Characterizing gene tree conflict in plastome-inferred phylogenies

Joseph F Walker [Corresp., Equal first author, 1], **Nathanael Walker-Hale** [Equal first author, 2], **Oscar M Vargas** [3], **Drew A Larson** [4], **Gregory W Stull** [Corresp. 5]

[1] Sainsbury Laboratory (SLCU), University of Cambridge, Cambridge, United Kingdom

[2] Department of Plant Sciences, University of Cambridge, Cambridge, Cambridgeshire, United Kingdom

[3] University of California, Santa Cruz, Santa Cruz, United States

[4] University of Michigan - Ann Arbor, Ann Arbor, Michigan, United States

[5] Department of Botany, Smithsonian Institution, Washington, United States

Corresponding Authors: Joseph F Walker, Gregory W Stull
Email address: jfw52@cam.ac.uk, gwstull@gmail.com

Evolutionary relationships among plants have been inferred primarily using chloroplast data. To date, no study has comprehensively examined the plastome for gene tree conflict. Using a broad sampling of angiosperm plastomes, we characterize gene tree conflict among plastid genes at various time scales and explore correlates to conflict (e.g., evolutionary rate, gene length, molecule type). We uncover notable gene tree conflict against a backdrop of largely uninformative genes. We find alignment length and tree length are strong predictors of concordance, and that nucleotides outperform amino acids. Of the most commonly used markers, *matK,* greatly outperforms *rbcL*; however, the rarely used gene *rpoC2* is the top-performing gene in every analysis. We find that *rpoC2* reconstructs angiosperm phylogeny as well as the entire concatenated set of protein-coding chloroplast genes. Our results suggest that longer genes are superior for phylogeny reconstruction. The alleviation of some conflict through the use of nucleotides suggests that stochastic and systematic error is likely the root of most of the observed conflict, but further research on biological conflict within plastome is warranted given documented cases of heteroplasmic recombination. We suggest that researchers should filter genes for topological concordance when performing downstream comparative analyses on phylogenetic data, even when using chloroplast genomes.

Characterizing gene tree conflict in plastome-inferred phylogenies

Joseph F. Walker[1,5*], Nathanael Walker-Hale[2*], Oscar M. Vargas[3], Drew A. Larson[4] and Gregory W. Stull[5,6]

[1] Sainsbury Laboratory (SLCU), University of Cambridge, Cambridge, UK
[2]Department of Plant Sciences, University of Cambridge, Cambridge, Cambridgeshire, United Kingdom
[3]University of California, Santa Cruz, Santa Cruz, United States
[4]University of Michigan - Ann Arbor, Ann Arbor, Michigan, United States
[5] Department of Botany, Smithsonian Institution, Washington, DC 20013, USA

[6]Authors for correspondence: J. F. Walker (jfw52@cam.ac.uk) and G. W. Stull (gwstull@gmail.com)

*Equal contribution

ABSTRACT

Evolutionary relationships among plants have been inferred primarily using chloroplast data. To date, no study has comprehensively examined the plastome for gene tree conflict. Using a broad sampling of angiosperm plastomes, we characterize gene tree conflict among plastid genes at various time scales and explore correlates to conflict (e.g., evolutionary rate, gene length, molecule type). We uncover notable gene tree conflict against a backdrop of largely uninformative genes. We find alignment length and tree length are strong predictors of concordance, and that nucleotides outperform amino acids. Of the most commonly used markers, *matK,* greatly outperforms *rbcL*; however, the rarely used gene *rpoC2* is the top-performing gene in every analysis. We find that *rpoC2* reconstructs angiosperm phylogeny as well as the entire concatenated set of protein-coding chloroplast genes. Our results suggest that longer genes are superior for phylogeny reconstruction. The alleviation of some conflict through the use of nucleotides suggests that stochastic and systematic error is likely the root of most of the observed

conflict, but further research on biological conflict within plastome is warranted given documented cases of heteroplasmic recombination. We suggest that researchers should filter genes for topological concordance when performing downstream comparative analyses on phylogenetic data, even when using chloroplast genomes.

Key words: Angiosperms, chloroplast, gene tree conflict, *rpoC2*, phylogenomics, plastome, *matK*, *rbcL*.

INTRODUCTION

Chloroplast data have been the most prominent source of information for plant phylogenetics, largely due to the ease with which chloroplast genes can be sequenced, assembled, and analyzed (Palmer, 1985; Taberlet et al., 1991). The majority of broad-scale phylogenetic studies on plants have used chloroplast genes (e.g., Chase et al., 1993; Soltis et al., 2000, 2011), and the resulting phylogenies have been used for countless comparative studies examining ancestral states, historical biogeography, and other evolutionary patterns. While older studies relied mostly on targeted genes such as *rbcL* and *matK*, recent advances in DNA sequencing have drastically increased the ease and affordability of whole-chloroplast genome (i.e., plastome) sequencing (Moore et al., 2006; Cronn et al., 2008, 2012; Stull et al., 2013; Uribe-Convers et al., 2014), increasing the number of studies employing plastome-scale data for phylogenetic and comparative analyses (e.g., Jansen et al., 2007; Moore et al., 2007, 2010; Ruhfel et al., 2014; Stull et al., 2015; Gitzendanner et al., 2018; Hong-Tao et al., 2019). Nonetheless, the utility of plastid genes, as well as the entire plastome, is ultimately determined by the extent to which they reflect ‘true’ evolutionary relationships (i.e., the ‘species tree’) of the lineages in question (Doyle, 1992).

Phylogenomic conflict (i.e., the presence of conflicting relationships among gene trees in a genomic dataset) is now recognized as a nearly ubiquitous feature of nuclear phylogenomic studies (Smith et al., 2015). Gene tree conflict is commonly attributed to biological causes such as incomplete lineage sorting, hybridization, and gene duplication and loss (Maddison, 1997; Galtier and Daubin, 2008; Smith et al., 2015; Walker et al., 2017; Vargas et al., 2017). The genes within the plastome, however, are generally thought to be free of such biological sources of

conflict. This is because the plastome is typically uniparentally inherited (maternally in angiosperms, and at times paternally in conifers: Mogensen, 1996) and undergoes a unique form of recombination that is not expected to result in conflicting gene histories within the plastome (Palmer, 1983; Bendich, 2004; Walker et al. 2015).

In angiosperms, nonmaternal inheritance of the chloroplast has been reported (e.g., Smith, 1989; McCauley et al., 2007), and several surveys of pollen in flowering plants (Corriveau and Coleman, 1988; Zhang et al., 2003) have documented plastid DNA in up to 18% of the species examined, indicating potential for biparental inheritance and heteroplasmy (i.e., the presence of two or more different plastomes in a single organism, cell, or organelle). Heteroplasmy, which has been documented in multiple angiosperm species (e.g., Johnson and Palmer, 1988; Lee et al., 1988; Hansen et al., 2007; Carbonell-Caballero et al., 2015), creates an opportunity for heteroplasmic recombination, which could result in gene tree conflict in plastome-inferred phylogenies. Unlike animal organelles, where there is little evidence for heteroplasmic recombination (White et al. 2008), and thus a strong justification for analyzing the mitochondrial genome as a single unit (Richards et al. 2018), heteroplasmic recombination (both intra- and interspecific) has been invoked by multiple studies to explain discordance in various clades (Huang et al., 2001; Marshall et al., 2001; Erixxon and Oxelman, 2008; Bouillé et al., 2011). However, only a few recent studies have documented clear evidence of this phenomenon (Sullivan et al., 2017; Sancho et al., 2018). Beyond heteroplasmy, sharing of genes between the chloroplast and nuclear or mitochondrial genomes remains another potential source of biological conflict (Martin et al., 1998; Martin, 2003; Stegemann et al., 2003; Rice et al., 2013; Straub et al., 2013; Smith, 2014), although transfer back to the plastome would be required to generate conflict and this has rarely been found (Smith, 2014). Even though biological conflict in the plastome generally seems rare, the full extent of intra-plastome conflict has yet to be evaluated. Quantifying its extent is of high importance given that most studies assume no conflict and treat the plastome as a single locus (Wolfe and Randle, 2004).

Aside from biological sources of conflict, there also remain significant potential sources of systematic and stochastic error that could contribute to conflict across the plastome (e.g., Burleigh and Mathews, 2007a,b). Chloroplast data are used at various time scales, and the accumulation of substitutions over long periods of evolutionary time increases the probability of encountering systematic error due to saturation (Rodríguez-Ezpeleta et al., 2007; Philippe et al.,

2011). Conflict has been demonstrated among different functional groups of genes (Liu et al., 2012), among different regions of the plastome (Davis et al., 2013; Walker et al., 2014), as well as among individual genes (e.g., Shepherd et al., 2008, Foster et al. 2018; Gonçalves et al. 2019). The rate of chloroplast evolution as a whole has been examined (and compared with the nuclear and mitochondrial genomes; Wolfe, 1987), and rate variation within the chloroplast—especially across the three major regions of the genome, i.e., the large single-copy (LSC) region, the small single-copy (SSC) region, and the inverted repeats (IRa, IRb)—has been explored to help determine the markers useful for phylogenetic inference at different time scales (e.g., Graham and Olmstead, 2000; Shaw et al., 2005, 2007, 2014). However, no study has comprehensively examined gene tree conflict within the plastome to better characterize the extent and sources of conflict, and to identify the plastid genes most concordant with our current understanding of angiosperm phylogeny (e.g., Soltis et al., 2011; Wickett et al., 2014; Zeng et al., 2014; Gitzendanner et al., 2018; but see Logacheva et al., 2007 for a preliminary investigation of the concordance of individual plastid genes with angiosperm phylogeny).

Here we use phylogenomic tools to characterize the extent of conflict among plastid genes as a function of evolutionary rate, rate variation among species, sequence length, and data type (i.e., nucleotides vs. amino acids) at varying time scales across angiosperms. Our results show that plastid genes, when analyzed in combination, perform agree with many community consensus relationships; however, across plastid genes there is notable gene tree conflict. The number of conflicting genes at each node is often comparable to the number of concordant genes; however, the majority of plastid genes are uninformative for most nodes when considering support and thus unlikely to positively mislead studies. Information content (gene length and molecule type, i.e., nucleotides vs amino acids) and evolutionary rate were positive predictors of concordance, suggesting that most observed gene-tree conflict is a consequence of spurious inferences from insufficient information. However, some nodes across angiosperm phylogeny show numerous strongly supported conflicting genes; this rare event indicates an area for future exploration. This may result from systematic error or could be the product of a biological source of gene-tree conflict (e.g., heteroplasmic recombination, horizontal gene transfer). We also document the performance of individual genes at recapitulating angiosperm phylogeny, finding the seldom-used gene *rpoC2* to outperform commonly used genes (e.g., *rbcL*, *matK*) in all cases, consistent with previous work highlighting the utility of this gene (Logacheva

et al., 2007). Although as a whole the chloroplast infers a topology similar to currently hypothesized angiosperm relationships, the conflict underlying it can influence branch lengths, and therefore dating analyses, comparative methods, and other work that relies on branch lengths could benefit from gene filtering procedures. Our results provide insight into the extent and sources of intraplastome conflict.

MATERIALS AND METHODS

***Data acquisition and sampling*—**Complete plastome coding data (both nucleotide and amino acid) were downloaded from NCBI for 53 taxa: 51 angiosperm ingroups and two gymnosperm outgroups (*Ginkgo Biloba* and *Podocarpus lambertii*; Table S1). Our sampling scheme was designed to capture all major angiosperm lineages (e.g., Soltis et al., 2011) thus allowing accurate inference of the angiosperm phylogeny, while also including denser sampling for nested clades in Asterales. This allowed us to evaluate the extent of gene tree conflict at different evolutionary levels/time scales, from species-level relationships in *Diplostephium* (Asteraceae) to the ordinal-level relationships defining the backbone of angiosperm phylogeny.

***Data preparation, alignment, and phylogenetic inference*—**All scripts and data used may be found on GitHub (https://github.com/jfwalker/ChloroplastPhylogenomics). Orthology was determined based upon the annotations of protein-coding genes on Genbank; this resulted in near-complete gene occupancy apart from instances of gene loss or reported pseudogenization. We did not use non-coding regions, as proper orthology at deep time scales can be difficult to assess due to gene re-arrangements and inversions; this is in part why most deep phylogenetic analyses of angiosperms using plastomes have excluded non-coding data (e.g., Moore et al., 2010; Ruhfel et al., 2014).

The amino acid and nucleotide data were aligned using Fast Statistical Alignment (FSA v1.15.9; Bradley et al., 2009) with the default settings for peptide and the setting "--noanchored" for nucleotide. FSA has been shown to be one of the top-performing alignment programs (Bradley et al., 2009; Redelings, 2014), and does not rely upon a guide tree for sequence alignment, alleviating downstream bias (Bradley et al., 2009; Boyce et al., 2015; Chatzou et al., 2018). A maximum likelihood (ML) tree was then inferred for each gene using RAxMLv.8.2.4 (Stamatakis, 2014), with the PROTGAMMAAUTO and GTR+G models of evolution used for

158 the amino acid and nucleotide data, respectively. For each dataset, we conducted 200 rapid
159 bootstrap replicates. The alignments were also concatenated into supermatrices and partitioned
160 by gene using the phyx (v0.99) program pxcat (Brown et al*.,* 2017).
161 The nucleotide and amino acid supermatrices were then each used to infer concatenated
162 'plastome' trees using the GTR+G and the PROTGAMMAAUTO models, respectively, as
163 implemented in RAxML. To complement the model inference performed by RAxML from the
164 AUTO feature, we also used IQ-TREE's (v1.5.4) (Nguyen et al., 2014) built-in model selection
165 process (Kalyaanamoorthy et al., 2017) on the partitioned data. We did not partition by codon
166 position because the lengths of most of the plastid genes, especially when divided into three
167 smaller partitions, would have been insufficient to inform an evolutionary model.
168 Recent work has advocated the use of multispecies coalescent methods when analyzing sets of
169 gene trees from the chloroplast (Gonçalves et al., 2019). As such we reanalyzed the nucleotide
170 data using the Maximum Quartets Support Species Tree method as implemented in Astral.v5.6.3
171 (Zhang et al., 2018) with default settings.
172
173
174 ***Reference phylogenies*—**For the conflict analyses, described below, we created several reference
175 trees against which the gene trees were mapped. Primarily, we used a topology based on Soltis et
176 al. (2011) and where possible that of Wickett et al. (2014), and—for species-level relationships
177 within Asteraceae—Vargas et al. (2017). We refer to this as our 'accepted tree', or AT (Fig 1.).
178 We also created a reference tree based on the recent plastome phylogeny of Gitzendanner et al.
179 (2018), which we refer to as the 'Gitzendanner tree', or GT. The GT is highly concordant with
180 the AT, but several deep (and commonly contentious) nodes differ, e.g., the placement of
181 Buxales and ordinal relationships in lamiids (here, we only sampled Lamiales, Gentianales, and
182 Solanales). In the latter case, all three possible combinations of lamiids taxa were assessed
183 (discussed further below in Assessment of Supported Conflict). Although it is difficult to
184 determine which of these trees better represents species relationships of angiosperms, they both
185 serve as useful frameworks for revealing conflict among genes. Here, we primarily focus on
186 analyses using the AT, but conflict analyses using the GT were also conducted to ensure that our
187 overall results were consistent across reference trees.
188

***Analysis of conflict*—**All gene trees were rooted on the outgroups using the phyx program pxrr (Brown et al., 2017), in a ranked fashion ("-r") in the order *Podocarpus lambertii* then *Ginkgo biloba*; this was performed to account for any missing genes in either of the taxa. In all cases the trees were rooted with at least one gymnosperm. Gene tree conflict in the data was identified using the bipartition method as implemented in phyparts v.0.0.1 (Smith et al. 2015), with the gene trees from each data set (amino acid and nucleotide) mapped against the AT and GT references trees. The concordance analyses were performed using both a support cutoff (at ≥70% bootstrap support, i.e., moderate support as suggested by Hillis and Bull, 1993) and no support cutoff. When the support cutoff is used, any gene tree node with less than 70% bootstrap support is regarded as uninformative for the reference node in question (i.e., it is uncertain whether the gene tree node is in conflict or concordance with the reference node); when no support cutoff is used, the gene tree node is evaluated as conflicting or concordant regardless of the support value. However, in both cases, a gene is considered uninformative if a taxon relevant to a particular node/relationship is missing from the gene dataset. For example, consider the reference tree ((A,B),C); if a gene tree is missing taxon B, but contains the relationship (A,C), then the gene tree is concordant with the node ((A,B),C) but uninformative regarding the node (A,B). See Smith et al. (2015) for an in-depth explanation of this method. To examine patterns of concordance and conflict among all gene trees, we calculated the pairwise Robison-Foulds distance (Robinson and Foulds, 1981) (RF) between each gene tree and the AT, concatenated topology, and ASTRAL tree (see below). We created a network of trees based on RF distance using the R package igraph v1.2.4.1 in R v3.6.1 (Csardi and Nepusz, 2006; R Core Team, 2018), using trees as nodes, inverse RF distance as edge weights, and the Fruchterman-Reingold algorithm to lay out the graph (Fruchterman and Reingold, 1991).

Levels of conflict were examined across the entire tree and within particular time intervals of the evolutionary history of angiosperms. For the latter case, the ~150 Ma of crown angiosperm evolution were divided into five time-intervals (30 Ma each): 150 to 121 mya, 120 to 91 mya, 90 to 61 mya, 60 to 31 mya, and 30 to 0 mya. The nodes across the tree were binned into the time intervals based on their inferred ages (ages > 20 Ma from Magallón *et al*., 2015; ages < 20 Ma from Vargas et al., 2017 and Roquet et al., 2009). At each time interval, the proportion of concordant nodes for each gene was calculated (for all the nodes falling within that time interval). This allowed us to assess the level(s) of divergence at which each gene is most

220 informative. We also determined concordance levels of the 'plastome' trees (against the
221 reference) within each of these time intervals.
222
223 ***Assessments of Individual Plastid Genes***—Using the phyx program pxlstr (Brown et al., 2017),
224 we calculated summary statistics for each gene alignment and corresponding gene tree: number
225 of included species (as gene alignments could be missing a species from pseudogenization or
226 gene loss), alignment length, tree length (a measure of gene evolutionary rate), and root-to-tip
227 variance (a measure of rate variation across the phylogeny). Alignment length and tree length
228 represent complementary measures of a gene's information content. All else being equal, longer
229 genes should harbor more phylogenetic signal. Tree length serves as a model-informed metric
230 because all sites, and not just those which are variable or parsimony informative, are used in
231 estimation of model distances. Because our study uses maximum likelihood methods, we chose
232 this model-informed metric as opposed to aforementioned metrics. Levels of concordance of
233 each gene tree with the reference trees were then assessed by tabulating the number of nodes
234 concordant between the gene tree and the given reference tree (e.g., the AT). The number of
235 concordant nodes (in **Fig. 2** treated as a proportion of total nodes available to support and in **Fig.**
236 **3** based on total nodes) was used as a measure of a gene's ability to accurately reconstruct the
237 angiosperm phylogeny.
238
239 ***Predictors of Concordance—***We examined the statistical relationships between gene tree
240 concordance (using concordance data from the AT-based conflict analysis) and alignment length,
241 tree length, and root-to-tip variance. We considered each node as a trial and analyzed the
242 concordance (success) or conflict (failure) in aggregate, such that each gene was a single
243 observation with x successes in n trials. Because of variable sampling and loss of some genes, n
244 was variable between observations. We analyzed the relationship between probability of
245 concordance and alignment length, tree length, and root-to-tip variance using logistic regression
246 of aggregate binomial trials with the function glm() in R v3.4.4 (R Core Team, 2018). Binomial
247 models were generally characterized by high residual deviance, and we thus allowed for
248 overdispersion by fitting quasibinomial logistic regressions (using 'family = quasibinomial()' in
249 R v3.4.4). We fitted models for both amino acid and nucleotide data, both not considering and

considering support, and for both alignment methods. We also fitted models for amino acid data inferred only under WAG.

We modeled gene performance as a function of length, tree length, and root-to-tip variation, and as a function of each predictor individually. Because it is possible that apparent relationships between alignment length and concordance may reflect signal from gene information content per alignment site, we also modelled gene performance as a function of length and tree length (as a proxy of gene information content and rate, respectively, noted above), to assess the relationship between alignment length and gene performance after controlling for variation associated with rate.

Investigation of model fits on full datasets by studentized residuals, leverage and Cook's distance values revealed that several observations were highly influential. Therefore, we also conducted analyses on reduced datasets to investigate the influence of these observations. In amino acid datasets, we excluded *rpl22* and *rpl32*, which were probably influential due to their high tree length values, and *ycf1* and *ycf2*, which were probably influential based on their long alignments. In nucleotide datasets, we excluded *clpP* which had poor performance relative to tree and alignment length, *rps15*, which had relatively long tree length, and *ycf1* and *ycf2* for the same reasons as previously. Combined analyses of alignment length and tree length were not subject to influence driven by high root-to-tip variance values, and hence reduced datasets had fewer genes removed. In this case, we excluded only *ycf1* and *ycf2*. Regression results were summarized in tables using the R package Stargazer (Hlavac, 2018).

***Saturation Analyses*—**We also performed saturation analyses on all the chloroplast genes to determine if they were capable of inferring deep divergences (Phillipe et al, 1994). In this analysis the pairwise distances of the sequences are compared to the model-corrected distances as a means of identifying if the real genetic distances may be properly estimated. The saturation analysis was performed on each gene and each codon position and thus the data were realigned by codons for this analysis. The amino acid alignments were used to guide the codon alignments using the program pxaa2cdn, from the phyx package. Four of the genes did not match their corresponding amino acid sequence in its entirety and these were left out of this analysis, since the codons could not be properly aligned. This discrepancy is likely due to errors in the GenBank submission (the amino acid and nucleotide data do not perfectly correspond). Saturation was

assessed by determining the observed number of differences between sequences compared to the inferred number of substitutions. This analysis was performed using the "dist.dna" and "dist.corrected" functions in the R package "ape" (Paradis et al. 2018), with the F84 model of evolution (Felsenstein and Churchill, 1996) used for the correction.

***Comparison of genomic regions*—**We assessed the utility of the three major plastome regions—the Large Single Copy (LSC) region, Small Single Copy (SSC), and the Inverted Repeat (IR) region—for reconstructing angiosperm phylogeny in two ways. First, we constructed ML phylogenies by concatenating all genes found in each region (as described above for each of the plastome trees inferred using all genes). We then calculated the number of concordant nodes with the AT for the resulting genomic regions trees. Second, using the concordance levels of each individual gene (described above), we created a plastome diagram (with genes arranged according to their genomic position) showing the concordance levels of each gene at the five different time scales discussed above (Fig. S1); this permits a qualitative visual assessment of the general concordance levels of each genomic region at each time slice.

***Assessment of supported conflict*—**We compared four contentious regions of the AT to a series of alternative relationships. The tested alternative relationships were present in either the GT (i.e., the plastome phylogeny of Gitzendanner et al., 2018), our plastome phylogenies (with ≥70 BS for the node in question; we refer to this tree below as the 'plastome tree', or PT), and/or three or more of the gene trees from our nucleotide dataset (with ≥70 BS for the node in question). These alternative relationships pertain to: (1) the placements of *Buxus* and *Trochodendron*; (2) relationships among *Acnistus* (Solanales), *Coffea* (Gentianales), and *Olea* (Lamiales); (3) the placements of *Jacobaea* and *Artemisia* in Asteraceae; and (4) the placements of *Galinsoga*, *Ageratina*, and *Parthenium* in Asteraceae. To test these alternative relationships, we used a modified version of the Maxmimum Gene Wise Edge (MGWE) method (Walker et al., 2018), where instead of a defined "TREE SET", we used a constraint tree for each alternative hypothesis. The likelihood for every gene was calculated across each constraint, thereby forcing the relationship in question to be the same. Similar to MGWE, this allowed for conflict across the rest of the topology; however, this provides a greater amount of tree space to be explored

(similar to Smith et al., 2018a). This method creates likelihood scores solely based on each alternative relationship, making it robust to gene tree conflict outside of the node in question. The modification of the MGWE method has been implemented in the EdgeTest.py program of the package PHylogenetic Analysis Into Lineages (https://github.com/jfwalker/PHAIL).

Assessment of influence of alignment method—Because initial homology statements at the alignment stage can strongly influence the outcome of phylogenetic inference (Morrison, 2009; Morrison et al*.,* 2015), we repeated the conflict analysis using the same parameters as above with two alternate sets of alignments. In one case we re-inferred using MAFFTv2.271 (Katoh and Standley, 2013) with the setting "—auto —max_iterate 1000". In the second case we cleaned the original FSA gene alignments using pxclsq (Brown *et al.* 2017) for a minimum of 30% column occupancy in order to remove gap-heavy and ambiguously aligned sites, and RAxMLv8 was used.

Assessment of influence of model— To test if the inferred model of evolution influenced the inference, we re-ran the tree inference using the same methods above on the uncleaned FSA amino acid datasets using only the WAG model of evolution (Wheelan and Goldman, 2001). The WAG model was not found to be the optimal amino acid model for any gene in the analysis, thus this allowed us to assess the influence of assigning the same sub-optimal model across all partitions.

RESULTS

Influence of alignment and model— The use of MAFFT as opposed to FSA as an alignment tool did not produce a qualitatively different result (Tables S2-7). Furthermore, our results were largely not qualitatively influenced by using an alignment that had been cleaned for 30% minimum occupancy (Tables S2-6 and Tables S8-9). Quasibinomial regression results for amino acids aligned with FSA and cleaned for 30% minimum occupancy were not significant for all predictors, but this model had an abnormally high estimated dispersion parameter (almost 17 times greater than the next largest estimate), suggesting that p-values inferred under this model were unreliable (Tables S2-5). When running a sub-optimal model of evolution for all genes (the

WAG model) we recapitulated the results of our model-tested analysis (Tables S2-5 and Table S10). The differences among all these analyses were occasionally off by one in terms of number of concordant nodes; however, no results qualitatively changed: ranking of genes in terms of concordance, inference from regression models and detection of strongly-supported conflict were largely unchanged. Therefore, for the remainder of the results and discussion we reference the uncleaned FSA alignment (Tables S11-12).

Patterns of conflicting chloroplast signal—The sampling for this experiment allowed us to examine conflict at multiple timescales (the values for each analysis may be found in Tables S6-12), with roughly 10 divergences occurring in each 30 Ma bin. Our 'plastome' trees, inferred from the concatenated gene sets (nucleotide and amino acid), were highly concordant with the AT (Fig. 1 and Fig. S2-3). Without considering support, the gene trees showed high levels of conflict across the different analyses. This pattern is similar to that seen in nuclear phylogenomic datasets (Salichos and Rokas, 2013; Smith et al., 2015; Shen et al., 2016). When considering support, the majority of the plastid genes were uninformative for practically all nodes in the phylogeny (Fig. 1; Tables S11-12); i.e., they had bootstrap support below 70% (moderate support) for that particular relationship (whether in conflict or concordance).

There was no obvious relationship between the amount of gene tree conflict and evolutionary scale (i.e., conflict was relatively evenly distributed across shallow and deeper nodes/time scales; Figs. 1, 2). Although the greatest degree of gene tree concordance with the AT appeared in the nodes with inferred ages between 90–61 Ma (ages based on Magallón et al., 2015), these nodes typically still contained at least 50% uninformative gene trees (Fig. 1). Genome location also showed weak correspondence to conflict/concordance (Fig. S1; Table S13); discussed further below. Rather than timescale and genome location, data type (nucleotide vs amino acid) had a much greater impact on the prevalence of conflict (Figs. 2, 3), with the amino acid dataset generally showing higher levels of gene tree conflict. When factoring in support (BS $\geq$70% cutoff), the amino acid data set showed even less concordance with the AT (as more genes were considered uninformative due to low BS support). The integration of support also decreased concordance of the nucleotide data set with the AT, but proportionally less. See Fig. S2-3 for conflict analyses showing the amino acid and nucleotide gene trees mapped onto the AT (the nucleotide results from Fig. S2 are also shown in Fig. 1).

To identify how many different models of amino acid evolution underlie the genes in the plastome, we tested each gene against the candidate set of amino acid models in IQ-TREE and RAxML. We found that a wide range of evolutionary models best fit the individual genes—rather than just a single model that best fits every gene (Table S12). Many of the models were not designed specifically for plastome data, and cpRev, which was designed for plastome data, was only the best fit for 19 of the 79 genes based upon the IQ-TREE model test.

To examine relationships between gene characteristics and levels of concordance/conflict, we calculated the following statistics for each gene: alignment length (a measure of gene information content), tree length (a measure of evolutionary rate), and root-to-tip variance (a measure of variation in evolutionary rate across the tree), all of which can influence properties of phylogenetic inference (Shen et al. 2016). This information is presented for each gene across both data types in the supplementary information (Tables S11-12). We used logistic multiple regression to test relationships between concordance and gene characteristics (Tables S14–21). We found that alignment length had a significant positive multiplicative relationship with odds of concordance across both datasets (Fig. 2, Tables S14-17). Tree length and root-to-tip variance had significant positive and negative multiplicative relationships, respectively, in both datasets (Tables S14-15). Notably, excluding highly influential observations did not affect the direction or significance of most relationships (Tables S16-17), indicating that the results were not due to the influence of these data points. However, this did increase the estimated impact of alignment length on log odds by a factor of 10, probably due to the high leverage exerted by the very long alignments of *ycf1* and *ycf2*. In models including only alignment length and tree length, both predictors had significant positive multiplicative relationships with odds of concordance even when the other was included in the model (Fig. 2, Tables S18-21).

We examined patterns of conflict among individual genes by calculating the pairwise RF distance and producing a network graph (Fig. S4). Most genes had large pairwise RF distances to all other genes. We observed a general cluster of genes that were less distant from one another and from the AT and inferred species trees, including *rpoC2* and *ndhF*, that appeared to correlate with alignment length (Fig. S1). In general, the tree set is characterized by high levels of discordance among gene trees as well as between gene trees and the AT and inferred species trees.

Genomic patterns of concordance/conflict—In terms of number of nodes concordant with the AT, the LSC and SSC regions were roughly comparable, with both outperforming the IR regardless of whether BS support was considered (Table S13). This pattern held across time periods, with the LSC and SSC regions having more concordant nodes than the IR at every time slice. The tree lengths of the LSC and SSC regions (1.82 and 2.05, respectively) were also considerably larger than that of the IR (0.98). The concatenated gene alignment lengths of the LSC , SSC, and IR were 73,422 bp, 10,395 bp, and 19,314 bp, respectively. The genome diagram of concordance (Fig. S1) does not show any striking patterns among the different genomic regions; however, it is notable that the majority of the LSC region is discordant (or uninformative) except for a few highly informative genes (namely, *rpoC2* and *matK*).

Performance of individual plastid genes—Across all analyses, *rpoC2* showed the highest levels of concordance with the AT (Fig. 2, Tables S11-12); in general, when support was not considered, it performed at least as well as the 79 concatenated genes in reconstructing the AT (Fig. 3). The commonly used genes *ndhF* and *matK* generally scored among the best-performing plastid genes (in terms of number of concordant nodes), while *rbcL*, the other most commonly used gene, performing relatively poorly (Fig. 2). The *matK* alignment is ~250 bp longer than the *rbcL* alignment; the best performing gene, *rpoC2* (alignment length 4,660 bp), is one of the longest plastid genes. However, the notably long region *ycf1*—the unaligned sequence of which is ~5,400 bp (Dong et al., 2015)—did not perform as well as *rpoC2*. In this study, the alignment length of *ycf1* was 21,696 bp (vs. 4,660 bp for *rpoC2*). In several cases it performed toward the top; however, it was never the top-performing gene in terms of number of concordant nodes (Tables S11-12). Despite the high levels of observed conflict overall, we found that every node of the AT was supported by at least one gene. Thus, to varying degrees, all relationships of the AT are found within the plastome gene tree set.

Saturation analyses—None of the genes analyzed showed significant signatures of saturation (Fig. S5). The uncorrected genetic distances and the corrected genetic distances appeared to be roughly the same, and this trend was true for the first, second, and third codon position. Four genes (*ndhD*, *psbL*, *rpl16,* and *rpl2*) were unable to be properly analyzed for saturation as the

435 nucleotide data uploaded to GenBank were missing the codons for 2 to 3 amino acids in some
436 taxa resulting in an improper codon alignment.
437
438 ***Analysis of well-supported conflict—***We found that, out of the four major contentious
439 relationships in the AT, the edge-based analysis only supports one (Table 1) compared to the
440 alternative hypotheses examined (i.e., the GT, the plastome trees, and the gene trees). In the case
441 of *Buxus* (node 1 in Fig. 1), this analysis supports *Buxus* as sister to Trochodendraceae (Table 1),
442 which is the topology found in the GT and four gene trees: *rbcL*, *petL*, *rps2*, *rps14* (Table 2). In
443 the case of lamiids relationships (node 2), the data support *Olea* (Lamiales) sister to *Acnistus*
444 (Solanales) + *Coffea* (Gentianales), which is the relationship present in the GT and 15 gene trees
445 (Tables 1 and 2). For node 3 (the placements of *Jacobaea* and *Artemisa*), the data support the AT
446 topology, yet no individual gene trees provided ≥70 BS support for this relationship; however,
447 *rpoC2* matched the GT topology with strong support (Tables 1 and 2). Lastly, for node 4 (the
448 placement of *Galinsoga*), the data collectively support *Parthenium argentatum* sister to
449 *Galinsoga* (our plastome tree, PT), with two gene trees (*petA*, *psaJ*) providing strong support for
450 this topology; however, 16 genes (Table 2) supported a dominant alternative (*Ageratina* sister to
451 *Galinsoga quadriradiata*).
452
453 **DISCUSSION**
454 Following the typical assumptions of chloroplast inheritance, we would expect all genes in the
455 plastomes to share the same evolutionary history. We would also expect all plastid genes to show
456 similar patterns of conflict when compared to non-plastid inferred phylogenies. Furthermore, the
457 amino acid and nucleotide plastid data we used should show the same conflicting/concordant
458 relationships against given reference trees (e.g., the AT or GT). Our results, however, discussed
459 below, frequently conflict with these common assumptions about chloroplast inheritance and
460 evolutionary history.
461
462 ***Conflicting topologies inferred from the chloroplast genome—***In general, the 'plastome'
463 topologies inferred from nucleotide and amino acid alignments showed high levels of
464 concordance with the AT (Fig. 1; Tables S11-12). While the genes within the chloroplast
465 genome are largely uninformative for most nodes of the phylogeny, a number of genes exhibited

well-supported conflict (Fig. 1). In general, there appears to be no relationship between evolutionary scale and amount of gene tree conflict (i.e., conflict generally does not appear to correlate with divergence time in angiosperms). Instead, the extent of conflict/concordance had a stronger relationship with the molecule type analyzed. The amino-acid dataset showed the highest levels of gene tree conflict (Fig. 3, Table S12), and the nucleotide dataset had about half the amount of gene tree conflict found in the amino acid data (Fig. 3, Table S11). Support from edge-based analyses (a topology test that is robust to gene tree conflict at nodes other than the one under consideration) for the GT over the AT suggest that at least some signal for conflicting relationships is derived from underlying gene tree conflict when resolving species relationships using all genes. Positive relationships between alignment length, tree length and concordance also suggest that stochastic error is at play, because shorter or slower-evolving genes harbor less signal to accurately infer relationships. Pairwise comparisons between chloroplast genes indicated that they did not appear to share the same evolutionary history (Fig. S4). Even longer genes were relatively discordant with one another (for example *rpoC2* and *ycf1* had an RF distance of 27). Against a backdrop of largely uninformative genes, long, influential sequences might drive the results of concatenated analyses (Shen et al., 2016). Indeed, *rpoC2* and *ndhF* had small distances to the AT, concatenated topology, and to one another. Because the AT is a synthetic topology based on analyses of all three plant genomes, we do not consider this to be a potential bias (i.e. the AT reflecting the signal of these genes), but instead indicative of our dataset converging on the signal of the true relationships. In our analysis, however, longer genes still had largely discordant trees.

The superior performance of (coding) nucleotide data compared to amino acid data possibly stems from the relatively greater information content of nucleotides (i.e., longer alignments). Assuming there is not a significant amount of missing/indel data (with the exception of *ycf1* and *2*), longer alignments should result in better-informed models of evolution, all else being equal, aided by both parsimony-informative and -uninformative characters (Yang, 1998). However, inherent differences in amino acid and nucleotide models might also explain differences in performance. The nucleotide data were analyzed under the GTR model, where substitution rates among bases are individually estimated; in contrast, for the empirical amino acid substitution models investigated here, substitution rates are pre-estimated, as the number of estimated parameters for changes among 20 states is extremely large. Additionally, although the

plastome has been treated as a single molecule for designing amino acid models of evolution (Adachi et al*.,* 2000), a wide variety of amino acid models (some of which were designed for viruses, such as flu or HIV) were inferred to be the best for different plastid genes (Table S12). This might be the result of the different methods implemented in RAxML vs. IQ-TREE for model testing, the different available models (the best-fitting model for some genes in the RAxML analysis was the unpublished matrix DUMMY2), or the lack of sufficient information (because of gene length) to inform the model. However, the most important point is that, based on the amount of information present in each gene, the chloroplast is inferred to evolve under significantly different models of evolution. Given the highly pectinate structure of the AT, phylogenetic inference in this case should rely heavily on the model for the likelihood calculations. While in some cases (e.g., the shortest plastid genes) limited amounts genetic information might be inherently insufficient, in others, improvements in amino acid modeling might lead to great improvements in phylogenetic inference; such has been suggested for animal mitochondrial data (Richards et al., 2018). At least insofar as the best-fitting model adequately models the dynamics of sequence evolution, conflict would be expected to decrease relative to a sub-optimal model if systematic error (through inadequate modeling) was the main driver of our results. However, we found that our results were recapitulated when analyzed using only the WAG model, suggesting that this was not the case. These analyses are not exhaustive, because they can only testify to the impact of systematic error resulting from the variation among the model set considered here and do not consider the adequacy of these models. Even though two genes analyzed under two different models produce the same tree, there could still be common assumptions underlying both models that are violated by the dataset, leading to systematic error. Additionally, two genes may share the best-fitting model and produce the same tree, but that model may be misspecified, leading to error. Although not the focus of our study, future examinations of chloroplast phylogenomics could benefit from tests of model adequacy (Brown, 2014; Duchêne et al., 2018; Chen et al., 2019).

We expect that at deeper time scales, nucleotides (of coding regions) may begin to experience saturation and thus information loss due to increased noise, at which point amino acids (with 20 states) would begin to outperform nucleotides. However, the time scale of angiosperm evolution does not appear great enough to result in nucleotide saturation (at least for the genes sampled here, given that no genes appeared to exhibit significant levels of saturation;

Fig. S5), indicating that nucleotides outperform amino acids for phylogenetic analysis of plastomes across angiosperms. Future work, with a broader plastome sampling across green plants, will be necessary to determine the evolutionary scale at which amino acids become more informative than nucleotides for phylogenetic inference.

Taxon sampling may also affect the accuracy of phylogenetic inference, and it is possible that some of the incongruence observed here is due to the influence of missing taxa (Zwickl and Hillis, 2002; Heath, Hedtke and Hillis, 2008). There is debate about the relative impact of more sequences versus more taxa (Rosenberg and Kumar, 2001; Zwickl and Hillis, 2002; Rokas and Carroll, 2005). However, our dataset sampled most major angiosperm lineages, and the total analysis was highly concordant with the AT and the GT. Furthermore, patterns of conflict in this dataset reflect those detected in more targeted sampling efforts (Foster et al*.,* 2018). Taken together, this suggests that taxon sampling did not play a large role in our results.

Analyses of well-supported conflict—The use of a reference topology (AT) afforded us the ability to examine how individual genes within the plastomes agree/conflict with our current, generally accepted hypothesis of angiosperm phylogeny. However, several contentious areas of angiosperm phylogeny remain unresolved, with different topologies recovered across different analyses (e.g., the AT, GT, and PT show small differences in several parts of the tree). Thus, we used edge-based analyses to compare these alternative topologies.

The placement of *Buxus* (Buxales) is a salient example, with different analyses resolving different placements. Our edge-based analyses found *Buxus* sister to *Trochodendron*/Trochdendrales (i.e., the GT topology) to have the highest likelihood score (Table 1), and this placement was recovered in four plastid genes (*rbcL*, *petL*, *rps2*, *rps14*) with strong support. However, five plastid genes recovered, with strong support, *Buxus* and *Trochodendron* successively sister to the core eudicots (i.e., the AT topology): *ndhD*, *rps4*, *matK*, *psbB*, and *rpoC2*. Relationships among core lamiid orders are another intriguing example of gene tree conflict. Core lamiid relationships have been notoriously problematic (Refulio-Rodriguez and Olmstead, 2014; Stull et al., 2015), and previous phylogenetic studies have recovered every possible combination of the constituent orders (Boraginales, Gentianales, Lamiales, Solanales). While our sampling does not include Boraginaceae, we tested alternative relationships among the three remaining orders. The relationship

559 (Lamiales(Gentianales,Solanales)), which is present in the GT, received the highest likelihood
560 support and is found in 15 gene trees (Tables 1 and 2). However, the alternatives were each
561 recovered with strong support by two plastid genes: *psaA* and *atpH* (Gentianales sister to the
562 rest), and *psbB* and *petA* (Solanales sister to the rest). Asteraceae also exhibited several examples
563 of strongly supported conflict (Tables 1 and 2), but sampling differences across the trees
564 examined make comparisons more difficult.
565 It is difficult to determine the causes of the observed instances of strongly supported
566 conflict, which can be found at most nodes in the phylogeny (Fig. 1, Tables S11-12). Based on
567 the present analyses, we cannot conclude whether this conflict is biological (e.g., genes with
568 different evolutionary histories in a single genome) or a consequence of systematic, stochastic, or
569 other analytical errors (e.g., modeling error, lack of information or misalignment). We have
570 endeavored to control for these factors. For example, it is possible that alignment errors could be
571 responsible for instances of strongly supported conflict (Richards et al., 2018). However,
572 amongst genes with strongly supported conflict, no obvious misalignment was present, and
573 recapitulation of the results when testing multiple alignment methods and a cleaning procedure
574 suggests misalignment, if present, does not explain the patterns observed here. Unlike in that
575 study (Richards et al. 2018) supported by a lack of evidence for recombination in animal
576 mitochondria (White et a. 2008), there are reasons to believe that biological gene tree conflict
577 could occur in the chloroplast genome (Sancho et al., 2018), but we stress that we make no
578 attempt to fully characterize this. In parallel to our work, Gonçalves et al. (2019) have analyzed
579 the prevalence of conflict across the plastome and have advocated for the use of both
580 concatenation and multispecies coalescent methods (MSC) for inferring full plastome
581 relationships. Our own analysis of the data with the summary coalescent method ASTRAL
582 (Zhang et al., 2018), demonstrates that the different methods result in different relationships (Fig.
583 1). Specifically, we found much of the conflict co-occurs with conflict from the full plastome
584 supermatrices, with the exception of the generic level resolutions of *Diplostephium*. We note that
585 a core assumption of the MSC is the lack of linkage amongst loci, likely violated by the
586 chloroplast in all but a small number of cases (Edwards et al., 2016; Gonçalves et al., 2019). See
587 Gonçalves et al. (2019) for an overview of coalescent methods applied to plastome data. With
588 our data, no analyses we performed fully explained the level of conflict observed and we suggest
589 that the possibility of biological conflict deserves further exploration, especially given the recent

590 documented instances of inter-plastome recombination (Sullivan et al., 2017; Sancho et al.,
591 2018) and exchange between the chloroplast and other genomes (Martin et al., 1998; Martin
592 2003; Straub et al., 2013).
593
594 ***Implication for chloroplast phylogenomics*—**These examples above are not meant to represent
595 an exhaustive examination of conflicting relationships within angiosperms. Instead, they are
596 intended to highlight several instances of strongly supported conflict within the plastome. These
597 results nevertheless have important implications for plastid phylogenomics, suggesting that
598 concatenated analyses should be performed with caution given the notable presence of gene-tree
599 conflict. The majority of nodes show at least one strongly supported conflicting gene, and many
600 nodes have roughly equivalent numbers of conflicting and concordant genes against the
601 reference tree (Fig. 1). Because of the character of the chloroplast, it has been common for
602 authors to assume that gene tree incongruence occurs due to stochastic error, and thus perform
603 concatenation analyses on the whole chloroplast to increase phylogenetic signal. One of the key
604 assumptions of concatenation methods is that all genes share the same underlying tree. While the
605 signal of stochastic error (e.g., longer genes being more concordant) supports this approach, our
606 detection of supported conflict suggests that this approach may be less valid than previously
607 assumed. Concatenated analyses with extensive underlying conflict can yield problematic results
608 in terms of both topology and branch lengths (Brown and Thomson, 2018; Shen et al., 2017;
609 Walker et al*.,* 2018). Particularly against a backdrop of largely uninformative genes, a few genes
610 exhibiting conflict due to stochastic error, inadequate modeling, or biological sources can drive
611 the analysis toward the wrong topology (Brown and Thomson, 2018; Shen et al*.,* 2017; Walker et
612 al., 2018). We suggest that authors should proceed with caution when assuming that all
613 chloroplast genes share the same tree.
614
615 ***Utility of individual plastid genes for future studies*—**Previous studies have laid a strong
616 framework for determining the utility of chloroplast regions at various phylogenetic scales. For
617 example, work by Shaw et al. (2005, 2007, 2014) highlighted non-coding DNA regions useful
618 for shallow evolutionary studies, while Graham and Olmstead (2000) explored protein-coding
619 genes useful for reconstructing deep relationships in angiosperms. Here, we expand upon
620 previous work by using a novel phylogenomic approach, allowing us examining the concordance

of individual protein-coding plastid genes with all nodes of the accepted angiosperm phylogeny (AT; as well as slight variations thereof, e.g., the GT). We paid special attention to *matK* and *rbcL*, given their historical significance for plant systematics (e.g., Donoghue *et al.*, 1992; Chase *et al.*, 1993; Hilu *et al.*, 2003). We find that *rbcL* performs relatively well in recapitulating the AT— however, *matK* performs considerably better (i.e., it generally has more nodes concordant with the AT; Tables S11-12). This is likely due to a strong positive correlation between alignment length and number of concordant nodes, as noted above (Fig. 2); *matK* has a longer alignment/gene length than *rbcL*.

The gene *ycf1* has been found to be a useful marker in phylogenetics (e.g., Neubig *et al.*, 2008; Neubig and Abbot, 2010; Thomson *et al.*, 2018) and barcoding (Dong et al. 2015), and here we find that it generally performs above average. The alignment of *ycf1* is abnormally long and contains a large number of gaps, and this is likely due to its position spanning the boundary of the IR and the SSC, an area known to fluctuate greatly in size. This variability likely contributes to the value of *ycf1* as a marker for 'species-level' phylogenetics and barcoding. However, the performance of *ycf1* does not scale with its alignment length. In terms of concordance, we find that *matK* performs roughly equally as well if not slightly better than *ycf1* (Tables S11-12). This might in part be a consequence of *ycf1* being missing/lacking annotation from some species, preventing us from analyzing its concordance/conflict with certain nodes. Nevertheless, our results add to the body of evidence supporting *ycf1* as a generally useful plastid region. However, we found *rpoC2* to outperform all other plastid regions in every case (Tables S11-12), and its alignment length (4,660 bp) is ~1/5th the length of *ycf1*, easing the computational burden of using alignment tools such as FSA and BAli-Phy (which would struggle with a region as long as *ycf1*), as well as divergence dating and tree-building programs such as BEAST (Redelings and Suchard 2006; Drummond and Rambaut 2007).

In our analyses, when BS support is not considered, *rpoC2* performed at least as well if not better than using the concatenation of all chloroplast genes (Fig. 3; Tables S11-12**)**. When support is considered (Fig. 3; Tables S11-12**)**, *rpoC2* still remains the best-performing gene, but it performs slightly worse than the concatenation of all chloroplast genes (in terms of number of supported nodes concordant with the AT). The utility of *rpoC2* likely stems from its notable length, resulting in a wealth of useful phylogenetic information. Biologically, even long genes may be functionally constrained such that not all sites are variable, such as *rbcL*, which

652 performed relatively poorly here (Kellogg and Juliano, 1997). However, were this to be the case
653 for *rpoC2*, its longer length relative to *rbcL* should be expected to counter this effect. In light of
654 our results, *rpoC2* should be a highly attractive coding region for future studies, as it generally
655 recapitulates the plastome phylogeny while allowing more proper branch length inferences
656 (given that conflicting signal among multiple genes can result in problematic branch length
657 estimates: Mendes and Hahn, 2016). Recent work (Salichos and Rokas, 2013; Smith et al.,
658 2018b; Steenwyk et al., 2018) indicates that filtering for a smaller set of highly informative
659 genes might yield more accurate results in various phylogenetic applications (e.g., topology
660 estimation, divergence dating). Given this, *rpoC2* might be particularly useful for comparative
661 analyses requiring accurate branch length estimates (i.e., the majority of comparative analyses).
662 Use of *rpoC2* alone (instead of the entire plastome) would also allow for more complex,
663 computationally expensive models to be implemented. Furthermore, focused sequencing of
664 *rpoC2* would increase compatibility of datasets from different studies, facilitating subsequent
665 comprehensive, synthetic analyses. Although the performance *rpoC2* may be dataset dependent,
666 our results support its utility at multiple levels, in terms of both time scales and sampling. It is
667 important to note, however, that we did not include non-coding regions in our study, which
668 would likely outperform coding regions at shallow phylogenetic levels. As previously noted,
669 however, it is generally too difficult to use non-coding regions across broad phylogenetic scales
670 (e.g., angiosperms) because of complications related to homology assessment and alignment.
671
672 ***Genomic patterns of concordance/conflict—***Several previous studies (e.g., Jian *et al*., 2008;
673 Moore *et al.*, 2011) have highlighted the Inverted Repeat (IR) as a valuable plastid region for
674 deep-level phylogenetic analyses, attributing its utility to its relatively slow rate of evolution,
675 resulting in less homoplasy and minimal saturation. Our results contradict this notion, suggesting
676 that the coding sequences of the IR alone perform poorly compared to the LSC and SSC coding
677 regions for reconstructing angiosperm phylogeny (Fig. S1 and Table S13). It is important to note
678 that there are important differences between our study and earlier studies on the IR (e.g., Jian et
679 al., 2008; Moore et al., 2011). For one, we did not include the ribosomal RNA genes, which are
680 highly conserved; thus if the conserved nature of the IR (or at least portions of it) is the basis of
681 its utility, then this might explain the poor performance in the current study. Our saturation
682 analyses, however, (described above) did not reveal any genes to have significant saturation

683 issues at the scale of angiosperm evolution. This calls into question the idea that the conserved
684 genes of the IR would make it superior for reconstruction of angiosperm phylogeny. Instead, it is
685 possible that the non-coding regions of the IR (which we did not include here) are highly
686 informative for angiosperm phylogeny. While the non-coding regions of the LSC and SSC
687 regions have been extensively examined for use as phylogenetic markers (Shaw *et al*., 2005,
688 2007, 2014), the non-coding regions of the IR have been underexplored. Among the IR genes
689 examined here, *ycf2* (which is exceptionally long) showed the greatest levels of concordance
690 (Fig. S1 and Tables S11-12), underscoring the idea that longer genes are generally more useful
691 for phylogeny reconstruction.
692 Another important difference between our study and Moore et al. (2011) is that we only
693 partitioned our data by gene region, while Moore et al. (2011) explored various partitioning
694 strategies (including codon positions and different combinations of genes). It is clear that
695 sequences within the plastomes follow various different models of molecular evolution (as
696 shown above in the results section ***Patterns of conflicting chloroplast signal***). Exploration of
697 more complicated partitioning schemes—which is a time-consuming process (Lanfear *et al*.,
698 2012; Kainer and Lanfear 2015), and beyond the scope of this study—might generally improve
699 plastome-inferred phylogenies.
700
701 ***Conclusions—***We find notable levels of gene-tree conflict within the plastome, at all levels of
702 angiosperm phylogeny, highlighting the necessity of future research into the causes of plastome
703 conflict: do some genes share different evolutionary histories, or is systematic error (e.g., poor
704 modeling) the source of the observed conflict? With rare exceptions, most genes are largely
705 uninformative and, when analyzed as a whole, are therefore unlikely to mislead researchers.
706 However, our results call into question the appropriateness of concatenating plastid genes for
707 phylogeny reconstruction without some level of scrutiny. Even if the concatenated topology is
708 largely congruent with the 'true' species tree of a given group, underlying conflict can mislead
709 branch length estimation (Mendes and Hahn, 2016). Therefore, we recommend that comparative
710 methods and/or dating analyses that rely on branch lengths use filtering procedures similar to
711 those proposed for nuclear data (Salichos and Rokas, 2013; Smith et al., 2018b; Steenwyk et al.,
712 2018).

We find alignment length and molecule type (nucleotide vs amino acids) to be the strongest correlate with plastid gene performance—i.e., longer genes and nucleotides (which have greater information content than amino acids) generally show higher levels of concordance with our accepted hypothesis of angiosperm phylogeny. This also shows that as a single unit, *matK*, *rbcL* and the other commonly used marker *ndhF* are not up to the task alone and if possible should be replaced by longer genes such as *rpoC2, ycf1, and ycf2*. Our findings support the notion that longer genes (with greater information content) are generally superior for phylogeny reconstruction, in line with theoretical predictions (Yang, 1998) and should be prioritized when developing phylogenetic studies.

ACKNOWLEDGEMENTS

We thank S. A. Smith and M. J. Moore for reading an early draft of the manuscript and proving helpful suggestions toward its improvement. We also thank A. Alverson, Liming Cai, two anonymous reviewers and S. Edwards for helpful comments that improved the manuscript. JFW was supported by a Rackham Predoctoral Fellowship. G.W.S. was supported by an NSF Postdoctoral Fellowship (NSF DBI grant 1612032). O.M.V was supported by NSF FESD 1338694 and DEB 1240869. D.A.L. was supported by NSF DEB grant 1458466.

LITERATURE CITED

Adachi, J., P. J. Waddell, W. Martin, and M. Hasegawa. 2000. Plastid genome phylogeny and a model of amino acid substitution for proteins encoded by chloroplast DNA. *Journal of Molecular Evolution* 50: 348–358.

Bendich, A. J. 2004. Circular chloroplast chromosomes: the grand illusion. *The Plant Cell* **16**: 1661–1666.

Birky, C. W., Jr. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences* 92: 11331–11338.

Boyce, Kieran, Fabian Sievers, and Desmond G. Higgins. "Instability in progressive multiple sequence alignment algorithms." Algorithms for molecular biology 10, no. 1 (2015): 26.

Bradley, R. K., A. Roberts, M. Smoot, S. Juvekar , J. Do, C. Dewey, I. Holmes, and L. Pachter. 2009. Fast statistical alignment. *PLoS Computational Biology* 5: e1000392.

744 Brown, J. M. 2014. Detection of implausible phylogenetic inferences using posterior predictive
745 assessment of model fit. *Systematic Biology* 63. 334-348.
746 Brown, J. M., and R. C. Thomson. 2017. Bayes factors unmask highly variable information
747 content, bias, and extreme influence in phylogenomic analyses. *Systematic Biology* 66: 517–
748 530.
749 Brown, J. W., J. F. Walker, and S. A. Smith. 2017. Phyx: phylogenetic tools for unix.
750 *Bioinformatics* 33: 1886–1888.
751 Burleigh, J. G., and S. Mathews. 2007a. Assessing among-locus variation in the inference of
752 seed plant phylogeny. *International Journal of Plant Sciences* 168: 111–124.
753 Burleigh, J. G., and S. Mathews. 2007b. Assessing systematic error in the inference of seed plant
754 phylogeny. *International Journal of Plant Sciences* 168: 125–135.
755 Carbonell-Caballero, J., R. Alonso, V. Ibañez, J. Terol, M. Talon, and J. Dopazo. 2015. A
756 Phylogenetic Analysis of 34 Chloroplast Genomes Elucidates the Relationships between Wild
757 and Domestic Species within the Genus Citrus. *Molecular Biology and Evolution* 32: 2015–
758 2035.
759 Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA,
760 Hills HG, Qiu YL, Kron KA. Phylogenetics of seed plants: an analysis of nucleotide
761 sequences from the plastid gene rbcL. Annals of the Missouri Botanical Garden. 1993 Jan
762 1:528-80.Chatzou, M., Floden, E. W., Di Tommaso, P., Gascuel, O., & Notredame, C. (2018).
763 Generalized bootstrap supports for phylogenetic analyses of protein sequences incorporating
764 alignment uncertainty. Systematic Biology, syx096.
765 Chen, W., Kenney, T., Bielawski, J., Gu, H. 2019. Testing adequacy for DNA substitution
766 models. *BMC Bioinformatics* 20. 349.
767 Corriveau, J. L., and A. W. Coleman. 1988. Rapid screening method to detect potential
768 biparental inheritance of plastid DNA and results for over 200 angiosperms. *American*
769 *Journal of Botany* 75: 1443–1458.
770 Cronn, R., A. Liston, M. Parks, D. S. Gernandt, R. Shen, and T. Mockler. 2008. Multiplex
771 sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis
772 technology. *Nucleic acids research* 36: e122–e122.

773 Cronn, R., B. J. Knaus, A. Liston, P. J. Maughan, M. Parks, J. V. Syring, and J. Udall. 2012.
774 Targeted enrichment strategies for next-generation plant biology. *American Journal of Botany*
775 99: 291–311.
776 Csardi, G., Nepusz, T. 2006. The igraph software package for complex network research.
777 *Interjournal* Complex Systems. 1695.
778 Davis, Jerrold I., Joel R. Mcneal, Craig F. Barrett, Mark W. Chase, James I. Cohen, Melvin R.
779 Duvall, Thomas J. Givnish, Sean W. Graham, Gitte Petersen, J. Chris Pires, Ole Seberg,
780 Dennis W. Stevenson and Jim Leebens-Mack. "Contrasting Patterns of Support among Plastid
781 Genes and Genomes for Major Clades of the Monocotyledons." Chapter. In *Early Events in*
782 *Monocot Evolution*, edited by Paul Wilkin and Simon J. Mayo. Systematics Association
783 Special Volume Series. Cambridge: Cambridge University Press, 2013.
784 Dong, W., C. Xu, C. Li, J. Sun, Y. Zuo, S. Shi, T. Cheng, J. Guo, and S. Zhou. 2015. *ycf1*, the
785 most promising plastid DNA barcode of land plants. *Scientific Reports* 5: 8348.
786 Donoghue, M. J., R. G. Olmstead, J. F. Smith, and J. D. Palmer. 1992. Phylogenetic relationships
787 of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 79: 333–
788 345.
789 Doyle, J. J. 1992. Gene trees and species trees: molecular systematics as one-character
790 taxonomy. *Systematic Botany* 17: 144–163.
791 Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling
792 trees. *BMC Evolutionary Biology* **7**: 214.
793 Duchêne, D. A., Duchêne, S., Ho, S. Y. W. 2018. Differences in performance among test
794 statistics for assessing phylogenomic model adequacy. *Genome Biology and Evolution* 10.
795 1375-1388.
796 Felsenstein, Joseph, and Gary A. Churchill. "A Hidden Markov Model approach to variation
797 among sites in rate of evolution." Molecular biology and evolution 13, no. 1 (1996): 93-104.
798 Foster, C.S., Henwood, M.J. and Ho, S.Y., 2018. Plastome sequences and exploration of tree-
799 space help to resolve the phylogeny of riceflowers (Thymelaeaceae: Pimelea). Molecular
800 phylogenetics and evolution.
801 Galtier, N., and V. Daubin. 2008. Dealing with incongruence in phylogenomic analyses.
802 *Philosophical Transactions of the Royal Society B: Biological Sciences* 363: 4023–4029.

803 Gitzendanner, M. A., P. S. Soltis, G. K. Wong, B. R. Ruhfel, and D. E. Soltis. 2018. Plastid
804 phylogenomic analysis of green plants: a billion years of evolutionary history. *American*
805 *Journal of Botany* 105: 291–301.
806 Gonçalves, Deise JP, Beryl B. Simpson, Edgardo M. Ortiz, Gustavo H. Shimizu, and Robert K.
807 Jansen. "Incongruence between gene trees and species trees and phylogenetic signal variation
808 in plastid genes." Molecular phylogenetics and evolution (2019).
809 Graham, S. W., and R. G. Olmstead. 2000. Utility of 17 chloroplast genes for inferring the
810 phylogeny of the basal angiosperms. *American Journal of Botany* 87: 1712–1730.
811 Fruchterman, T. M. J., Reingold, E. M. 1991. Graph drawing by force-directed placement.
812 *Software: Practice and Experience* 21. 1129-1164.
813 Heath, Tracy A., Shannon M. Hedtke, and David M. Hillis. Taxon sampling and the accuracy of
814 phylogenetic analyses. *J Syst Evol* 46, no. 3 (2008): 239-257.
815 Hensen, A.K., L. K. Escobar, L. E. Gilbert, and R. K. Jansen. 2007. Paternal, maternal, and
816 bipaternal inheritance of the chloroplast genome in *Passiflora* (Passifloraceae): implications
817 for phylogenetic studies. *American Journal of Botany* 94: 42–46.
818 Hillis, D. M., J. J. Bull. An empirical test of bootstrapping as a method for assessing confidence
819 in phylogenetic analysis. *Systematic Biology* 42: 182-192.
820 Hilu, K. W., T. Borsch, K. Müller, D. E. Soltis, P. S. Soltis, V. Savolainen, M. W. Chase, M. P.
821 Powell, L. A. Alice, R. Evans. 2003. Angiosperm phylogeny based on *matK* sequence
822 information. *American Journal of Botany* 90: 1758–1776.
823 Hlavac, Marek. "Stargazer: Well-formatted regression and summary statistics tables." R package
824 version 5, no. 2 (2015).
825 Jansen, R. K., Z. Cai, L. A. Raubeson, H. Daniell, C. W. dePamphilis, J. Leebens-Mack, K. F.
826 Müller, M. Guisinger-Bellian, R. C. Haberle, and A. K. Hansen. 2007. Analysis of 81 genes
827 from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale
828 evolutionary patterns. *Proceedings of the National Academy of Sciences* 104: 19369–19374.
829 Jian, S., P. S. Soltis, M. A. Gitzendanner, M. J. Moore, R. Li, T. A. Hendry, Y. L. Qiu, A.
830 Dhingra, C. D. Bell, and D. E. Soltis. 2008. Resolving an ancient, rapid radiation in
831 Saxifragales. *Systematic Biology* 57: 38–57.
832 Johnson, L. B., and J. D. Palmer. 1988. Heteroplasmy of chloroplast DNA in Medicago. *Plant*
833 *Molecular Biology* 12: 3–11.

834 Kainer, D., and R. Lanfear. 2015. The effects of partitioning on phylogenetic inference.
835 *Molecular Biology and Evolution* 32: 1611–1627.
836 Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017.
837 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 14:
838 587–589.
839 Katoh, Kazutaka, and Daron M. Standley. "MAFFT multiple sequence alignment software
840 version 7: improvements in performance and usability." Molecular biology and evolution 30,
841 no. 4 (2013): 772-780.
842 Kellogg, Elizabeth A., and Nickolas D. Juliano. "The structure and function of RuBisCO and
843 their implications for systematic studies." *American journal of botany* 84, no. 3 (1997): 413-
844 428.
845 R. Lanfear, B. Calcott, S. Y. Ho, and S. Guindon. 2012. PartitionFinder: combined selection of
846 partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology*
847 *and Evolution* 29: 1695–1701.
848 Lee, D. J., T. K. Blake, and S. E. Smith. 1988. Biparental inheritance of chloroplast DNA and the
849 existence of heteroplasmic cells in alfalfa. *Theoretical and Applied Genetics* 76: 545–549.
850 Li, Hong-Tao, Ting-Shuang Yi, Lian-Ming Gao, Peng-Fei Ma, Ting Zhang, Jun-Bo Yang,
851 Matthew A. Gitzendanner, Peter W Fritsch, Jie Cai, Yang Luo, Hong Wang, Michelle van der
852 Bank, Shu-Dong Zhang, Qing-Feng Wang, Jian Wang, Zhi-Rong Zhang, Chao-Nan Fu, Jing
853 Yang, Peter M. Hollingsworth, Mark W. Chase, Douglas E. Soltis, Pamela S. Soltis and De-
854 Zhu Li. "Origin of angiosperms and the puzzle of the Jurassic gap." *Nature plants* 5, no. 5
855 (2019): 461.
856 Liu, J., Z. C. Qi, Y. P. Zhao, C. X. Fu CX, and Q. Y. Jenny Xiang. 2012. Complete cpDNA
857 genome sequence of Smilax china and phylogenetic placement of Liliales—Influences of
858 gene partitions and taxon sampling. *Molecular Phylogenetics and Evolution* 64: 545–562.
859 Logacheva, M. D., A. A. Penin, T. H. Samigullin, C. M. Vallejo-Roman, and A. S. Antonov.
860 "Phylogeny of flowering plants by the chloroplast genome sequences: in search of a “lucky
861 gene”." Biochemistry (Moscow) 72, no. 12 (2007): 1324-1330.
862 Maddison, W. P. 1997. Gene trees in species trees. *Systematic Biology* 46: 523–536.

Magallón, S., S. Gómez-Acevedo, L. L. Sánchez-Reyes, T. Hernández-Hernández. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* **207**: 437–453.

Martin, W. 2003. Gene transfer from organelles to the nucleus: frequent and in big chunks. *Proceedings of the National Academy of Sciences* 100: 8612–8614.

Martin, W., B. Stoebe, V. Goremykin, S. Hapsmann, M. Hasegawa, and K. V. Kowallik. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393: 162–165.

McCauley, D. E., A. K. Sundby, M. F. Bailey MF, M E. Welch. 2007. Inheritance of chloroplast DNA is not strictly maternal in Silene vulgaris (Caryophyllaceae): evidence from experimental crosses and natural populations. *American Journal of Botany* **94**: 1333–1337.

Mendes, F. K., and M. H. Hahn. 2016. Gene tree discordance causes apparent substitution rate variation. *Systematic Biology* 65: 711–721.

Mogensen, H. L. 1996. The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany* 83: 383–404

Moore, M. J., A. Dhingra, P. S. Soltis, R. Shaw, W. G. Farmerie, K. M. Folta, and D. E. Soltis. 2006. Rapid and accurate pyrosequencing of angiosperm plastid genomes. *BMC Plant Biology* **6**: 17.

Moore, M. J., C. D. Bell, P. S. Soltis, and D. E. Soltis. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences* 104: 19363–19368.

Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences* 107: 4623–4628.

Moore, M. J., N. Hassan, M. A. Gitzendanner, R. A. Bruenn, M. Croley, A. Vandeventer, J. W. Horn, A. Dhingra, S. F. Brockington, and M. Latvis. 2011. Phylogenetic analysis of the plastid inverted repeat for 244 species: insights into deeper-level angiosperm relationships from a long, slowly evolving sequence region. *International Journal of Plant Sciences* 172: 541–558.

Morrison, David A. "Why would phylogeneticists ignore computerized sequence alignment?." *Systematic biology* 58, no. 1 (2009): 150-158.

Morrison, David A., Matthew J. Morgan, and Scot A. Kelchner. "Molecular homology and multiple-sequence alignment: an analysis of concepts and practice." *Australian Systematic Botany* 28, no. 1 (2015): 46-62.

Neubig, K. M., and J. R. Abbott. 2010. Primer development for the plastid region *ycf1* in Annonaceae and other magnoliids. *American Journal of Botany* 97: e52–e55.

Neubig, K. M., W. M. Whitten, B. S. Carlsward, M. A. Blanco, L. Endara, N. H. Williams, and M. J. Moore. 2009. Phylogenetic utility of ycf1 in orchids: a plastid gene more variable than *matK*. *Plant Systematics and Evolution* **277**: 75–84.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.

Palmer, J. D. 1983. Chloroplast DNA exists in two orientations. *Nature* 301: 92–93.

Palmer, J. D. 1985. Comparative organization of chloroplast genomes. *Annual review of genetics* 19: 325–354.

Paradis, Emmanuel, and Klaus Schliep. "ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R." *Bioinformatics* 35, no. 3 (2018): 526-528.

Philippe, Hervé, Ulf Sörhannus, Anne Baroin, Roland Perasso, Françoise Gasse, and André Adoutte. "Comparison of molecular and paleontological data in diatoms suggests a major gap in the fossil record." Journal of Evolutionary Biology 7, no. 2 (1994): 247-265.

Philippe, H.,H. Brinkmann, D. V. Lavrov, D. T. J. Littlewood, M. Manuel, G. Wörheide, and D. Baurain. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biology* 9: e1000602.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Redelings, B. 2014. Erasing errors due to alignment ambiguity when estimating positive selection. *Molecular Biology and Evolution* 31: 1979–1993.

Refulio-Rodriguez, N. F., and R. G. Olmstead. 2014. Phylogeny of Lamiidae. *American Journal of Botany* 101: 287–299.

Danny W. Rice, Andrew J. Alverson, Aaron O. Richardson, Gregory J. Young, M. Virginia Sanchez-Puerta, Jérôme Munzinger, Kerrie Barry, Jeffrey L. Boore, Yan Zhang, Claude W. dePamphilis, Eric B. Knox, Jeffrey D. Palmer. "Horizontal transfer of entire genomes via

924 mitochondrial fusion in the angiosperm Amborella." Science 342, no. 6165 (2013): 1468-
925 1473.
926 Richards, E. J., J. M. Brown, A. J. Barley, R. A. Chong, R. C. Thomson. 2018. Variation across
927 mitochondrial gene trees provides evidence for systematic error: How much gene tree
928 variation is biological? *Systematic Biology* doi:10.1093/sysbio/syy013.
929 Robinson, D. F., Foulds, L. R. 1981. Comparison of phylogenetic trees. *Mathematical*
930 *Biosciences* 53: 131-147.
931 Rodríguez-Ezpeleta, N., H. Brinkmann, G. Burger, A. J. Roder, M. W. Gray, H. Philippe, and B.
932 F. Lang. 2007. Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids
933 and cercozoans. *Current Biology* 17: 1420–1425.
934 Rokas, A., B. L. Williams, N. King, S. B. Carroll. 2003. Genome-scale approaches to resolving
935 incongruence in molecular phylogenies. *Nature* 425: 798–804.
936 Rokas, Antonis, and Sean B. Carroll. More genes or more taxa? The relative contribution of gene
937 number and taxon number to phylogenetic accuracy. *Molecular biology and evolution* 22, no.
938 5 (2005): 1337-1344.
939 Roquet, C., I. Sanmartín, N. Garcia-Jacas, L. Sáez, A. Susanna, N. Wikström, and J. J. Aldasoro.
940 2009. Reconstructing the history of Campanulaceae with a Bayesian approach to molecular
941 dating and dispersal–vicariance analyses. *Molecular Phylogenetics and Evolution* 52: 575–
942 587.
943 Rosenberg, Michael S., and Sudhir Kumar. Incomplete taxon sampling is not a problem for
944 phylogenetic inference. *Proceedings of the National Academy of Sciences* 98, no. 19 (2001):
945 10751-10756.
946 Ruhfel, B. R., M. A. Gitzendanner, P. S. Soltis, D. E. Soltis, and J. G. Burleigh. 2014. From
947 algae to angiosperms–inferring the phylogeny of green plants (*Viridiplantae*) from 360 plastid
948 genomes. *BMC Evolutionary Biology* 14: 23.
949 Salichos, Leonidas, and Antonis Rokas. "Inferring ancient divergences requires genes with
950 strong phylogenetic signals." Nature 497, no. 7449 (2013): 327.
951 Sancho, R., C. P. Cantalapiedra, D. López-Alvarez, S. P. Gordon, J. P. Vogel, P. Catalán, B.
952 Contreras-Moreira. 2018. Comparative plastome genomics and phylogenomics of
953 *Brachypodium*: flowering time signatures, introgression and recombination in recently
954 diverged ecotypes. *New Phytologist* 218: 1631–1644.

955 Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, C. T. Winder,
956 E. E. Schilling, R. L. Small. 2005. The tortoise and the hare II: relative utility of 21 noncoding
957 chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–
958 166.
959 Shaw, J., E. B. Lickey, E. E. Schilling, R. L. Small. 2007. Comparison of whole chloroplast
960 genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the
961 tortoise and the hare III. *American Journal of Botany* 94: 275–288.
962 Shaw, J., H. L. Shafer, O. R. Leonard, M. J. Kovach, M. Schorr, and A. B. Morris. 2014.
963 Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic
964 inferences in angiosperms: the tortoise and the hare IV. *American Journal of Botany* 101:
965 1987–2004.
966 Shen, Xing-Xing, Leonidas Salichos, and Antonis Rokas. "A genome-scale investigation of how
967 sequence, function, and tree-based gene properties influence phylogenetic inference."
968 Genome biology and evolution 8, no. 8 (2016): 2565-2580.
969 Shen, X.-X., C. T. Hittinger, and A. Rokas. 2017. Contentious relationships in phylogenomic
970 studies can be driven by a handful of genes. *Nature Ecology & Evolution* 1: 126.
971 Shepherd, L. D., B. R. Holland, L. P. Perrie. 2008. Conflict amongst chloroplast DNA sequences
972 obscures the phylogeny of a group of *Asplenium* ferns. *Molecular Phylogenetics and*
973 *Evolution* 48: 176–187.
974 Smith, David Roy. "Mitochondrion-to-plastid DNA transfer: it happens." New Phytologist 202,
975 no. 3 (2014): 736-738.
976 Smith, S. E. 1989. Biparental inheritance of organelles and its implications in crop improvement.
977 *Plant Breeding Reviews* 6: 361–393.
978 Smith, S. A., M. J. Moore, J. W. Brown, and Y. Yang. 2015. Analysis of phylogenomic datasets
979 reveals conflict, concordance, and gene duplications with examples from animals and plants.
980 *BMC Evolutionary Biology* 15: 150.
981 Stephen A. Smith, Nathanael Walker-Hale, Joseph F. Walker, Joseph W. Brown. Nested
982 phylogenetic conflicts, combinability, and deep phylogenomics in plants. bioRxiv 371930;
983 doi: https://doi.org/10.1101/371930
984 DOUGLAS E. SOLTIS, PAMELA S. SOLTIS, MARK W. CHASE, MARK E. MORT, DIRK
985 C. ALBACH, MICHAEL ZANIS, VINCENT SAVOLAINEN, WILLIAM H. HAHN, SARA

B. HOOT, MICHAEL F. FAY, MICHAEL AXTELL, SUSAN M. SWENSEN, LINDA M. PRINCE, W. JOHN KRESS, KEVIN C. NIXON, JAMES S. FARRIS, Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences, *Botanical Journal of the Linnean Society*, Volume 133, Issue 4, August 2000, Pages 381–46.

Douglas E. Soltis Stephen A. Smith Nico Cellinese Kenneth J. Wurdack David C. Tank Samuel F. Brockington Nancy F. Refulio-Rodriguez Jay B. Walker Michael J. Moore Barbara S. Carlsward Charles D. Bell Maribeth Latvis Sunny Crawley Chelsea Black Diaga Diouf Zhenxiang Xi Catherine A. Rushworth Matthew A. Gitzendanner Kenneth J. Sytsma Yin-Long Qiu Khidir W. Hilu Charles C. Davis Michael J. Sanderson Reed S. Beaman Richard G. Olmstead Walter S. Judd Michael J. Donoghue Pamela S. Soltis. Angiosperm phylogeny: 17 genes, 640 taxa. American journal of botany. 2011 Apr;98(4):704-30.

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

Steenwyk, Jacob L., Xing-Xing Shen, Abigail L. Lind, Gustavo G. Goldman, and Antonis Rokas. "A robust phylogenomic timetree for biotechnologically and medically important fungi from Aspergillaceae (Eurotiomycetes, Ascomycota)." *bioRxiv*(2018): 370429.

Stegemann, S., S. Hartmann, S. Ruf, and R. Bock. 2003. High-frequencing gene transfer from the chloroplast genome to the nucleus. *Proceedings of the National Academy of Sciences* 100: 8828–8833.

Straub, S, C, K,, Cronn, R. C., Edwards, C., Fishbein, M,, and A. Liston. 2013. Horizontal Transfer of DNA from the Mitochondrial to the Plastid Genome and Its Subsequent Evolution in Milkweeds (Apocynaceae). *Genome Biology and Evolution* 5: 1872–1885.

Stull, G. W., M. J. Moore, V. S. Mandala, N. A. Douglas, H. R. Kates, X. Qi, S. F. Brockington, P. S. Soltis, D. E. Soltis, and M. A. Gitzendanner. 2013. A targeted enrichment strategy for massively parallel sequencing of angiosperm plastid genomes. *Applications in Plant Sciences* 1: 1200497.

Stull, G. W., R. Duno de Stefano, D. E. Soltis, P. S. Soltis. 2015. Resolving basal lamiid phylogeny and the circumscription of Icacinaceae with a plastome-scale data set. *American Journal of Botany 102*: 1794–1813.

Suchard, M. A., and B. D. Redelings. 2006. BAli-Phy: simultaneous Bayesian inference of alignment and phylogeny. *Bioinformatics* 22: 2047–2048.

1017 Sullivan, A. R., B. Schiffthaler, S. L. Thompson, N. R. Street, and X. R. Wang. 2017.
1018 Interspecific plastome recombination reflects ancient reticulate evolution in *Picea* (Pinaceae).
1019 *Molecular Biology and* Evolution 34: 1689–1701.
1020 Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of
1021 three non-coding regions of chloroplast DNA. *Plant molecular biology* 17: 1105–1109.
1022 Thomson, A. M., O. M. Vargas, and C. W. Dick. 2018. Complete plastome sequences from
1023 *Bertholletia excelsa* and 23 related species yield informative markers for Lecythidaceae.
1024 *Applications in Plant Sciences* 6: e1151.
1025 Uribe-Convers, S., J. R. Duke, M. J. Moore, D. C. Tank. 2014. A long PCR-based approach for
1026 DNA enrichment prior to next-generation sequencing for systematic studies. *Applications in*
1027 *Plant Sciences* 2: 1300063.
1028 Vargas, O. M., E. M. Ortiz, and B. B. Simpson. 2017. Conflicting phylogenomic signals reveal a
1029 pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae:
1030 Diplostephium). *New Phytologist* 214: 1736–1750.
1031 Walker, J. F., M. J. Zanis, N. C. Emery. 2014. Comparative analysis of complete chloroplast
1032 genome sequence and inversion variation in *Lasthenia burkei* (Madieae, Asteraceae)."
1033 *American Journal of Botany* 101: 722–729.
1034 Walker, J. F., R. K. Jansen, M. J. Zanis, N. C. Emery. 2015. Sources of inversion variation in the
1035 small single copy (SSC) region of chloroplast genomes. *American Journal of Botany* 102:
1036 1751–1752.
1037 Walker, J. F., Y. Yang, M. J. Moore, J. Mikenas, A. Timoneda, S. F. Brockington, and S. A.
1038 Smith. 2017. Widespread paleopolyploidy, gene tree conflict, and recalcitrant relationships
1039 among the carnivorous Caryophyllales. *American Journal of Botany* 104: 858–867.
1040 Walker, J.F., Brown, J.W. and Smith, S.A., 2018. Analyzing contentious relationships and outlier
1041 genes in phylogenomics. *Systematic biology* 67: 916–924.
1042 Whelan, Simon, and Nick Goldman. "A general empirical model of protein evolution derived
1043 from multiple protein families using a maximum-likelihood approach." Molecular biology
1044 and evolution 18, no. 5 (2001): 691-699.
1045 White, Daniel James, Jonci Nikolai Wolff, Melanie Pierson, and Neil John Gemmell. "Revealing
1046 the hidden complexities of mtDNA inheritance." *Molecular Ecology* 17, no. 23 (2008): 4925-
1047 4942.

Norman J. Wickett, Siavash Mirarab, Nam Nguyen, Tandy Warnow, Eric Carpenter, Naim Matasci, Saravanaraj Ayyampalayam, Michael S. Barker, J. Gordon Burleigh, Matthew A. Gitzendanner, Brad R. Ruhfel, Eric Wafula, Joshua P. Der, Sean W. Graham, Sarah Mathews, Michael Melkonian, Douglas E. Soltis, Pamela S. Soltis, Nicholas W. Miles, Carl J. Rothfels, Lisa Pokorny, A. Jonathan Shaw, Lisa DeGironimo, Dennis W. Stevenson, Barbara Surek, Juan Carlos Villarreal, Béatrice Roure, Hervé Philippe, Claude W. dePamphilis, Tao Chen, Michael K. Deyholos, Regina S. Baucom, Toni M. Kutchan, Megan M. Augustin, Jun Wang, Yong Zhang, Zhijian Tian, Zhixiang Yan, Xiaolei Wu, Xiao Sun, Gane Ka-Shu Wong, James Leebens-Mack. Phylotranscriptomic analysis of the origin and early diversification of land plants. Proceedings of the National Academy of Sciences. 2014 Nov 11;111(45):E4859-68.

Wolfe, K. H., W.-H. Li, and P. M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences* 84: 9054–9058.

Wolfe A. D., and C. P. Randle. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: implications for plant molecular systematics. *Systematic Botany* 29:1011–1020.

Yang, Z. 1998. On the best evolutionary rate for phylogenetic analysis. *Systematic Biology* 47: 125–133.

Zeng, L., Q. Zhang, R. Sun, H. Kong, N. Zhang, H. Ma. 2014. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. *Nature Communications* 5: 4956.

Zhang, Chao, Maryam Rabiee, Erfan Sayyari, and Siavash Mirarab. "ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees." BMC bioinformatics 19, no. 6 (2018): 153.

Zhang, Q., Y. Liu, and Sodmergen. 2003. Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. *Plant and Cell Physiology* 44: 941–951.

Zwickl, Derrick J., and David M. Hillis. "Increased taxon sampling greatly reduces phylogenetic error." *Systematic Biology* 51, no. 4 (2002): 588-598.

Figure Legends

Figure 1. Summary of chloroplast conflict against the reference phylogeny of angiosperms. Green, orange, and purple boxes indicate where the amino acid, nucleotide-inferred, or Maximum Quartet Support Species Tree plastome trees conflict with the reference phylogeny. Pie charts depict the amount of gene tree conflict observed in the nucleotide analysis, with the blue, green, red and gray slices representing, respectively, the proportion of gene trees concordant, conflicting (supporting a single main alternative topology), conflicting (supporting various alternative topologies), and uninformative (BS < 70 or missing taxon) at each node in the species tree. The dashed lines represent 30 myr time intervals (positioned based on Magallon *et al.* 2015 and Vargas *et al.* 2017) used to bin nodes for examinations of conflict at different levels of divergence.

Figure 2. Gene tree concordance/conflict at varying time scales. Each diagram represents a different molecule type and shows the proportion of concordance each gene exhibits at the five time slices shown in Fig. 1: (1) 150–120 mya, (2) 120–90 mya, (3) 90–60 mya, (4) 60–30 mya, (5) 30–0 mya and C is the concordance summed over all the time scales. The individual genes are scaled by length of alignment; however, *ycf1* and *ycf2* are cut to approximately the length of *rpoC2* due to their extremely long alignments. **A** Results from amino acid data considering only nodes with bootstrap support ≥ 70%. **B** Results from nucleotide data considering only nodes with bootstrap support ≥ 70%. The plots along the bottom show relationships between gene concordance levels and alignment length and tree length, excluding outlying genes (see **Methods**). Each point represents the proportion of concordance considering only nodes with bootstrap support ≥ 70%. Red lines show the predicted values from logistic regression and asterisks give the p-value of the relationship from univariate logistic regression, *** = $p < 0.001$. **C** Logistic regression of concordant nodes from amino acid data against alignment length. **D** Logistic regression of concordant nodes from amino acid data against tree length. **E** Logistic regression of concordant nodes from nucleotide data against alignment length. **F** Logistic regression of concordant nodes from nucleotide data against tree length.

Figure 3. Histograms depicting number of concordant edges each gene tree contains compared to the reference phylogeny (i.e., the 'accepted tree, AT). Histograms are binned by number of concordant nodes, bar heights give the number of genes in each bin. Commonly used markers (*matK*, *ndhF*, *rbcL*, and *ycf1*) are labeled on the graph, along with the most concordant gene (*rpoC2)* and the number of concordant nodes for the complete chloroplast (CC) compared to the AT. **A** Concordance in the nucleotide dataset not considering bootstrap support. **B** Concordance in the amino acid dataset not considering bootstrap support. **C** Concordance in the nucleotide dataset considering bootstrap support ≥ 70%. **D** Concordance in the amino acid dataset considering bootstrap support ≥ 70%.

Figure 1

Figure 1. Summary of chloroplast conflict against the reference phylogeny of angiosperms.

Green, orange, and purple boxes indicate where the amino acid, nucleotide-inferred, or Maximum Quartet Support Species Tree plastome trees conflict with the reference phylogeny. Pie charts depict the amount of gene tree conflict observed in the nucleotide analysis, with the blue, green, red and gray slices representing, respectively, the proportion of gene trees concordant, conflicting (supporting a single main alternative topology), conflicting (supporting various alternative topologies), and uninformative (BS < 70 or missing taxon) at each node in the species tree. The dashed lines represent 30 myr time intervals (positioned based on Magallon *et al.* 2015 and Vargas *et al.* 2017) used to bin nodes for examinations of conflict at different levels of divergence.

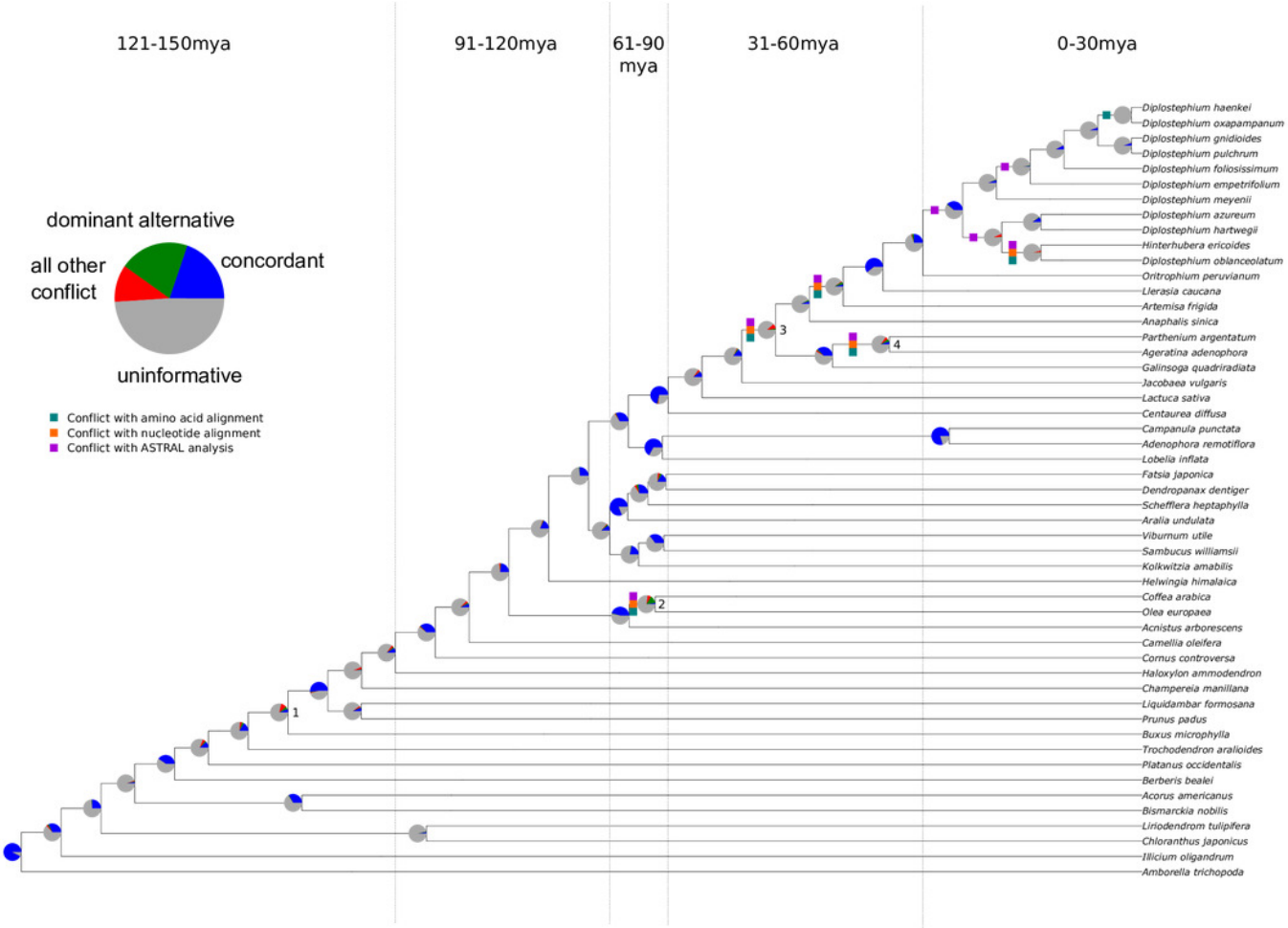
121-150mya
91-120mya
61-90 mya
31-60mya
0-30mya
dominant alternative
all other conflict
concordant
uninformative
Conflict with amino acid alignment
Conflict with nucleotide alignment
Conflict with ASTRAL analysis
1
2
3
4
Diplostephium haenkei
Diplostephium oxapampanum
Diplostephium gnidioides
Diplostephium pulchrum
Diplostephium foliosissimum
Diplostephium empetrifolium
Diplostephium meyenii
Diplostephium azureum
Diplostephium hartwegii
Hinterhubera ericoides
Diplostephium oblanceolatum
Oritrophium peruvianum
Llerasia caucana
Artemisa frigida
Anaphalis sinica
Parthenium argentatum
Ageratina adenophora
Galinsoga quadriradiata
Jacobaea vulgaris
Lactuca sativa
Centaurea diffusa
Campanula punctata
Adenophora remotiflora
Lobelia inflata
Fatsia japonica
Dendropanax dentiger
Schefflera heptaphylla
Aralia undulata
Viburnum utile
Sambucus williamsii
Kolkwitzia amabilis
Helwingia himalaica
Coffea arabica
Olea europaea
Acnistus arborescens
Camellia oleifera
Cornus controversa
Haloxylon ammodendron
Champereia manillana
Liquidambar formosana
Prunus padus
Buxus microphylla
Trochodendron aralioides
Platanus occidentalis
Berberis bealei
Acorus americanus
Bismarckia nobilis
Liriodendrom tulipifera
Chloranthus japonicus
Illicium oligandrum
Amborella trichopoda

Figure 2

Figure 2. Gene tree concordance/conflict at varying time scales.

Each diagram represents a different molecule type and shows the proportion of concordance each gene exhibits at the five time slices shown in Fig. 1: (1) 150–120 mya, (2) 120–90 mya, (3) 90–60 mya, (4) 60–30 mya, (5) 30–0 mya and C is the concordance summed over all the time scales. The individual genes are scaled by length of alignment; however, *ycf1* and *ycf2* are cut to approximately the length of *rpoC2* due to their extremely long alignments. **A)** Results from amino acid data considering only nodes with bootstrap support ≥ 70%. **B)** Results from nucleotide data considering only nodes with bootstrap support ≥ 70%. The plots along the bottom show relationships between gene concordance levels and alignment length and tree length, excluding outlying genes (see **Methods**). Each point represents the proportion of concordance considering only nodes with bootstrap support ≥ 70%. Red lines show the predicted values from logistic regression and asterisks give the p-value of the relationship from univariate logistic regression, *** = $p < 0.001$. **C)** Logistic regression of concordant nodes from amino acid data against alignment length. **D)** Logistic regression of concordant nodes from amino acid data against tree length. **E)** Logistic regression of concordant nodes from nucleotide data against alignment length. **F)** Logistic regression of concordant nodes from nucleotide data against tree length.

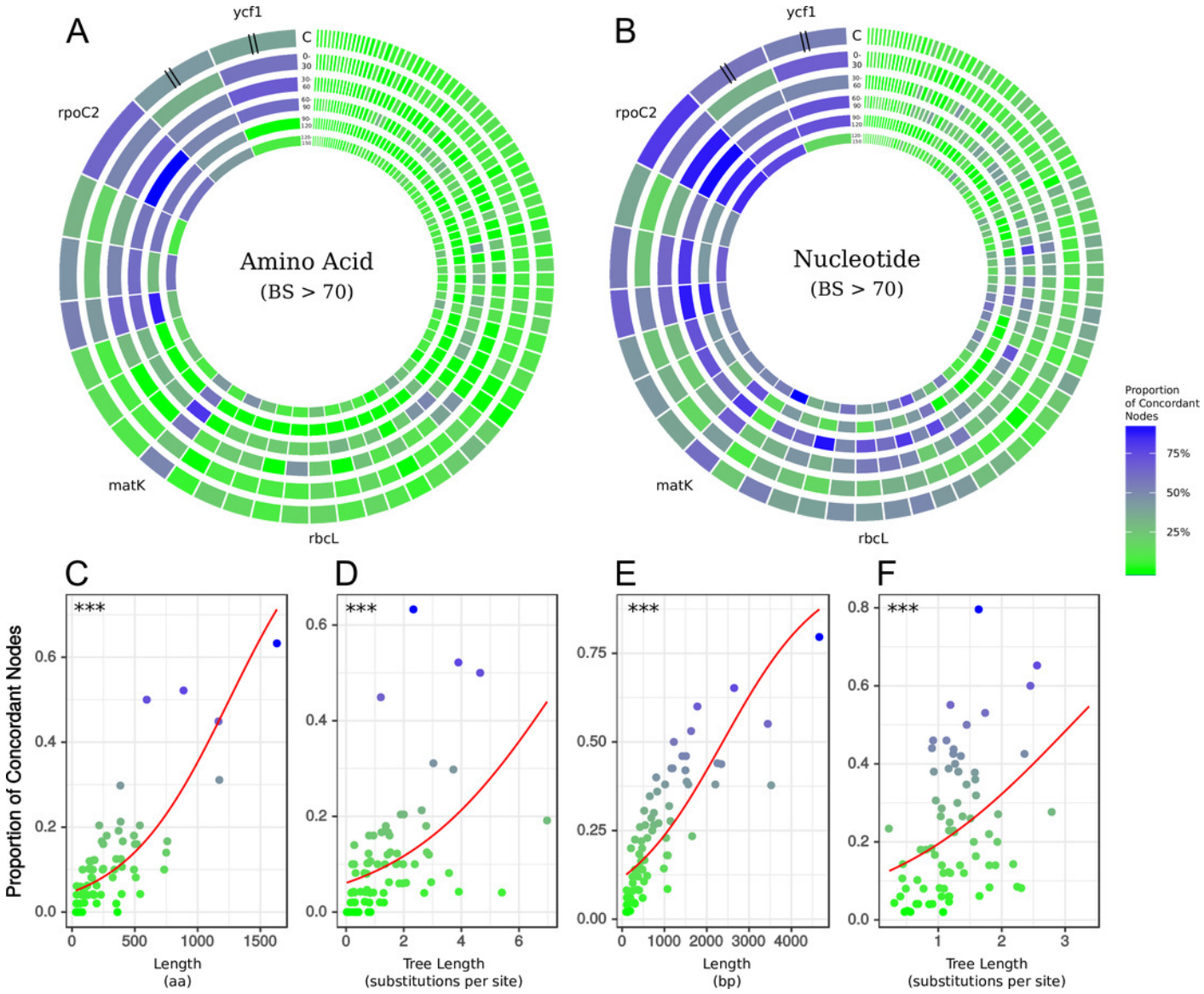
A
ycf1
C
0-30
30-60
60-90
90-120
120-150
rpoC2
Amino Acid
(BS > 70)
matK
rbcL
B
ycf1
C
rpoC2
Nucleotide
(BS > 70)
matK
rbcL
Proportion of Concordant Nodes
75%
50%
25%
C
D
E
F
Proportion of Concordant Nodes
***
Length (aa)
Tree Length (substitutions per site)
Length (bp)
Tree Length (substitutions per site)

Figure 3

Figure 3. Histograms depicting number of concordant edges each gene tree contains compared to the reference phylogeny (i.e., the 'accepted tree, AT).

Histograms are binned by number of concordant nodes, bar heights give the number of genes in each bin. Commonly used markers (*matK*, *ndhF*, *rbcL*, and *ycf1*) are labeled on the graph, along with the most concordant gene (*rpoC2)* and the number of concordant nodes for the complete chloroplast (CC) compared to the AT. **A)** Concordance in the nucleotide dataset not considering bootstrap support. **B)** Concordance in the amino acid dataset not considering bootstrap support. **C)** Concordance in the nucleotide dataset considering bootstrap support ≥ 70%. **D)** Concordance in the amino acid dataset considering bootstrap support ≥ 70%.

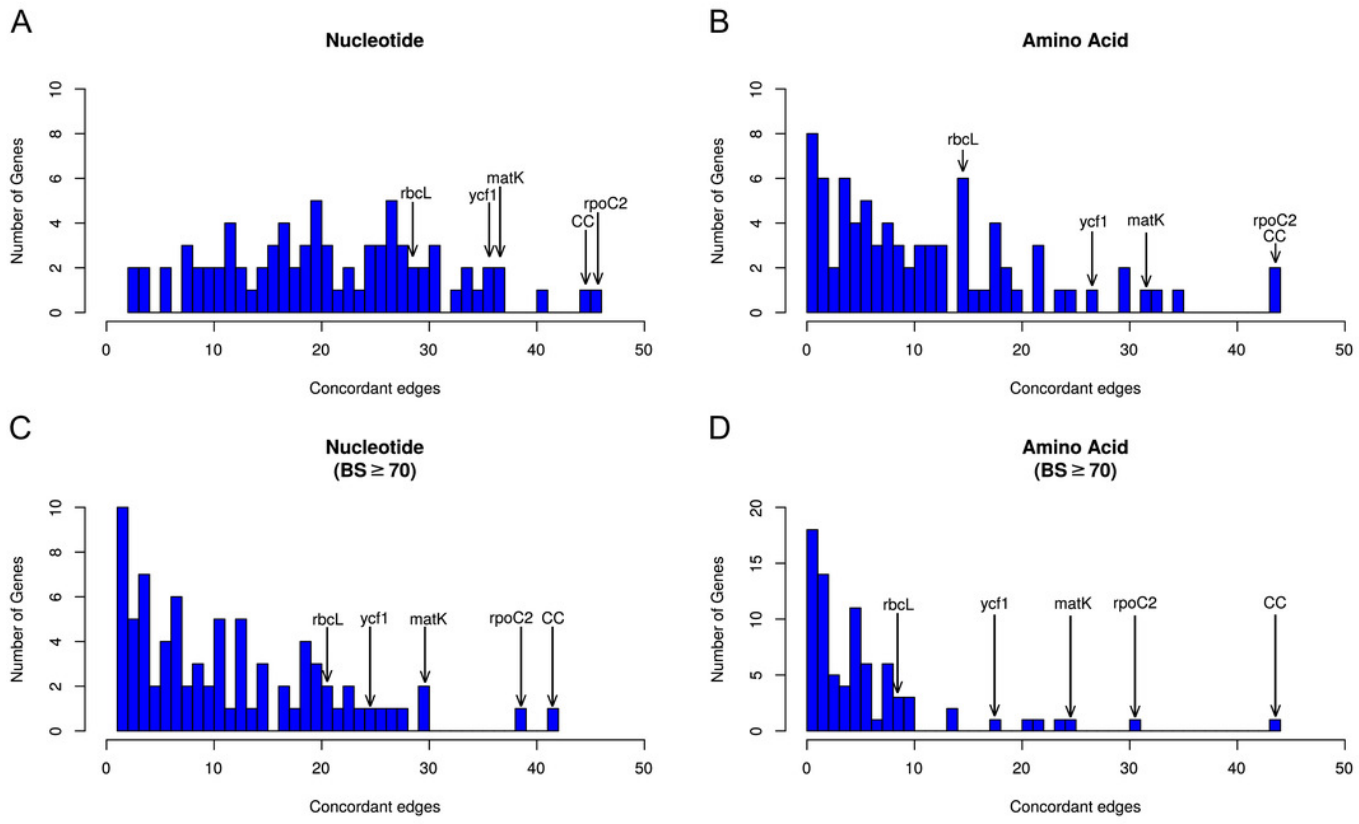

Table 1(on next page)

Table 1. Log-Likelihood scores from the edge-based analysis.

Log-Likelihood scores rounded to the nearest whole number for alternative resolutions of contentious nodes based on edge-based analyses using a modification of the Maximum Gene Wise Edge (MGWE) method (Walker et al. 2018). The best-supported relationship (highest log-likelihood score) for each case is presented in **bold**.

Table 1. Log-Likelihood scores rounded to the nearest whole number for alternative resolutions of contentious nodes based on edge-based analyses using a modification of the Maxmimum Gene Wise Edge (MGWE) method (Walker et al. 2018). The best-supported relationship (highest log-likelihood score) for each case is presented in **bold**.

Contentious relationships	GT	AT	CC	Alternative topology (>3 genes supporting)
Buxus placement (node 10)	**-618242**	-618371	-618371	**-618242**
Lamiid relationships (node 19)	**-618203**	-618235	**-618203**	**-618203**
Jacobaea and *Artemisia* placements (node 34)	-618535	**-618506**	N/A (BS < 70)	N/A
Galinsoga placement (node 36)	NA	-618308	**-618174**	-618319

Table 2(on next page)

Table 2. Contentious relationships tested using edge-based analyses.

Individual genes supporting (with ≥70 BS) the alternative topologies examined for each contentious relationship.

Table 2. Individual genes supporting (with ≥70 BS) the alternative topologies examined for each contentious relationship.

Contentious relationships	GT	AT	CC	Alternative topology (>3 genes supporting)
Buxus placement (node 10)	*rbcL, petL, rps2, rps14*	*ndhD, rps4, matK, psbB, rpoC2*	*ndhD, rps4, matK, psbB, rpoC2*	*rbcL, petL, rps2, rps14*
Lamiid relationships (node 19)	*ndhK, rbcL, rps11, rps7, ycf4, rpoC2, rps4, ndhA, ndhJ, ccsA, rps2, petB, matK, ndhI, atpB*	*psbB, petA*	*ndhK, rbcL, rps11, rps7, ycf4, rpoC2, rps4, ndhA, ndhJ, ccsA, rps2, petB, matK, ndhI, atpB*	*ndhK, rbcL, rps11, rps7, ycf4, rpoC2, rps4, ndhA, ndhJ, ccsA, rps2, petB, matK, ndhI, atpB*
Jacobaea and *Artemisia* placements (node 34)	*rpoC2*	None	N/A	N/A