological process other than the background lineage sorting, including confounding variables, such as introgression, high heterogeneity of evolutionary rate, heterogeneous base composition, etc. (Pease et al., 2018).

The ΔGLS values show the higher proportions of phylogenetic signal for [Mayacaceae, (Eriocaulaceae, (Xyridaceae, (graminids, restiids)))], and the maximum is 47.5%, but the other two topologies have about 50%. Previous studies have suggested that removing problematic sequences would avoid artifacts to some extent and contribute to the robustness of phylogenetic results (Goremykin et al., 2010; Parks et al., 2012; Yang et al., 2021), although identifying these sequences is difficult (Som, 2015). We removed the outlier genes but found that they have little effect on resolving these long-debated relationships between Mayacaceae, Eriocaulaceae, and Xyridaceae. Recently, Doyle (2021) stressed that the plastid genome should be treated as a single unit for phylogenetic analyses. However, some studies indicate that it is not proper to treat the plastid genome as a single unit particularly when the evidence of recombination is present (Gonçalves et al., 2020; Daniell et al., 2021; Yang et al., 2021). Several studies have shown that the conflicts within the plastid genome (Walker et al., 2019; Zhang R. et al., 2020; Yang et al., 2021), and our results also support this point. However, the cause of the plastid conflict has not yet been determined. Stochastic errors, such as those associated with rapid radiation and limited phylogenetic signals may explain the majority of the observed conflicts (Zhang R. et al., 2020; Yang et al., 2021). In plastid phylogenetic analyses, the vast majority of plastid loci are generally uninformative, and a few genes with strong signals will largely determine the phylogenetic resolution as shown here (**Supplementary Figure 27**; Walker et al., 2019; Zhang R. et al., 2020; Yang et al., 2021). The biological cause for the conflict of plastids could be heterogeneous recombination and a gene transfer between genomes (Walker et al., 2019; Zhang R. et al., 2020; Daniell et al., 2021; Yang et al., 2021). However, we did not find any clues for these factors playing a role here. Instead, the phylogenetic tree of the Poales harbors numerous contrasting short and long branches, meaning highly heterogeneous plastome evolutionary rates among these families of the Poales, and this could be one of the reasons for causing conflicts within the

plastid genomes (Barrett et al., 2013, Barrett et al., 2016). The species of Poales have a variety of habitats, and ten families of them grow in swamp or wet habitats, among which Typhaceae and Mayacaceae are typical aquatic plants (Linder and Rudall, 2005). The plastome of aquatic plants may have a more complex structure, and the deletion and inversion were found in a large number of aquatic plants, e.g., *Eleocharis* (Cyperaceae) and *Najas flexilis* (Willd.) Rostk. and W.L.E. Schmidt (Hydrocharitaceae) (Peredo et al., 2013; Lee et al., 2020). This may also contribute to the phylogenetic conflict. In addition, the photosynthetic pathways of Poales are also diverse, including all known three pathways of C_3, C_4, and Crassulacean acid metabolism (CAM) (Linder and Rudall, 2005). This is a potential explanation for the high heterogeneity of substitution rates heterogeneity, which ultimately results in phylogenetic conflict (Barrett et al., 2016).

Conflicts Among Three Plant Genomes

The phylogenetic conflict between organelles (plastid and mitochondrial) and nuclear genes has been reported in various taxa, such as coralline red algae, *Lachemilla* (Rosaceae), and *Pterocarya* (Juglandaceae) (Lee et al., 2018; Morales-Briones et al., 2018; Mu et al., 2020). The plastid and mitochondrial genomes have a strong conflict in our study. The mitochondrial data have more missing data and lower coverage than the plastid data, which may be one of the factors causing the conflict between mitochondrial and plastid phylogenies. Another possible explanation is that the mitochondrial genomes of Poales have undergone extensive horizontal gene transfer between nuclear and plastid genomes, which is typical in land plants (Folk et al., 2016). In our study, we found that the systematic position of Centrolepidaceae is extraordinary and the branches are very long (**Figure 2** and **Supplementary Figures 19–23**). This phenomenon is also observed by Folk et al. (2016), who indicates that the foreign DNA of the nucleus may be the cause for the large difference in branch length and location in mitochondrial analysis.

Meanwhile, strong conflicts are also detected between the organelle genomes and nuclear genes. The phylogenetic relationships of the three clades/assembly (cyperid, restiid, and xyrid) are different among the three genomes. This may be caused by insufficient sampling of nuclear data. Three families of Centrolepidaceae, Rapateaceae, and Eriocaulaceae are not sampled in our nuclear data, and the sampling distribution is uneven between each family. For example, the Xyridaceae have only one species included, which may be the reason for the odd phylogenetic position of this family. There is also a potential biological source of the incongruence among plastid, mitochondrial, and nuclear loci (Sun et al., 2015). The factors could be incomplete lineage sorting, hybridization, lateral transfer of organellar genomes, plastome capture, and polyploidy (Stegemann et al., 2012; Lee et al., 2018; Morales-Briones et al., 2018).

In particular, polyploidy is prevalent in plant groups and all the families of Poales have experienced polyploidization events in their evolutionary history (McKain et al., 2016; Van de Peer et al., 2017; Morales-Briones et al., 2018; Wu et al., 2020; Guo et al., 2021). Here, the sampled taxa of *Centrolepis aristata* (R.Br.)

Roem. and Schult. (restiid) and the species of Eriocaulaceae (xyrid) are probably hexaploid (Šmarda et al., 2014), and *Juncus bufonius* L. (cyperid) is octoploid (Kubešová et al., 2010). Meanwhile, hybridization is ubiquitous in the green plant (Soltis and Soltis, 2009; Triplett et al., 2010), and the combination of hybridization and polyploid events (Soltis and Soltis, 2009) is another usual cause of phylogenetic conflict. The conflicts of three genomic data reflected in the branch lengths (Jost et al., 2021) indicate that the molecular evolution rate between different genomes is highly heterogeneous with certain families of the Poales experiencing an accelerated rate of sequence evolution (Guisinger et al., 2010; Barrett et al., 2013, Barrett et al., 2016). The fast-evolving sites are more likely to be saturated and prone to the accumulation of non-phylogenetic signals (Rodríguez-Ezpeleta et al., 2007), and thus leading to topological conflicts among the three genomes. In short, the conflict of three genomic data might be due to a combination of the missing data in mitochondria, polyploid history, heterogeneity of molecular evolution rate, and sparse sampling of nuclear data, and these factors deserve to be explored in detail in future studies. Furthermore, due to the limitations of the organelle genome, the nuclear genomic data will be used to finally resolve the phylogenetic relationships of Poales, so as to improve the understanding of the cause of phylogenetic conflict in this order.

CONCLUSION

With a broad taxon sampling, we used plastid and mitochondria genomes to infer the phylogenetic tree of Poales and found the long-standing controversial nodes of Poales mainly caused by extensive conflict across genomic compartments. For the xyrid assembly, we found it paraphyletic, and its relationship with the three families, Eriocaulaceae, Mayacaceae, and Xyridaceae within the Poales is still not fully resolved. Our study has not only revealed phylogenetic conflicts within the plastid genomes, but also extensive conflicts among the plastid, mitochondrial and nuclear data in the Poales. Many factors, such as the missing data of mitochondrion, insufficient nuclear sampling, rapid radiation, heterogeneity of molecular evolution rate, and allopolyploidy by hybridization are potentially involved in generating these conflicts in the Poales.

REFERENCES

Alves, M., Trovó, M., Forzza, R. C., and Viana, P. (2015). Overview of the systematics and diversity of Poales in the Neotropics with emphasis on the Brazilian flora. *Rodriguésia* 66, 305–328. doi: 10.1590/2175-7860201566203

Angiosperm Phylogeny Group [APG] (1998). An ordinal classification for the families of flowering plants. *Ann. Missouri Bot. Gard.* 85, 531–553. doi: 10.2307/2992015

Angiosperm Phylogeny Group IV [APG IV], Chase, M. W., Christenhusz, M. J. M., Fay, M. F., Byng, J. W., Judd, W. S., et al. (2016). An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* 181, 1–20. doi: 10.1111/boj.12385

Baker, W. J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., et al. (2021). A comprehensive phylogenomic platform for exploring the angiosperm tree of life. *Syst. Biol.* syab035. doi: 10.1093/sysbio/syab035 [Epub ahead of print].

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

P-FM, D-ZL, and HW conceptualized the study. HW, P-FM, J-BY, and J-XL were involved in data generation and methodology. HW worked on data analysis and visualization. HW, P-FM, D-ZL, J-BY, and J-XL reviewed and revised the first draft of the manuscript and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.824672/full#supplementary-material

Barrett, C. F., Baker, W. J., Comer, J. R., Conran, J. G., Lahmeyer, S. C., Leebens-Mack, J. H., et al. (2016). Plastid genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. *New Phytol.* 209, 855–870. doi: 10.1111/nph.13617

Barrett, C. F., Davis, J. I., Leebens-Mack, J., Conran, J. G., and Stevenson, D. W. (2013). Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics* 29, 65–87. doi: 10.1111/j.1096-0031.2012.00418.x

Bock, D. G., Kane, N. C., Ebert, D. P., and Rieseberg, L. H. (2014). Genome skimming reveals the origin of the Jerusalem Artichoke tuber crop species: neither from Jerusalem nor an artichoke. *New Phytol.* 201, 1021–1030. doi: 10.1111/nph.12560

Bouchenak-Khelladi, Y., Muasya, A. M., and Linder, H. P. (2014). A revised evolutionary history of Poales: origins and diversification. *Bot. J. Linn. Soc.* 175, 4–16. doi: 10.1111/boj.12160

Bremer, K. (2002). Gondwanan evolution of the grass alliance of families (Poales). *Evolution* 56, 1374–1387. doi: 10.1111/j.0014-3820.2002.tb01451.x

Briggs, B. G., Marchant, A. D., and Perkins, A. J. (2014). Phylogeny of the restiid clade (Poales) and implications for the classification of Anarthriaceae, Centrolepidaceae and Australian Restionaceae. *Taxon* 63, 24–46. doi: 10.12705/631.1

Brown, J. M., and Thomson, R. C. (2016). Bayes factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. *Syst. Biol.* 66, 517–530. doi: 10.1093/sysbio/syw101

Brown, W. J. (2019). *Phyx v1.01*. Available online at: https://github.com/FePhyFoFum/phyx/releases/tag/v1.01 (accessed September 27, 2019).

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

Chase, M. W., Soltis, D. E., Soltis, P. S., Rudall, P. J., Fay, M. F., Hahn, W. H., et al. (2000). "Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification," in *Monocots: Systematics and Evolution*, eds K. L. Wilson and D. A. Morrison (Collingwood, ON: CSIRO), 3–16. doi: 10.1201/9781315166339-2

Chase, M., Fay, M., Devey, D. S., Maurin, O., Rønsted, N., Davies, T., et al. (2006). Multigene analyses of monocot relationships. *Aliso* 22, 63–75. doi: 10.5642/ALISO.20062201.06

Chen, L.-Y., VanBuren, R., Paris, M., Zhou, H., Zhang, X., Wai, C. M., et al. (2019). The bracteatus pineapple genome and domestication of clonally propagated crops. *Nat. Genet.* 51, 1549–1558. doi: 10.1038/s41588-019-0506-8

Christenhusz, M. J., and Byng, J. W. (2016). The number of known plants species in the world and its annual increase. *Phytotaxa* 261, 201–217.

Christin, P.-A., Salamin, N., Muasya, A. M., Roalson, E. H., Russier, F., and Besnard, G. (2008). Evolutionary switch and genetic convergence on rbcL following the evolution of C4 photosynthesis. *Mol. Biol. Evol.* 25, 2361–2368. doi: 10.1093/molbev/msn178

Cummings, M. P. (2004). "PAUP* "Phylogenetic analysis using parsimony (and other methods)"," in *Dictionary of Bioinformatics and Computational Biology*, eds J. M. Hancock and M. J. Zvelebil (Hoboken, NJ: Wiley).

Daniell, H., Jin, S., Zhu, X. G., Gitzendanner, M. A., Soltis, D. E., and Soltis, P. S. (2021). Green giant—a tiny chloroplast genome with mighty power to produce high-value proteins: history and phylogeny. *Plant Biotechnol. J.* 19, 430–447. doi: 10.1111/pbi.13556

Darshetkar, A. M., Datar, M. N., Tamhankar, S., Li, P., and Choudhary, R. K. (2019). Understanding evolution in Poales: insights from Eriocaulaceae plastome. *PLoS One* 14:e0221423. doi: 10.1371/journal.pone.0221423

Davis, C. C., Xi, Z., and Mathews, S. (2014). Plastid phylogenomics and green plant phylogeny: almost full circle but not quite there. *BMC Biol.* 12:11. doi: 10.1186/1741-7007-12-11

Davis, J. I., Stevenson, D. W., Petersen, G., Seberg, O., Campbell, L. M., Freudenstein, J. V., et al. (2004). A phylogeny of the monocots, as inferred from rbcL and atpA sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. *Syst. Bot.* 29, 467–510. doi: 10.1600/0363644041744365

Degnan, J. H., and Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332–340. doi: 10.1016/j.tree.2009.01.009

Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375. doi: 10.1038/nrg1603

Doyle, J. J. (2021). Defining coalescent genes: Theory meets practice in organelle phylogenomics. *Syst. Biol.* syab053. doi: 10.1093/sysbio/syab053

Duvall, M. R., Clegg, M. T., Chase, M. W., Clark, W. D., Kress, W. J., Hills, H. G., et al. (1993). Phylogenetic hypotheses for the monocotyledons constructed from rbcL sequence data. *Ann. Missouri Bot. Gard.* 80, 607–619. doi: 10.2307/2399849

Folk, R. A., Mandel, J. R., and Freudenstein, J. V. (2016). Ancestral gene flow and parallel organellar genome capture result in extreme phylogenomic discord in a lineage of angiosperms. *Syst. Biol.* 66, 320–337. doi: 10.1093/sysbio/syw083

Gao, L., Su, Y.-J., and Wang, T. (2010). Plastid genome sequencing, comparative genomics, and phylogenomics: current status and prospects. *J. Syst. Evol.* 48, 77–93. doi: 10.1111/j.1759-6831.2010.00071.x

Givnish, T. J., Ames, M., McNeal, J. R., McKain, M. R., Steele, P. R., Depamphilis, C. W., et al. (2010). Assembling the tree of the monocotyledons: plastome

sequence phylogeny and evolution of Poales1. *Ann. Missouri Bot. Gard.* 97, 584–616. doi: 10.3417/2010023

Givnish, T. J., Zuluaga, A., Spalink, D., Soto Gomez, M., Lam, V. K. Y., Saarela, J. M., et al. (2018). Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. *Am. J. Bot.* 105, 1888–1910. doi: 10.1002/ajb2.1178

Gonçalves, D. J. P., Simpson, B. B., Ortiz, E. M., Shimizu, G. H., and Jansen, R. K. (2019). Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. *Mol. Phylogenet. Evol.* 138, 219–232. doi: 10.1016/j.ympev.2019.05.022

Gonçalves, D. J., Jansen, R. K., Ruhlman, T. A., and Mandel, J. R. (2020). Under the rug: abandoning persistent misconceptions that obfuscate organelle evolution. *Mol. Phylogenet. Evol.* 151:106903. doi: 10.1016/j.ympev.2020.106903

Goremykin, V. V., Nikiforova, S. V., and Bininda-Emonds, O. R. P. (2010). Automated removal of noisy data in phylogenomic analyses. *J. Mol. Evol.* 71, 319–331. doi: 10.1007/s00239-010-9398-z

Gouy, M., Guindon, S., and Gascuel, O. (2009). SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27, 221–224. doi: 10.1093/molbev/msp259

Guisinger, M. M., Chumley, T. W., Kuehl, J. V., Boore, J. L., and Jansen, R. K. (2010). Implications of the plastid genome sequence of Typha (Typhaceae, Poales) for understanding genome evolution in *Poaceae. J. Mol. Evol.* 70, 149–166. doi: 10.1007/s00239-009-9317-3

Guo, C., Ma, P. F., Yang, G. Q., Ye, X. Y., Guo, Y., Liu, J. X., et al. (2021). Parallel ddRAD and genome skimming analyses reveal a radiative and reticulate evolutionary history of the temperate bamboos. *Syst. Biol.* 70, 756–773. doi: 10.1093/sysbio/syaa076

Guo, Y. L., and Ge, S. (2004). The utility of mitochondrial nad1 intron in phylogenetic study of Oryzeae with reference to the systematic position of Porteresia. *Acta Phytotaxon. Sin.* 42, 333–344.

Harpak, A., Lan, X., Gao, Z., and Pritchard, J. K. (2017). Frequent nonallelic gene conversion on the human lineage and its effect on the divergence of gene duplicates. *Proc. Natl. Acad. Sci. U S A.* 114, 12779–12784. doi: 10.1073/pnas.1708151114

Hertweck, K. L., Kinney, M. S., Stuart, S. A., Maurin, O., Mathews, S., Chase, M. W., et al. (2015). Phylogenetics, divergence times and diversification from three genomic partitions in monocots. *Bot. J. Linn. Soc.* 178, 375–393. doi: 10.1111/boj.12260

Hochbach, A., Linder, H. P., and Röser, M. (2018). Nuclear genes, matK and the phylogeny of the Poales. *Taxon* 67, 521–536. doi: 10.12705/673.5

Janssen, T., and Bremer, K. (2004). The age of major monocot groups inferred from 800+rbcL sequences. *Bot. J. Linn. Soc.* 146, 385–398. doi: 10.1111/j.1095-8339.2004.00345.x

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C. W., Yi, T.-S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21:241. doi: 10.1186/s13059-020-02154-5

Johnson, M. (2020). Phypartspiecharts. https://github.com/mossmatters/phyloscripts/tree/master/phypartspiecharts (accessed November 2, 2021).

Jost, M., Samain, M.-S., Marques, I., Graham, S. W., and Wanke, S. (2021). Discordant phylogenomic placement of Hydnoraceae and Lactoridaceae within Piperales using data from all three genomes. *Front. Plant. Sci.* 12:642598. doi: 10.3389/fpls.2021.642598

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

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Kapli, P., Yang, Z., and Telford, M. J. (2020). Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* 21, 428–444. doi: 10.1038/s41576-020-0233-0

Katoh, K., Kuma, K.-I., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33, 511–518. doi: 10.1093/nar/gki198

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kellogg, E., and Linder, H. (1995). "Phylogeny of Poales," in *Monocotyledons: Systematics and Evolution*, eds P.J. Rudall, P.J. Cribb, D.F. Cutler, and C.J. Humphries (London: Royal Botanic Gardens, Kew), 511-542.

Kubešová, M., Moravcová, L., Suda, J., Jarošík, V., and Pyšek, P. (2010). Naturalized plants have smaller genomes than their non-invading relatives: a flow cytometric analysis of the Czech alien flora. *Preslia* 82, 81–96.

Larson, D. A., Walker, J. F., Vargas, O. M., and Smith, S. A. (2020). A consensus phylogenomic approach highlights paleopolyploid and rapid radiation in the history of Ericales. *Am. J. Bot.* 107, 773–789. doi: 10.1002/ajb2.1469

Lee, C., Ruhlman, T. A., and Jansen, R. K. (2020). Unprecedented intraindividual structural heteroplasmy in eleocharis (Cyperaceae. Poales) Plastomes. *Genome Biol. Evol.* 12, 641–655. doi: 10.1093/gbe/evaa076

Lee, J. M., Song, H. J., Park, S. I., Lee, Y. M., Jeong, S. Y., Cho, T. O., et al. (2018). Mitochondrial and plastid genomes from coralline red algae provide insights into the incongruent evolutionary histories of organelles. *Genome Biol. Evol.* 10, 2961–2972. doi: 10.1093/gbe/evy222

Leseberg, C. H., and Duvall, M. R. (2009). The complete chloroplast genome of Coix lacryma-jobi and a comparative molecular evolutionary analysis of plastomes in Cereals. *J. Mol. Evol.* 69, 311–318. doi: 10.1007/s00239-009-9275-9

Li, H.-T., Yi, T.-S., Gao, L.-M., Ma, P.-F., Zhang, T., Yang, J.-B., et al. (2019). Origin of angiosperms and the puzzle of the Jurassic gap. *Nat. Plants* 5, 461–470. doi: 10.1038/s41477-019-0421-0

Linder, H. P., and Rudall, P. J. (1993). The megagametophyte in anarthria (Anarthriaceae, Poales) and its implications for the phylogeny of the Poales. *Am. J. Bot.* 80, 1455–1464. doi: 10.2307/2445675

Linder, H. P., and Rudall, P. J. (2005). Evolutionary history of Poales. *Annu. Rev. Ecol. Evol. Syst.* 36, 107–124.

Ma, P. F., Vorontsova, M. S., Nanjarisoa, O. P., Razanatsoa, J., Guo, Z. H., Haevermans, T., et al. (2017). Negative correlation between rates of molecular evolution and flowering cycles in temperate woody bamboos revealed by plastid phylogenomics. *BMC Plant Biol.* 17:260. doi: 10.1186/s12870-017-1199-8

Ma, P. F., Zhang, Y. X., Zeng, C. X., Guo, Z. H., and Li, D. Z. (2014). Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (poaceae). *Syst. Biol.* 63, 933–950. doi: 10.1093/sysbio/syu054

McKain, M. R., Tang, H., McNeal, J. R., Ayyampalayam, S., Davis, J. I., et al. (2016). A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. *Genome Biol. Evol.* 8, 1150–1164. doi: 10.1093/gbe/evw060

Michelangeli, F. A., Davis, J. I., and Stevenson, D. W. (2003). Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *Am. J. Bot.* 90, 93–106. doi: 10.3732/ajb.90.1.93

Morales-Briones, D. F., Liston, A., and Tank, D. C. (2018). Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus Lachemilla (Rosaceae). *New Phytol.* 218, 1668–1684. doi: 10.1111/nph.15099

Mu, X.-Y., Tong, L., Sun, M., Zhu, Y.-X., Wen, J., Lin, Q.-W., et al. (2020). Phylogeny and divergence time estimation of the walnut family (Juglandaceae) based on nuclear RAD-Seq and chloroplast genome data. *Mol. Phylogenet. Evol.* 147:106802. doi: 10.1016/j.ympev.2020.106802

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Norman, J. E., and Gray, M. W. (2001). A complex organization of the gene encoding cytochrome oxidase subunit 1 in the mitochondrial genome of the dinoflagellate, crypthecodinium cohnii: homologous recombination generates two different cox1 open reading frames. *J. Mol. Evol.* 53, 351–363. doi: 10.1007/s002390010225

Parks, M., Cronn, R., and Liston, A. (2012). Separating the wheat from the chaff: mitigating the effects of noise in a plastome phylogenomic data set from *Pinus* L. (Pinaceae). *BMC Evol. Biol.* 12:100. doi: 10.1186/1471-2148-12-100

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., and Smith, S. A. (2018). Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. *Am. J. Bot.* 105, 385–403. doi: 10.1002/ajb2.1016

Peredo, E. L., King, U. M., and Les, D. H. (2013). The plastid genome of *Najas flexilis*: adaptation to submersed environments is accompanied by the complete loss of the NDH complex in an aquatic angiosperm. *PLoS One* 8:e68591. doi: 10.1371/journal.pone.0068591

Perrotta, G., Grienenberger, J. M., and Gualberto, J. M. (2002). Plant mitochondrial rps2 genes code for proteins with a C-terminal extension that is processed. *Plant Mol. Biol.* 50, 523–533. doi: 10.1023/a:1019878212696

Qiu, Y.-L., Li, L., Wang, B., Xue, J.-Y., Hendry, T. A., Li, R.-Q., et al. (2010). Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *J. Syst. Evol.* 48, 391–425. doi: 10.1111/j.1759-6831.2010.00097.x

Qu, X. J., Moore, M. J., Li, D. Z., and Yi, T. S. (2019). PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* 15:50. doi: 10.1186/s13007-019-0435-7

Rambaut, A. (2012). Figtree v1.4.4. https://github.com/rambaut/figtree/releases/tag/v1.4.4 (accessed November 26, 2018).

Rodríguez-Ezpeleta, N., Brinkmann, H., Roure, B., Lartillot, N., Lang, B. F., and Philippe, H. (2007). Detecting and overcoming systematic errors in genome-scale phylogenies. *Syst. Biol.* 56, 389–399. doi: 10.1080/10635150701397643

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Saarela, J. M., Burke, S. V., Wysocki, W. P., Barrett, M. D., Clark, L. G., Craine, J. M., et al. (2018). A 250 plastome phylogeny of the grass family (Poaceae): topological support under different data partitions. *PeerJ* 6:e4299. doi: 10.7717/peerj.4299

Shen, X.-X., Hittinger, C. T., and Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat. Ecol. Evol.* 1:0126. doi: 10.1038/s41559-017-0126

Shen, X.-X., Steenwyk, J. L., and Rokas, A. (2021). Dissecting incongruence between concatenation- and quartet-based approaches in phylogenomic data. *Syst. Biol.* 70, 997–1014. doi: 10.1093/sysbio/syab011

Šmarda, P., Bureš, P., Horová, L., Leitch, I. J., Mucina, L., Pacini, E., et al. (2014). Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proc. Natl. Acad. Sci. U S A.* 111, E4096–E4102. doi: 10.1073/pnas.1321152111

Smith, S. A., Moore, M. J., Brown, J. W., and Yang, Y. (2015). Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evol. Biol.* 15:150. doi: 10.1186/s12862-015-0423-0

Soltis, D. E., Smith, S. A., Cellinese, N., Wurdack, K. J., Tank, D. C., Brockington, S. F., et al. (2011). Angiosperm phylogeny: 17 genes, 640 taxa. *Am. J. Bot.* 98, 704–730. doi: 10.3732/ajb.1000404

Soltis, P. S., and Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* 60, 561–588. doi: 10.1146/annurev.arplant.043008.092039

Som, A. (2015). Causes, consequences and solutions of phylogenetic incongruence. *Brief. Bioinform.* 16, 536–548. doi: 10.1093/bib/bbu015

Stamatakis, A. (2015). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Stegemann, S., Keuthe, M., Greiner, S., and Bock, R. (2012). Horizontal transfer of chloroplast genomes between plant species. *Proc. Natl. Acad. Sci. U S A.* 109, 2434–2438. doi: 10.1073/pnas.1114076109

Stevens, P. F. (2001). *Angiosperm Phylogeny Website, Version 14*. Available online at: http://www.mobot.org/MOBOT/Research/APweb/ (accessed May, 22, 2021)

Sun, M., Soltis, D. E., Soltis, P. S., Zhu, X., Burleigh, J. G., and Chen, Z. (2015). Deep phylogenetic incongruence in the angiosperm clade Rosidae. *Mol. Phylogenet. Evol.* 83, 156–166. doi: 10.1016/j.ympev.2014.11.003

Sun, X. Z. (1992). "Typhaceae," in *Flora of China Section 13*, ed. Flora of China Editorial Committee (Beijing: Science Press).

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. doi: 10.1093/nar/gkl315

Talavera, G., Lukhtanov, V., Pierce, N. E., and Vila, R. (2021). DNA barcodes combined with multi-locus data of representative taxa can generate reliable higher-level phylogenies. *Syst. Biol.* syab038. doi: 10.1093/sysbio/syab038 [Epub ahead of print].

Tian, X., and Li, D. Z. (2002). Application of DNA sequences in plant phylogenetic study. *Acta Bot. Yunnanica* 24, 170–184.

Triplett, J. K., Oltrogge, K. A., and Clark, L. G. (2010). Phylogenetic relationships and natural hybridization among the North American woody bamboos

(Poaceae: Bambusoideae: Arundinaria). *Am. J. Bot.* 97, 471–492. doi: 10.3732/ajb.0900244

Vallée, G. C., Muñoz, D. S., and Sankoff, D. (2016). Economic importance, taxonomic representation and scientific priority as drivers of genome sequencing projects. *BMC Genomics.* 17, 782. doi: 10.1186/s12864-016-3100-9

Van de Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411–424. doi: 10.1038/nrg.2017.26

Vargas, O. M., Ortiz, E. M., and Simpson, B. B. (2017). Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: Diplostephium). *New Phytol.* 214, 1736–1750. doi: 10.1111/nph.14530

Walker, J. F., Walker-Hale, N., Vargas, O. M., Larson, D. A., and Stull, G. W. (2019). Characterizing gene tree conflict in plastome-inferred phylogenies. *PeerJ* 7:e7747. doi: 10.1101/512079

Wei, W., Youliang, Z., Li, C., Yuming, W., Zehong, Y., and Ruiwu, Y. (2005). PCR-RFLP analysis of cpDNA and mtDNA in the genus Houttuynia in some areas of China. *Hereditas* 142, 24–32. doi: 10.1111/j.1601-5223.2005.01704.x

Wen, J., Liu, J., Ge, S., Xiang, Q.-Y., and Zimmer, E. A. (2015). Phylogenomic approaches to deciphering the tree of life. *J. Syst. Evol.* 53, 369–370. doi: 10.1111/jse.12175

Wendel, J. F., Schnabel, A., and Seelanan, T. (1995). An unusual ribosomal DNA sequence from gossypium gossypioides reveals ancient, cryptic, intergenomic introgression. *Mol. Phylogenet. Evol.* 4, 298–313. doi: 10.1006/mpev.1995.1027

Wu, S., Han, B., and Jiao, Y. (2020). Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. *Mol. Plant.* 13, 59–71. doi: 10.1016/j.molp.2019.10.012

Xi, Z., Ruhfel, B. R., Schaefer, H., Amorim, A. M., Sugumaran, M., Wurdack, K. J., et al. (2012). Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proc. Natl. Acad. Sci. U S A.* 109, 17519–17524. doi: 10.1073/pnas.1205818109

Yang, Y.-Y., Qu, X.-J., Zhang, R., Stull, G. W., and Yi, T.-S. (2021). Plastid phylogenomic analyses of Fagales reveal signatures of conflict and ancient chloroplast capture. *Mol. Phylogenet. Evol.* 163:107232. doi: 10.1016/j.ympev.2021.107232

Yu, L., and Zhang, Y. P. (2006). Phylogenomics–an attractive avenue to reconstruct "tree of life". *Hereditas* 28, 1445–1450. doi: 10.1360/yc-006-1445

Zeng, L., Zhang, N., and Ma, H. (2014). Advances and challenges in resolving the angiosperm phylogeny. *Biodiversity Sci.* 22:21. doi: 10.3724/SP.J.1003.2014.13189

Zeng, L., Zhang, N., Zhang, Q., Endress, P. K., Huang, J., and Ma, H. (2017). Resolution of deep eudicot phylogeny and their temporal diversification using nuclear genes from transcriptomic and genomic datasets. *New Phytol.* 214, 1338–1354. doi: 10.1111/nph.14503

Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinf.* 19:153. doi: 10.1186/s12859-018-2129-y

Zhang, D., Gao, F., Jakovliæ, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096

Zhang, R., Wang, Y.-H., Jin, J.-J., Stull, G. W., Bruneau, A., Cardoso, D., et al. (2020). Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst. Biol.* 69, 613–622. doi: 10.1093/sysbio/syaa013

Zhang, S.-D., Jin, J.-J., Chen, S.-Y., Chase, M. W., Soltis, D. E., Li, H.-T., et al. (2017). Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol.* 214, 1355–1367. doi: 10.1111/nph.14461

Zhao, F., Chen, Y. P., Salmaki, Y., Drew, B. T., Wilson, T. C., Scheen, A. C., et al. (2021). An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biol.* 19:2. doi: 10.1186/s12915-020-00931-z

Zou, X. H., and Ge, S. (2008). Conflicting gene trees and phylogenomics. *J. Syst. Evol.* 46:795. doi: 10.3724/SP.J.1002.2008.08081