Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales

Zhenxiang Xi[a], Brad R. Ruhfel[a,b], Hanno Schaefer[a,c], André M. Amorim[d], M. Sugumaran[e], Kenneth J. Wurdack[f], Peter K. Endress[g], Merran L. Matthews[g], Peter F. Stevens[h], Sarah Mathews[i,1], and Charles C. Davis[a,1]

[a]Department of Organismic and Evolutionary Biology, Harvard University Herbaria, Cambridge, MA 02138; [b]Department of Biological Sciences, Eastern Kentucky University, Richmond, KY 40475; [c]Biodiversitaet der Pflanzen, Technische Universitaet Muenchen, D-85354 Freising, Germany; [d]Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Ilhéus, 45.662-900, Bahia, Brazil; [e]Rimba Ilmu Botanic Garden, Institute of Biological Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia; [f]Department of Botany, Smithsonian Institution, Washington, DC 20013; [g]Institute of Systematic Botany, University of Zurich, CH-8008 Zurich, Switzerland; [h]Department of Biology, University of Missouri, St. Louis, MO 63166; and [i]Arnold Arboretum, Harvard University, Boston, MA 02131

Edited by Robert K. Jansen, University of Texas, Austin, TX, and accepted by the Editorial Board September 11, 2012 (received for review April 6, 2012)

The angiosperm order Malpighiales includes ~16,000 species and constitutes up to 40% of the understory tree diversity in tropical rain forests. Despite remarkable progress in angiosperm systematics during the last 20 y, relationships within Malpighiales remain poorly resolved, possibly owing to its rapid rise during the mid-Cretaceous. Using phylogenomic approaches, including analyses of 82 plastid genes from 58 species, we identified 12 additional clades in Malpighiales and substantially increased resolution along the backbone. This greatly improved phylogeny revealed a dynamic history of shifts in net diversification rates across Malpighiales, with bursts of diversification noted in the Barbados cherries (Malpighiaceae), cocas (Erythroxylaceae), and passion flowers (Passifloraceae). We found that commonly used a priori approaches for partitioning concatenated data in maximum likelihood analyses, by gene or by codon position, performed poorly relative to the use of partitions identified a posteriori using a Bayesian mixture model. We also found better branch support in trees inferred from a taxon-rich, data-sparse matrix, which deeply sampled only the phylogenetically critical placeholders, than in trees inferred from a taxon-sparse matrix with little missing data. Although this matrix has more missing data, our a posteriori partitioning strategy reduced the possibility of producing multiple distinct but equally optimal topologies and increased phylogenetic decisiveness, compared with the strategy of partitioning by gene. These approaches are likely to help improve phylogenetic resolution in other poorly resolved major clades of angiosperms and to be more broadly useful in studies across the Tree of Life.

Malpighiales are one of the most surprising clades discovered in broad molecular phylogenetic studies of the flowering plants (1–3). The order contains ~16,000 species and 42 families (2, 3) that exhibit remarkable morphological and ecological diversity. A few examples include cactus-like succulents (Euphorbiaceae), epiphytes (Clusiaceae), holoparasites (Rafflesiaceae), submerged aquatics (Podostemaceae), and wind-pollinated trees (temperate Salicaceae). The order is ecologically important: species in Malpighiales constitute up to 40% of the understory tree diversity in tropical rain forests worldwide (4). They also include many economically important species, such as Barbados nut (*Jatropha curcas* L., Euphorbiaceae), cassava (*Manihot esculenta* Crantz, Euphorbiaceae), castor bean (*Ricinus communis* L., Euphorbiaceae), coca (*Erythroxylum coca* Lam., Erythroxylaceae), flax (*Linum usitatissimum* L., Linaceae), the poplars (*Populus* spp., Salicaceae), and the rubber tree (*Hevea brasiliensis* Müll. Arg., Euphorbiaceae). Partially for this reason, genomic resources for Malpighiales are growing at a rapid pace and include whole-genome sequencing projects completed or near completion for Barbados nut (5), cassava, castor bean (6), flax, and poplar (7). Thus, a resolved phylogeny of Malpighiales is critical not only for evolutionary, ecological, developmental, and genomic investigations of flowering plants, but also for crop improvement.

Despite substantial progress in resolving the angiosperm Tree of Life during the last 20 y (1, 8–12), phylogenetic relationships within Malpighiales remain poorly resolved. Molecular studies (1, 4) using multiple gene regions from the plastid, mitochondrial, and nuclear genomes have confirmed the monophyly of Malpighiales and its component families with a high degree of confidence but have identified only a handful of well-supported multifamily clades. The most recent analysis by Wurdack and Davis (3) included 13 genes, totaling 15,604 characters, sampled across all three genomes from 144 Malpighiales. Their results indicated that all families are monophyletic, but interrelationships among the 16 major subclades remained unresolved. The difficulty in determining these deep relationships may result from the rapid rise of the order during the mid-Cretaceous (4).

We used phylogenomic approaches to resolve relationships within Malpighiales to provide a framework for studying their tempo and mode of diversification. Our core data set included 82 genes sampled from the plastomes of 58 species, 48 of which were newly sequenced for this study. We combined this core data set with the previously described taxon-rich data set of Wurdack and Davis (3). Our results greatly improve phylogenetic resolution within Malpighiales, highlight the value of a unique partitioning strategy for phylogenomic analyses, and reveal a dynamic history of shifts in net diversification rates across the order.

Results and Discussion

Taxon and Gene Sampling. Our core data set, the ***82-gene*** matrix, included 58 taxa (48 are newly sequenced; *SI Appendix*, Table S2) and 82 plastid genes common to most angiosperms (72,828 characters; 17% of the cells in the matrix were gaps or missing data; each taxon was represented by an average of 86% of the 82 genes; *SI Appendix*, Tables S2 and S3). The taxa were carefully selected to capture the basal nodes within deeply diverged families, such as Centroplacaceae and Euphorbiaceae (4);

Author contributions: Z.X., B.R.R., H.S., K.J.W., S.M., and C.C.D. designed research; Z.X., B.R.R., A.M.A., M.S., M.L.M., and C.C.D. performed research; Z.X., B.R.R., H.S., P.K.E., P.F.S., S.M., and C.C.D. analyzed data; and Z.X., B.R.R., H.S., K.J.W., P.K.E., S.M., and C.C.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. R.K.J. is a guest editor invited by the Editorial Board.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. JX661767–JX665032).

[1]To whom correspondence may be addressed. E-mail: cdavis@oeb.harvard.edu or smathews@oeb.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205818109/-/DCSupplemental.

they represent 39 of the 42 families of Malpighiales (excluding Lophopyxidaceae, Malesherbiaceae, and Rafflesiaceae; *SI Appendix, SI Materials and Methods*) and relevant outgroups. To obtain the most comprehensive phylogenetic tree for the order, we used the ***82-gene*** matrix as a scaffold to which we added the existing taxon-rich but character-sparse ***13-gene*** matrix (186 taxa; 15,574 characters; 15% missing data) (3) to create our ***combined-incomplete*** matrix (Table 1). This matrix included 191 taxa and 91 genes (82 plastid genes, six mitochondrial genes, and three nuclear genes; 81,259 characters; 64% missing data). We also created a ***combined-complete*** matrix by reducing the taxon sampling in the ***combined-incomplete*** matrix to match the taxon sampling of the ***82-gene*** matrix. This greatly reduced the percentage of missing data cells in our alignment from 64% to 12%. The ***combined-complete*** matrix included 58 taxa and 91 genes (81,117 characters). Finally, we reanalyzed the ***13-gene*** matrix alone to determine the phylogenetic impact of adding the ***82-gene*** matrix. Each of the four matrices was analyzed using four different data partitioning strategies that are described below.

Relationships in Malpighiales. Our analyses produced a well-resolved phylogeny of Malpighiales (Fig. 1; relationships of outgroups provided in *SI Appendix*, Fig. S1). The maximum likelihood (ML) and Bayesian trees inferred from the ***combined-incomplete*** matrix are congruent with trees from the remaining three matrices (i.e., ***82-gene***, ***combined-complete***, and ***13-gene***; *SI Appendix*, Figs. S1–S17 and S26–S28), using 75 ML bootstrap percentage (BP; as calculated using the standard bootstrap option in RAxML) and 0.95 Bayesian posterior probability (PP) thresholds. The 16 subclades of Malpighiales whose interrelationships were previously unresolved with respect to one another are resolved into three well-supported (>80 BP, 1.0 PP) clades. Moreover, we find comparable or greatly improved support for previously identified clades (3) and moderate to strong support for the 12 additional clades we identified (Fig. 1). Six of these clades were supported with >80 BP, 1.0 PP; two with ≥70 BP, >0.60 PP; and one with >60 BP, >0.95 PP. Importantly, each of the 12 clades is also united by morphological features (summarized in Table 2).

Clade 1 (85 BP, 1.0 PP) includes two major subclades: the euphorbioids (clade 4) and Humiriaceae + the parietal clade *sensu* Wurdack and Davis (3) (clade 7). Surprisingly, the euphorbioid clade (64 BP, 0.61 PP) reunites most of the former Euphorbiaceae (including Euphorbiaceae, Peraceae, Phyllanthaceae, and Picrodendraceae but excluding Putranjivaceae) (13, 14) with the well-supported (96 BP, 1.0 PP) linoid clade (clade 6; Ixonanthaceae + Linaceae) we identified. Within the euphorbioids, the linoids are well-supported (clade 5; 84 BP, 1.0 PP) as sister to the phyllanthoids (Phyllanthaceae + Picrodendraceae; 100 BP, 1.0 PP). The second major subclade identified here, clade 7 (62 BP, 0.79 PP), includes Humiriaceae and the parietal clade (100 BP, 1.0 PP). Within the parietal clade, Goupiaceae is sister to Violaceae (clade 9; 75 BP, 0.62 PP). Also within the parietal clade, (Malesherbiaceae (Passifloraceae + Turneraceae)) is sister to the salicoids [clade 8; (Lacistemataceae (Samydaceae (Salicaceae + Scyphostegiaceae)))); 96 BP, 1.0 PP].

Clade 2 (83 BP, 1.0 PP) includes three subclades in a trichotomy. Its first major subclade, clade 10 (70 BP, 0.81 PP), includes the previously identified (6, 15) clusioid clade [((Bonnetiaceae + Clusiaceae) (Calophyllaceae (Hypericaceae + Podostemaceae))); 100 BP, 1.0 PP] plus their sister group the ochnoids [(Ochnaceae (Medusagynaceae + Quiinaceae)); 100 BP, 1.0 PP]. The second subclade in clade 2 is the recently identified (3) rhizophoroids [(Ctenolophonaceae (Erythroxylaceae + Rhizophoraceae)); 100 BP, 1.0 PP]. The third subclade is the pandoids (clade 11; Irvingiaceae + Pandaceae; 64 BP, 0.97 PP).

Clade 3 (81 BP, 1.0 PP) consists of four subclades, three of which have been previously identified (3), in a polytomy. These four subclades are the chrysobalanoids [(Balanopaceae ((Chrysobalanaceae + Euphroniaceae) (Dichapetalaceae + Trigoniaceae))); 100 BP, 1.0 PP], the malpighioids [clade 12; (Centroplacaceae (Elatinaceae + Malpighiaceae)); 63 BP, 0.51 PP], the putranjivoids (Lophopyxidaceae + Putranjivaceae; 100 BP, 1.0 PP), and Caryocaraceae.

Improved Phylogenetic Resolution Results from a Posteriori Data Partitioning and Better Taxon Sampling. Several previous phylogenomic studies of angiosperms have applied a single substitution matrix in ML analyses to multiple-gene concatenated data sets (OnePart; e.g., refs. 8, 16, and 17). More recently, to better accommodate evolutionary rate heterogeneity across different characters, alignments have been partitioned a priori by gene (GenePart; e.g.,

Table 1. Characteristics of the four matrices and statistics of the best-scoring ML trees inferred from each of the four partitioning strategies

Matrix	Taxa/characters/ missing data %	Partitioning strategy	No. of partitions	Log-likelihood	AICc	ΔAICc	Coverage density	Fraction of triples	*D*	*d*	Terrace size
82-gene	58/72,828/17%	OnePart	1	−689042	1,378,328	166,322	1.00	1.00	1.00	1.00	1
		GenePart	82	−680357	1,362,435	150,429	0.88	1.00	1.00	1.00	1
		CodonPart	4	−680281	1,360,860	148,854	1.00	1.00	1.00	1.00	1
		MixtPart	13	**−605772**	**1,212,006**	**0**	1.00	1.00	1.00	1.00	1
Combined-complete	58/81,117/12%	OnePart	1	−739270	1,478,784	193,023	1.00	1.00	1.00	1.00	1
		GenePart	91	−728235	1,458,355	172,594	0.88	1.00	1.00	1.00	1
		CodonPart	4	−730551	1,461,401	175,640	1.00	1.00	1.00	1.00	1
		MixtPart	15	**−642632**	**1,285,761**	**0**	1.00	1.00	1.00	1.00	1
Combined-incomplete	191/81,259/64%	OnePart	1	−892791	1,786,362	234,881	1.00	1.00	1.00	1.00	1
		GenePart	91	−879681	1,761,794	210,313	0.36	0.93	0.00	0.97	14,025
		CodonPart	4	−883407	1,767,647	216,166	1.00	1.00	1.00	1.00	1
		MixtPart	20	**−775178**	**1,551,481**	**0**	1.00	1.00	1.00	1.00	1
13-gene	186/15,574/15%	OnePart	1	−292212	585,198	47,256	1.00	1.00	1.00	1.00	1
		GenePart	13	−288145	577,294	39,352	0.93	1.00	1.00	1.00	1
		CodonPart	4	−289988	580,807	42,865	1.00	1.00	1.00	1.00	1
		MixtPart	14	**−268460**	**537,942**	**0**	1.00	1.00	1.00	1.00	1

The MixtPart partitioning strategy is highlighted in bold to indicate that it produced the best log-likelihood and AICc values in each case. The fraction of triples is the fraction of all possible triples of taxa for which sequence data are present for at least some partitions of the matrix; *D* is the probability that a pattern of taxon coverage is decisive for a random binary tree; *d* is the average number of edges distinguishable on a random binary tree (further details in *SI Appendix, Materials and Methods*).

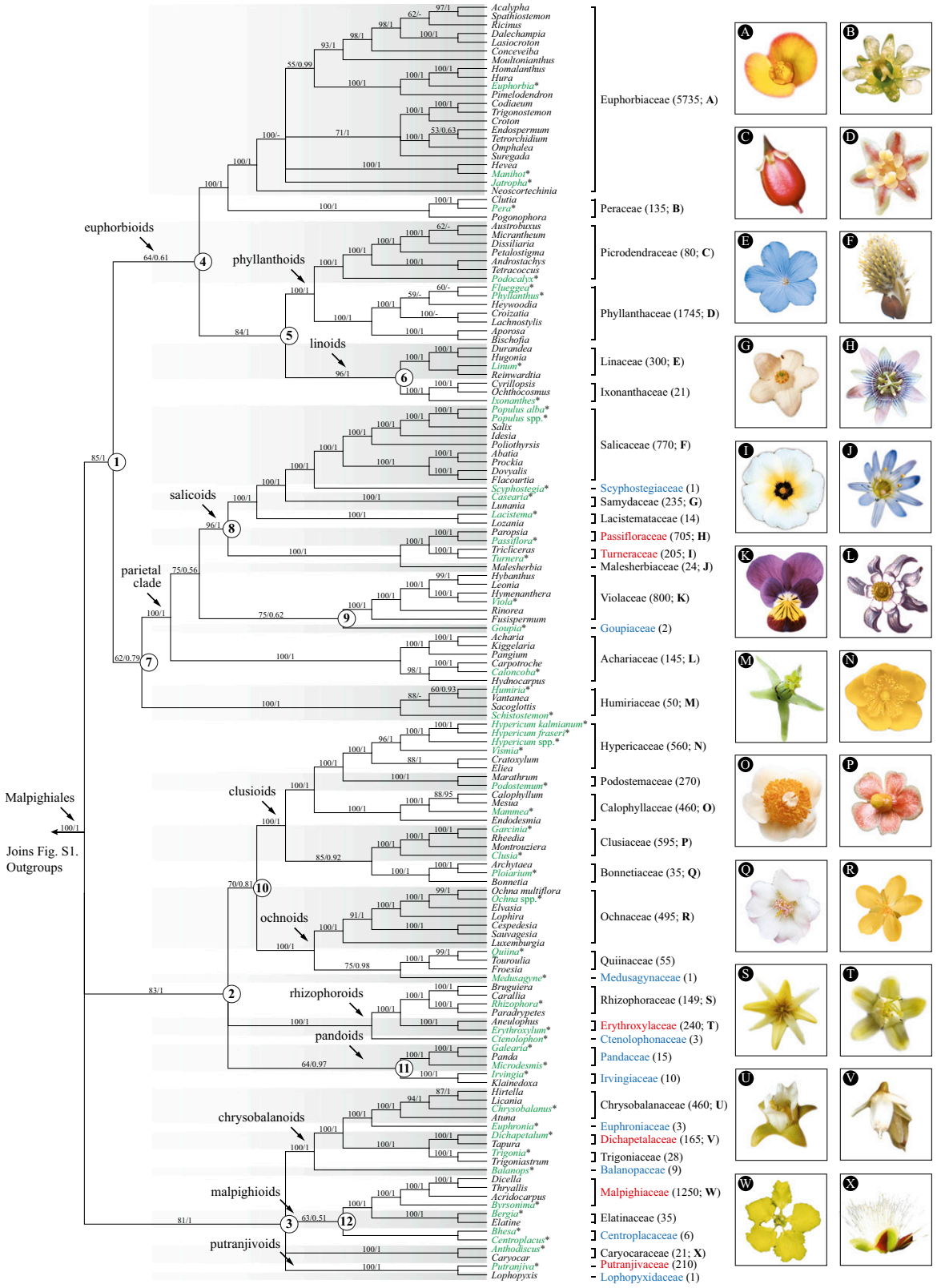

Fig. 1. ML 50% majority-rule bootstrap consensus tree of Malpighiales inferred from the ***combined-incomplete*** matrix using the MixtPart partitioning strategy. ML BPs/Bayesian PPs are indicated above each branch; a hyphen indicates that the node is not present in the Bayesian 50% majority-rule consensus tree. The 12 additional clades we identified are designated using the numbers corresponding to those in Table 2. Taxa included in the ***82-gene*** matrix are highlighted in green and marked with asterisks; spp. = composite terminals compiled from multiple closely related species; the approximate number of accepted species for each family is given in parentheses to the right; lettered photographs depicting representative species from all included families are shown to the right. Clades exhibiting a shift in net species diversification are highlighted in red (acceleration) and blue (deceleration). More detailed results of diversification analyses are provided in the text and *SI Appendix*. Outgroup relationships are shown in *SI Appendix*, Fig. S1.

Table 2. Morphological features for the 12 additional clades we identified in Malpighiales

Clade	Morphological features
1	Androgynophore; ovules mostly crassinucellar
2	Tendency to incompletely tenuinucellar ovules
3	Tendency to (oblique) floral monosymmetry in chrysobalanoids and Malpighiaceae; tendency to bulging of ovaries; placentation mostly axile; inner integument thicker than outer
4	Tendency to unisexual, trimerous flowers with reduced petals (not in Ixonanthaceae and Linaceae); petals, if present, often contort; placentation mostly axile; ovules 2 (more rarely 1) per carpel; antitropous, pendant, with obturator; inner integument thicker than outer
5	Tendency to false septa in carpels; placentation axile; ovules 2 per carpel, antitropous, pendant, with obturator; inner integument thicker than outer
6	Flowers bisexual; mostly diplostemonous; carpels with false septa
7	Flowers mostly haplostemonous (not in Humiriaceae); carpels often 3 (5 in *Humiria*); placentation parietal (not in Humiriaceae); ovules often more than 2 per carpel, crassinucellar, without endothelium; seeds often with aril (not in Humiriaceae)
8	Corona present in some families; placentation parietal; ovules mostly more than 2 per carpel, crassinucellar; seeds often with aril
9	Flowers haplostemonous; anthers with conspicuous appendages; nectary, if present, at outer base of stamens; ovules more than 2 per carpel
10	Petals often contort; mostly polystemonous; placentation mostly axile; ovules often incompletely tenuinucellar, with endothelium
11	Placentation axile; ovule 1 per carpel, antitropous
12	Placentation axile; ovules crassinucellar, without endothelium; sepals persistent in fruit

Data compiled from *SI Appendix*, SI refs. 24 and 28–34. Clades are labeled in Fig. 1 accordingly.

refs. 11 and 12) or by codon position (CodonPart; e.g., refs. 11 and 18). The GenePart approach creates a partition for each gene and estimates the substitution rate matrix parameters separately for each partition, resulting in up to 83 partitions for many plastid data sets. The CodonPart approach partitions characters according to codon position, with a fourth partition added for noncoding regions (if present). These partitioning strategies are somewhat arbitrary, assuming for example that all third codon positions evolve rapidly or that gene boundaries define a class of sites that are expected to share a similar model of molecular evolution. As an alternative, we explored the use of an a posteriori partitioning strategy for ML analyses based on the partitions inferred from Bayesian searches of the matrix using a mixture model approach (19). Using a reversible-jump implementation, the Bayesian mixture model estimates the number of substitution rate matrices that best fit an alignment by allowing the fitting of multiple rate matrices to each character separately (20). We used this approach to find the optimal number of partitions for each matrix and then defined the characters in each class as a partition for subsequent ML analyses (MixtPart).

Using MixtPart, we found that the optimal number of partitions was 13 for the ***82-gene*** matrix, 15 for the ***combined-complete*** matrix, and 20 for the ***combined-incomplete*** matrix. Thus, in all cases, defining the partitions on the basis of the mixture model search reduced the number of partitions substantially (from 82 for the ***82-gene*** and from 91 for the two combined matrices using GenePart). Notably, our results show that using MixtPart substantially improved the likelihood of the best-scoring ML tree as measured by the corrected Akaike information criterion (AICc) (21) for all four matrices (Table 1). For example, compared with the GenePart approach, the MixtPart approach increased the AICc values by 7–12%. MixtPart also outperformed the OnePart, GenePart, and CodonPart approaches with respect to improving the branch support as measured by BP values. To compare these BP values among trees with different taxon sets, the bipartition trees inferred from the ***combined-incomplete*** and ***13-gene*** matrices (*SI Appendix*, Figs. S10–S17) were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices (*SI Appendix*, Figs. S18–S25). This revealed that the use of MixtPart resulted in an increase in mean BP values by 5–11% (Fig. 2 and *SI Appendix*, Table S4) and most strikingly a mean increase in BP values by 20–49% for the 12 clades we identified (Fig. 3). It should be noted that the addition of our ***82-gene*** matrix alone was insufficient to resolve the deeper nodes of Malpighiales. Although it was helpful [e.g., mean BP values increased by 13% when comparing between the ***82-*** vs. ***13-gene*** MixtPart analyses (Fig. 2)], only 4 of these 12 clades were supported with >50 BP using OnePart, GenePart, and CodonPart, vs. 10 clades that were resolved with >50 BP using MixtPart (Fig. 3). This indicates that the use of MixtPart results in substantial improvement.

Our results also suggest that for the commonly used partitioning strategies, particularly for OnePart and GenePart, increased taxon sampling improves branch support, regardless of

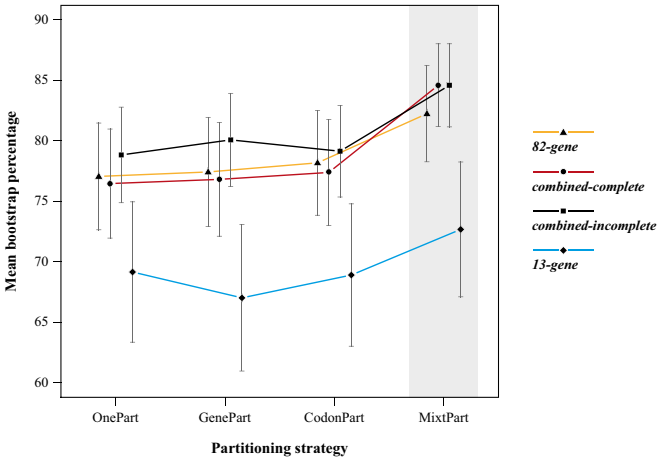

Fig. 2. Mean ML BPs of the bipartition trees inferred using ML for each of the four matrices and four partitioning strategies. SEs around the means are indicated, and the MixtPart partitioning strategy is highlighted in gray. The bipartition trees inferred from the ***combined-incomplete*** and ***13-gene*** matrices were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices.

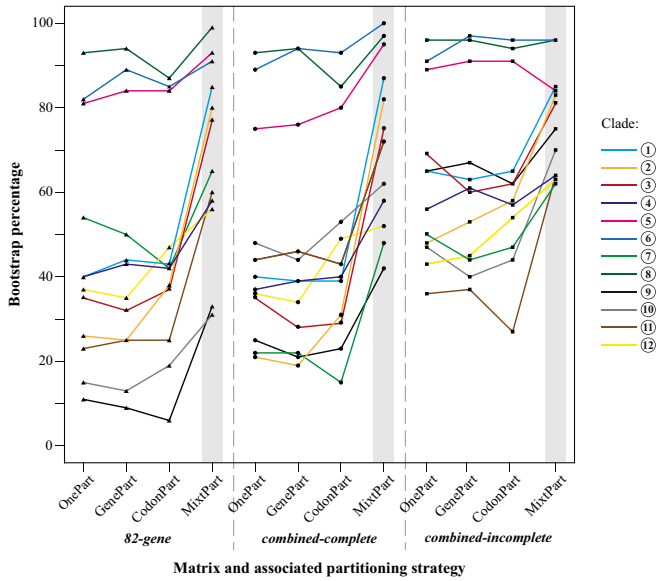

Fig. 3. ML BPs of the 12 additional clades we identified in Malpighiales (Fig. 1) inferred from three matrices and four partitioning strategies. The MixtPart partitioning strategy is highlighted in gray.

the increase in missing data. For example, despite its much higher percentage of missing data (64% vs. 12%), analyses of the 191-taxon ***combined-incomplete*** matrix yielded a better-supported phylogeny than the 58-taxon ***combined-complete*** matrix: the mean BP values increased by 3% and 4% for OnePart and GenePart, respectively (Fig. 2). Although this improvement might seem relatively small when comparing mean BP values, it is much more impressive for the 12 clades we identified, which showed an average increase of BP values by 34% and 36% for the OnePart and GenePart analyses, respectively (Fig. 3). These results provide empirical support for conclusions that increased taxon sampling improves phylogenetic accuracy (22–24), even when the amount of missing data increases (25–27). Theoretical studies (e.g., refs. 28 and 29) have shown that it is the number of complete characters rather than the number of empty cells that determines the impact of missing data on phylogenetic accuracy. The improved branch support we observed in trees from the ***combined-incomplete*** matrix likely results from our strategic scaffold approach, in which we ensured that critical nodes were deeply sampled for most characters. A similar scaffold approach was advocated by Wiens et al. (30) and more recently applied using large amounts of genomic data to successfully resolve relationships of butterflies and moths (31).

Despite the apparent successes of the scaffold approach, recent studies (32, 33) have shown that partial taxon coverage (whereby sequences from some partitions are missing for some taxa) can result in a vast terrace of phylogenetic trees that have different topologies but the same optimality score. In cases where complete taxon coverage for a partition is achieved the data set is expected to be decisive for all trees (32), and under these circumstances the problem of terraces does not arise (33). This is likely to be rare for large phylogenomic data sets, however, which sacrifice completeness for the additional taxa and characters. This problem was most clearly illustrated in the recent analysis of a 298-taxon grass data set with 34% missing data that produced a terrace including 61 million optimal trees (33). We found that different partitioning strategies induced different patterns of taxon coverage. Notably, the use of GenePart reduced taxon coverage density in all cases, and in the case of the ***combined-incomplete*** matrix it resulted in a pattern of taxon coverage that was indecisive and the best-scoring ML tree was on a terrace of 14,025 trees, whereas the use of MixtPart was decisive for all trees (Table 1). Despite this lack of decisiveness, the BUILD tree (i.e., the Adams consensus of the 14,025 trees) includes only four polytomies, all of which are restricted to subfamily relationships (*SI Appendix*, Fig. S29). Thus, our scaffold approach yielded a matrix that is resilient to reduced coverage density. Together our results suggest that there may be cases, depending on the patterns of taxon coverage, in which GenePart would be a poor choice for partitioning concatenated matrices.

Patterns of Species Diversification in Malpighiales. Studies of diversification patterns across angiosperms have not previously detected shifts in net species diversification rates (speciation minus extinction) in Malpighiales (34, 35), possibly because a well-resolved, taxon-dense phylogeny was not available for the order. We used our 191-taxon, ***combined-incomplete*** matrix to test the hypothesis that net diversification rates have been constant throughout the history of Malpighiales. This matrix was originally constructed to include the deepest phylogenetic splits within each family (3, 4) and is an excellent foundation for exploring the tempo of evolution in the order. We first used the approach implemented in MEDUSA (36) to detect shifts of diversification rate using a time-calibrated Malpighiales phylogeny (*SI Appendix*, Fig. S30) that accounts for unsampled taxonomic diversity (*SI Appendix*, Fig. S31). This method sequentially adds break points to a multirate birth-and-death model fitting the given branch lengths and terminal diversities until subsequent break points do not improve the AICc values. Using MEDUSA we found significant decelerations in five clades (Goupiaceae, Lophopyxidaceae, Medusagynaceae, Scyphostegiaceae, and Irvingiaceae + Pandaceae) and acceleration in one clade (Passifloraceae + Turneraceae) (Fig. 1 and *SI Appendix*, Fig. S31).

Additionally, we used a method that models diversification as a stochastic, time-homogeneous birth-and-death process (34). This method does not use the phylogeny directly but considers stem or crown group ages within clades of interest and the survival of each lineage to the present. The results were similar to those from the phylogeny-based MEDUSA analysis, with the main difference being the detection of an additional four rate decelerations and four accelerations. Assuming a constant birth-and-death model, eight clades (Balanopaceae, Centroplacaceae, Ctenolophonaceae, Euphroniaceae, Goupiaceae, Lophopyxidaceae, Medusagynaceae, and Scyphostegiaceae) experienced decelerations, and five clades (Dichapetalaceae, Erythroxylaceae, Malpighiaceae, Passifloraceae, and Putranjivaceae) experienced accelerations (Fig. 1 and *SI Appendix*, Fig. S32).

These overlapping results, together with a well-resolved phylogeny, provide an improved foundation for exploring the mechanisms that have led to such substantial diversity within Malpighiales. In some cases (e.g., Malpighiaceae and Passifloraceae), specialized plant–pollinator mutualisms (37–39) may account for all or part of their exceptional diversification rates. These and other hypotheses can now be tested in more detailed studies of phylogeny, morphology, ecology, and biogeography.

Conclusions

Our phylogeny of Malpighiales provides a critical context for future comparative studies of plant species that are economically and ecologically important. Although the increasing ease of genome-scale sampling may render moot the long-standing argument about whether it is better to add taxa or characters (40), the question remains important. Given the amount of biodiversity remaining to be discovered, described, and classified, the goal should be to maximize taxonomic sampling for phylogenetic study, but to do so in the most effective way possible. Our analyses confirm that one efficient and economical way to resolve difficult clades is to construct a scaffold using phylogenetically critical placeholders sampled for many characters augmented by many

more taxa sampled for a modest number of characters. Most importantly, our analyses indicate that searching with a Bayesian mixture models leads to an optimal, a posteriori data partitioning strategy, which not only improves the branch support of phylogenetic trees but also minimizes the impact of missing data on phylogenetic decisiveness. Its use is likely to help resolve several remaining poorly resolved, major clades of angiosperms (e.g., Euasterids I and II and Ericales) (12) and to be more broadly useful in studies across the Tree of Life.

Materials and Methods

See *SI Appendix, SI Materials and Methods* for details on plastome sequencing, sequence alignment, and analyses of phylogenetic decisiveness, divergence time, and species diversification.

Phylogenetic Analyses. Bayesian and ML analyses were performed on four matrices (Table 1) as described above. The Bayesian analyses were implemented with the parallel version of BayesPhylogenies v2.0 (19) using a reversible-jump implementation of the mixture model as described by Venditti et al. (20). This approach allows the fitting of multiple models of sequence evolution to each character in an alignment without a priori partitioning. Two independent Markov chain Monte Carlo (MCMC) analyses were performed, and the consistency of stationary-phase likelihood values and estimated parameter values was determined using Tracer v1.5. We ran each MCMC analysis for 10 million generations, sampling trees and parameters every 1,000 generations. Bayesian PPs were determined by building a 50% majority-rule consensus tree from two MCMC analyses after discarding the 20% burn-in generations (Fig. 1 and *SI Appendix*, Figs. S1 and S26–S28).

The ML analyses were conducted using RAxML v7.2.8 (41) with the GTR+Γ model. The best-scoring ML tree was obtained for each matrix using the rapid hill-climbing algorithm (41), and 1,000 bootstrap replicates were estimated using the standard bootstrap option. The BPs were summarized from all 1,000 bootstrap trees, and the bipartition tree was obtained by mapping these BPs to the best-scoring ML tree (*SI Appendix*, Figs. S2–S17) (42). We used four different partitioning strategies for our data analyses described above in *Results and Discussion*: OnePart (single data partition), GenePart (partitioned by gene), CodonPart (partitioned by codon), and MixtPart (described below). For the MixtPart approach, the data partitions identified in the Bayesian analyses were extracted from the output using a custom Perl script (*SI Appendix, SI Script*). This script selected the partition with the highest probability for each character. The matrices were then partitioned accordingly in RAxML.

ACKNOWLEDGMENTS. We thank D. Barua, J. Beaulieu, M. Clements, R. Cronn, M. Ethier, D. Goldman, M. Guisinger-Bellian, R. Jansen, M. Kent, M. McMahon, A. Meade, M. Moore, M. Sanderson, A. Stamatakis, and members of the C.C.D. and S.M. laboratories for technical assistance. This work was supported by Brazil Conselho Nacional de Desenvolvimento Científico e Tecnológico Grant 563548/10-0 (to A.M.A.), Swiss National Science Foundation Grant 129804 (to P.K.E.), US National Science Foundation (NSF) Assembling the Tree of Life Grants DEB-0622764 and DEB-1120243 (to C.C.D.), and NSF Doctoral Dissertation Enhancement Project Grant OISE-0936076 (to C.C.D. and B.R.R.).

1. Chase MW, et al. (1993) Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann Mo Bot Gard* 80:528–580.
2. APG (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Bot J Linn Soc* 141:399–436.
3. Wurdack KJ, Davis CC (2009) Malpighiales phylogenetics: Gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *Am J Bot* 96(8):1551–1570.
4. Davis CC, Webb CO, Wurdack KJ, Jaramillo CA, Donoghue MJ (2005) Explosive radiation of Malpighiales supports a mid-cretaceous origin of modern tropical rain forests. *Am Nat* 165(3):E36–E65.
5. Sato S, et al. (2011) Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L. *DNA Res* 18(1):65–76.
6. Chan AP, et al. (2010) Draft genome sequence of the oilseed species *Ricinus communis*. *Nat Biotechnol* 28(9):951–956.
7. Tuskan GA, et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313(5793):1596–1604.
8. Jansen RK, et al. (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci USA* 104(49):19369–19374.
9. Moore MJ, Bell CD, Soltis PS, Soltis DE (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc Natl Acad Sci USA* 104 (49):19363–19368.
10. Wang H, et al. (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proc Natl Acad Sci USA* 106(10):3853–3858.
11. Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE (2010) Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc Natl Acad Sci USA* 107(10):4623–4628.
12. Soltis DE, et al. (2011) Angiosperm phylogeny: 17 genes, 640 taxa. *Am J Bot* 98(4): 704–730.
13. Cronquist A (1988) *The Evolution and Classification of Flowering Plants* (New York Botanical Garden, Bronx, NY), 2nd Ed.
14. Webster GL (1994) Classification of the Euphorbiaceae. *Ann Mo Bot Gard* 81:3–32.
15. Ruhfel BR, et al. (2011) Phylogeny of the clusioid clade (Malpighiales): Evidence from the plastid and mitochondrial genomes. *Am J Bot* 98(2):306–325.
16. Cai ZQ, et al. (2006) Complete plastid genome sequences of *Drimys*, *Liriodendron*, and *Piper*: Implications for the phylogenetic relationships of magnoliids. *BMC Evol Biol* 6: 77.
17. Hansen DR, et al. (2007) Phylogenetic and evolutionary implications of complete chloroplast genome sequences of four early-diverging angiosperms: *Buxus* (Buxaceae), *Chloranthus* (Chloranthaceae), *Dioscorea* (Dioscoreaceae), and *Illicium* (Schisandraceae). *Mol Phylogenet Evol* 45(2):547–563.
18. Moore MJ, et al. (2011) Phylogenetic analysis of the plastid inverted repeat for 244 species: Insights into deeper-level angiosperm relationships from a long, slowly evolving sequence region. *Int J Plant Sci* 172:541–558.
19. Pagel M, Meade A (2004) A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst Biol* 53(4):571–581.
20. Venditti C, Meade A, Pagel M (2008) Phylogenetic mixture models can reduce node-density artifacts. *Syst Biol* 57(2):286–293.
21. Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. *Biometrika* 76:297–307.
22. Pollock DD, Zwickl DJ, McGuire JA, Hillis DM (2002) Increased taxon sampling is advantageous for phylogenetic inference. *Syst Biol* 51(4):664–671.
23. Zwickl DJ, Hillis DM (2002) Increased taxon sampling greatly reduces phylogenetic error. *Syst Biol* 51(4):588–598.
24. Hedtke SM, Townsend TM, Hillis DM (2006) Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Syst Biol* 55(3):522–529.
25. McMahon MM, Sanderson MJ (2006) Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. *Syst Biol* 55(5):818–836.
26. Heath TA, Hedtke SM, Hillis DM (2008) Taxon sampling and the accuracy of phylogenetic analyses. *J Syst Evol* 46:239–257.
27. Burleigh JG, Hilu KW, Soltis DE (2009) Inferring phylogenies with incomplete data sets: a 5-gene, 567-taxon analysis of angiosperms. *BMC Evol Biol* 9:61.
28. Wiens JJ (2003) Missing data, incomplete taxa, and phylogenetic accuracy. *Syst Biol* 52 (4):528–538.
29. Wiens JJ (2005) Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? *Syst Biol* 54(5):731–742.
30. Wiens JJ, Fetzner JW, Parkinson CL, Reeder TW (2005) Hylid frog phylogeny and sampling strategies for speciose clades. *Syst Biol* 54(5):778–807.
31. Cho S, et al. (2011) Can deliberately incomplete gene sample augmentation improve a phylogeny estimate for the advanced moths and butterflies (Hexapoda: Lepidoptera)? *Syst Biol* 60(6):782–796.
32. Sanderson MJ, McMahon MM, Steel M (2010) Phylogenomics with incomplete taxon coverage: The limits to inference. *BMC Evol Biol* 10:155.
33. Sanderson MJ, McMahon MM, Steel M (2011) Terraces in phylogenetic tree space. *Science* 333(6041):448–450.
34. Magallón S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. *Evolution* 55(9):1762–1780.
35. Smith SA, Beaulieu JM, Stamatakis A, Donoghue MJ (2011) Understanding angiosperm diversification using small and large phylogenetic trees. *Am J Bot* 98(3): 404–414.
36. Alfaro ME, et al. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci USA* 106(32):13410–13414.
37. Anderson WR (1979) Floral conservatism in neotropical Malpighiaceae. *Biotropica* 11: 219–223.
38. Neff JL (2003) The passionflower bee: *Anthemurgus passiflorae*. *J Newsl Passiflora Soc Int* 13:7–9.
39. Zhang W, Kramer EM, Davis CC (2010) Floral symmetry genes and the origin and maintenance of zygomorphy in a plant-pollinator mutualism. *Proc Natl Acad Sci USA* 107(14):6388–6393.
40. Graybeal A (1998) Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst Biol* 47(1):9–17.
41. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690.
42. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57(5):758–771.

SI Appendix for "Phylogenomics and *a posteriori* data partitioning resolve the Cretaceous angiosperm radiation Malpighiales"

Xi et al.

SI Materials and Methods

Plastid isolation, Amplification, and Sequencing. Plastids were isolated from ~20 grams of fresh leaf material using the sucrose gradient centrifugation protocols of Jansen et al. (1). DNA extracted from purified plastids was amplified with the REPLI-g Midi Kit (Qiagen, Inc.). When we were unable to acquire fresh material, total genomic DNA was extracted from silica-dried leaf material following the protocols of Wurdack and Davis (2). In these instances, the plastid genome was amplified with 32 pairs of primers (Table S5) using long-range PCR amplification. These primers were designed from the plastid genome sequences of *Arabidopsis thaliana*, *Calycanthus floridus*, *Cucumis sativus*, *Jatropha curcas*, *Manihot esculenta*, *Medicago truncatula*, *Nicotiana tabacum*, *Populus alba*, *Populus trichocarpa*, and *Vitis vinifera* using Primaclade (3).

Lophopyxidaceae and Malesherbiaceae were not sampled for the ***82-gene*** matrix because appropriate samples were unavailable. However, these two families have been very well-placed as members of recently derived clades that we included here (2, 4). Thus, their exclusion should not affect our results. Additionally, it was not possible to include the holoparasitic Rafflesiaceae in the ***82-gene*** matrix because its plastid genome appears to have been lost entirely (5). The current phylogenetic placement of Rafflesiaceae as sister to Euphorbiaceae is largely based on slowly evolving mitochondrial genes (2, 6). Rafflesiaceae were placed as sister to Euphorbiaceae + Peraceae when analyzed here with the ***combined-incomplete*** matrix, which is similar to earlier analyses using parsimony (2). However, this placement is likely to be unreliable because including a taxon with such a strong bias in sampled genes (i.e., specifically, slowly evolving mitochondrial genes present versus more rapidly evolving plastid genes absent) is likely to cause spurious phylogenetic placements (7).

For each taxon, a library was prepared from five micrograms of amplified DNA and then indexed for Illumina multiplex sequencing following the

protocols of Bentley et al. (8). All libraries were sequenced on the Genome Analyzer II (Illumina, Inc.) at the FAS Center for Systems Biology at Harvard University.

Sequence Assembly and Alignment. Multiplexed Illumina reads were first sorted by indices, and then assembled *de novo* in Velvet v1.0.9 (9). To avoid sequencing errors, and potential contamination from mitochondrial and/or nuclear genomes, the coverage cut-off was set to 25 to remove low coverage nodes in Velvet. The assembled contigs were annotated against the *Manihot esculenta* plastid genome (10) using BLASTX v2.2.23 (11). The gene coding sequences were extracted from the BLASTX output. Although a variable number of plastid genes were recovered for each taxon, either due to natural variation in gene content or due to failure of our long-range PCR amplifications, most taxa have excellent coverage of the 82 plastid genes included in our analyses (mean=86%, median=95%; Tables S2 and S3). Amino acid sequences from all taxa were queried against themselves using BLASTP v2.2.23 (11). BLASTP hits with an *e*-value $\leq 10^{-5}$ were passed to MCL v08-312 (12) for Markov clustering. The amino acid sequences for each of these cluster groups were aligned with MAFFT v6.624 (13), and trimmed to exclude poorly aligned positions using the default settings in Gblocks v0.91b (14). The protein-coding nucleotide sequences for each cluster group were aligned based on corresponding amino acid alignments using PAL2NAL v12.2 (15) to ensure the correct reading frame.

Phylogenetic Decisiveness and Tree Terraces. Phylogenetic decisiveness and tree terraces were determined for each matrix and associated partitioning strategies following refs. 16 and 17, respectively. *D* (the probability that a pattern of taxon coverage is decisive for a random binary tree) and *d* (the average number of edges distinguishable on a random binary tree) were estimated from 1,000 replicate simulated random trees using the equiprobable option in PAUP* (18). *D* is used to determine phylogenetic decisiveness: when *D* equals one the matrix is decisive for all trees; when less than one it is decisive for some or none of the trees. *d* is a measure of "partial" decisiveness that allows us to determine if the internal edges of a random binary tree can be distinguished by the associated

taxon coverage (16). A high *d* indicates that there is a better chance to resolve most edges of the tree. We also calculated the fraction of triples, which is the fraction of all possible triples of taxa for which sequence data are present for at least some partitions of the matrix, and a value of one is a necessary condition for decisiveness for all trees (16). Finally, to determine the tree terrace sizes, we first generated pruned subtrees for each partition that were derived from the best-scoring ML tree in which all taxa were included. The set of pruned subtrees was further reduced to a set of rooted triplets, and all binary parent trees that were compatible with these triplets were counted and summarized in Table 1. Additionally, to better visualize all the tree topologies on a single terrace, we constructed the BUILD tree (i.e., the Adams consensus tree; Fig. S29) by summarizing all binary parent trees according to ref. 19, and most recently employed by ref. 17.

Divergence Time Estimates. A Bayesian method implemented in BEAST v1.6.1 (20) with an uncorrelated lognormal (UCLN) model was used to estimate divergence times within Malpighiales. The ***13-gene*** matrix was used for this purpose to avoid the effect of missing data and to reduce the computational burden. Divergence times were estimated under a GTR+Γ model of nucleotide substitution across all characters, and the tree topology was constrained with the best-scoring ML tree inferred from the ***combined-incomplete*** matrix, due to its greater resolution. The MCMC analysis was initiated from an ultrametric tree created using PATHd8 (21) with branch lengths that satisfied the priors on all divergence times. All fossils (16 total; Table S6) were treated as minimum age constraints, and the root node was set to a uniform distribution between 120 Ma and 135 Ma following ref. 22. We modeled all fossil constraints as a lognormal distribution with a mean of 1.5 and standard deviation of 0.5.

For the MCMC analysis, we ran two independent chains for 20 million generations sampling trees and parameters every 1,000 generations. Tracer v1.5 was used to check for convergence of the model likelihood and parameters between the two chains. The resulting log files were combined from both chains after discarding the 20% burn-in generations using LogCombiner v1.6.1, and results were considered reliable once the effective sampling size (ESS) for all

parameters exceeded 200. The mean node ages and 95% highest posterior density interval (HPD) of divergence time estimates were then summarized using TreeAnnonator v1.6.1 (Fig. S30 and Table S7).

Species Diversification Analyses. A family-level chronogram for Malpighiales was created by pruning the 191-taxon chronogram described above (Fig. S30) such that each monophyletic family is represented by a single taxon (Fig. S31). An incorrect clade age will result in under or over estimates of diversification rates if the true clade age is older or younger, respectively. However, we are confident that we have found reliable crown group ages for most of the included families. The one exception is Podostemaceae where our sampling does not capture the earliest divergence within this family. Crown group age for Podostemaceae was therefore taken from a study focusing on the clusioids (23), where a much more comprehensive sampling properly spanned the crown group node of this family. Species numbers for each family were obtained from the Angiosperm Phylogeny Website (24) to accommodate unsampled standing taxonomic diversity. The family-level chronogram with the species richness for each family was used as the input data (Fig. S31).

To model patterns of diversification, we first applied the MEDUSA method (25) as implemented in Turbo-MEDUSA v0.1 with the default settings (i.e., up to 20 piecewise models were used to fit our family-level chronogram, and AICc was used to determine whether larger models fit the data significantly better). Secondly, to apply the approach of Magallón and Sanderson (26) we calculated net diversification rates using the age estimates of all Malpighiales families included in our divergence time analyses (with the modified age for Podostemaceae described above). We plotted clade age and clade size for Malpighiales families on a semilog scale and compared them with expected values under low (ε=0.0) and high (ε=0.9) rates of extinction using GEIGER v1.3-1 (27) in R. The background rate of speciation was calculated separately for the three major clades identified in Malpighiales (clade1: 0.08046 for ε=0.0 and 0.06494 for ε=0.9; clade 2: 0.06799 for ε=0.0 and 0.05250 for ε=0.9; clade 3: 0.06561

for ε=0.0 and 0.05012 for ε=0.9) to detect more local shifts in net diversification rates (Fig. S32).

Morphological Features. The morphological features supporting the 12 additional clades identified in our analyses (Fig. 1 and Table 2) were compiled from refs. 24 and 28-34.

SI References

1. Jansen RK, et al. (2005) Methods for obtaining and analyzing whole chloroplast genome sequences. *Methods Enzymol* 395:348-384.
2. Wurdack KJ, Davis CC (2009) Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *Am J Bot* 96:1551-1570.
3. Gadberry MD, Malcomber ST, Doust AN, Kellogg EA (2005) Primaclade - a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics* 21:1263-1264.
4. Davis CC, Webb CO, Wurdack KJ, Jaramillo CA, Donoghue MJ (2005) Explosive radiation of Malpighiales supports a mid-Cretaceous origin of tropical rain forests. *Am Nat* 165:E36-E65.
5. Nickrent DL, Ouyang Y, Duff RJ, dePamphilis CW (1997) Do nonasterid holoparasitic flowering plants have plastid genomes? *Plant Mol Biol* 34:717-729.
6. Davis CC, Latvis M, Nickrent DL, Wurdack KJ, Baum DA (2007) Floral gigantism in Rafflesiaceae. *Science* 315:1812.
7. Wiens JJ (2003) Incomplete taxa, incomplete characters, and phylogenetic accuracy: is there a missing data problem? *J Vert Paleontol* 23:297-310.
8. Bentley DR, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456:53-59.
9. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821-829.
10. Daniell H, et al. (2008) The complete nucleotide sequence of the cassava (*Manihot esculenta*) chloroplast genome and the evolution of *atpF* in

Malpighiales: RNA editing and multiple losses of a group II intron. *Theor Appl Genet* 116:723-737.

11. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402.
12. Enright AJ, van Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30:1575-1584.
13. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059-3066.
14. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540-552.
15. Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 34:W609-W612.
16. Sanderson MJ, McMahon MM, Steel M (2010) Phylogenomics with incomplete taxon coverage: the limits to inference. *BMC Evol Biol* 10:155.
17. Sanderson MJ, McMahon MM, Steel M (2011) Terraces in phylogenetic tree space. *Science* 333:448-450.
18. Swofford DL (2002) *PAUP*. Phylogenetic analysis using parsimony (*and Other Methods)*. 4th edition. (Sinauer Associates, Sunderland, MA).
19. Aho AV, Sagiv Y, Szymanski TG, Ullman JD (1981) Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions. *SIAM J Comput* 10:405-421.
20. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4:e88.
21. Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating divergence times in large phylogenetic trees. *Syst Biol* 56:741-752.
22. Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-revisited. *Am J Bot* 97:1296-1303.

23. Ruhfel BR (2011) Systematics and biogeography of the clusioid clade (Malpighiales). Doctor of Philosophy (Harvard University, Cambridge, MA).
24. Stevens PF (2001 onwards) Angiosperm Phylogeny Website. http://www.mobot.org/MOBOT/research/APweb/.
25. Alfaro ME, et al. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci USA* 106:13410-13414.
26. Magallón S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. *Evolution* 55:1762-1780.
27. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: investigating evolutionary radiations. *Bioinformatics* 24:129-131.
28. Sutter DM, Forster PI, Endress PK (2006) Female flowers and systematic position of Picrodendraceae (Euphorbiaceae s.l., Malpighiales). *Pl Syst Evol* 261:187-215.
29. Matthews ML, Endress PK (2008) Comparative floral structure and systematics in Chrysobalanaceae s.l. (Chrysobalanaceae, Dichapetalaceae, Euphroniaceae, Trigoniaceae; Malpighiales). *Bot J Linn Soc* 157:249-309.
30. Endress PK (2011) Angiosperm ovules: diversity, development, evolution. *Ann Bot* 107:1465-1489.
31. Matthews ML, Endress PK (2011) Comparative floral structure and systematics in Rhizophoraceae, Erythroxylaceae and the potentially related Ctenolophonaceae, Linaceae, Irvingiaceae and Caryocaraceae (Malpighiales). *Bot J Linn Soc* 166:331-416.
32. Tobe H, Raven PH (2011) Embryology of the Irvingiaceae, a family with uncertain relationships among the Malpighiales. *J Plant Res* 124:577-591.
33. Zhang W, Kramer EM, Davis CC (2012) Similar genetic mechanisms underlie the parallel evolution of floral phenotypes. *PLoS ONE* 7:e36033.
34. Matthews ML, Amaral MCE, Endress PK (in press) Comparative floral structure and systematics in Ochnaceae s.l. (Ochnaceae, Quiinaceae and Medusagynaceae; Malpighiales). *Bot J Linn Soc*.

SI Script

Perl script that was used to extract the data partitions identified in the Bayesian analyses for the MixtPart approach

```
#!/user/bin/perl

use strict;
use warnings;
use Cwd;

my $dir = getcwd;

##############################################################################################
###########

&Usage();


die "\n\nCannot find <SiteBreakDown> in current working directory <$dir>!\n\nPlease
require the program <SiteBreakDown> for BayesPhylogenies from Dr. Andrew
Meade\n(a.meade\@reading.ac.uk) at University of Reading\n\n\n" unless -e
"SiteBreakDown";


if ((defined $ARGV[0]) && ($ARGV[0] =~ /^-+?\w+$/)) {
        die "\nUsage: perl $0 SiteBySite.txt #Rates #Patterns #Branch_Sets\n\n";
}


die "\nUsage: perl $0 SiteBySite.txt #Rates #Patterns #Branch_Sets\n\n" unless defined
$ARGV[0] && $ARGV[1] && $ARGV[2] && $ARGV[3];

my ($INFILE, $nrate, $npattern, $nbranch) = @ARGV[0,1,2,3];
die "\nCannot find <$INFILE>!\n\n" unless -e $INFILE;


system "chmod +x SiteBreakDown";
system "./SiteBreakDown $INFILE $nrate $npattern $nbranch > $INFILE.out";

##############################################################################################
###########

my @PATTERN;
my ($nchar, $site, $j);
my $i = 0;

open FH1, "<", "$INFILE.out";
while (<FH1>) {
        if (/^Matrix: Patterns Lh/) {
                $i = 1;
        }
        elsif ($i == 0) {
                next;
        }
        elsif (/^[01]\./) {
                $site++;
                chomp;

                my @TEMP = split /\t/;
                die "\n<$_>!\n\n" unless scalar(@TEMP) == $npattern;

                ##################################

                my ($pattern, $max) = (0, 0);


                for ($j = 0; $j < $npattern; $j++) {
                        if ($TEMP[$j] > $max) {
```

```
                                ($pattern, $max) = ($j, $TEMP[$j]);
                        }
                }

                push @{$PATTERN[$pattern]}, $site;
        }
        elsif (/^(\d+?)\s+?(\d+)$/) {
                ($nchar, $npattern, $site) = ($1, $2, 0);
        }
}
close FH1;
die "\nIncorrect file: <$INFILE>!\n\n" unless ($site == $nchar) && (scalar(@PATTERN) ==
$npattern);

##############################################################################################
###########

open FH2, ">", "Partitions.txt";

for ($i = 0; $i < $npattern; $i++) {

        print FH2 "DNA, Partition", $i+1, " = ";

        for ($j = 0; $j < $#{$PATTERN[$i]}; $j++) {
                print FH2 "$PATTERN[$i][$j],";
        }

        print FH2 "$PATTERN[$i][-1]\n";
}
close FH2;

##############################################################################################
###########

print "\n# chars: $nchar\n";
print "# patterns: $npattern\n\n";


for ($i = 0; $i < $npattern; $i++) {
        print "# chars in Partition", $i+1, ": ", scalar(@{$PATTERN[$i]}), "\n";
}

print "\n";

##############################################################################################
###########

sub Usage {


print <<EOF;


##############################################################################################
#########
#
#
# This is a Perl script for extracting the partitions of an alignment found by
BayesPhylogenies, #
# and outputing the partitions in the format required for subsequent RAxML analyses.
#
#
#
##############################################################################################
#########

EOF

}
```

Table S1. Overview of phylogenetic trees inferred for each of the matrices and partitioning strategies in this study.

Matrix	Taxa/characters missing data%	Phylogenetic method	Partitioning strategy	Bipartition tree	Pruned bipartition tree	50% majority-rule consensus tree
82-gene	58/72,828 17.0%	Maximum likelihood	OnePart	Fig. S2	-	-
		Maximum likelihood	GenePart	Fig. S3	-	-
		Maximum likelihood	CodonPart	Fig. S4	-	-
		Maximum likelihood	MixtPart	Fig. S5	-	-
		Bayesian	Mixture model	-	-	Fig. S26
combined-complete	58/81,117 11.9%	Maximum likelihood	OnePart	Fig. S6	-	-
		Maximum likelihood	GenePart	Fig. S7	-	-
		Maximum likelihood	CodonPart	Fig. S8	-	-
		Maximum likelihood	MixtPart	Fig. S9	-	-
		Bayesian	Mixture model	-	-	Fig. S27
combined-incomplete	191/81,259 63.7%	Maximum likelihood	OnePart	Fig. S10	Fig. S18	-
		Maximum likelihood	GenePart	Fig. S11	Fig. S19	-
		Maximum likelihood	CodonPart	Fig. S12	Fig. S20	-
		Maximum likelihood	MixtPart	Fig. S13	Fig. S21	Figs. 1 and S1
		Bayesian	Mixture model	-	-	Figs. 1 and S1
13-gene	186/15,574 14.6%	Maximum likelihood	OnePart	Fig. S14	Fig. S22	-
		Maximum likelihood	GenePart	Fig. S15	Fig. S23	-
		Maximum likelihood	CodonPart	Fig. S16	Fig. S24	-
		Maximum likelihood	MixtPart	Fig. S17	Fig. S25	-
		Bayesian	Mixture model	-	-	Fig. S28

Table S2. List of 58 taxa included in the ***82-gene*** matrix. Newly sequenced taxa are highlighted in bold. The plastid genome coverage was estimated for each taxon based on the assembled and annotated plastid genome of *Manihot esculenta*. (bp = base pair; gDNA = genomic DNA)

Taxon	Family	Illumina library source	Illumina read length (bp)	Velvet longest contig (bp)	Velvet N50 (bp)	Plastid genome coverage	No. of plastid genes sampled (82 total)	GenBank accession number(s)
***Anthodiscus peruanus* Baill.**	Caryocaraceae	gDNA	100	5,875	1,035	77%	54	JX661767, JX661840, JX661927, JX661968, JX662049, JX662091, JX662132, JX662178, JX662263, JX662305, JX662344, JX662604, JX662649, JX662693, JX662736, JX662780, JX662827, JX662936, JX662982, JX663016, JX663088, JX663130, JX663178, JX663263, JX663306, JX663350, JX663394, JX663459, JX663535, JX663564, JX663611, JX663658, JX663705, JX663748, JX663788, JX663896, JX663944, JX664031, JX664190, JX664236, JX664281, JX664326, JX664367, JX664516, JX664562, JX664607, JX664680, JX664707, JX664755, JX664835, JX664878, JX664920, JX664985
Arabidopsis thaliana (L.) Heynh.	Brassicaceae	-	-	-	-	-	82	AP000423
***Averrhoa carambola* L.**	Oxalidaceae	Plastid	54	18,172	1,409	98%	80	JX661801, JX661841, JX661886, JX661928, JX661969, JX662005, JX662050, JX662092, JX662133, JX662179, JX662223, JX662264, JX662306, JX662345, JX662385, JX662422, JX662455, JX662487, JX662527, JX662568, JX662605, JX662650, JX662694, JX662737, JX662781, JX662828, JX662864, JX662901, JX662937, JX663017, JX663060, JX663089, JX663131, JX663179, JX663223, JX663264, JX663307, JX663351, JX663395, JX663460, JX663505, JX663536, JX663565, JX663612, JX663659, JX663706, JX663749, JX663789, JX663829, JX663862, JX663897, JX663945, JX663991, JX664032, JX664077, JX664115, JX664150, JX664191, JX664237, JX664282, JX664327, JX664368, JX664410, JX664445, JX664483, JX664517, JX664563, JX664608, JX664648, JX664681, JX664708, JX664756, JX664800, JX664836, JX664879, JX664921, JX664968, JX664986

***Balanops pachyphylla* Baill. ex Guillaumin**	Balanopaceae	gDNA	100	5,596	1,520	84%	65	JX661768, JX661842, JX661887, JX661929, JX661970, JX662006, JX662051, JX662093, JX662134, JX662180, JX662224, JX662265, JX662307, JX662346, JX662488, JX662528, JX662606, JX662651, JX662695, JX662738, JX662782, JX662829, JX662865, JX662938, JX662983, JX663018, JX663090, JX663132, JX663180, JX663265, JX663308, JX663352, JX663396, JX663461, JX663537, JX663566, JX663613, JX663660, JX663707, JX663750, JX663790, JX663898, JX663946, JX663992, JX664033, JX664078, JX664151, JX664192, JX664238, JX664283, JX664328, JX664369, JX664518, JX664564, JX664609, JX664682, JX664709, JX664757, JX664801, JX664837, JX664880, JX664922, JX664987
***Bergia texana* (Hook.) Seub. ex Walp.**	Elatinaceae	gDNA	100	5,769	904	75%	56	JX661769, JX661802, JX661843, JX661930, JX661971, JX662052, JX662094, JX662135, JX662181, JX662225, JX662266, JX662529, JX662607, JX662652, JX662696, JX662739, JX662783, JX662866, JX662939, JX663091, JX663133, JX663181, JX663224, JX663266, JX663397, JX663462, JX663538, JX663567, JX663614, JX663661, JX663708, JX663791, JX663863, JX663899, JX663947, JX663993, JX664034, JX664079, JX664193, JX664284, JX664329, JX664370, JX664446, JX664519, JX664565, JX664610, JX664683, JX664710, JX664838, JX664881, JX664923, JX664988
***Bhesa* sp.**	Centroplacaceae	Plastid	54	18,796	2,817	99%	82	JX661770, JX661803, JX661844, JX661888, JX661931, JX661972, JX662007, JX662053, JX662095, JX662136, JX662182, JX662226, JX662267, JX662308, JX662347, JX662386, JX662423, JX662456, JX662489, JX662530, JX662569, JX662608, JX662653, JX662697, JX662740, JX662784, JX662830, JX662867, JX662902, JX662940, JX662984, JX663019, JX663061, JX663092, JX663134, JX663182, JX663225, JX663267, JX663309, JX663353, JX663398, JX663436, JX663463, JX663506, JX663539, JX663568, JX663615, JX663662, JX663709, JX663751, JX663792, JX663830, JX663864, JX663900, JX663948, JX663994, JX664035, JX664080, JX664116, JX664152, JX664194, JX664239, JX664285, JX664330, JX664371, JX664411, JX664447, JX664484, JX664520, JX664566, JX664611, JX664649, JX664684, JX664711, JX664758, JX664802, JX664839, JX664882, JX664924, JX664969, JX664989
***Byrsonima crassifolia* (L.) Kunth**	Malpighiaceae	gDNA	100	4,562	1,273	64%	42	JX661771, JX661804, JX661845, JX661889, JX661932, JX662096, JX662137, JX662183, JX662309, JX662348, JX662387, JX662609, JX662654, JX662741, JX662785, JX662831, JX662903, JX662941, JX663135, JX663183, JX663268, JX663464, JX663569, JX663616, JX663663, JX663710, JX663752, JX663793, JX663901, JX663949, JX664036, JX664195, JX664372, JX664412, JX664448, JX664521, JX664567, JX664712, JX664925, JX664970, JX664990

***Caloncoba echinata* (Oliv.) Gilg**	Achariaceae	Plastid	36	11,585	1,273	99%	82	JX661772, JX661805, JX661846, JX661890, JX661933, JX661973, JX662008, JX662054, JX662097, JX662138, JX662184, JX662227, JX662268, JX662310, JX662349, JX662388, JX662424, JX662457, JX662490, JX662531, JX662570, JX662610, JX662655, JX662698, JX662742, JX662786, JX662832, JX662868, JX662904, JX662942, JX662985, JX663020, JX663062, JX663093, JX663136, JX663184, JX663226, JX663269, JX663310, JX663354, JX663399, JX663437, JX663465, JX663507, JX663540, JX663570, JX663617, JX663664, JX663711, JX663753, JX663794, JX663831, JX663865, JX663902, JX663950, JX663995, JX664037, JX664081, JX664117, JX664153, JX664196, JX664240, JX664286, JX664331, JX664373, JX664413, JX664449, JX664485, JX664522, JX664568, JX664612, JX664650, JX664685, JX664713, JX664759, JX664803, JX664840, JX664883, JX664926, JX664991
Calycanthus floridus L.	Calycanthaceae	-	-	-	-	-	82	AJ428413
***Casearia nitida* Jacq.**	Samydaceae	Plastid	54	25,834	1,861	97%	78	JX661806, JX661847, JX661891, JX661934, JX661974, JX662009, JX662055, JX662098, JX662139, JX662185, JX662228, JX662269, JX662311, JX662350, JX662389, JX662425, JX662458, JX662491, JX662532, JX662571, JX662611, JX662699, JX662743, JX662787, JX662833, JX662869, JX662905, JX662943, JX662986, JX663021, JX663063, JX663094, JX663137, JX663185, JX663227, JX663270, JX663311, JX663355, JX663400, JX663466, JX663508, JX663571, JX663618, JX663665, JX663712, JX663754, JX663795, JX663832, JX663866, JX663903, JX663951, JX663996, JX664038, JX664082, JX664118, JX664154, JX664197, JX664241, JX664287, JX664332, JX664374, JX664414, JX664450, JX664486, JX664523, JX664569, JX664613, JX664651, JX664686, JX664714, JX664760, JX664804, JX664841, JX664884, JX664927, JX664992
***Centroplacus glaucinus* Pierre**	Centroplacaceae	gDNA	100	5,003	654	72%	57	JX661773, JX661848, JX661892, JX662010, JX662099, JX662140, JX662186, JX662229, JX662351, JX662390, JX662492, JX662612, JX662656, JX662700, JX662744, JX662788, JX662834, JX662906, JX662944, JX662987, JX663022, JX663138, JX663186, JX663228, JX663312, JX663356, JX663438, JX663467, JX663541, JX663572, JX663619, JX663666, JX663713, JX663755, JX663796, JX663904, JX663952, JX663997, JX664039, JX664083, JX664155, JX664198, JX664242, JX664375, JX664415, JX664524, JX664570, JX664715, JX664761, JX664805, JX664842, JX664885, JX664928, JX664971, JX664993

***Chrysobalanus icaco* L.**	Chrysobalanaceae	Plastid	54	16,177	1,273	91%	77	JX661807, JX661849, JX661893, JX661935, JX661975, JX662011, JX662056, JX662100, JX662141, JX662187, JX662230, JX662270, JX662312, JX662352, JX662391, JX662426, JX662459, JX662493, JX662533, JX662572, JX662613, JX662657, JX662701, JX662745, JX662789, JX662870, JX662907, JX662945, JX663023, JX663064, JX663095, JX663139, JX663187, JX663229, JX663271, JX663313, JX663357, JX663401, JX663439, JX663509, JX663573, JX663620, JX663667, JX663714, JX663756, JX663797, JX663833, JX663867, JX663905, JX663953, JX663998, JX664040, JX664084, JX664119, JX664156, JX664199, JX664243, JX664288, JX664333, JX664376, JX664416, JX664451, JX664487, JX664525, JX664571, JX664614, JX664652, JX664687, JX664716, JX664762, JX664806, JX664843, JX664886, JX664929, JX664994
***Clusia rosea* Jacq.**	Clusiaceae	Plastid	100	18,064	1,743	99%	82	JX661774, JX661808, JX661850, JX661894, JX661936, JX661976, JX662012, JX662057, JX662101, JX662142, JX662188, JX662231, JX662271, JX662313, JX662353, JX662392, JX662427, JX662460, JX662494, JX662534, JX662573, JX662614, JX662658, JX662702, JX662746, JX662790, JX662835, JX662871, JX662908, JX662946, JX662988, JX663024, JX663065, JX663096, JX663140, JX663188, JX663230, JX663272, JX663314, JX663358, JX663402, JX663440, JX663468, JX663510, JX663542, JX663574, JX663621, JX663668, JX663715, JX663757, JX663834, JX663868, JX663906, JX663954, JX663999, JX664041, JX664085, JX664120, JX664157, JX664200, JX664244, JX664289, JX664334, JX664377, JX664417, JX664452, JX664488, JX664526, JX664572, JX664615, JX664653, JX664688, JX664717, JX664763, JX664807, JX664844, JX664887, JX664930, JX664995
***Ctenolophon englerianus* Mildbr.**	Ctenolophonaceae	gDNA	100	9,719	923	58%	42	JX661775, JX661809, JX661851, JX661895, JX662013, JX662102, JX662143, JX662189, JX662535, JX662615, JX662659, JX662703, JX662791, JX662872, JX662947, JX662989, JX663025, JX663141, JX663189, JX663315, JX663359, JX663469, JX663575, JX663622, JX663669, JX663907, JX663955, JX664201, JX664245, JX664290, JX664378, JX664527, JX664573, JX664718, JX664764, JX664845, JX664888, JX664931, JX664996
Cucumis sativus L.	Cucurbitaceae	-	-	-	-	-	76	AJ970307
***Dichapetalum zenkeri* Engl.**	Dichapetalaceae	gDNA	100	4,568	930	74%	56	JX661776, JX661852, JX661896, JX661937, JX662058, JX662103, JX662144, JX662190, JX662272, JX662314, JX662428, JX662461, JX662495, JX662536, JX662574, JX662616, JX662660, JX662704, JX662792, JX662873, JX662948, JX663066, JX663097, JX663142, JX663190, JX663231, JX663273, JX663403, JX663470, JX663576, JX663623, JX663670, JX663798, JX663835, JX663908, JX663956, JX664000, JX664042, JX664121, JX664158, JX664202, JX664291, JX664335, JX664379, JX664489, JX664528, JX664574, JX664616, JX664719, JX664846, JX664889, JX664932, JX664972, JX664997

***Elaeodendron orientale* Jacq.**	Celastraceae	Plastid	54	18,813	2,402	99%	81	JX661777, JX661810, JX661853, JX661897, JX661938, JX662014, JX662059, JX662104, JX662145, JX662191, JX662232, JX662273, JX662315, JX662354, JX662393, JX662429, JX662462, JX662496, JX662537, JX662575, JX662617, JX662661, JX662705, JX662747, JX662793, JX662836, JX662874, JX662909, JX662949, JX662990, JX663026, JX663067, JX663098, JX663143, JX663191, JX663232, JX663274, JX663316, JX663360, JX663404, JX663441, JX663471, JX663511, JX663543, JX663577, JX663624, JX663671, JX663716, JX663758, JX663799, JX663836, JX663869, JX663909, JX663957, JX664001, JX664043, JX664086, JX664122, JX664159, JX664203, JX664246, JX664292, JX664380, JX664418, JX664453, JX664490, JX664529, JX664575, JX664617, JX664654, JX664689, JX664720, JX664765, JX664808, JX664847, JX664890, JX664933, JX664973, JX664998
***Erythroxylum areolatum* L.**	Erythroxylaceae	Plastid	54	16,729	2,601	98%	78	JX661811, JX661854, JX661898, JX661939, JX662015, JX662060, JX662105, JX662146, JX662192, JX662233, JX662274, JX662316, JX662355, JX662394, JX662430, JX662463, JX662497, JX662538, JX662576, JX662618, JX662662, JX662706, JX662748, JX662794, JX662837, JX662875, JX662910, JX662950, JX662991, JX663027, JX663068, JX663099, JX663144, JX663192, JX663233, JX663275, JX663317, JX663361, JX663405, JX663442, JX663472, JX663512, JX663578, JX663625, JX663672, JX663717, JX663759, JX663800, JX663837, JX663870, JX663910, JX663958, JX664002, JX664044, JX664087, JX664123, JX664160, JX664204, JX664247, JX664293, JX664336, JX664381, JX664419, JX664454, JX664491, JX664530, JX664576, JX664618, JX664655, JX664690, JX664721, JX664766, JX664809, JX664848, JX664891, JX664934, JX664999
***Euphorbia maculata* L.**	Euphorbiaceae	Plastid	36	7,490	1,153	97%	78	JX661812, JX661855, JX661899, JX661940, JX661977, JX662016, JX662061, JX662106, JX662147, JX662193, JX662234, JX662275, JX662317, JX662356, JX662395, JX662431, JX662464, JX662498, JX662539, JX662577, JX662619, JX662663, JX662707, JX662749, JX662795, JX662838, JX662876, JX662911, JX662951, JX662992, JX663028, JX663069, JX663100, JX663145, JX663193, JX663234, JX663276, JX663318, JX663362, JX663406, JX663473, JX663513, JX663579, JX663626, JX663673, JX663718, JX663760, JX663801, JX663838, JX663871, JX663911, JX663959, JX664003, JX664045, JX664088, JX664124, JX664161, JX664205, JX664248, JX664294, JX664337, JX664382, JX664420, JX664455, JX664492, JX664531, JX664577, JX664619, JX664656, JX664691, JX664722, JX664767, JX664810, JX664849, JX664892, JX664935, JX665000

***Euphronia guianensis* (R.H.Schomb.) Hallier f.**	Euphroniaceae	gDNA	100	5,400	363	52%	43	JX661778, JX661813, JX661856, JX661941, JX662017, JX662062, JX662148, JX662276, JX662396, JX662499, JX662540, JX662664, JX662796, JX662877, JX662952, JX662993, JX663029, JX663101, JX663146, JX663277, JX663319, JX663407, JX663474, JX663544, JX663580, JX663627, JX663674, JX663912, JX664046, JX664162, JX664206, JX664249, JX664338, JX664532, JX664620, JX664723, JX664768, JX664936, JX665001
***Flueggea suffruticosa* (Pall.) Baill.**	Phyllanthaceae	Plastid	54	22,202	2,315	99%	81	JX661779, JX661814, JX661857, JX661900, JX661942, JX661978, JX662018, JX662063, JX662107, JX662149, JX662194, JX662235, JX662277, JX662318, JX662357, JX662397, JX662432, JX662465, JX662500, JX662541, JX662578, JX662620, JX662708, JX662750, JX662797, JX662839, JX662878, JX662912, JX662953, JX662994, JX663030, JX663070, JX663102, JX663147, JX663194, JX663235, JX663278, JX663320, JX663363, JX663408, JX663443, JX663475, JX663514, JX663545, JX663581, JX663628, JX663675, JX663719, JX663761, JX663802, JX663839, JX663872, JX663913, JX663960, JX664004, JX664047, JX664089, JX664125, JX664163, JX664207, JX664250, JX664295, JX664339, JX664383, JX664421, JX664456, JX664493, JX664533, JX664578, JX664621, JX664657, JX664692, JX664724, JX664769, JX664811, JX664850, JX664893, JX664937, JX665002
***Galearia maingayi* Hook. f.**	Pandaceae	Plastid	100	21,405	1,815	99%	81	JX661780, JX661815, JX661858, JX661901, JX661943, JX661979, JX662019, JX662064, JX662108, JX662150, JX662195, JX662236, JX662278, JX662319, JX662358, JX662398, JX662433, JX662466, JX662501, JX662542, JX662579, JX662621, JX662665, JX662709, JX662751, JX662798, JX662840, JX662879, JX662913, JX662954, JX662995, JX663031, JX663103, JX663148, JX663195, JX663236, JX663279, JX663321, JX663364, JX663409, JX663476, JX663515, JX663546, JX663582, JX663629, JX663676, JX663720, JX663762, JX663803, JX663840, JX663873, JX663914, JX663961, JX664005, JX664048, JX664090, JX664126, JX664164, JX664208, JX664251, JX664296, JX664340, JX664384, JX664422, JX664457, JX664494, JX664534, JX664579, JX664622, JX664658, JX664693, JX664725, JX664770, JX664851, JX664894, JX664938, JX664974, JX665003

***Garcinia mangostana* L.**	Clusiaceae	Plastid	36	7,507	779	92%	77	JX661816, JX661859, JX661902, JX661944, JX661980, JX662020, JX662065, JX662109, JX662151, JX662196, JX662237, JX662279, JX662320, JX662359, JX662399, JX662434, JX662467, JX662502, JX662543, JX662580, JX662622, JX662666, JX662710, JX662752, JX662799, JX662841, JX662880, JX662914, JX662955, JX662996, JX663032, JX663071, JX663104, JX663149, JX663196, JX663237, JX663280, JX663322, JX663365, JX663410, JX663583, JX663630, JX663677, JX663721, JX663763, JX663804, JX663841, JX663874, JX663915, JX663962, JX664006, JX664049, JX664091, JX664127, JX664165, JX664209, JX664252, JX664297, JX664341, JX664385, JX664458, JX664495, JX664535, JX664580, JX664623, JX664659, JX664694, JX664726, JX664771, JX664812, JX664852, JX664895, JX664939, JX665004
***Goupia glabra* Aubl.**	Goupiaceae	gDNA	100	4,485	256	38%	30	JX661781, JX661903, JX662021, JX662110, JX662238, JX662667, JX662800, JX662997, JX663150, JX663238, JX663366, JX663477, JX663631, JX663678, JX663916, JX663963, JX664253, JX664298, JX664386, JX664727, JX664772, JX664896, JX664940, JX664975, JX665005
***Humiria balsamifera* (Aubl.) J.St.-Hil.**	Humiriaceae	gDNA	100	5,736	175	51%	45	JX661782, JX661904, JX661945, JX662022, JX662111, JX662152, JX662360, JX662668, JX662711, JX662753, JX662801, JX662842, JX662881, JX662998, JX663033, JX663151, JX663239, JX663281, JX663323, JX663367, JX663444, JX663478, JX663584, JX663632, JX663679, JX663764, JX663805, JX663917, JX663964, JX664050, JX664254, JX664299, JX664387, JX664536, JX664728, JX664773, JX664853, JX664897, JX664941, JX664976, JX665006
***Hypericum (Triadenum) fraseri* (Spach) Steudel**	Hypericaceae	Plastid	100	7,140	1,271	88%	75	JX661817, JX661860, JX661905, JX661981, JX662023, JX662066, JX662153, JX662197, JX662239, JX662280, JX662321, JX662361, JX662400, JX662435, JX662468, JX662503, JX662544, JX662581, JX662623, JX662669, JX662712, JX662754, JX662802, JX662843, JX662915, JX662956, JX663034, JX663072, JX663105, JX663152, JX663197, JX663240, JX663282, JX663324, JX663368, JX663411, JX663479, JX663516, JX663585, JX663633, JX663680, JX663722, JX663765, JX663842, JX663875, JX663918, JX663965, JX664007, JX664051, JX664092, JX664128, JX664166, JX664210, JX664255, JX664300, JX664342, JX664388, JX664423, JX664459, JX664496, JX664537, JX664581, JX664624, JX664660, JX664729, JX664774, JX664813, JX664854, JX664898, JX664942, JX665007

***Hypericum kalmianum* L.**	Hypericaceae	Plastid	100	11,168	1,131	92%	79	JX661783, JX661818, JX661861, JX661906, JX661946, JX661982, JX662024, JX662067, JX662112, JX662154, JX662198, JX662240, JX662281, JX662322, JX662362, JX662401, JX662436, JX662469, JX662504, JX662545, JX662582, JX662624, JX662713, JX662755, JX662803, JX662844, JX662882, JX662916, JX662957, JX662999, JX663035, JX663073, JX663106, JX663153, JX663198, JX663241, JX663283, JX663325, JX663369, JX663412, JX663480, JX663517, JX663547, JX663586, JX663634, JX663681, JX663723, JX663766, JX663843, JX663876, JX663919, JX663966, JX664008, JX664052, JX664093, JX664129, JX664167, JX664211, JX664256, JX664301, JX664343, JX664389, JX664424, JX664460, JX664497, JX664538, JX664582, JX664625, JX664661, JX664730, JX664775, JX664814, JX664855, JX664943, JX665008
***Hypericum perforatum* L.**	Hypericaceae	Plastid	100	15,860	2,018	83%	78	JX661784, JX661819, JX661862, JX661907, JX661947, JX661983, JX662025, JX662068, JX662113, JX662155, JX662199, JX662241, JX662282, JX662323, JX662363, JX662402, JX662437, JX662470, JX662505, JX662546, JX662583, JX662625, JX662670, JX662714, JX662756, JX662804, JX662845, JX662917, JX662958, JX663036, JX663107, JX663154, JX663199, JX663242, JX663284, JX663326, JX663370, JX663413, JX663481, JX663518, JX663548, JX663587, JX663635, JX663682, JX663724, JX663767, JX663806, JX663844, JX663877, JX663920, JX663967, JX664009, JX664053, JX664094, JX664130, JX664168, JX664212, JX664257, JX664302, JX664344, JX664425, JX664461, JX664498, JX664539, JX664583, JX664626, JX664662, JX664731, JX664776, JX664815, JX664856, JX664899, JX664944, JX665009
***Irvingia malayana* Oliv. ex Benn.**	Irvingiaceae	Plastid	100	18,380	1,594	99%	81	JX661785, JX661820, JX661863, JX661908, JX661948, JX661984, JX662026, JX662069, JX662114, JX662156, JX662200, JX662242, JX662283, JX662324, JX662364, JX662403, JX662438, JX662471, JX662506, JX662547, JX662584, JX662626, JX662671, JX662715, JX662757, JX662805, JX662846, JX662883, JX662918, JX662959, JX663037, JX663074, JX663108, JX663155, JX663200, JX663243, JX663285, JX663327, JX663371, JX663414, JX663445, JX663482, JX663519, JX663549, JX663588, JX663636, JX663683, JX663725, JX663768, JX663807, JX663845, JX663878, JX663921, JX663968, JX664010, JX664054, JX664095, JX664131, JX664169, JX664213, JX664258, JX664303, JX664345, JX664390, JX664426, JX664462, JX664499, JX664540, JX664584, JX664627, JX664663, JX664695, JX664732, JX664777, JX664816, JX664857, JX664900, JX664945, JX665010

***Ixonanthes* sp.**	Ixonanthaceae	Plastid	54	17,190	2,821	97%	78	JX661821, JX661864, JX661909, JX661949, JX661985, JX662027, JX662070, JX662115, JX662157, JX662201, JX662243, JX662284, JX662325, JX662365, JX662404, JX662439, JX662472, JX662507, JX662548, JX662585, JX662627, JX662672, JX662716, JX662758, JX662806, JX662847, JX662884, JX662919, JX662960, JX663000, JX663038, JX663075, JX663109, JX663156, JX663201, JX663286, JX663328, JX663372, JX663415, JX663483, JX663520, JX663550, JX663589, JX663637, JX663684, JX663726, JX663769, JX663808, JX663846, JX663879, JX663922, JX663969, JX664011, JX664055, JX664096, JX664132, JX664170, JX664214, JX664259, JX664304, JX664346, JX664391, JX664427, JX664463, JX664500, JX664541, JX664585, JX664628, JX664664, JX664696, JX664733, JX664778, JX664817, JX664858, JX664901, JX664946, JX665011
Jatropha curcas L.	Euphorbiaceae	-	-	-	-	-	81	AJ428413
***Lacistema robustum* Schnizl.**	Lacistemataceae	Plastid	54	15,115	1,228	98%	79	JX661786, JX661822, JX661865, JX661910, JX661950, JX661986, JX662028, JX662071, JX662116, JX662158, JX662202, JX662244, JX662285, JX662326, JX662366, JX662405, JX662440, JX662473, JX662508, JX662549, JX662586, JX662628, JX662673, JX662717, JX662759, JX662807, JX662885, JX662920, JX662961, JX663001, JX663039, JX663076, JX663110, JX663157, JX663202, JX663244, JX663287, JX663329, JX663373, JX663416, JX663446, JX663484, JX663521, JX663590, JX663638, JX663685, JX663727, JX663770, JX663809, JX663847, JX663880, JX663923, JX663970, JX664012, JX664056, JX664097, JX664133, JX664171, JX664215, JX664260, JX664305, JX664347, JX664392, JX664428, JX664464, JX664501, JX664542, JX664586, JX664629, JX664665, JX664697, JX664734, JX664779, JX664818, JX664859, JX664902, JX664947, JX665012
***Linum usitatissimum* L.**	Linaceae	Plastid	54	11,608	1,133	91%	78	JX661787, JX661823, JX661866, JX661911, JX661951, JX661987, JX662029, JX662072, JX662159, JX662203, JX662245, JX662286, JX662327, JX662367, JX662406, JX662441, JX662474, JX662509, JX662550, JX662587, JX662629, JX662674, JX662718, JX662760, JX662808, JX662848, JX662921, JX662962, JX663002, JX663040, JX663077, JX663111, JX663158, JX663203, JX663245, JX663288, JX663330, JX663374, JX663417, JX663447, JX663485, JX663522, JX663551, JX663591, JX663639, JX663686, JX663728, JX663771, JX663810, JX663848, JX663881, JX663924, JX663971, JX664013, JX664057, JX664098, JX664134, JX664172, JX664216, JX664261, JX664306, JX664348, JX664393, JX664429, JX664465, JX664502, JX664543, JX664587, JX664666, JX664735, JX664780, JX664819, JX664860, JX664903, JX664948, JX665013

***Mammea americana* L.**	Calophyllaceae	Plastid	100	17,847	3,577	99%	80	JX661824, JX661867, JX661912, JX661952, JX661988, JX662030, JX662073, JX662117, JX662160, JX662204, JX662246, JX662287, JX662328, JX662368, JX662407, JX662442, JX662475, JX662510, JX662551, JX662588, JX662630, JX662675, JX662719, JX662761, JX662809, JX662849, JX662886, JX662922, JX662963, JX663003, JX663041, JX663078, JX663112, JX663159, JX663204, JX663246, JX663289, JX663331, JX663375, JX663418, JX663448, JX663486, JX663523, JX663592, JX663640, JX663687, JX663729, JX663772, JX663811, JX663849, JX663882, JX663925, JX663972, JX664014, JX664058, JX664099, JX664135, JX664173, JX664217, JX664262, JX664307, JX664349, JX664394, JX664430, JX664466, JX664503, JX664544, JX664588, JX664630, JX664667, JX664698, JX664736, JX664781, JX664820, JX664861, JX664904, JX664949, JX664977, JX665014
Manihot esculenta Crantz	Euphorbiaceae	-	-	-	-	-	82	EU117376
Medicago truncatula Gaertn.	Fabaceae	-	-	-	-	-	77	AC093544
***Medusagyne oppositifolia* Baker**	Medusagynaceae	gDNA	100	5,222	650	89%	71	JX661788, JX661825, JX661868, JX661913, JX661953, JX661989, JX662031, JX662074, JX662118, JX662161, JX662205, JX662247, JX662288, JX662329, JX662369, JX662511, JX662589, JX662631, JX662676, JX662720, JX662762, JX662810, JX662850, JX662964, JX663042, JX663113, JX663160, JX663205, JX663247, JX663290, JX663332, JX663376, JX663419, JX663449, JX663487, JX663552, JX663593, JX663641, JX663688, JX663730, JX663773, JX663812, JX663926, JX663973, JX664015, JX664059, JX664100, JX664136, JX664174, JX664218, JX664263, JX664308, JX664350, JX664395, JX664467, JX664545, JX664589, JX664631, JX664699, JX664737, JX664782, JX664821, JX664862, JX664905, JX664950, JX664978, JX665015
***Microdesmis casearifolia* Planch.**	Pandaceae	Plastid	100	14,558	1,636	99%	82	JX661789, JX661826, JX661869, JX661914, JX661954, JX661990, JX662032, JX662075, JX662119, JX662162, JX662206, JX662248, JX662289, JX662330, JX662370, JX662408, JX662443, JX662476, JX662512, JX662552, JX662590, JX662632, JX662677, JX662721, JX662763, JX662811, JX662851, JX662887, JX662923, JX662965, JX663004, JX663043, JX663079, JX663114, JX663161, JX663206, JX663248, JX663291, JX663333, JX663377, JX663420, JX663450, JX663488, JX663524, JX663553, JX663594, JX663642, JX663689, JX663731, JX663774, JX663813, JX663850, JX663883, JX663927, JX663974, JX664016, JX664060, JX664101, JX664137, JX664175, JX664219, JX664264, JX664309, JX664351, JX664396, JX664431, JX664468, JX664504, JX664546, JX664590, JX664632, JX664668, JX664700, JX664738, JX664783, JX664863, JX664906, JX664951, JX664979, JX665016
Nicotiana tabacum L.	Solanaceae	-	-	-	-	-	81	BA000042

***Ouratea* sp.**	Ochnaceae	Plastid	54	18,120	1,991	93%	79	JX661827, JX661870, JX661915, JX661955, JX661991, JX662033, JX662076, JX662120, JX662163, JX662207, JX662249, JX662290, JX662331, JX662371, JX662409, JX662444, JX662477, JX662513, JX662553, JX662591, JX662633, JX662678, JX662722, JX662764, JX662812, JX662852, JX662888, JX662924, JX662966, JX663044, JX663080, JX663115, JX663162, JX663207, JX663249, JX663292, JX663334, JX663378, JX663421, JX663489, JX663525, JX663595, JX663643, JX663690, JX663732, JX663775, JX663814, JX663851, JX663884, JX663928, JX663975, JX664017, JX664061, JX664102, JX664138, JX664176, JX664220, JX664265, JX664310, JX664352, JX664397, JX664432, JX664469, JX664505, JX664547, JX664591, JX664633, JX664669, JX664701, JX664739, JX664784, JX664822, JX664864, JX664907, JX664952, JX664980, JX665017
***Passiflora ciliata* Aiton**	Passifloraceae	Plastid	100	13,122	928	92%	79	JX661790, JX661828, JX661871, JX661916, JX661956, JX661992, JX662034, JX662077, JX662164, JX662208, JX662250, JX662291, JX662332, JX662372, JX662410, JX662445, JX662478, JX662514, JX662554, JX662592, JX662634, JX662679, JX662723, JX662765, JX662813, JX662889, JX662925, JX662967, JX663045, JX663081, JX663116, JX663163, JX663208, JX663250, JX663293, JX663335, JX663379, JX663422, JX663451, JX663490, JX663526, JX663554, JX663596, JX663644, JX663691, JX663733, JX663776, JX663815, JX663852, JX663885, JX663929, JX663976, JX664018, JX664062, JX664103, JX664139, JX664177, JX664221, JX664266, JX664311, JX664353, JX664398, JX664433, JX664470, JX664506, JX664548, JX664592, JX664634, JX664670, JX664740, JX664785, JX664823, JX664865, JX664908, JX664953, JX665018
***Pera bumeliifolia* Griseb.**	Peraceae	Plastid	54	16,716	2,550	96%	82	JX661791, JX661829, JX661872, JX661917, JX661957, JX661993, JX662035, JX662078, JX662121, JX662165, JX662209, JX662251, JX662292, JX662333, JX662373, JX662411, JX662446, JX662479, JX662515, JX662555, JX662593, JX662635, JX662680, JX662724, JX662766, JX662814, JX662853, JX662890, JX662926, JX662968, JX663005, JX663046, JX663082, JX663117, JX663164, JX663209, JX663251, JX663294, JX663336, JX663380, JX663423, JX663452, JX663491, JX663527, JX663555, JX663597, JX663645, JX663692, JX663734, JX663777, JX663816, JX663853, JX663886, JX663930, JX663977, JX664019, JX664063, JX664104, JX664140, JX664178, JX664222, JX664267, JX664312, JX664354, JX664399, JX664434, JX664471, JX664507, JX664549, JX664593, JX664635, JX664671, JX664702, JX664741, JX664786, JX664824, JX664866, JX664909, JX664954, JX664981, JX665019

***Phyllanthus urinaria* L.**	Phyllanthaceae	Plastid	54	14,323	1,615	97%	79	JX661830, JX661873, JX661918, JX661958, JX661994, JX662036, JX662079, JX662122, JX662166, JX662210, JX662252, JX662293, JX662334, JX662374, JX662412, JX662447, JX662480, JX662516, JX662556, JX662594, JX662636, JX662681, JX662725, JX662767, JX662815, JX662854, JX662891, JX662927, JX662969, JX663006, JX663047, JX663083, JX663118, JX663165, JX663210, JX663252, JX663295, JX663337, JX663381, JX663424, JX663453, JX663492, JX663528, JX663598, JX663646, JX663693, JX663735, JX663778, JX663817, JX663854, JX663887, JX663931, JX663978, JX664020, JX664064, JX664105, JX664141, JX664223, JX664268, JX664313, JX664355, JX664400, JX664435, JX664472, JX664508, JX664550, JX664594, JX664636, JX664672, JX664742, JX664787, JX664825, JX664867, JX664910, JX664955, JX664982, JX665020
***Ploiarium* sp.**	Bonnetiaceae	Plastid	100	12,438	1,547	99%	82	JX661792, JX661831, JX661874, JX661919, JX661959, JX661995, JX662037, JX662080, JX662123, JX662167, JX662211, JX662253, JX662294, JX662335, JX662375, JX662413, JX662448, JX662481, JX662517, JX662557, JX662595, JX662637, JX662726, JX662768, JX662816, JX662855, JX662892, JX662928, JX662970, JX663007, JX663048, JX663084, JX663119, JX663166, JX663211, JX663253, JX663296, JX663338, JX663382, JX663425, JX663454, JX663493, JX663529, JX663556, JX663599, JX663647, JX663694, JX663736, JX663779, JX663818, JX663855, JX663888, JX663932, JX663979, JX664021, JX664065, JX664106, JX664142, JX664179, JX664224, JX664269, JX664314, JX664356, JX664401, JX664436, JX664473, JX664509, JX664551, JX664595, JX664637, JX664673, JX664703, JX664743, JX664788, JX664826, JX664868, JX664911, JX664956, JX664983, JX665021
***Podocalyx loranthoides* Klotzsch**	Picrodendraceae	gDNA	100	5,854	1,197	80%	61	JX661793, JX661832, JX661875, JX661960, JX661996, JX662038, JX662081, JX662124, JX662168, JX662212, JX662295, JX662376, JX662414, JX662518, JX662638, JX662682, JX662727, JX662769, JX662817, JX662856, JX662929, JX662971, JX663049, JX663120, JX663167, JX663212, JX663297, JX663339, JX663383, JX663426, JX663494, JX663557, JX663600, JX663648, JX663695, JX663737, JX663780, JX663819, JX663933, JX663980, JX664066, JX664180, JX664225, JX664270, JX664315, JX664357, JX664402, JX664437, JX664474, JX664552, JX664596, JX664638, JX664744, JX664789, JX664869, JX664912, JX664957, JX665022

***Podostemum ceratophyllum* Michx.**	Podostemaceae	Plastid	100	17,091	1,795	86%	79	JX661794, JX661833, JX661876, JX661920, JX661961, JX661997, JX662039, JX662082, JX662169, JX662213, JX662254, JX662296, JX662336, JX662377, JX662415, JX662449, JX662482, JX662519, JX662558, JX662596, JX662639, JX662683, JX662728, JX662770, JX662818, JX662857, JX662893, JX662930, JX662972, JX663008, JX663050, JX663085, JX663121, JX663168, JX663213, JX663254, JX663298, JX663340, JX663384, JX663427, JX663455, JX663495, JX663530, JX663558, JX663601, JX663649, JX663696, JX663738, JX663781, JX663820, JX663856, JX663889, JX663934, JX663981, JX664022, JX664067, JX664107, JX664143, JX664181, JX664226, JX664271, JX664316, JX664358, JX664403, JX664438, JX664475, JX664510, JX664553, JX664597, JX664639, JX664674, JX664745, JX664790, JX664827, JX664870, JX664913, JX664958, JX665023
Populus alba L.	Salicaceae	-	-	-	-	-	80	AP008956
Populus trichocarpa (Torr. & Gray)	Salicaceae	-	-	-	-	-	80	EF489041
***Putranjiva roxburghii* Wall.**	Putranjivaceae	Plastid	36	3,044	740	72%	69	JX661877, JX661921, JX661998, JX662040, JX662083, JX662170, JX662214, JX662255, JX662297, JX662337, JX662378, JX662416, JX662450, JX662520, JX662559, JX662597, JX662640, JX662684, JX662729, JX662771, JX662819, JX662858, JX662894, JX662931, JX662973, JX663009, JX663051, JX663122, JX663169, JX663214, JX663255, JX663341, JX663385, JX663428, JX663456, JX663496, JX663602, JX663650, JX663739, JX663782, JX663857, JX663890, JX663935, JX663982, JX664023, JX664068, JX664108, JX664144, JX664182, JX664227, JX664272, JX664317, JX664359, JX664439, JX664476, JX664511, JX664554, JX664598, JX664640, JX664675, JX664746, JX664791, JX664828, JX664871, JX664959, JX665024
***Quiina glaziovii* Engl.**	Quiinaceae	gDNA	100	4,881	1,442	73%	52	JX661795, JX661834, JX661878, JX661999, JX662041, JX662084, JX662125, JX662171, JX662215, JX662298, JX662560, JX662641, JX662685, JX662772, JX662820, JX662895, JX662974, JX663010, JX663052, JX663123, JX663170, JX663215, JX663299, JX663342, JX663386, JX663429, JX663497, JX663559, JX663603, JX663651, JX663697, JX663740, JX663821, JX663936, JX663983, JX664069, JX664228, JX664273, JX664318, JX664360, JX664404, JX664555, JX664599, JX664641, JX664704, JX664747, JX664792, JX664914, JX664960, JX664984, JX665025

***Rhizophora mangle* L.**	Rhizophoraceae	Plastid	100	12,267	1,860	98%	80	JX661796, JX661835, JX661879, JX661922, JX661962, JX662000, JX662042, JX662085, JX662126, JX662172, JX662216, JX662256, JX662299, JX662338, JX662379, JX662417, JX662451, JX662483, JX662521, JX662561, JX662598, JX662642, JX662686, JX662730, JX662773, JX662821, JX662859, JX662896, JX662932, JX662975, JX663053, JX663086, JX663124, JX663171, JX663216, JX663256, JX663300, JX663343, JX663387, JX663430, JX663457, JX663498, JX663531, JX663560, JX663604, JX663652, JX663698, JX663741, JX663783, JX663822, JX663858, JX663891, JX663937, JX663984, JX664024, JX664070, JX664109, JX664145, JX664183, JX664229, JX664274, JX664319, JX664361, JX664405, JX664440, JX664477, JX664512, JX664556, JX664600, JX664642, JX664676, JX664705, JX664748, JX664793, JX664829, JX664872, JX664915, JX664961, JX665026
***Schistostemon retusum* (Ducke) Cuatrec.**	Humiriaceae	gDNA	100	4,171	777	81%	65	JX661797, JX661880, JX661923, JX661963, JX662001, JX662043, JX662086, JX662127, JX662173, JX662217, JX662257, JX662300, JX662380, JX662522, JX662562, JX662643, JX662687, JX662731, JX662774, JX662822, JX662860, JX662897, JX662976, JX663011, JX663054, JX663125, JX663172, JX663217, JX663257, JX663301, JX663344, JX663388, JX663431, JX663499, JX663561, JX663605, JX663653, JX663699, JX663742, JX663784, JX663823, JX663938, JX663985, JX664025, JX664071, JX664184, JX664230, JX664275, JX664320, JX664362, JX664406, JX664478, JX664557, JX664601, JX664643, JX664706, JX664749, JX664794, JX664873, JX664916, JX664962, JX665027
***Scyphostegia borneensis* Stapf.**	Scyphostegiaceae	gDNA	100	6,225	321	78%	63	JX661798, JX661836, JX661881, JX662044, JX662128, JX662174, JX662218, JX662258, JX662339, JX662381, JX662418, JX662523, JX662563, JX662599, JX662644, JX662688, JX662732, JX662775, JX662823, JX662861, JX662898, JX662977, JX663012, JX663055, JX663173, JX663218, JX663258, JX663345, JX663389, JX663500, JX663606, JX663654, JX663700, JX663743, JX663824, JX663892, JX663939, JX663986, JX664026, JX664072, JX664110, JX664146, JX664185, JX664231, JX664276, JX664321, JX664407, JX664441, JX664479, JX664558, JX664602, JX664750, JX664795, JX664830, JX664874, JX664917, JX664963, JX665028
***Trigonia* sp.**	Trigoniaceae	gDNA	100	4,628	1,486	70%	53	JX661882, JX661964, JX662045, JX662087, JX662129, JX662219, JX662259, JX662301, JX662340, JX662524, JX662564, JX662600, JX662645, JX662689, JX662776, JX662899, JX662978, JX663013, JX663056, JX663126, JX663174, JX663219, JX663259, JX663302, JX663346, JX663390, JX663432, JX663458, JX663501, JX663607, JX663701, JX663744, JX663825, JX663940, JX663987, JX664027, JX664073, JX664111, JX664186, JX664232, JX664277, JX664322, JX664363, JX664408, JX664603, JX664644, JX664751, JX664796, JX664831, JX664918, JX664964, JX665029

***Turnera ulmifolia* L.**	Turneraceae	Plastid	54	25,114	1,025	81%	73	JX661799, JX661837, JX661883, JX661924, JX661965, JX662002, JX662046, JX662088, JX662175, JX662220, JX662260, JX662302, JX662341, JX662382, JX662419, JX662452, JX662484, JX662565, JX662601, JX662646, JX662690, JX662733, JX662777, JX662824, JX662933, JX662979, JX663057, JX663127, JX663175, JX663220, JX663260, JX663303, JX663347, JX663391, JX663433, JX663502, JX663532, JX663562, JX663608, JX663655, JX663702, JX663745, JX663785, JX663826, JX663859, JX663893, JX663941, JX663988, JX664028, JX664074, JX664112, JX664147, JX664187, JX664233, JX664278, JX664323, JX664364, JX664442, JX664480, JX664513, JX664559, JX664604, JX664645, JX664677, JX664752, JX664797, JX664832, JX664875, JX664965, JX665030
***Viola pubescens* Aiton**	Violaceae	Plastid	54	10,254	771	97%	80	JX661800, JX661838, JX661884, JX661925, JX661966, JX662003, JX662047, JX662089, JX662130, JX662176, JX662221, JX662261, JX662303, JX662342, JX662383, JX662420, JX662453, JX662485, JX662525, JX662566, JX662602, JX662647, JX662691, JX662734, JX662778, JX662825, JX662862, JX662900, JX662934, JX662980, JX663014, JX663058, JX663087, JX663128, JX663176, JX663221, JX663261, JX663304, JX663348, JX663392, JX663434, JX663503, JX663533, JX663563, JX663609, JX663656, JX663703, JX663746, JX663786, JX663827, JX663860, JX663894, JX663942, JX663989, JX664029, JX664075, JX664113, JX664148, JX664188, JX664234, JX664279, JX664324, JX664365, JX664409, JX664443, JX664481, JX664514, JX664560, JX664605, JX664646, JX664678, JX664753, JX664798, JX664833, JX664876, JX664919, JX664966, JX665031
***Vismia ferruginea* Kunth**	Hypericaceae	Plastid	100	9,790	2,021	83%	75	JX661839, JX661885, JX661926, JX661967, JX662004, JX662048, JX662090, JX662131, JX662177, JX662222, JX662262, JX662304, JX662343, JX662384, JX662421, JX662454, JX662486, JX662526, JX662567, JX662603, JX662648, JX662692, JX662735, JX662779, JX662826, JX662863, JX662935, JX662981, JX663015, JX663059, JX663129, JX663177, JX663222, JX663262, JX663305, JX663349, JX663393, JX663435, JX663504, JX663534, JX663610, JX663657, JX663704, JX663747, JX663787, JX663828, JX663861, JX663895, JX663943, JX663990, JX664030, JX664076, JX664114, JX664149, JX664189, JX664235, JX664280, JX664325, JX664366, JX664444, JX664482, JX664515, JX664561, JX664606, JX664647, JX664679, JX664754, JX664799, JX664834, JX664877, JX664967, JX665032
Vitis vinifera L.	Vitaceae	-	-	-	-	-	81	DQ424856

Table S3. Data characteristics for all 82 plastid genes, including the number of taxa sampled per gene, the number of total aligned characters per gene, and the percentage of gaps or missing data per gene.

Gene	No. of taxa (58 total)	Total aligned characters	Missing data%
accD	51	852	20.8%
atpA	49	1,509	20.9%
atpB	55	1,503	6.6%
atpE	56	426	15.7%
atpF	46	558	22.4%
atpH	48	243	19.9%
atpI	58	768	3.7%
ccsA	52	1,101	22.9%
cemA	48	717	24.9%
clpP	49	753	37.3%
matK	50	1,752	26.3%
ndhA	57	1,137	6.7%
ndhB	56	1,536	4.5%
ndhC	49	363	16.7%
ndhD	52	1,515	11.7%
ndhE	52	303	11.9%
ndhF	55	2,463	24.4%
ndhG	52	534	11.5%
ndhH	57	1,194	11.4%
ndhI	58	474	8.0%
ndhJ	53	477	9.4%
ndhK	56	726	19.0%
petA	50	975	19.2%
petB	58	639	4.8%
petD	54	492	8.2%
petG	43	111	26.0%
petL	42	93	27.6%
petN	39	87	33.5%
psaA	56	2,274	5.5%
psaB	56	2,202	3.5%
psaC	52	243	10.4%
psaI	50	111	14.9%
psaJ	41	132	29.9%
psbA	53	1,059	8.7%
psbB	56	1,524	5.2%

psbC	45	1,419	24.3%
psbD	50	1,059	20.5%
psbE	43	249	25.9%
psbF	43	117	26.3%
psbH	55	219	5.3%
psbI	44	108	25.2%
psbJ	50	120	16.4%
psbK	51	189	17.4%
psbL	50	114	15.6%
psbM	44	102	24.3%
psbN	55	129	5.2%
psbT	55	114	13.2%
psbZ	45	186	22.8%
rbcL	56	1,425	6.7%
rpl2	57	840	7.8%
rpl14	55	366	5.2%
rpl16	56	402	6.9%
rpl20	51	369	17.4%
rpl22	45	339	22.7%
rpl23	53	291	12.9%
rpl32	31	138	47.2%
rpl33	41	210	33.6%
rpl36	54	111	7.0%
rpoA	55	1,155	19.0%
rpoB	56	3,324	14.7%
rpoC1	58	2,193	9.8%
rpoC2	58	4,890	16.0%
rps2	58	729	5.1%
rps3	51	804	28.7%
rps4	53	702	21.7%
rps7	54	486	13.6%
rps8	55	426	11.8%
rps11	55	453	13.9%
rps12	57	381	14.8%
rps14	56	303	4.6%
rps15	41	282	34.1%
rps16	28	258	57.2%
rps18	54	309	34.6%
rps19	51	306	21.1%
rrn5	47	121	20.5%
rrn16	57	1,506	4.0%
rrn23	58	2,949	5.3%
rrn45	56	110	13.0%

ycf1	36	2,931	49.0%
ycf2	51	7,677	26.5%
ycf3	48	507	18.3%
ycf4	50	564	21.5%

Table S4. Maximum likelihood (ML) bootstrap percentages of all clades recovered in the 12 ML bipartition trees (Figs. S2-S9 and S18-S21), which were inferred from three matrices (***82-gene***, ***combined-complete***, and ***combined-incomplete***) and four partitioning strategies (OnePart, GenePart, CodonPart, and MixtPart) applied here.

Matrix	Partitioning	Node1	Node2	Node3	Node4	Node5	Node6	Node7	Node8	Node9	Node10	Node11	Node12	Node13	Node14	Node15	Node16	Node17	Node18	Node19	Node20
plastid-gene	OnePart	100	100	79	100	78	100	-	70	100	-	16	-	40	-	26	-	-	17	54	23
plastid-gene	GenePart	100	100	85	100	82	100	-	76	100	-	-	17	44	-	25	-	-	12	50	22
plastid-gene	CodonPart	100	100	70	100	67	100	-	51	100	11	-	-	43	-	38	-	-	22	42	28
plastid-gene	MixtPart	100	100	89	100	81	100	-	57	100	-	19	-	85	-	80	12	-	-	65	-
combined-complete	OnePart	100	100	91	100	85	100	-	64	100	-	27	-	40	-	21	-	9	-	22	-
combined-complete	GenePart	100	100	89	100	80	100	-	78	100	-	11	-	39	-	19	14	-	-	22	-
combined-complete	CodonPart	100	100	83	100	75	100	-	63	100	-	19	-	39	24	31	22	-	-	-	-
combined-complete	MixtPart	100	100	90	100	77	100	-	71	100	-	60	-	87	-	82	28	-	-	48	-
combined-incomplete	OnePart	100	100	80	100	89	100	-	44	100	-	45	-	65	-	48	-	19	-	50	-
combined-incomplete	GenePart	100	100	66	100	89	100	-	60	100	-	36	-	63	-	53	-	-	23	44	-
combined-incomplete	CodonPart	100	100	78	100	88	100	41	-	100	-	45	-	65	-	58	-	-	27	47	-
combined-incomplete	MixtPart	100	100	80	100	88	100	56	-	100	-	-	15	85	-	83	-	37	-	62	-

Matrix	Partitioning	Node21	Node22	Node23	Node24	Node25	Node26	Node27	Node28	Node29	Node30	Node31	Node32	Node33	Node34	Node35	Node36	Node37	Node38	Node39	Node40
plastid-gene	OnePart	-	35	-	27	100	-	39	-	100	40	-	-	93	-	22	98	100	100	81	100
plastid-gene	GenePart	-	32	-	21	100	-	45	-	100	43	-	19	94	-	-	97	100	100	84	100
plastid-gene	CodonPart	-	37	-	-	100	-	47	-	100	42	10	-	87	-	17	98	100	100	84	100
plastid-gene	MixtPart	31	77	-	68	100	-	15	-	100	58	-	-	99	-	14	100	100	100	93	100
combined-complete	OnePart	48	35	-	22	100	29	-	-	100	37	-	-	93	40	-	98	100	100	75	100
combined-complete	GenePart	44	28	-	-	100	-	-	-	100	39	-	18	94	-	-	99	100	100	76	100
combined-complete	CodonPart	53	29	-	-	100	36	-	-	100	40	-	-	85	-	-	99	100	100	80	100
combined-complete	MixtPart	62	75	-	62	100	-	-	-	100	58	-	-	97	44	8	100	100	100	95	100
combined-incomplete	OnePart	47	69	-	-	100	-	-	63	100	56	15	-	96	-	-	100	100	100	89	100
combined-incomplete	GenePart	-	60	24	-	100	30	-	64	100	61	-	-	96	-	-	100	100	100	91	100
combined-incomplete	CodonPart	-	62	26	-	100	-	-	59	100	57	20	-	94	-	-	100	100	100	91	100
combined-incomplete	MixtPart	70	81	-	48	100	-	-	75	100	64	-	-	96	-	17	100	100	100	84	100

Matrix	Partitioning	Node41	Node42	Node43	Node44	Node45	Node46	Node47	Node48	Node49	Node50	Node51	Node52	Node53	Node54	Node55	Node56	Node57	Node58	Node59	Node60
plastid-gene	OnePart	-	37	100	100	100	100	28	-	92	60	-	100	100	-	100	100	100	-	100	100
plastid-gene	GenePart	-	-	100	100	100	100	19	25	91	60	-	100	100	-	100	100	100	-	100	100
plastid-gene	CodonPart	-	47	100	100	100	100	26	-	93	60	-	100	100	-	100	100	100	-	100	100
plastid-gene	MixtPart	-	56	100	100	100	100	-	-	100	60	-	100	100	60	100	100	100	-	100	100
combined-complete	OnePart	-	36	100	100	100	100	-	-	98	-	38	100	100	44	100	100	100	-	100	100
combined-complete	GenePart	8	-	100	100	100	100	-	24	98	48	32	100	100	46	100	100	100	-	100	100
combined-complete	CodonPart	-	49	100	100	100	100	-	-	98	45	29	100	100	43	100	100	100	7	100	100
combined-complete	MixtPart	-	52	100	100	100	100	-	-	100	-	55	100	100	72	100	100	100	-	100	100
combined-incomplete	OnePart	-	43	100	100	100	79	-	-	100	52	-	99	100	36	100	100	100	27	100	100
combined-incomplete	GenePart	-	45	100	100	100	82	29	-	100	73	-	99	100	-	100	100	100	42	100	100
combined-incomplete	CodonPart	-	54	100	100	100	80	38	-	99	48	-	99	100	-	100	100	100	36	100	100
combined-incomplete	MixtPart	-	63	100	100	100	96	-	-	100	85	-	100	100	64	100	100	100	-	100	100

Matrix	Partitioning	Node61	Node62	Node63	Node64	Node65	Node66	Node67	Node68	Node69	Node70	Node71	Node72	Node73	Node74	Node75	Node76	Node77	Node78
plastid-gene	OnePart	100	100	100	98	100	61	-	64	100	100	-	100	97	82	-	44	100	100
plastid-gene	GenePart	100	100	100	99	100	73	-	75	100	100	-	100	97	89	-	60	100	100
plastid-gene	CodonPart	100	100	100	100	100	75	-	73	100	100	-	100	93	85	73	-	100	100
plastid-gene	MixtPart	100	100	100	98	100	31	-	48	100	100	-	100	100	91	-	68	100	100
combined-complete	OnePart	100	100	100	100	100	53	-	39	100	100	-	100	98	89	-	56	100	100
combined-complete	GenePart	100	100	100	100	100	63	-	78	100	100	-	100	99	94	-	71	100	100
combined-complete	CodonPart	100	100	100	100	100	52	-	64	100	100	-	100	95	93	60	-	100	100
combined-complete	MixtPart	100	100	100	100	100	-	46	-	100	100	42	100	100	100	-	73	100	100
combined-incomplete	OnePart	100	100	100	41	100	-	-	14	55	100	65	100	100	91	59	-	100	81
combined-incomplete	GenePart	100	100	100	35	100	-	-	19	62	100	67	100	99	97	-	75	100	90
combined-incomplete	CodonPart	100	100	100	100	100	-	-	49	50	100	62	100	99	96	-	40	100	77
combined-incomplete	MixtPart	100	100	100	34	100	-	-	14	60	100	75	100	100	96	-	75	100	100

Node number	Clade
Node1	(Achariaceae,Anthodiscus,Arabidopsis,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Calycanthus,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Cucurbitaceae,Dichapetalum,Elaeodendron,Erythroxylum,Euphorbia,Euphronia,Fabaceae,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Nicotiana,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia,Vitis)
Node2	(Achariaceae,Anthodiscus,Arabidopsis,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Cucurbitaceae,Dichapetalum,Elaeodendron,Erythroxylum,Euphorbia,Euphronia,Fabaceae,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Nicotiana,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia,Vitis)
Node3	(Achariaceae,Anthodiscus,Arabidopsis,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Cucurbitaceae,Dichapetalum,Elaeodendron,Erythroxylum,Euphorbia,Euphronia,Fabaceae,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia,Vitis)
Node4	(Achariaceae,Anthodiscus,Arabidopsis,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Cucurbitaceae,Dichapetalum,Elaeodendron,Erythroxylum,Euphorbia,Euphronia,Fabaceae,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia)
Node5	(Achariaceae,Anthodiscus,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Cucurbitaceae,Dichapetalum,Elaeodendron,Erythroxylum,Euphorbia,Euphronia,Fabaceae,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia)
Node6	(Achariaceae,Anthodiscus,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Dichapetalum,Elaeodendron,Erythroxylum,Euphorbia,Euphronia,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia)
Node7	(Achariaceae,Anthodiscus,Averrhoa,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Dichapetalum,Erythroxylum,Euphorbia,Euphronia,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia)
Node8	(Achariaceae,Anthodiscus,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Dichapetalum,Erythroxylum,Euphorbia,Euphronia,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia)
Node9	(Achariaceae,Anthodiscus,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Dichapetalum,Erythroxylum,Euphorbia,Euphronia,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Putranjiva,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola,Vismia)
Node10	(Achariaceae,Casearia,Clusia,Ctenolophon,Erythroxylum,Euphorbia,Flueggea,Galearia,Garcinia,Goupia,Humiria,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Ixonanthes,Jatropha,Lacistema,Linum,Mammea,Manihot,Medusagyne,Microdesmis,Ochnaceae,Passiflora,Pera,Phyllanthus,Ploiarium,Podocalyx,Podostemum,Populus_alba,Populus_spp.,Quiina,Rhizophoraceae,Schistostemon,Scyphostegia,Turnera,Viola,Vismia)
Node11	(Achariaceae,Anthodiscus,Balanops,Bergia,Bhesa,Byrsonima,Casearia,Centroplacus,Chrysobalanus,Dichapetalum,Euphorbia,Euphronia,Flueggea,Goupia,Humiria,Ixonanthes,Jatropha,Lacistema,Linum,Manihot,Passiflora,Pera,Phyllanthus,Podocalyx,Populus_alba,Populus_spp.,Putranjiva,Schistostemon,Scyphostegia,Trigonia,Turnera,Viola)
Node12	(Anthodiscus,Balanops,Bergia,Bhesa,Byrsonima,Centroplacus,Chrysobalanus,Clusia,Ctenolophon,Dichapetalum,Erythroxylum,Euphronia,Galearia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Mammea,Medusagyne,Microdesmis,Ochnaceae,Ploiarium,Podostemum,Putranjiva,Quiina,Rhizophoraceae,Trigonia,Vismia)
Node13	(Achariaceae,Casearia,Euphorbia,Flueggea,Goupia,Humiria,Ixonanthes,Jatropha,Lacistema,Linum,Manihot,Passiflora,Pera,Phyllanthus,Podocalyx,Populus_alba,Populus_spp.,Schistostemon,Scyphostegia,Turnera,Viola)
Node14	(Achariaceae,Casearia,Euphorbia,Flueggea,Goupia,Ixonanthes,Jatropha,Lacistema,Linum,Manihot,Passiflora,Pera,Phyllanthus,Podocalyx,Populus_alba,Populus_spp.,Scyphostegia,Turnera,Viola)
Node15	(Clusia,Ctenolophon,Erythroxylum,Galearia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Mammea,Medusagyne,Microdesmis,Ochnaceae,Ploiarium,Podostemum,Quiina,Rhizophoraceae,Vismia)
Node16	(Clusia,Ctenolophon,Erythroxylum,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Medusagyne,Ochnaceae,Ploiarium,Podostemum,Quiina,Rhizophoı
Node17	(Clusia,Galearia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Irvingia,Mammea,Medusagyne,Microdesmis,Ochnaceae,Ploiarium,Podostemum,Quiina,Vismia)
Node18	(Clusia,Ctenolophon,Erythroxylum,Galearia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Microdesmis,Ploiarium,Podostemum,Rhizophoraceae,Vismia)
Node19	(Achariaceae,Casearia,Goupia,Humiria,Lacistema,Passiflora,Populus_alba,Populus_spp.,Schistostemon,Scyphostegia,Turnera,Viola)
Node20	(Clusia,Ctenolophon,Erythroxylum,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Ploiarium,Podostemum,Rhizophoraceae,Vismia)
Node21	(Clusia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Medusagyne,Ochnaceae,Ploiarium,Podostemum,Quiina,Vismia)
Node22	(Anthodiscus,Balanops,Bergia,Bhesa,Byrsonima,Centroplacus,Chrysobalanus,Dichapetalum,Euphronia,Putranjiva,Trigonia)
Node23	(Clusia,Galearia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Microdesmis,Ploiarium,Podostemum,Vismia)
Node24	(Anthodiscus,Balanops,Bergia,Bhesa,Byrsonima,Centroplacus,Chrysobalanus,Dichapetalum,Euphronia,Trigonia)
Node25	(Achariaceae,Casearia,Goupia,Lacistema,Passiflora,Populus_alba,Populus_spp.,Scyphostegia,Turnera,Viola)
Node26	(Balanops,Bergia,Bhesa,Byrsonima,Centroplacus,Chrysobalanus,Dichapetalum,Euphronia,Trigonia)
Node27	(Achariaceae,Casearia,Lacistema,Passiflora,Populus_alba,Populus_spp.,Scyphostegia,Turnera,Viola)
Node28	(Casearia,Goupia,Lacistema,Passiflora,Populus_alba,Populus_spp.,Scyphostegia,Turnera,Viola)
Node29	(Clusia,Garcinia,Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Ploiarium,Podostemum,Vismia)
Node30	(Euphorbia,Flueggea,Ixonanthes,Jatropha,Linum,Manihot,Pera,Phyllanthus,Podocalyx)
Node31	(Anthodiscus,Balanops,Chrysobalanus,Dichapetalum,Euphronia,Putranjiva,Trigonia)
Node32	(Balanops,Bhesa,Centroplacus,Chrysobalanus,Dichapetalum,Euphronia,Trigonia)
Node33	(Casearia,Lacistema,Passiflora,Populus_alba,Populus_spp.,Scyphostegia,Turnera)
Node34	(Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Ploiarium,Podostemum,Vismia)
Node35	(Anthodiscus,Balanops,Chrysobalanus,Dichapetalum,Euphronia,Trigonia)
Node36	(Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Mammea,Podostemum,Vismia)
Node37	(Balanops,Chrysobalanus,Dichapetalum,Euphronia,Trigonia)
Node38	(Casearia,Lacistema,Populus_alba,Populus_spp.,Scyphostegia)
Node39	(Flueggea,Ixonanthes,Linum,Phyllanthus,Podocalyx)
Node40	(Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Podostemum,Vismia)
Node41	(Anthodiscus,Bergia,Byrsonima,Putranjiva)
Node42	(Bergia,Bhesa,Byrsonima,Centroplacus)
Node43	(Casearia,Populus_alba,Populus_spp.,Scyphostegia)
Node44	(Chrysobalanus,Dichapetalum,Euphronia,Trigonia)
Node45	(Euphorbia,Jatropha,Manihot,Pera)
Node46	(Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.,Vismia)
Node47	(Irvingia,Medusagyne,Ochnaceae,Quiina)
Node48	(Anthodiscus,Bergia,Byrsonima)
Node49	(Ctenolophon,Erythroxylum,Rhizophoraceae)
Node50	(Clusia,Garcinia,Ploiarium)
Node51	(Achariaceae,Goupia,Viola)
Node52	(Euphorbia,Jatropha,Manihot)
Node53	(Flueggea,Phyllanthus,Podocalyx)
Node54	(Galearia,Irvingia,Microdesmis)
Node55	(Hypericum_fraseri,Hypericum_kalmianum,Hypericum_spp.)
Node56	(Medusagyne,Ochnaceae,Quiina)
Node57	(Populus_alba,Populus_spp.,Scyphostegia)
Node58	(Anthodiscus,Putranjiva)
Node59	(Bergia,Byrsonima)
Node60	(Bhesa,Centroplacus)
Node61	(Erythroxylum,Rhizophoraceae)
Node62	(Chrysobalanus,Euphronia)
Node63	(Clusia,Garcinia)
Node64	(Cucurbitaceae,Fabaceae)
Node65	(Dichapetalum,Trigonia)
Node66	(Achariaceae,Viola)
Node67	(Euphorbia,Jatropha)
Node68	(Euphorbia,Manihot)
Node69	(Flueggea,Phyllanthus)
Node70	(Galearia,Microdesmis)
Node71	(Goupia,Viola)
Node72	(Humiria,Schistostemon)
Node73	(Hypericum_fraseri,Hypericum_kalmianum)
Node74	(Ixonanthes,Linum)
Node75	(Medusagyne,Ochnaceae)
Node76	(Medusagyne,Quiina)
Node77	(Passiflora,Turnera)
Node78	(Populus_alba,Populus_spp.)

Table S5. Newly designed primers for amplifying Malpighiales plastid genomes from genomic DNA using long-range PCR amplification.

Primer	Sequence (5′ to 3′)	Position in *Manihot esculenta* plastid genome	Gene
CP1-F	TGCCTTRATCCACTTGGCTACAT	50-72	*trnH*-GUG
CP1-R	GTACTCGGCTTTTAAGTGCGRCT	4311-4333	*trnK*-UUU
CP2-F	AGYCGCACTTAAAAGCCGAGTAC	4311-4333	*trnK*-UUU
CP2-R	CTTACAGCRGCTTGCCAAACA	8113-8133	*psbK*
CP3-F	TGTTTGGCAAGCYGCTGTAAG	8113-8133	*psbK*
CP3-R	CAGAATCAATTGGCAAGAGGTCAA	11523-11546	*atpA*
CP4-F	TTGACCTCTTGCCAATTGATTCTG	11523-11546	*atpA*
CP4-R	GTTATTGCTGCTGGRTTGGC	14165-14184	*atpH*
CP5-F	GCCAAYCCAGCAGCAATAAC	14165-14184	*atpH*
CP5-R	TGGTTTCMGAAGABGGAATGTC	17875-17896	*rpoC2*
CP6-F	GACATTCCVTCTTCKGAAACCA	17875-17896	*rpoC2*
CP6-R	TTGATGGRGATCAAATGGCTGTTC	22322-22345	*rpoC1*
CP7-F	GAACAGCCATTTGATCYCCATCAA	22322-22345	*rpoC1*
CP7-R	ATCCATTGACACAAATMGTTCATGGG	26659-26684	*rpoB*
CP8-F	CCCATGAACKATTTGTGTCAATGGAT	26659-26684	*rpoB*
CP8-R	TAGAGTCCACTYCTTCCCCA	29442-29461	*petN*
CP9-F	TGGGGAAGRAGTGGACTCTA	29442-29461	*petN*
CP9-R	GACGGACTGTAAATTCGTTGGC	32680-32701	*trnY*-GUA
CP10-F	GCCAACGAATTTACAGTCCGTC	32680-32701	*trnY*-GUA
CP10-R	GCACCATGAATRGCGCATAGC	35946-35966	*psbD*
CP11-F	GCTATGCGCYATTCATGGTGC	35946-35966	*psbD*
CP11-R	TGGTAGCTCGCAAGGCTCATAAC	39437-39459	*trnfM*-CAU
CP12-F	GTTATGAGCCTTGCGAGCTACCA	39437-39459	*trnfM*-CAU
CP12-R	AGGACTCAATGATTATTCGTTCGC	44539-44562	*psaA*
CP13-F	GCGAACGAATAATCATTGAGTCCT	44539-44562	*psaA*
CP13-R	CTTSAACCACTCGGCCATCTC	47631-47651	*trnS*-GGA
CP14-F	GAGATGGCCGAGTGGTTSAAG	47631-47651	*trnS*-GGA
CP14-R	TTCCTGTVGATGTVTATTTGCC	52345-52366	*ndhK*
CP15-F	GGCAAATABACATCBACAGGAA	52345-52366	*ndhK*
CP15-R	GACATTGAKCCACAAGAAGCTCAG	55534-55557	*atpE*
CP16-F	CTGAGCTTCTTGTGGMTCAATGTC	55534-55557	*atpE*
CP16-R	CGTCCYTCATTACGAGCTTKTACA	59343-59366	*rbcL*
CP17-F	TGTAMAAGCTCGTAATGARGGACG	59343-59366	*rbcL*
CP17-R	DGGCCCAGCAGAAATTRCT	63441-63459	*ycf4*
CP18-F	AGYAATTTCTGCTGGGCCH	63441-63459	*ycf4*
CP18-R	CAATGCARTTCATCCAACGATAA	68299-68321	*psbF*
CP19-F	TTATCGTTGGATGAAYTGCATTG	68299-68321	*psbF*

CP19-R	CTATAATCAATTCGATCCCCCGAT	72647-72670	*rps18*
CP20-F	ATCGGGGGATCGAATTGATTATAG	72647-72670	*rps18*
CP20-R	GTATGAACACGATACCAAGGCAA	77275-77297	*psbB*
CP21-F	TTGCCTTGGTATCGTGTTCATAC	77275-77297	*psbB*
CP21-R	CATACTTGCYGACCATCGATGAA	80738-80760	*petB*
CP22-F	TTCATCGATGGTCRGCAAGTATG	80738-80760	*petB*
CP22-R	GCGAGGAGCTGGATGAGAAG	86741-86760	*rpl16*
CP23-F	CTTCTCATCCAGCTCCTCGC	86741-86760	*rpl16*
CP23-R	TGCTCAGCAACAGTCGGACA	89774-89793	*rpl2*
CP24-F	TGTCCGACTGTTGCTGAGCA	89774-89793	*rpl2*
CP24-R	ACCTCRGACCAATCAATCGAATATT	94196-94220	*ycf2*
CP25-F	AATATTCGATTGATTGGTCYGAGGT	94196-94220	*ycf2*
CP25-R	CATTCAGTGACTTTGGCACTGGA	99291-99313	Spacer
CP26-F	TCCAGTGCCAAAGTCACTGAATG	99291-99313	Spacer
CP26-R	GCTGTCGGAGTAAAGGATCGTCA	103966-103988	*rps12*
CP27-F	TGACGATCCTTTACTCCGACAGC	103966-103988	*rps12*
CP27-R	TGCTAATGTGCCTTGGATGATCC	108149-108171	*ndhB*
CP28-F	GGATCATCCAAGGCACATTAGCA	108149-108171	*ndhB*
CP28-R	ACATCACTGCACTTCCACTTGACAC	113096-113120	*rrn4.5*
CP29-F	GTGTCAAGTGGAAGTGCAGTGATGT	113096-113120	*rrn4.5*
CP29-R	TTCGGCTCTDATACATGCTGCTAC	117660-117683	*ndhF*
CP30-F	GTAGCAGCATGTATHAGAGCCGAA	117660-117683	*ndhF*
CP30-R	GCYACTCGGACTCGAACCGAGAT	120468-120490	*trnL*-UAG
CP31-F	ATCTCGGTTCGAGTCCGAGTRGC	120468-120490	*trnL*-UAG
CP31-R	GAATTYGATAAATGCATTGCTTGTGA	125813-125838	*ndhI*
CP32-F	TCACAAGCAATGCATTTATCRAATTC	125813-125838	*ndhI*
CP32-R	TATGGCTHGGMCCTTTTATGG	129128-129148	*ndhH*

Table S6. Fossil age constraints used in divergence time estimates for Malpighiales. Source of age constraints presented below.

No.	Fossil taxon	Minimum age (Ma)	Most recent common ancestor
1	*Acalypha* type[a]	61.0	*Acalypha* and *Spathiostemon*
2	*Malvacipollis diversus*[b]	55.5	*Austrobuxus* and *Micrantheum*
3	*Balanops caledonica*[c]	23.8	*Balanops* and *Dichapetalum*
4	*Retisyncolporites angularis*[d]	55.5	*Anthodiscus* and *Caryocar*
5	*Casearia* type[e]	37.0	*Casearia* and *Lunania*
6	*Chrysobalanus* type[f]	49.0	*Chrysobalanus* and *Hirtella*
7	*Palaeoclusia chevalieri*[g]	89.0	*Clusia* and *Mammea*
8	*Ctenolophonidites costatus*[h,i]	66.0	*Aneulophus* and *Ctenolophon*
9	*Platydiscus peltatus*[j]	83.5	*Cephalotus* and *Eucryphia*
10	*Drypetes* type[k]	33.7	*Drypetes* and *Lophopyxis*
11	*Crepetocarpon perkinsii*[l,m]	40.0	*Homalanthus* and *Hura*
12	*Phyllanthus* type[n]	33.7	*Heywoodia* and *Phyllanthus*
13	*Pseudosalix handleyi*[o]	48.0	*Idesia* and *Salix*
14	*Perisyncolporites pokornyi*[p,q,r]	49.0	*Acridocarpus* and *Dicella*
15	Vitaceae type[s]	57.9	*Leea* and *Vitis*
16	Fabales type[t]	59.9	*Lotus* and *Prunus*

References

a. Sun M, Sun X, Zhao Y, Wang O, Li Z, Hu Z, Mei P (1989) in *Cenozoic paleobiota of the continental shelf of East China Sea*, ed Division of Comprehensive Studies on Ocean Geology of the Ministry of Geology and Minerals of P. R. China and Institute of Geology of the Chinese Academy of Geological Sciences (Geological Publishing House, Beijing), pp 6-111.

b. Martin HA (1981) Changing Cenozoic barriers and the Australian paleobotanical record. *Ann Missouri Bot Gard* 69:625-667.

c. Simpson JB (1961) The Tertiary pollen-flora of Mull and Ardnamurchan. *T Roy Soc Edin-Earth* 64:421-468.

d. Rull V (2000) Ecostratigraphic study of Paleocene and Early Eocene palynological cyclicity in northern South America. *PALAIOS* 15:14-24.

e. Graham A (1985) Studies in neotropical paleobotany. IV. The Eocene communities of Panama. *Ann Missouri Bot Gard* 72:504-534.

f. Wodehouse RP (1932) Tertiary pollen. I. Pollen of the living representatives of the Green River flora. *Bull Torrey Bot Club* 59:313-340.

g. Crepet WL, Nixon KC (1998) Fossil Clusiaceae from the Late Cretaceous (Turonian) of New Jersey and implications regarding the history of bee pollination. *Am J Bot* 85:1122-1133.
h. Edet JJ, Nyong EE (1994) Palynostratigraphy of Nkporo Shale exposures (late Campanian-Maastrichtian) on the Calabar Flank, SE Nigeria. *Rev Palaeobot Palyno* 80:131-147.
i. Schrank E (1994) Palynology of the Yesomma Formation in northern Somalia: a systematic study of pollen, spores and associated phytoplankton from the Late Cretaceous Palmae Province. *Palaeontogr Abt B* 231:63-112.
j. Schönenberger J, Friis EM, Matthews ML, Endress PK (2001) Cunoniaceae in the Cretaceous of Europe: evidence from fossil flowers. *Ann Bot* 88:423-437.
k. Gruas-Cavagnetto C, Anadón P. 1995. Une mangrove complexe dans le Bartonien du Bassin de l'Ebre (NE de l'Espagne). *Palaeontogr Abt B* 236:147-165.
l. Dilcher DL, Manchester SR (1988) Investigations of angiosperms from the Eocene of North America: a fruit belonging to the Euphorbiaceae. *Tertiary Res* 9:45-58.
m. Potter FW, Jr, Dilcher DL (1980) in *Biostratigraphy of fossil plants*, eds Dilcher DL, Taylor TN (Dowden, Hutchinson & Ross, Stroudsburg, PA), pp 211-225
n. Zaklinskaya ED (1978) Palynology of Paleogene Clay from DSDP Site 368, Cape Verde Rise. *Initial Rep Deep Sea* 41:933-937.
o. Boucher LD, Manchester SR, Judd WS (2003) An extinct genus of Salicaceae based on twigs with attached flowers, fruits, and foliage from the Eocene Green River Formation of Utah and Colorado, USA. *Am J Bot* 90:1389-1399.
p. Jaramillo CA, Dilcher DL (2001) Middle Paleogene palynology of central Colombia, South America: a study of pollen and spores from tropical latitudes. *Palaeontogr Abt B* 285:87-213.
q. Davis CC, Anderson WR, Donoghue MJ (2001) Phylogeny of Malpighiaceae: evidence from chloroplast *ndhF* and *trnL-F* nucleotide sequences. *Am J Bot* 88:1830-1846.
r. Davis CC, Fritsch PW, Bell CD, Mathews S (2004) High-latitude Tertiary migrations of an exclusively tropical clade: evidence from Malpighiaceae. *Int J Plant Sci* 165(suppl):S107-S121.

s. Collinson ME, Boulter MC, Holmes PL (1993) in *The fossil record*, vol. 2, ed Benton MJ (Chapman and Hall, London, UK), pp 809-841
t. Herendeen PS, Crane PR (1992) in *Advances in legume systematics*, part 4 (Royal Botanic Gardens, Kew, UK), pp57-68

Table S7. Estimated ages for major Malpighiales clades. Clade numbers refer to numbered nodes in Fig. 1 from the main text.

Clade name	Stem group		Crown group	
	Mean age (Ma)	95% highest posterior density interval (HPD)	Mean age	95% HPD
Malpighiales	117.0	123.1–111.6	109.2	113.1–106.1
Clade 1	108.6	112.2–105.3	107.9	111.8–104.5
Clade 2	109.2	113.1–106.1	106.6	110.9–101.3
Clade 3	108.6	112.2–105.3	107.0	111.4–102.6
Clade 4 (euphorbioids)	107.9	111.8–104.5	106.9	111.3–103.1
Clade 5	106.9	111.3–103.1	102.5	108.6–94.8
Clade 6 (linoids)	102.5	108.6–94.8	90.0	103.4–73.6
Clade 7	107.9	111.8–104.5	105.7	110.0–101.6
Clade 8 (salicoids)	97.1	102.9–90.6	94.4	100.5–87.5
Clade 9	97.1	102.9–90.6	92.0	100.8–81.9
Clade 10	105.2	109.6–99.9	101.5	106.7–96.4
Clade 11 (pandoids)	105.2	109.6–99.9	91.2	107.1–70.5
Clade 12 (malpighioids)	105.0	110.1–99.7	102.1	109.4–93.3
Chrysobalanoids	102.3	109.3–95.8	83.5	95.0–74.8
Clusioids	101.5	106.7–96.4	89.5	91.0–88.4
Ochnoids	101.5	106.7–96.4	77.8	90.5–61.3
Parietal clade	105.7	110.0–101.6	99.2	104.6–93.1
Phyllanthoids	102.5	108.6–94.8	94.0	101.6–86.5
Putranjivoids	107.0	111.4–102.6	34.2	35.8–33.1
Rhizophoroids	106.6	110.9–101.3	66.7	68.6–65.4
Achariaceae	99.2	104.6–93.1	65.0	86.2–48.2
Balanopaceae	83.5	94.9–74.8	-	-
Bonnetiaceae	82.1	89.6–69.6	52.6	66.8–35.7
Calophyllaceae	82.2	88.7–73.6	57.6	72.6–40.0
Caryocaraceae	102.3	109.3–95.8	55.8	57.2–54.8
Centroplacaceae	102.1	109.4–93.3	63.7	94.3–35.1
Chrysobalanaceae	66.2	74.9–60.3	52.9	57.7–49.9
Clusiaceae	82.1	89.6–69.6	52.9	72.4–30.4
Ctenolophonaceae	66.7	68.6–65.4	-	-
Dichapetalaceae	59.7	71.4–47.4	20.7	36.0–6.8
Elatinaceae	86.1	99.9–72.9	48.0	76.3–25.1
Erythroxylaceae	54.6	63.1–38.4	29.3	42.5–12.3
Euphorbiaceae	97.2	103.8–87.1	89.9	97.4–81.2
Euphroniaceae	66.2	74.9–60.3	-	-

Goupiaceae	92.0	100.8–81.9	-	-
Humiriaceae	105.7	110.0–101.6	20.7	32.1–10.4
Hypericaceae	69.7	78.4–59.3	53.7	63.0–42.5
Irvingiaceae	91.2	107.1–70.5	11.8	22.6–4.1
Ixonanthaceae	90.0	103.4–73.6	51.9	75.4–26.6
Lacistemataceae	87.1	94.2–80.5	19.8	49.3–2.6
Linaceae	90.0	103.4–73.6	39.5	54.8–28.9
Lophopyxidaceae	34.2	35.8–33.1	-	-
Malpighiaceae	86.1	99.9–72.9	59.8	69.0–52.5
Malesherbiaceae	62.8	80.3–47.7	-	-
Medusagynaceae	72.3	89.1–54.9	-	-
Ochnaceae	77.8	90.5–61.3	53.0	72.4–38.9
Pandaceae	91.2	107.1–70.5	47.0	72.7–23.3
Passifloraceae	50.0	66.8–35.5	26.6	42.6–11.6
Peraceae	97.2	103.8–87.1	63.5	93.4–32.1
Phyllanthaceae	94.0	101.6–86.5	81.2	93.7–58.8
Picrodendraceae	94.0	101.6–86.5	78.0	92.9–72.0
Podostemaceae	69.7	78.4–59.3	26.4	37.6–16.9
Putranjivaceae	34.2	35.8–33.1	-	-
Quiinaceae	72.3	89.1–54.9	18.1	38.7–4.3
Rhizophoraceae	54.6	63.1–38.4	41.7	54.5–30.4
Salicaceae	68.9	78.7–59.8	58.0	65.1–51.3
Samydaceae	79.2	87.0–72.8	37.4	38.8–36.3
Scyphostegiaceae	68.9	78.7–59.8	-	-
Trigoniaceae	59.7	71.4–47.4	31.6	49.6–12.6
Turneraceae	50.0	66.8–35.5	27.8	40.3–14.6
Violaceae	92.0	100.8–81.9	72.9	86.8–57.4

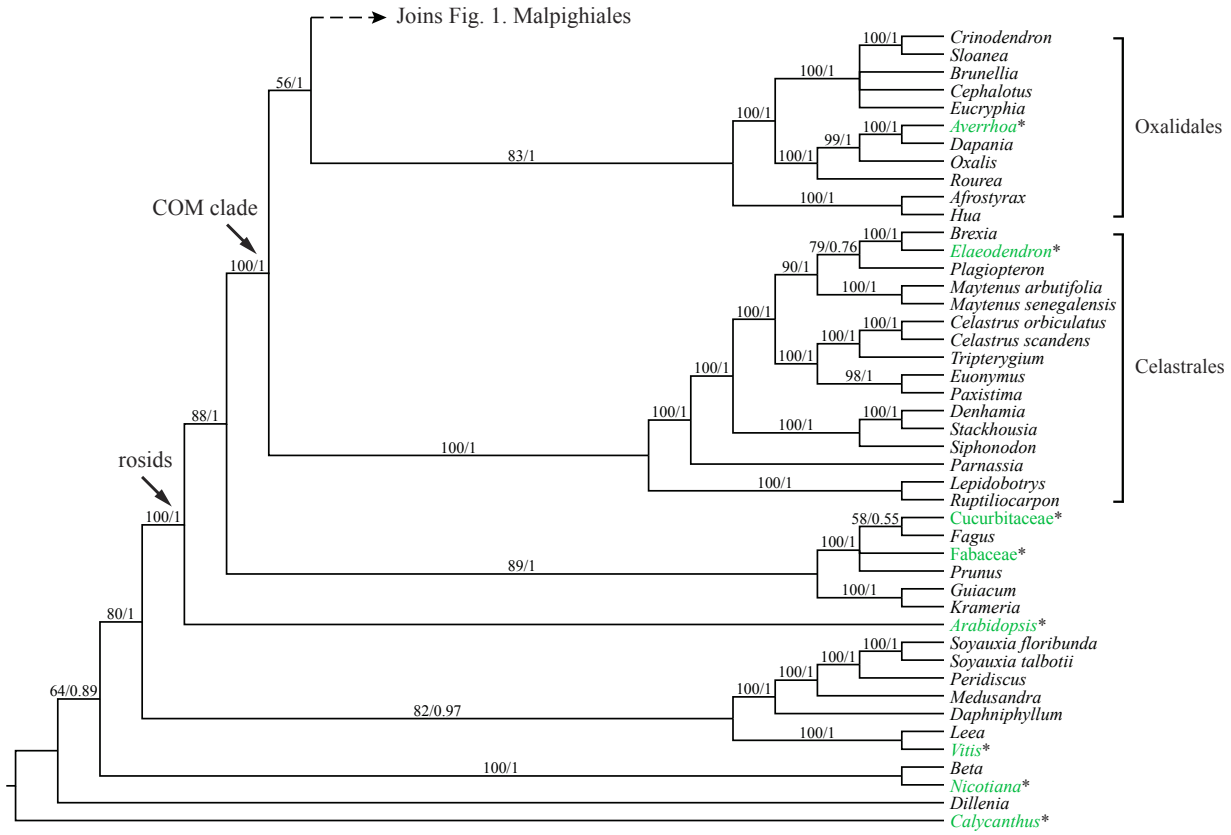

Fig. S1. Continuation of Fig. 1 from the main text. Maximum likelihood (ML) 50% majority-rule bootstrap consensus tree inferred from the ***combined-incomplete*** matrix using the MixtPart partitioning strategy. ML bootstrap percentages/Bayesian posterior probabilities are indicated above each branch. Taxa included in the ***82-gene*** matrix are highlighted in green and marked with asterisks.

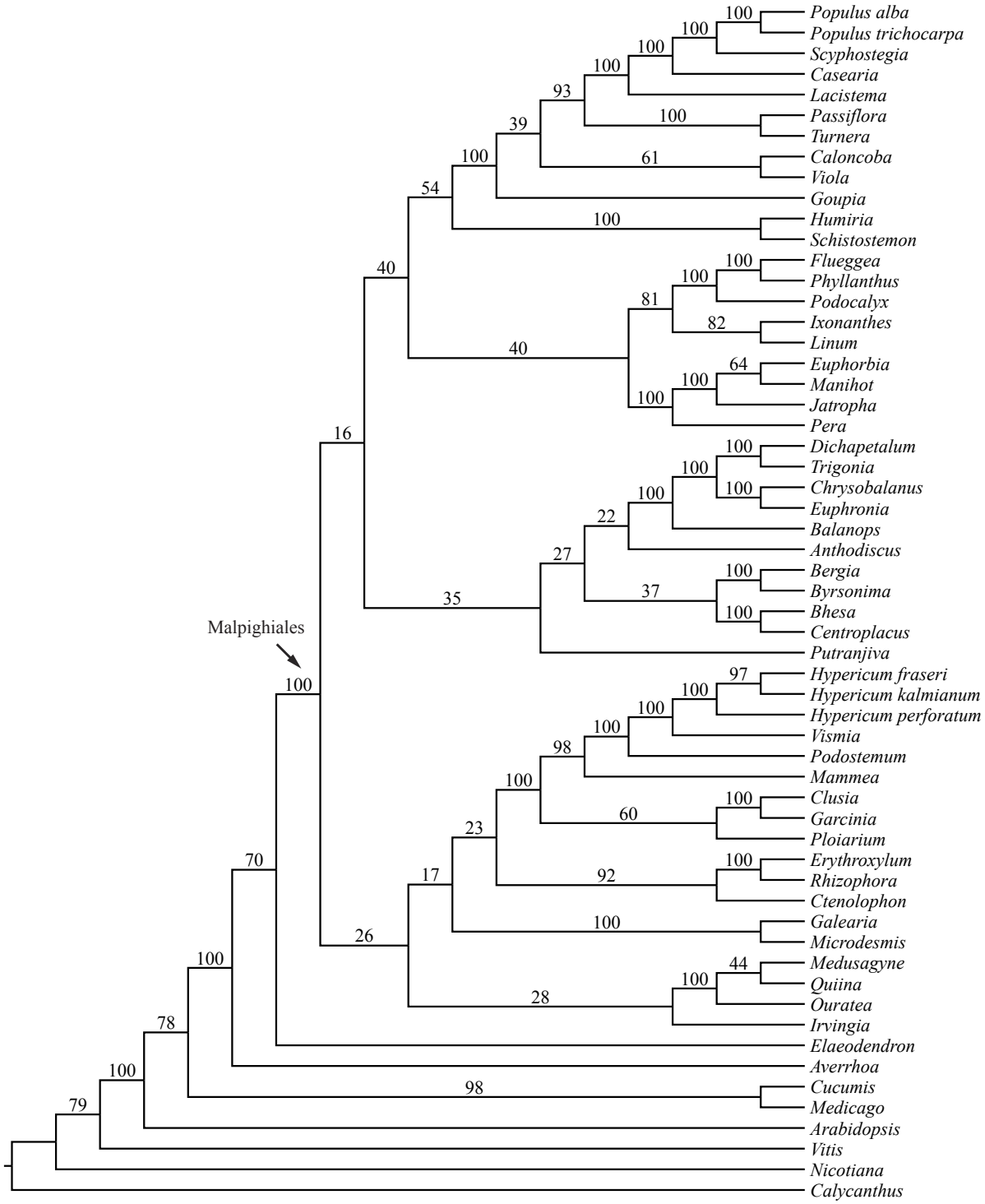

Fig. S2. The maximum likelihood (ML) bipartition tree inferred from the ***82-gene*** matrix using the OnePart partitioning strategy. ML bootstrap percentages are indicated above each branch.

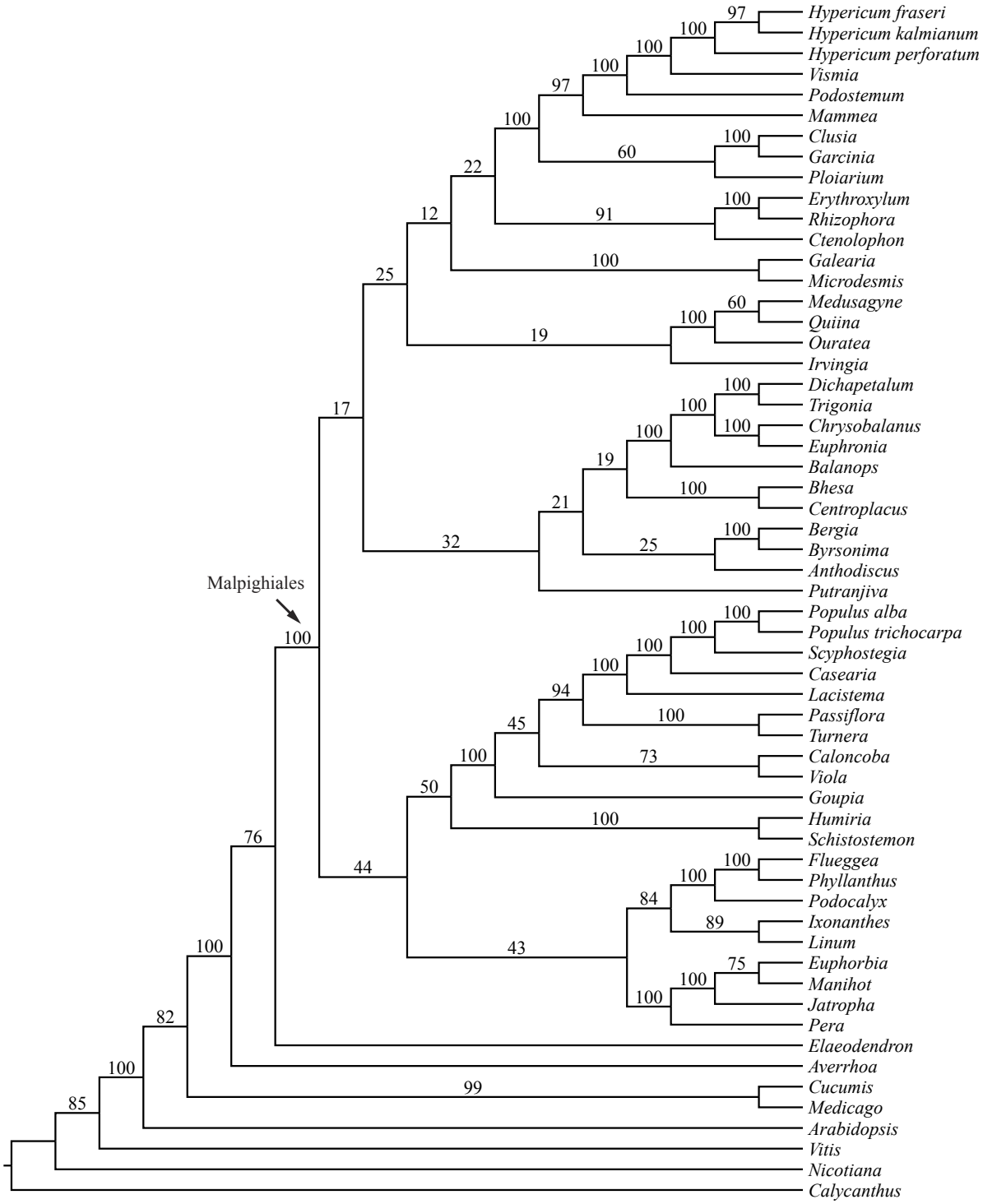

Fig. S3. The maximum likelihood (ML) bipartition tree inferred from the ***82-gene*** matrix using the GenePart partitioning strategy. ML bootstrap percentages are indicated above each branch.

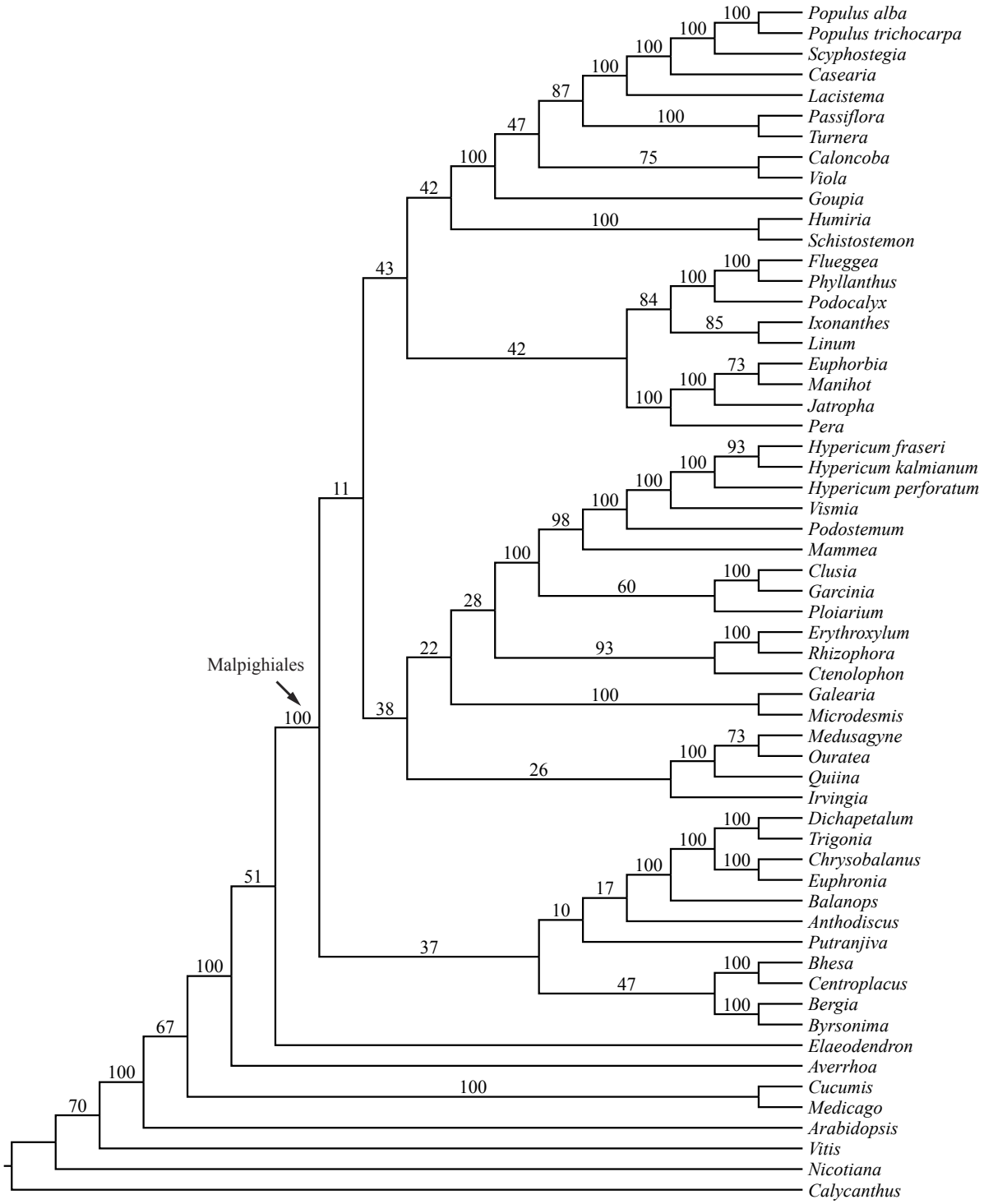

Fig. S4. The maximum likelihood (ML) bipartition tree inferred from the ***82-gene*** matrix using the CodonPart partitioning strategy. ML bootstrap percentages are indicated above each branch.

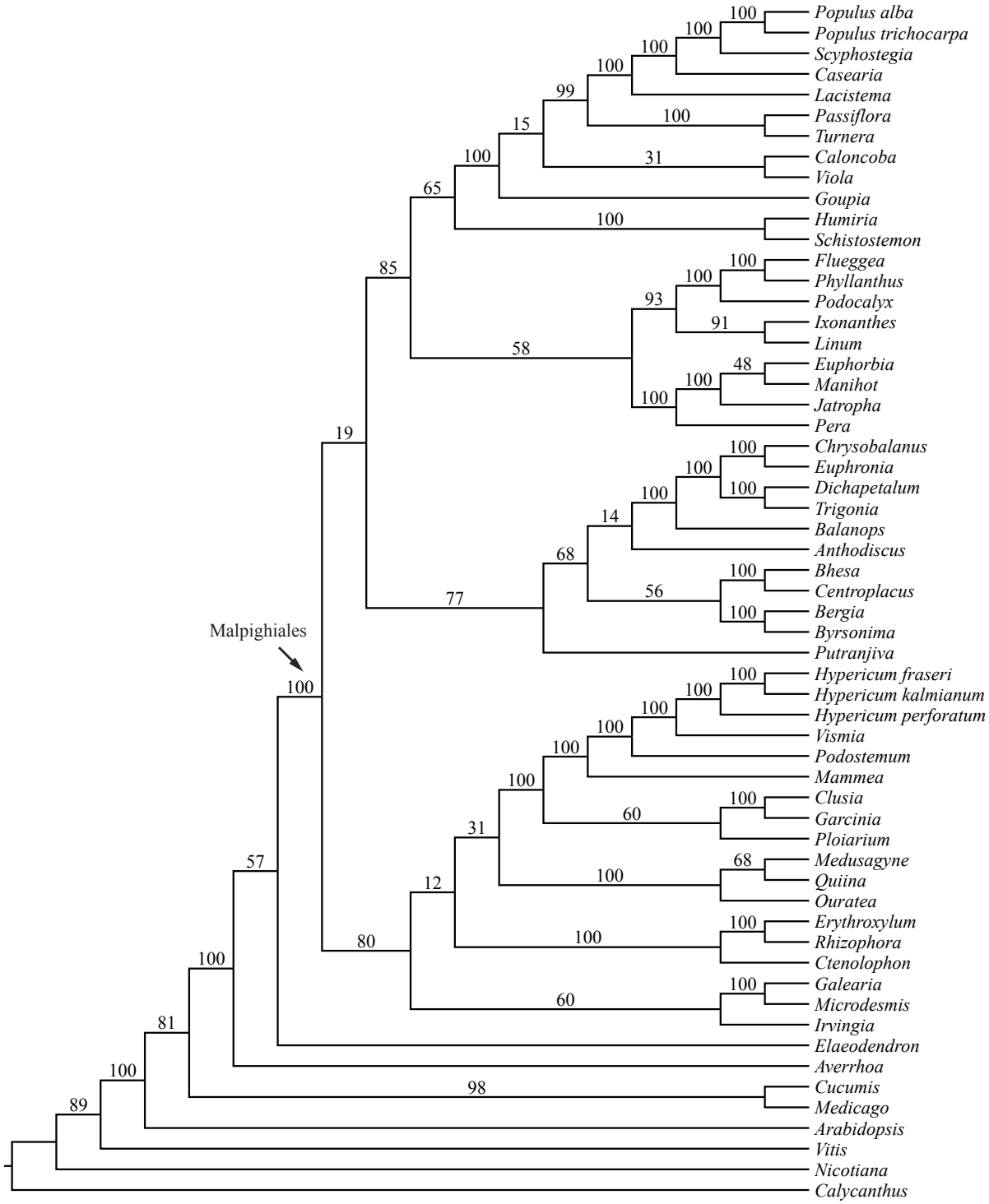

Fig. S5. The maximum likelihood (ML) bipartition tree inferred from the ***82-gene*** matrix using the MixtPart partitioning strategy. ML bootstrap percentages are indicated above each branch.

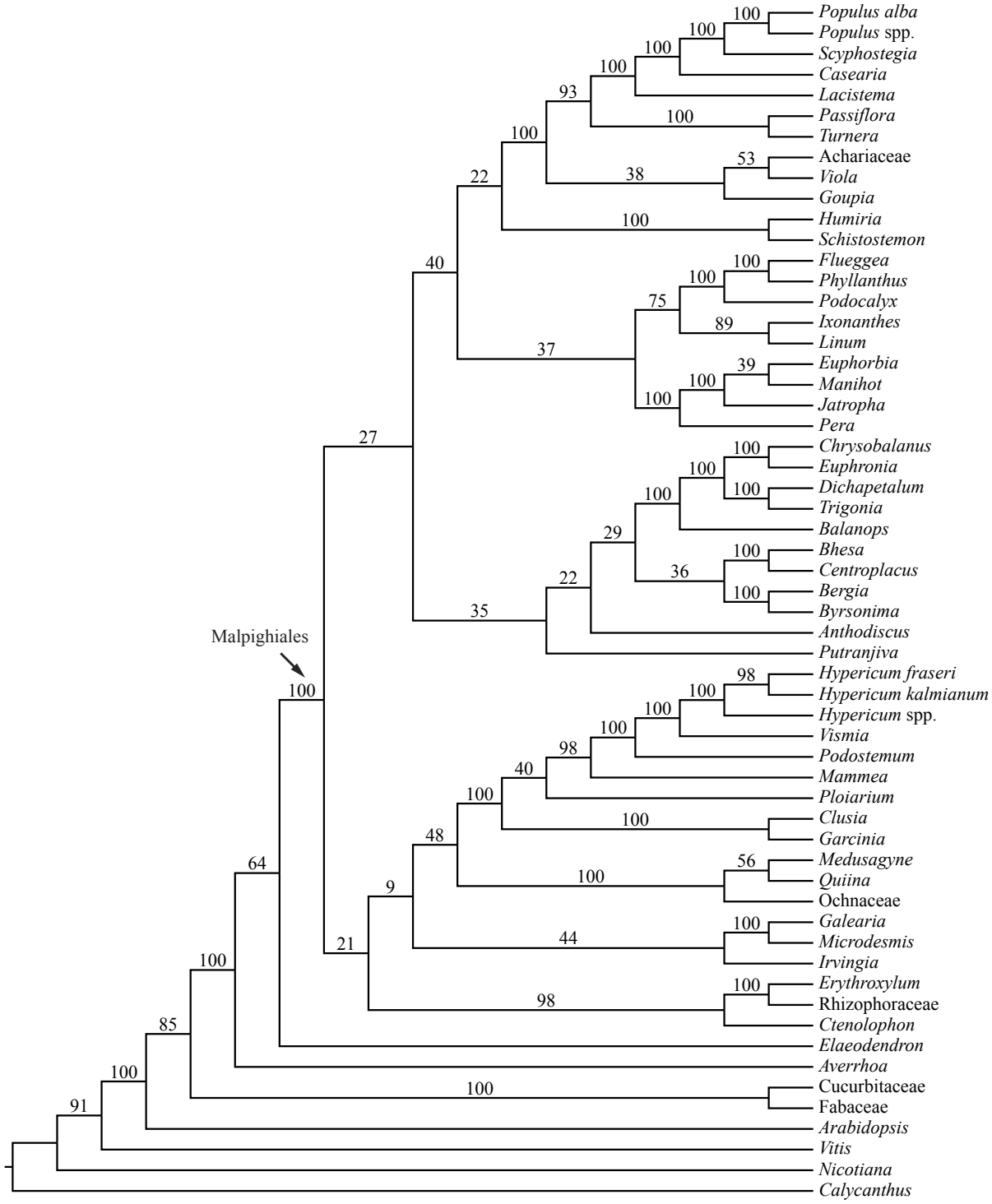

Fig. S6. The maximum likelihood (ML) bipartition tree inferred from the ***combined-complete*** matrix using the OnePart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

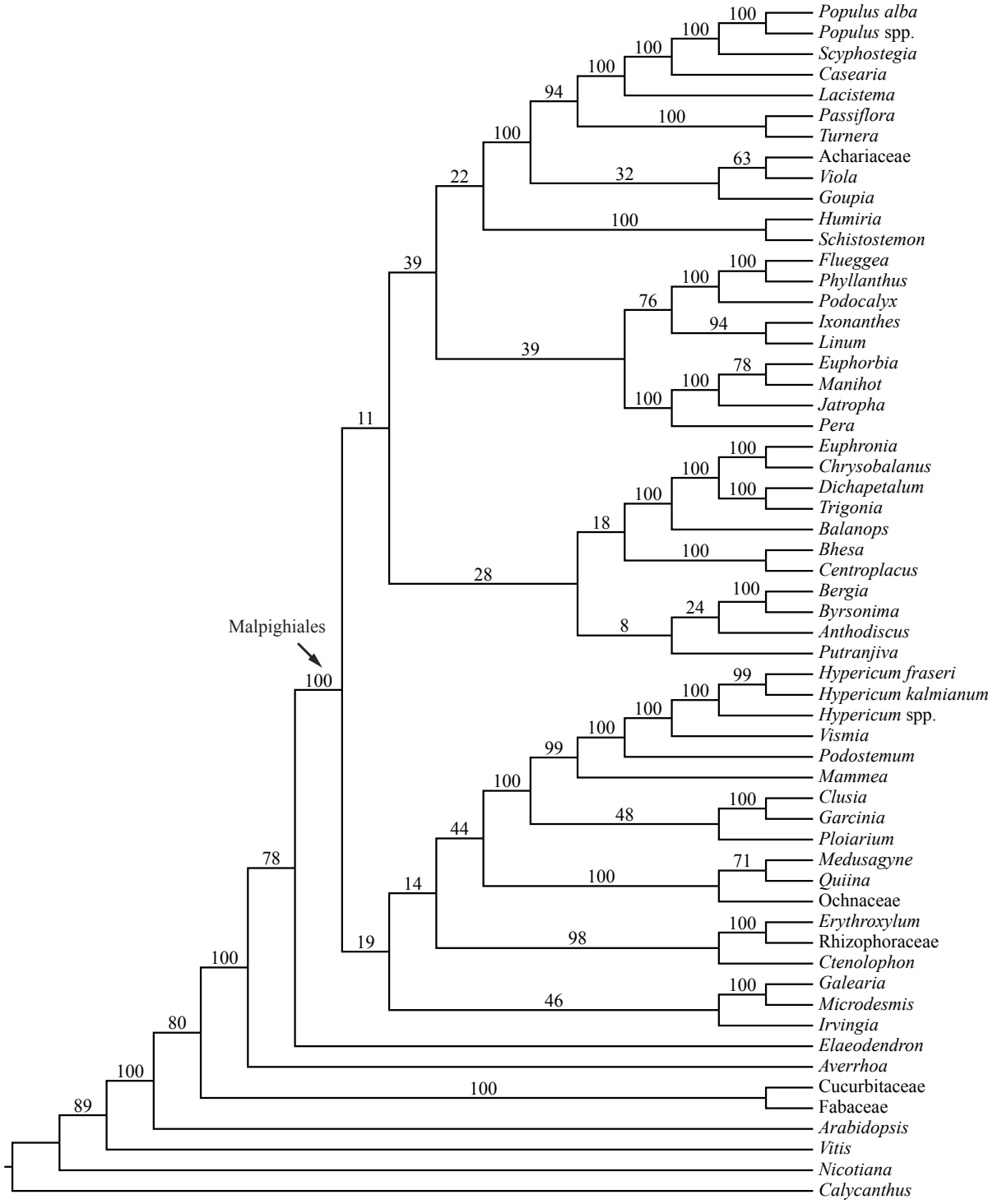

Fig. S7. The maximum likelihood (ML) bipartition tree inferred from the ***combined-complete*** matrix using the GenePart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

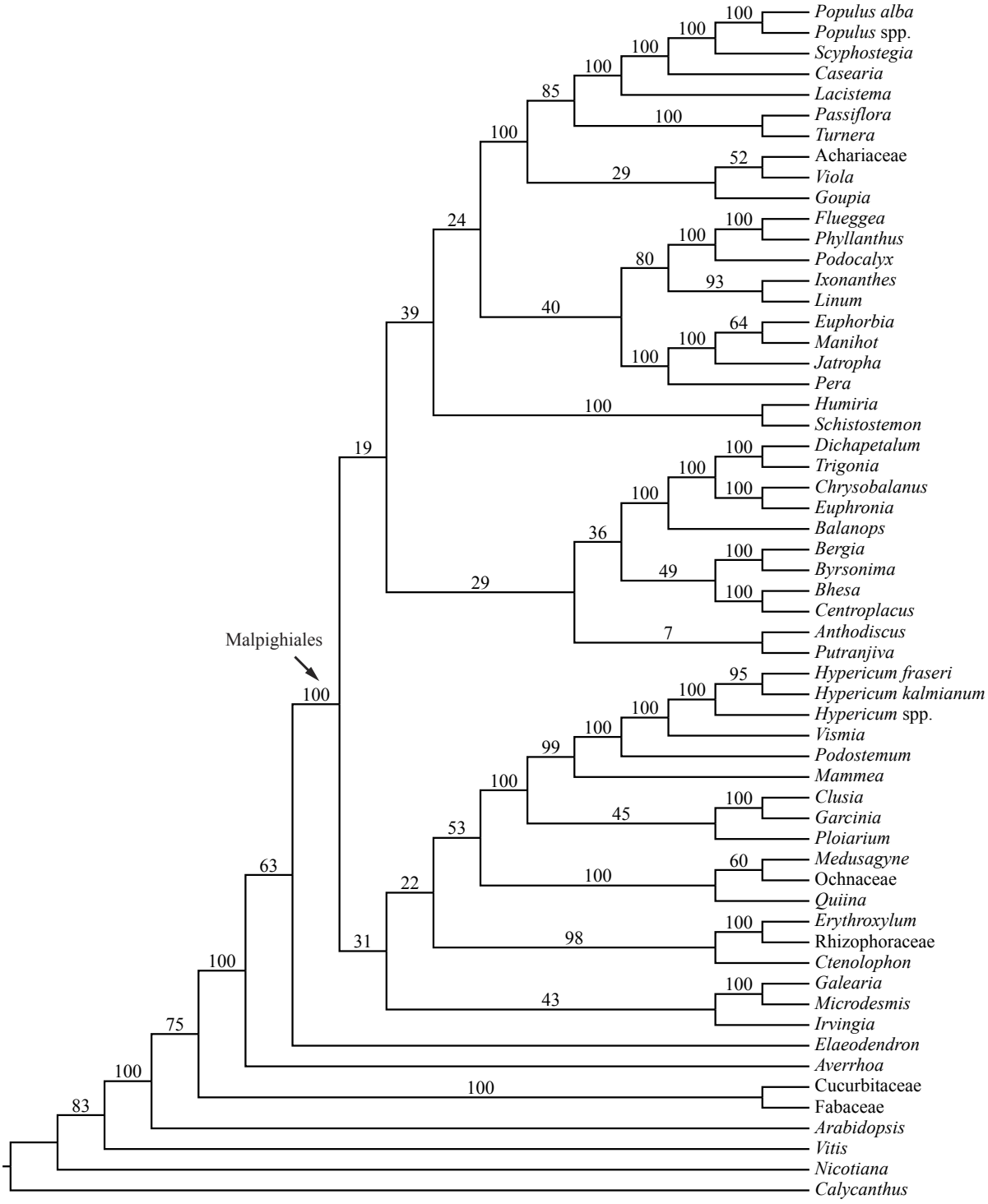

Fig. S8. The maximum likelihood (ML) bipartition tree inferred from the ***combined-complete*** matrix using the CodonPart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

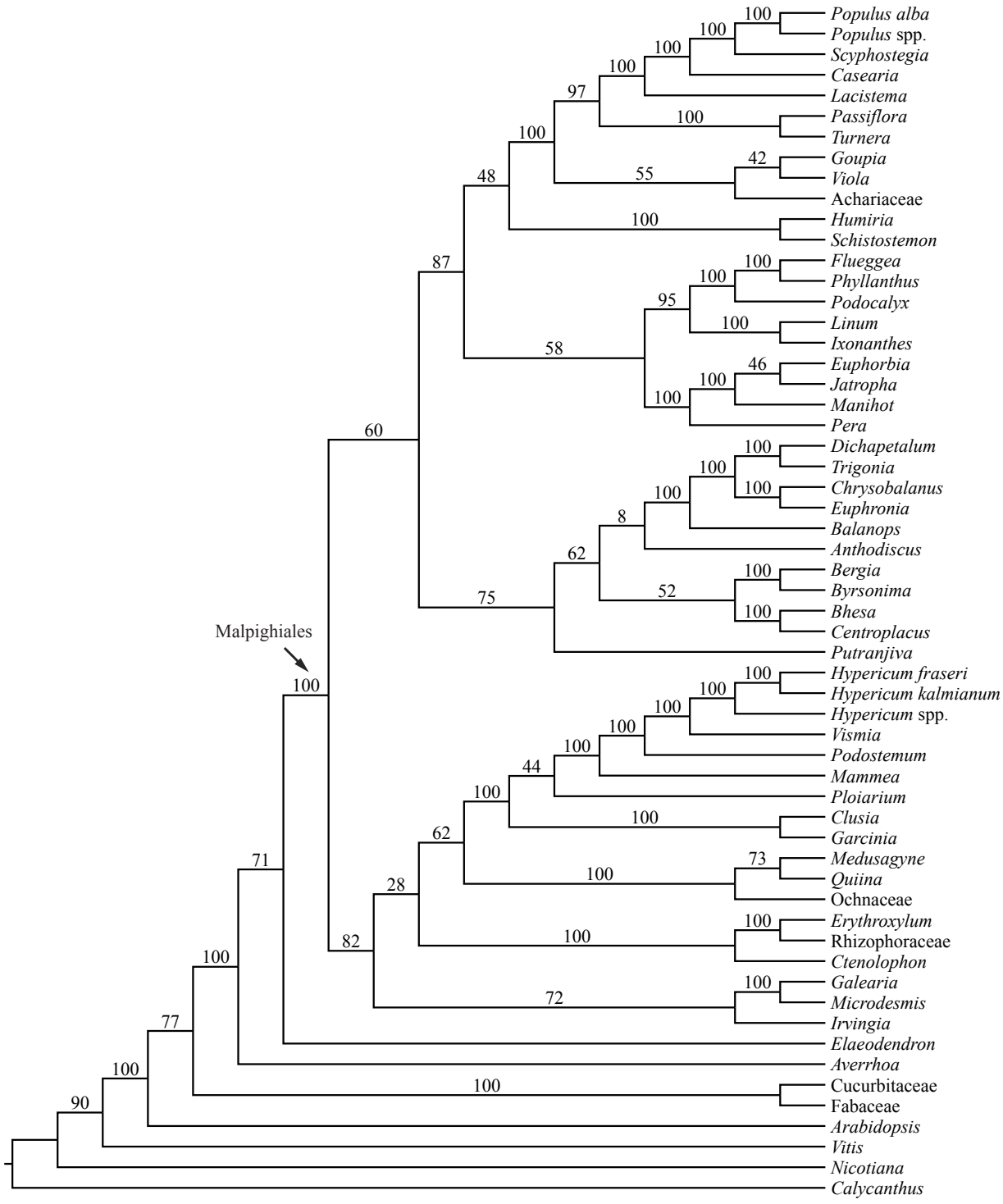

Fig. S9. The maximum likelihood (ML) bipartition tree inferred from the ***combined-complete*** matrix using the MixtPart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

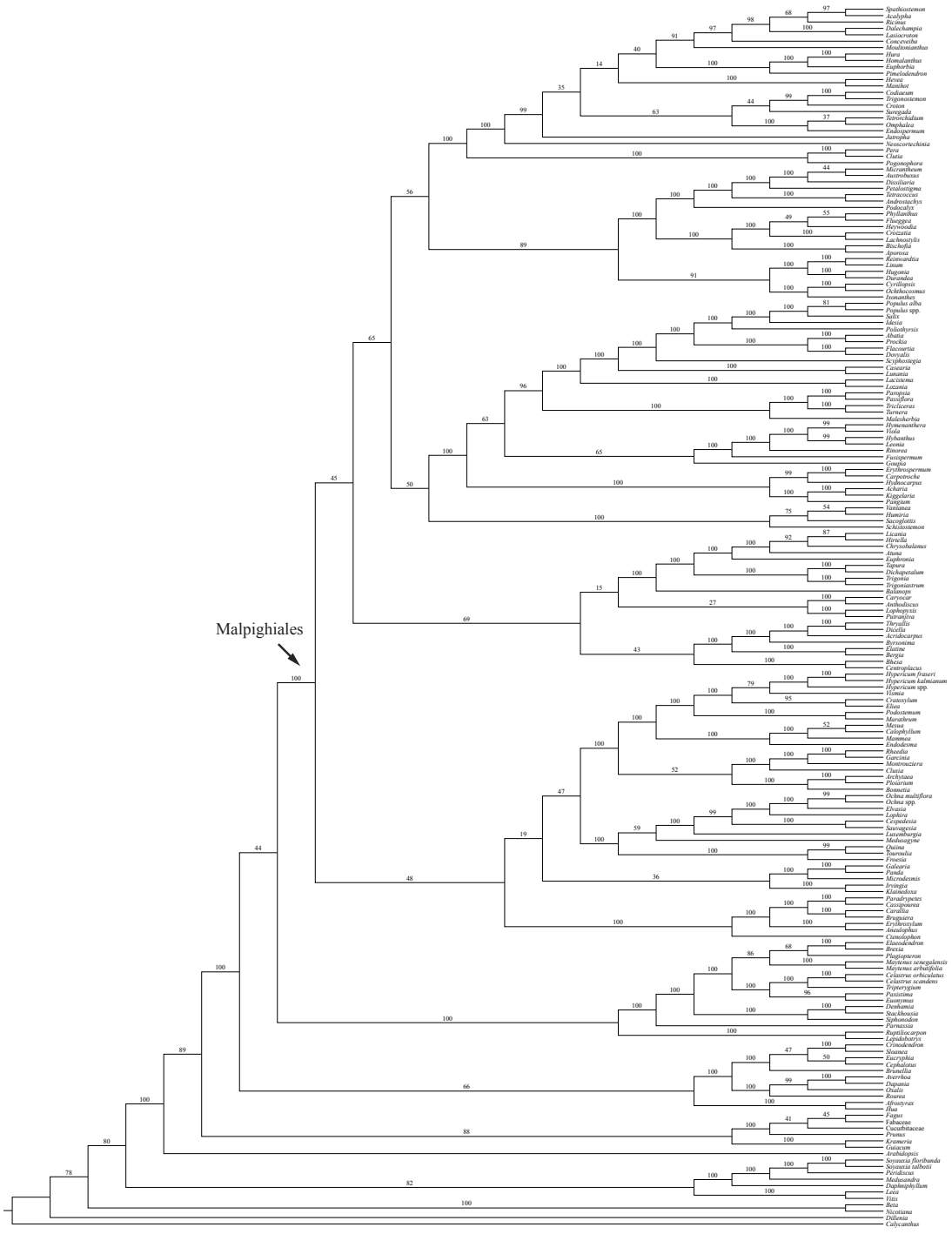

Fig. S10. The maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the OnePart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

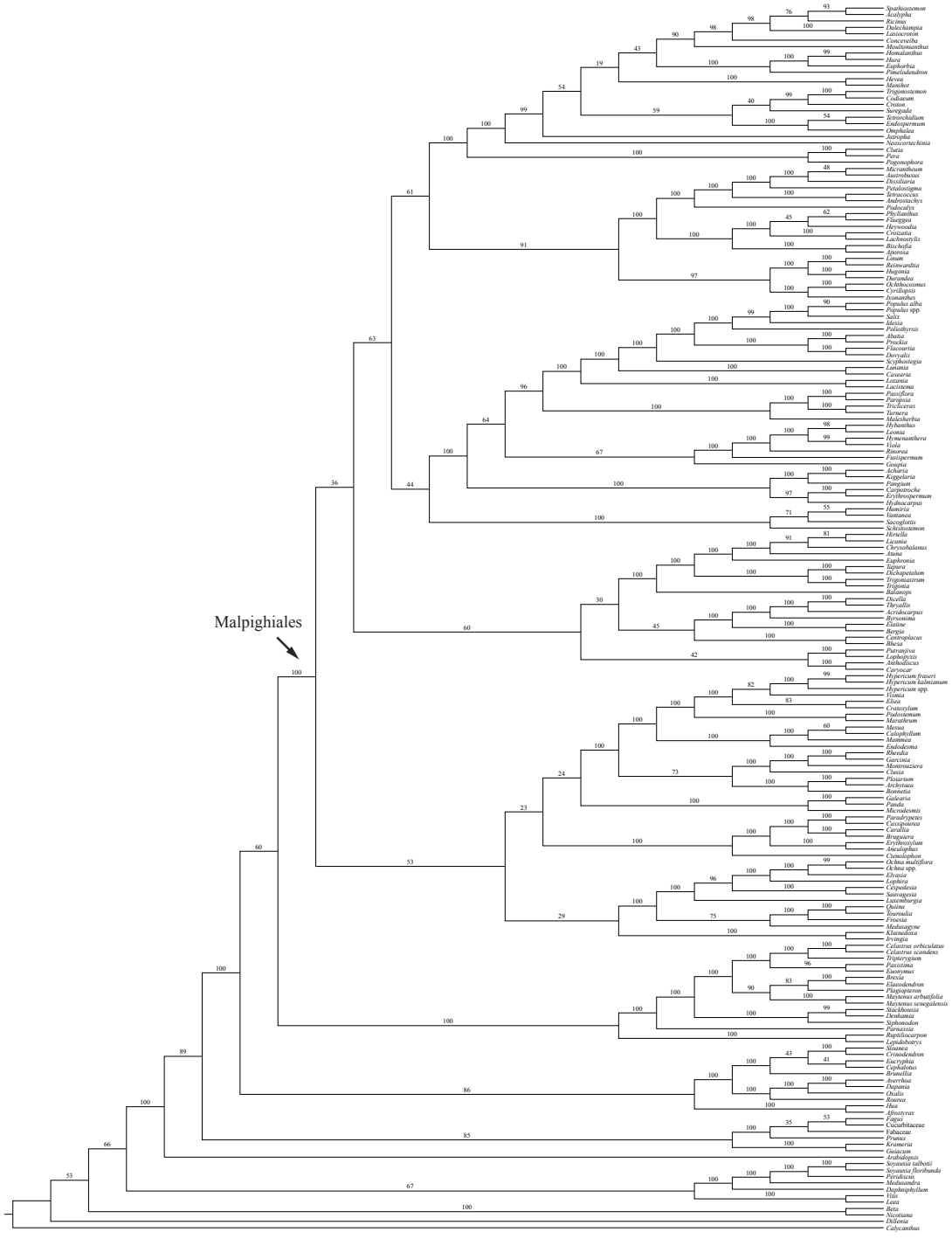

Fig. S11. The maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the GenePart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

Fig. S12. The maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the CodonPart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

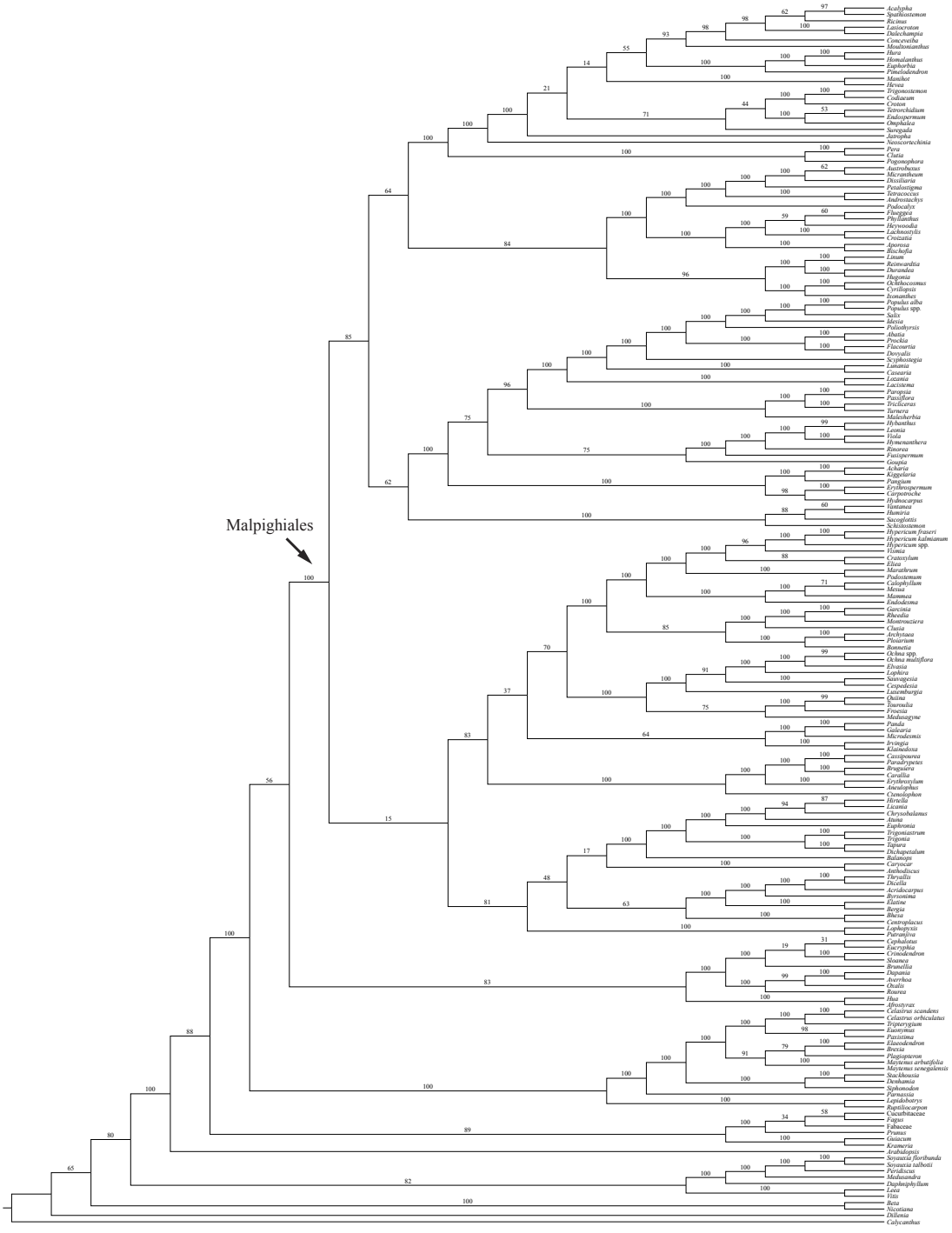

Fig. S13. The maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the MixtPart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

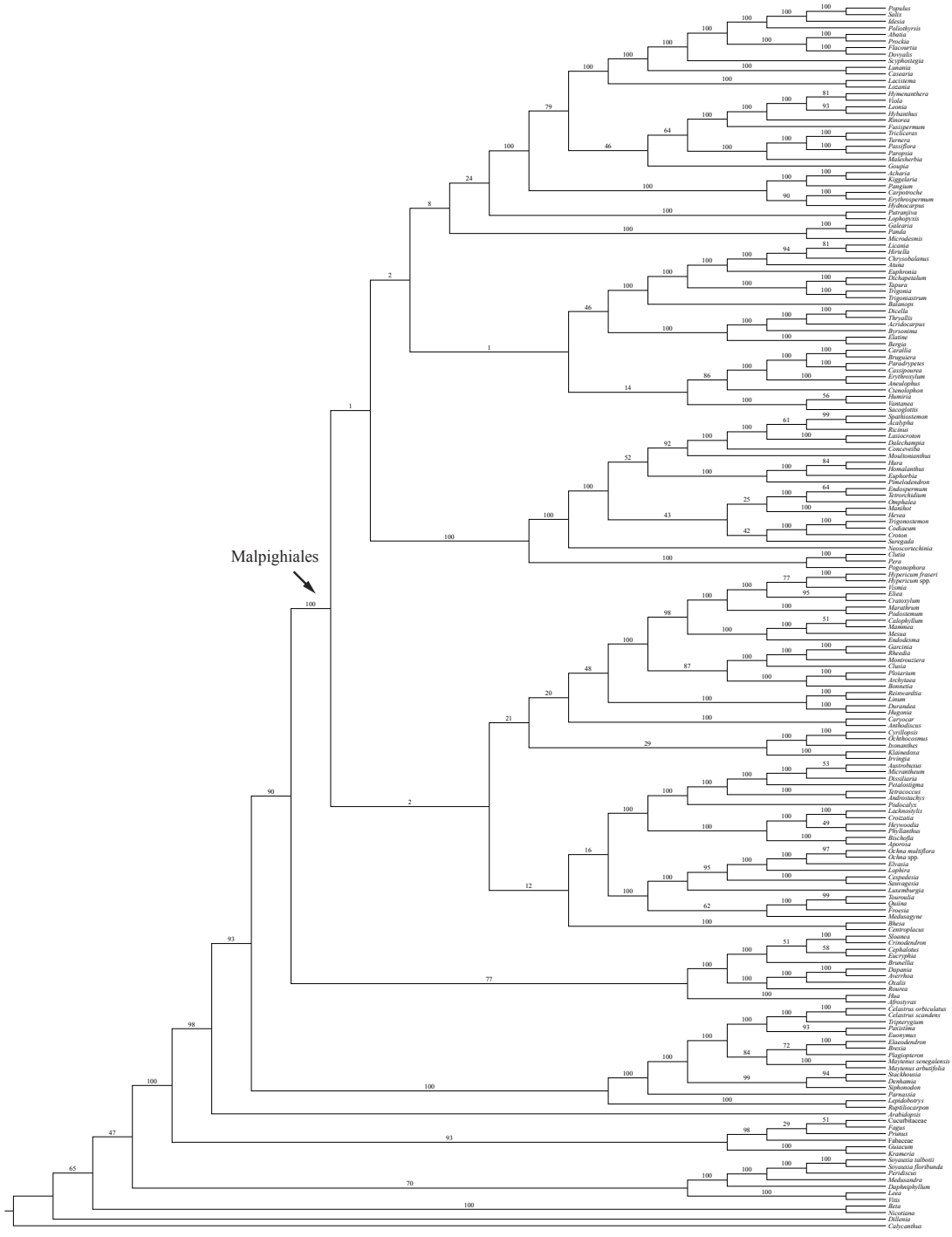

Fig. S14. The maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the OnePart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

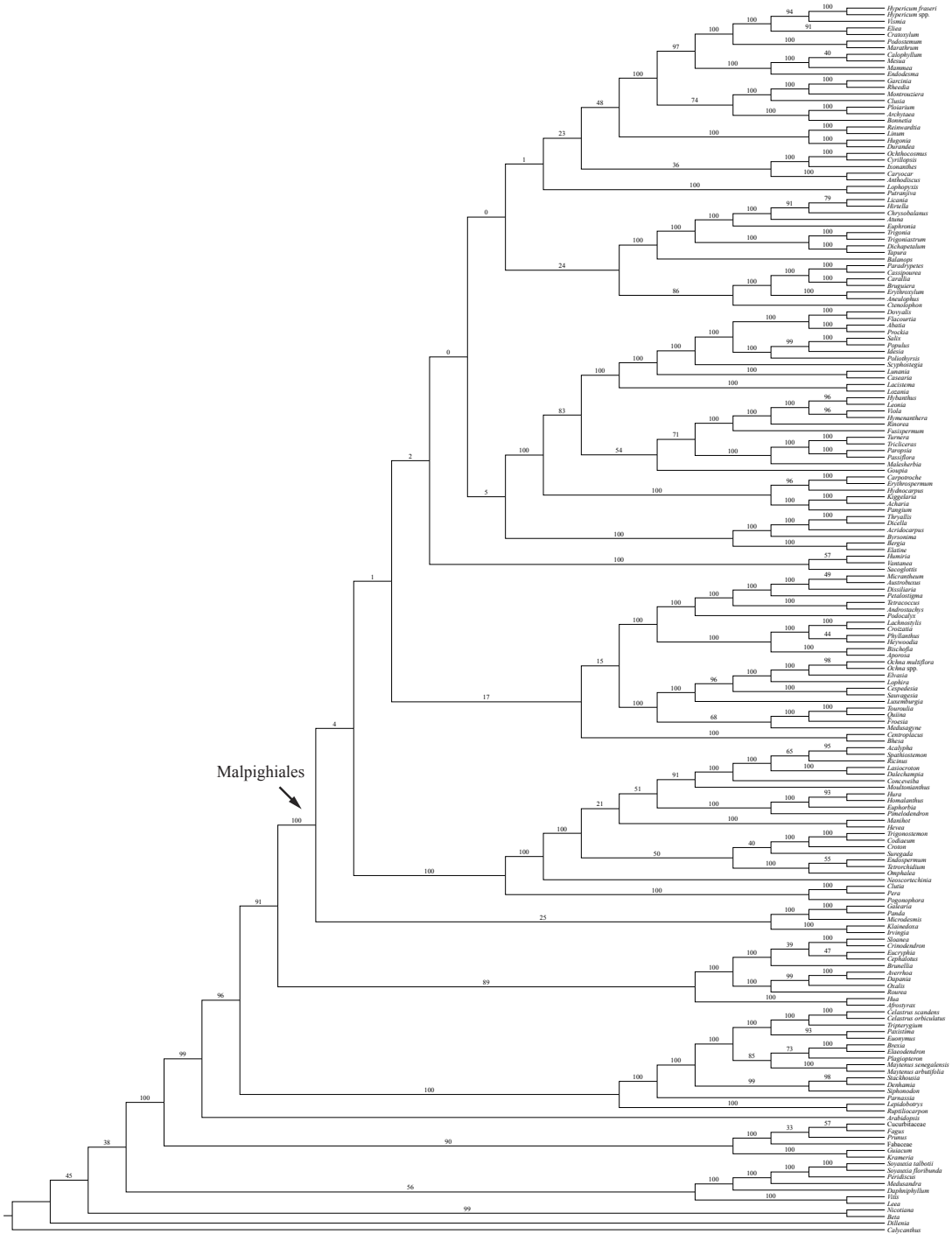

Fig. S15. The maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the GenePart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

Fig. S16. The maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the CodonPart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

Fig. S17. The maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the MixtPart partitioning strategy. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

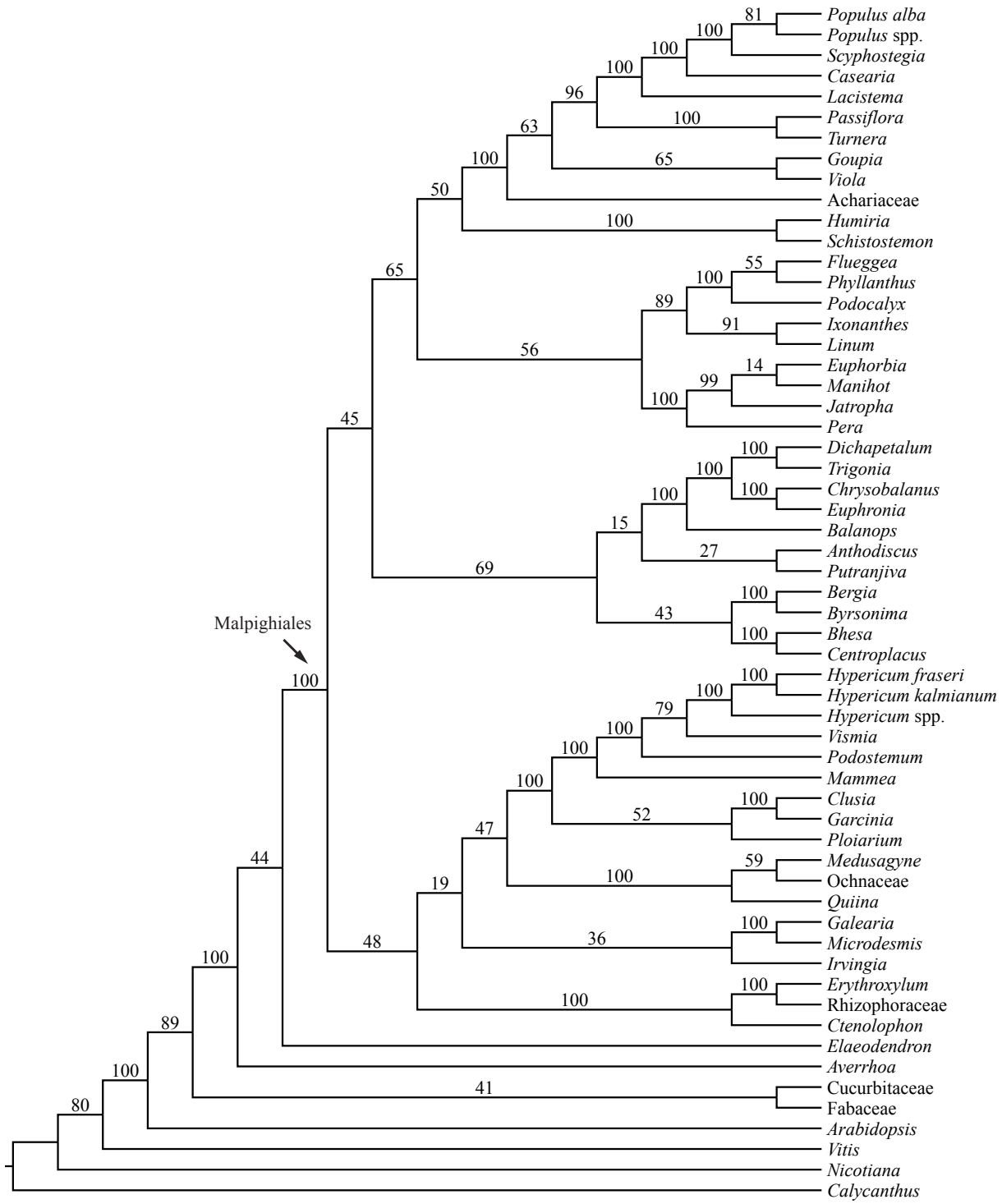

Fig. S18. The pruned maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the OnePart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

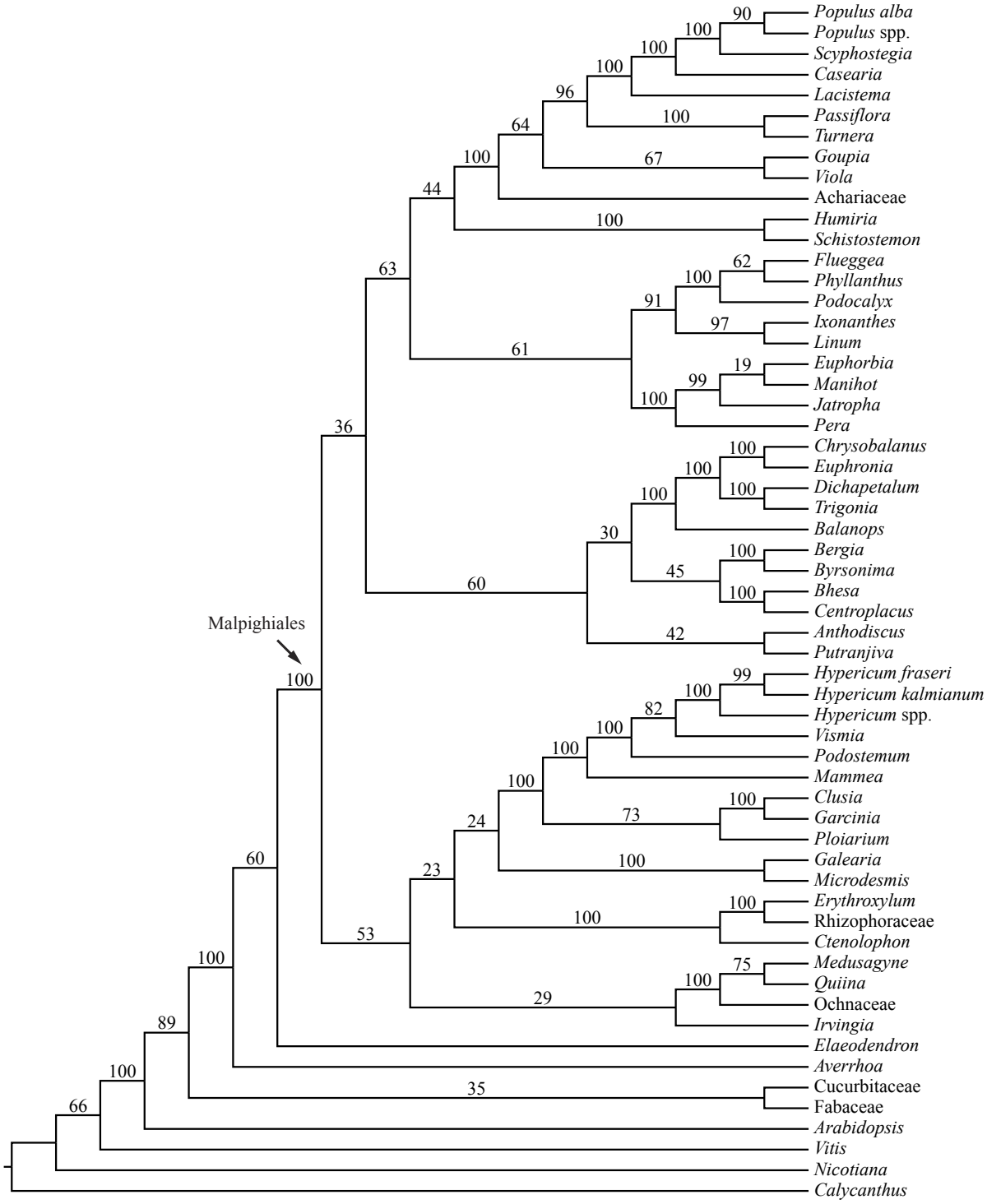

Fig. S19. The pruned maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the GenePart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

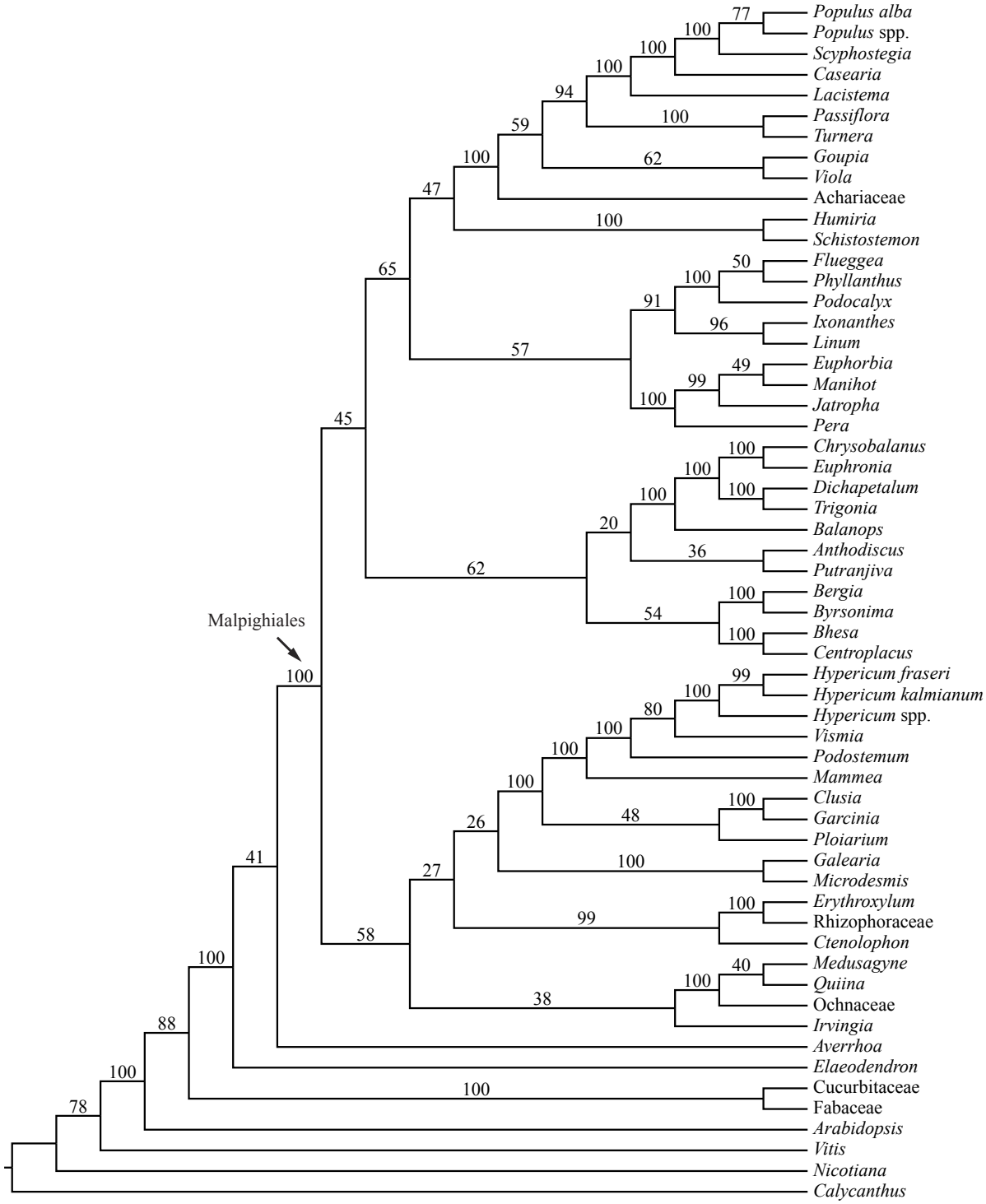

Fig. S20. The pruned maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the CodonPart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

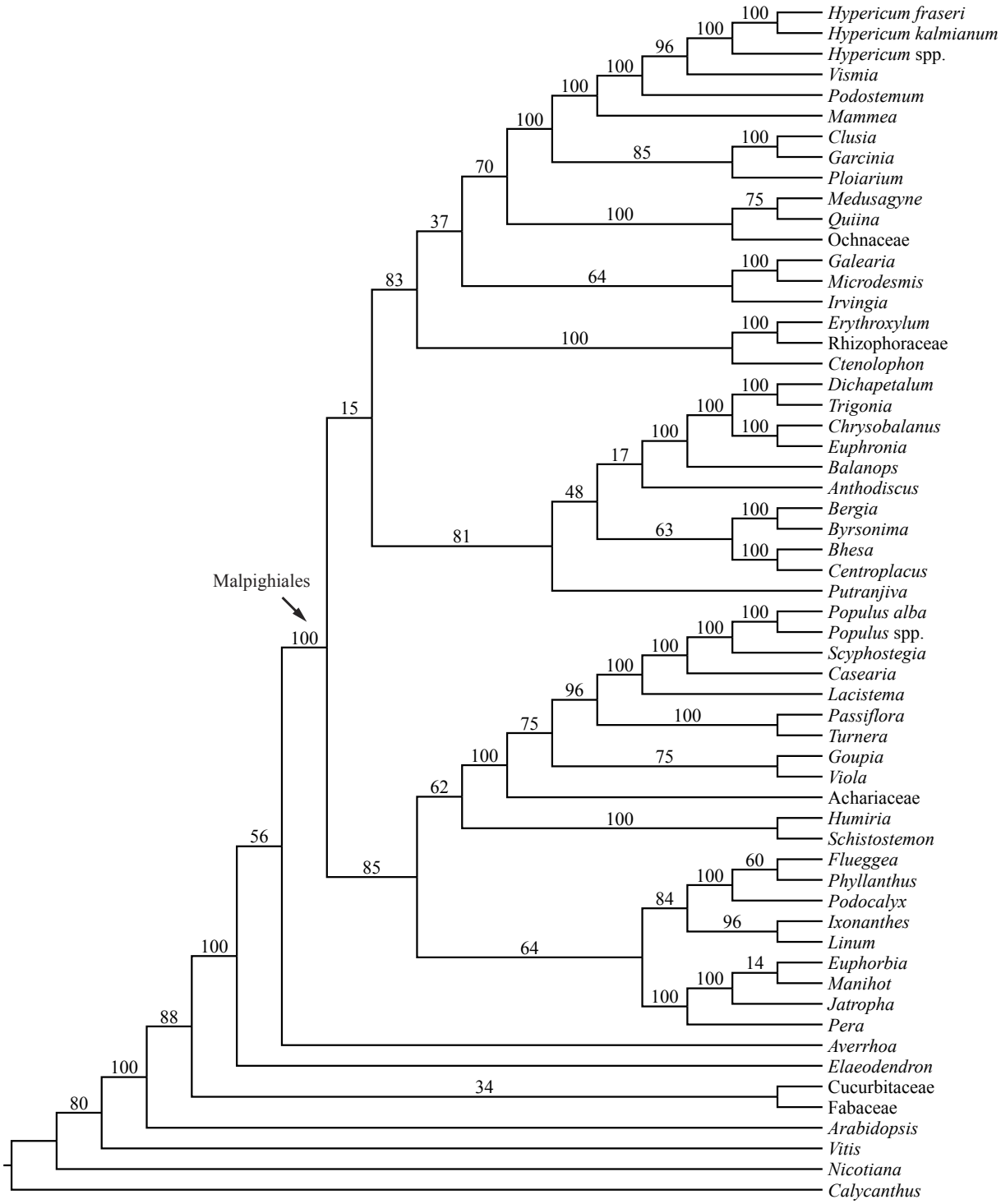

Fig. S21. The pruned maximum likelihood (ML) bipartition tree inferred from the ***combined-incomplete*** matrix using the MixtPart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

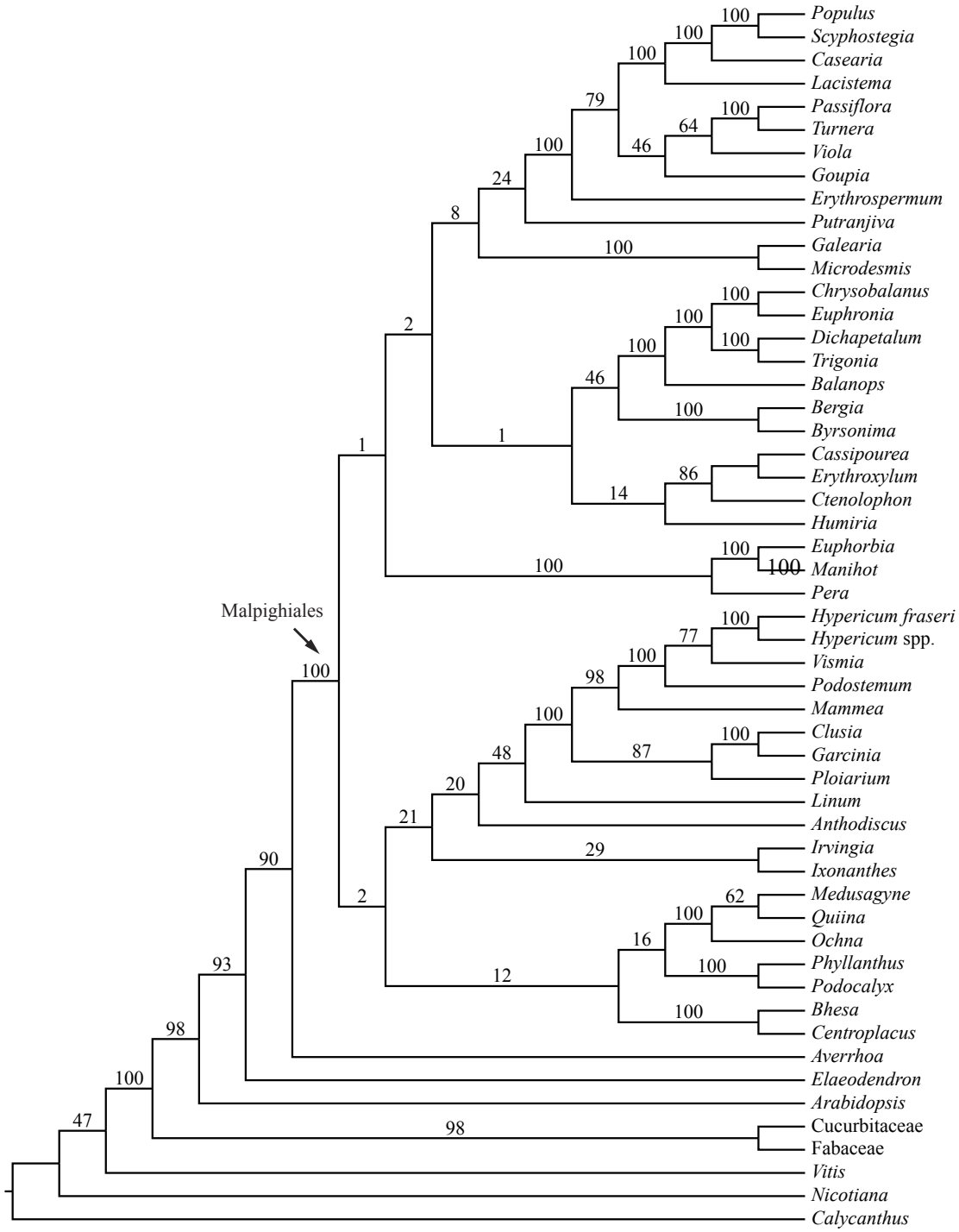

Fig. S22. The pruned maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the OnePart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

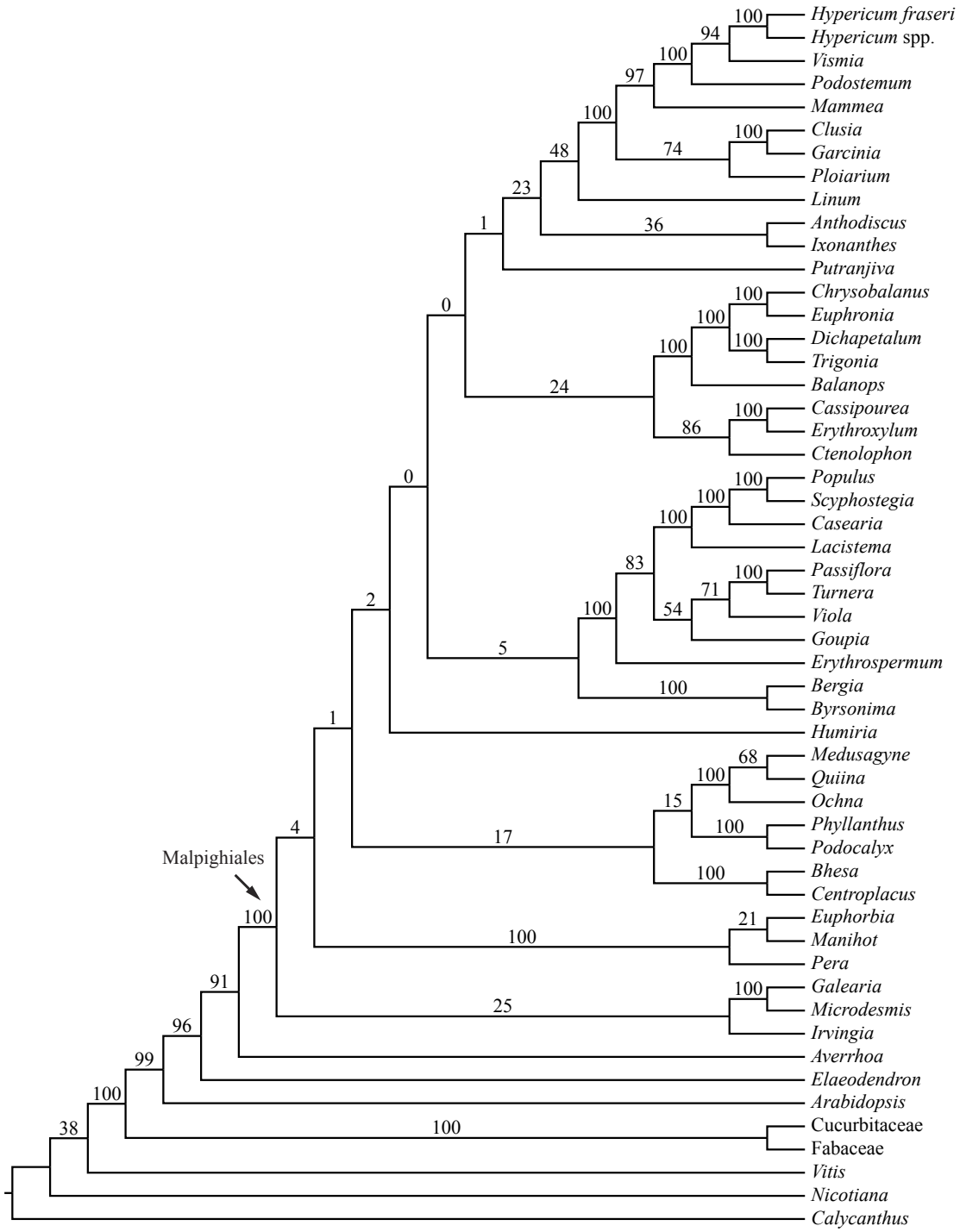

Fig. S23. The pruned maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the GenePart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

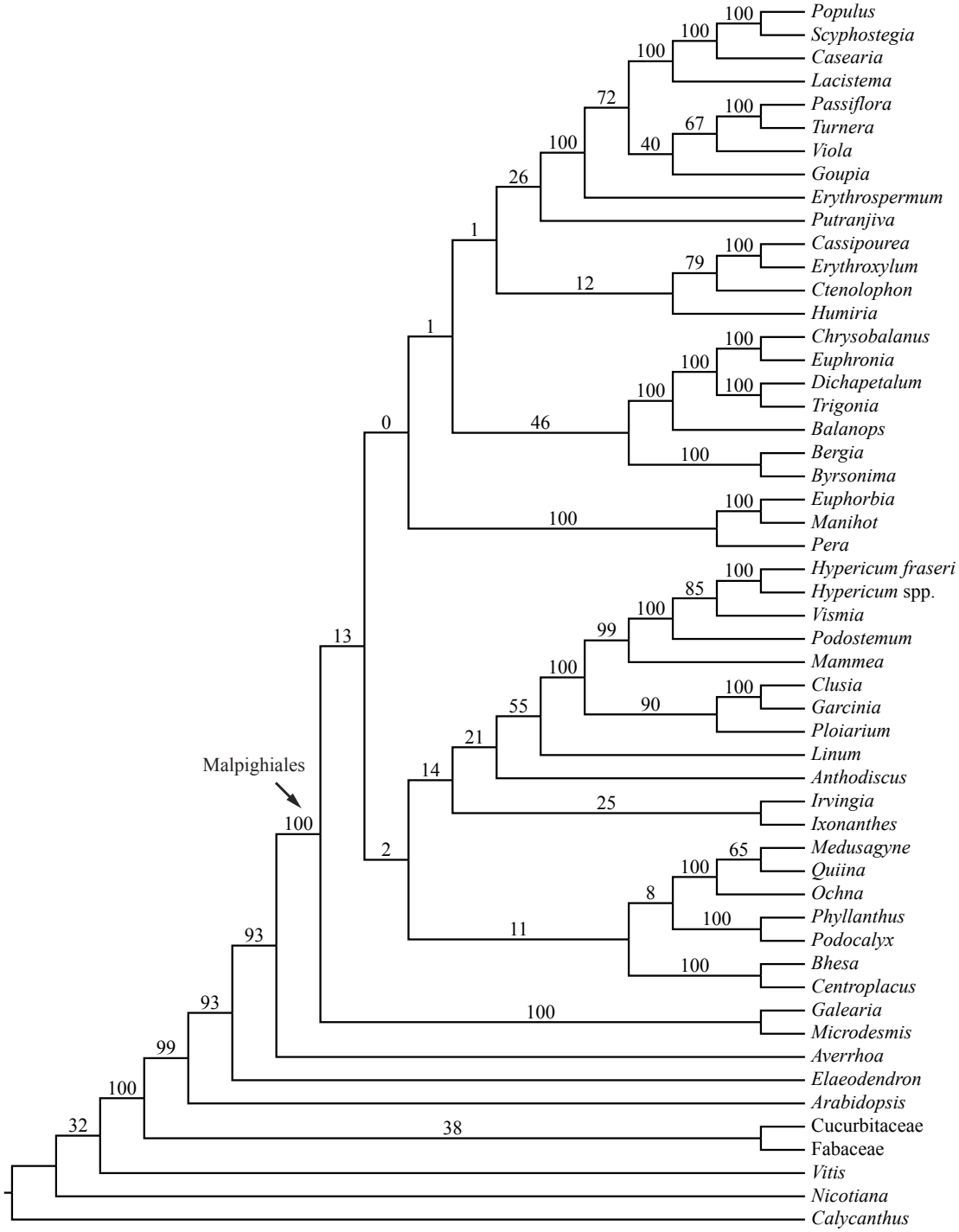

Fig. S24. The pruned maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the CodonPart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

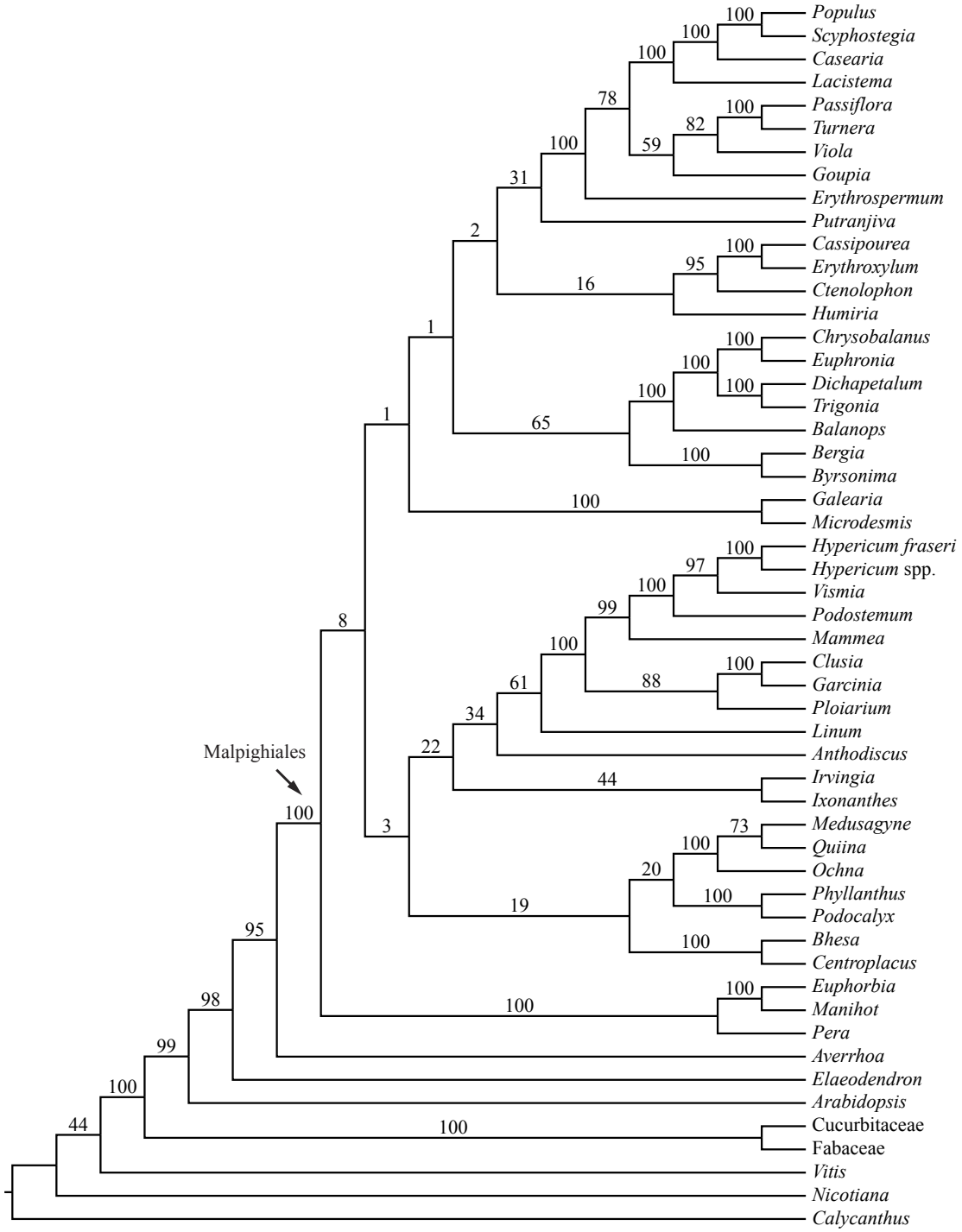

Fig. S25. The pruned maximum likelihood (ML) bipartition tree inferred from the ***13-gene*** matrix using the MixtPart partitioning strategy. Taxa were pruned to match the taxon sampling of the ***82-gene*** and ***combined-complete*** matrices. ML bootstrap percentages are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

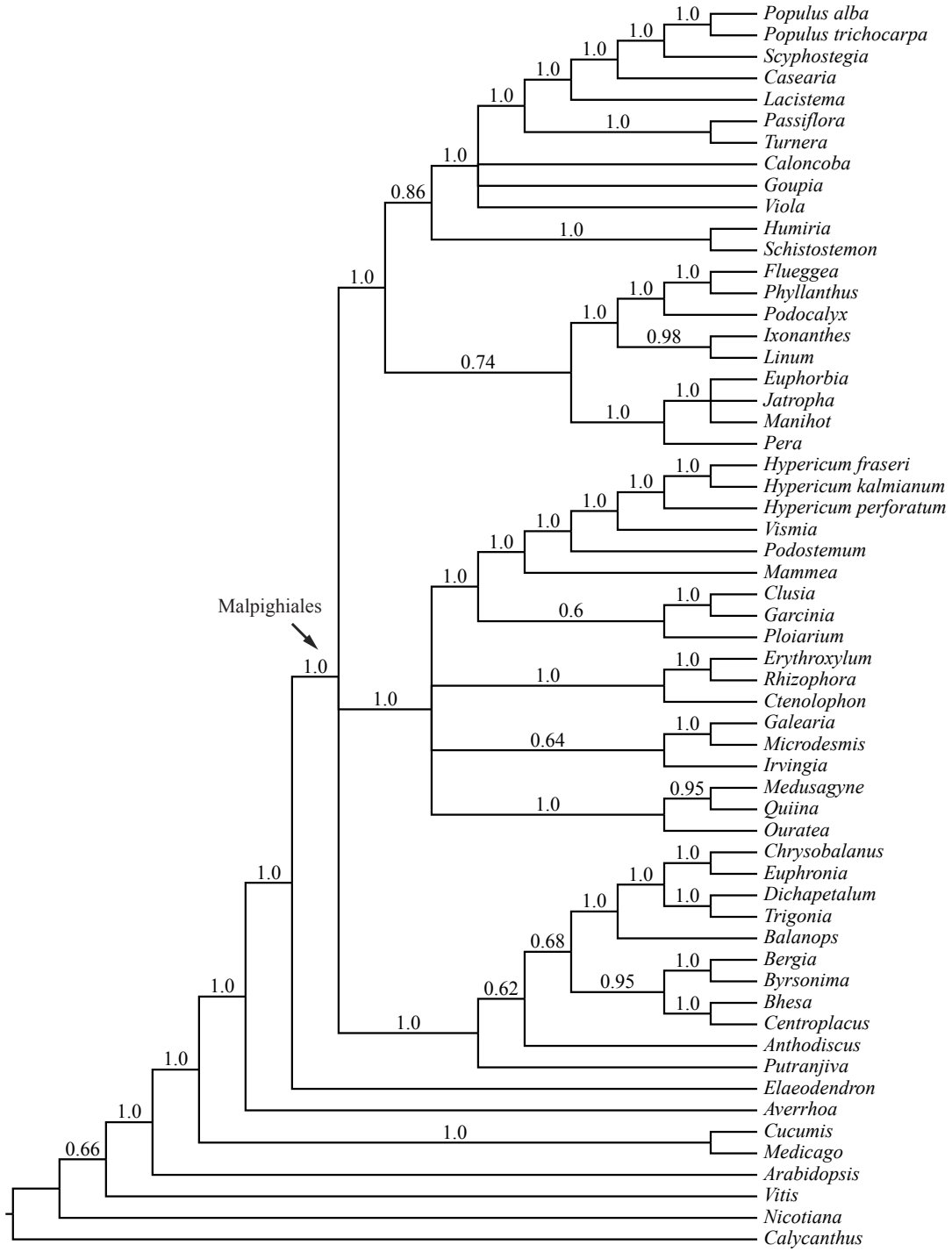

Fig. S26. Bayesian 50% majority-rule consensus tree inferred from the ***82-gene*** matrix. Bayesian posterior probabilities are indicated above each branch.

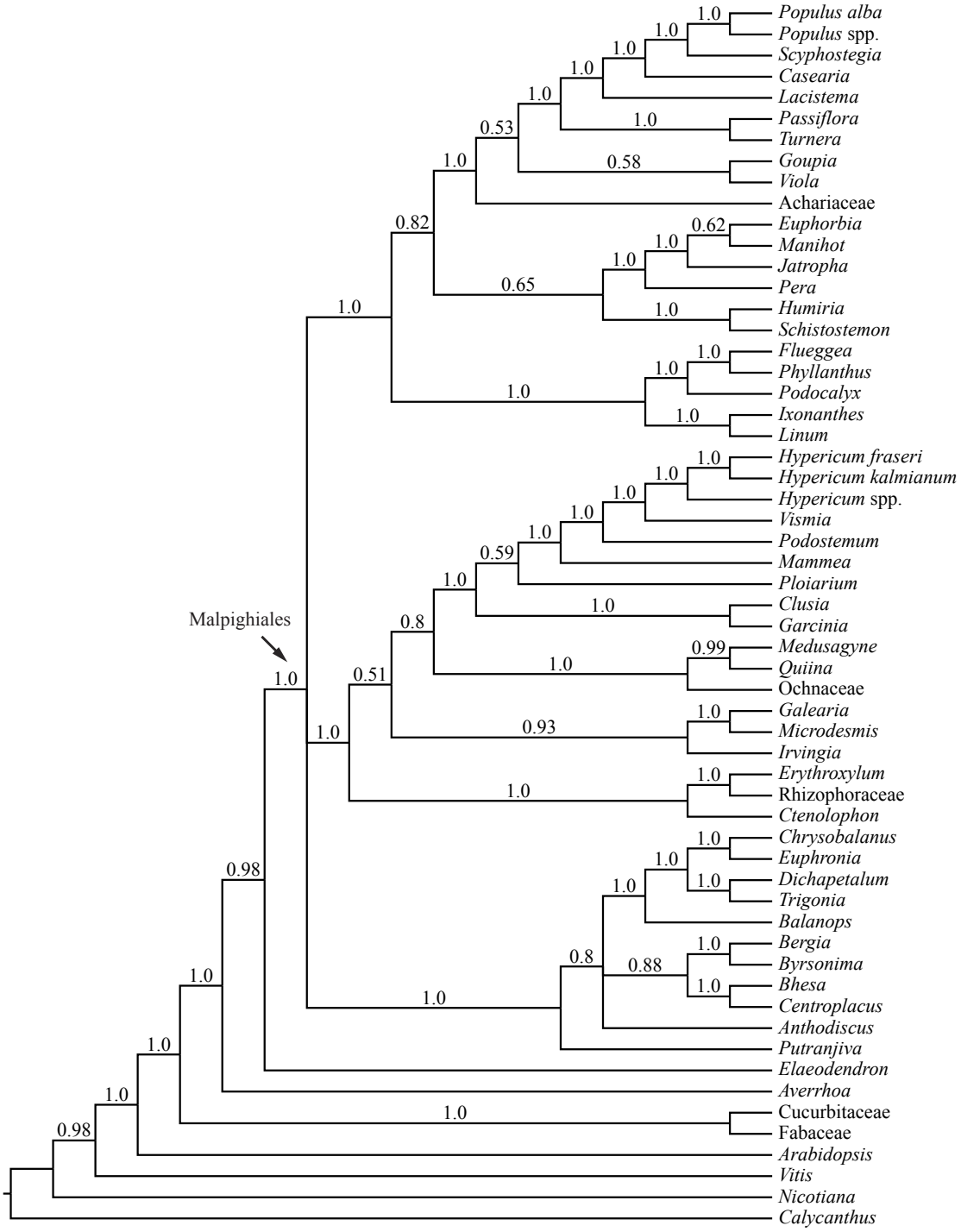

Fig. S27. Bayesian 50% majority-rule consensus tree inferred from the ***combined-complete*** matrix. Bayesian posterior probabilities are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

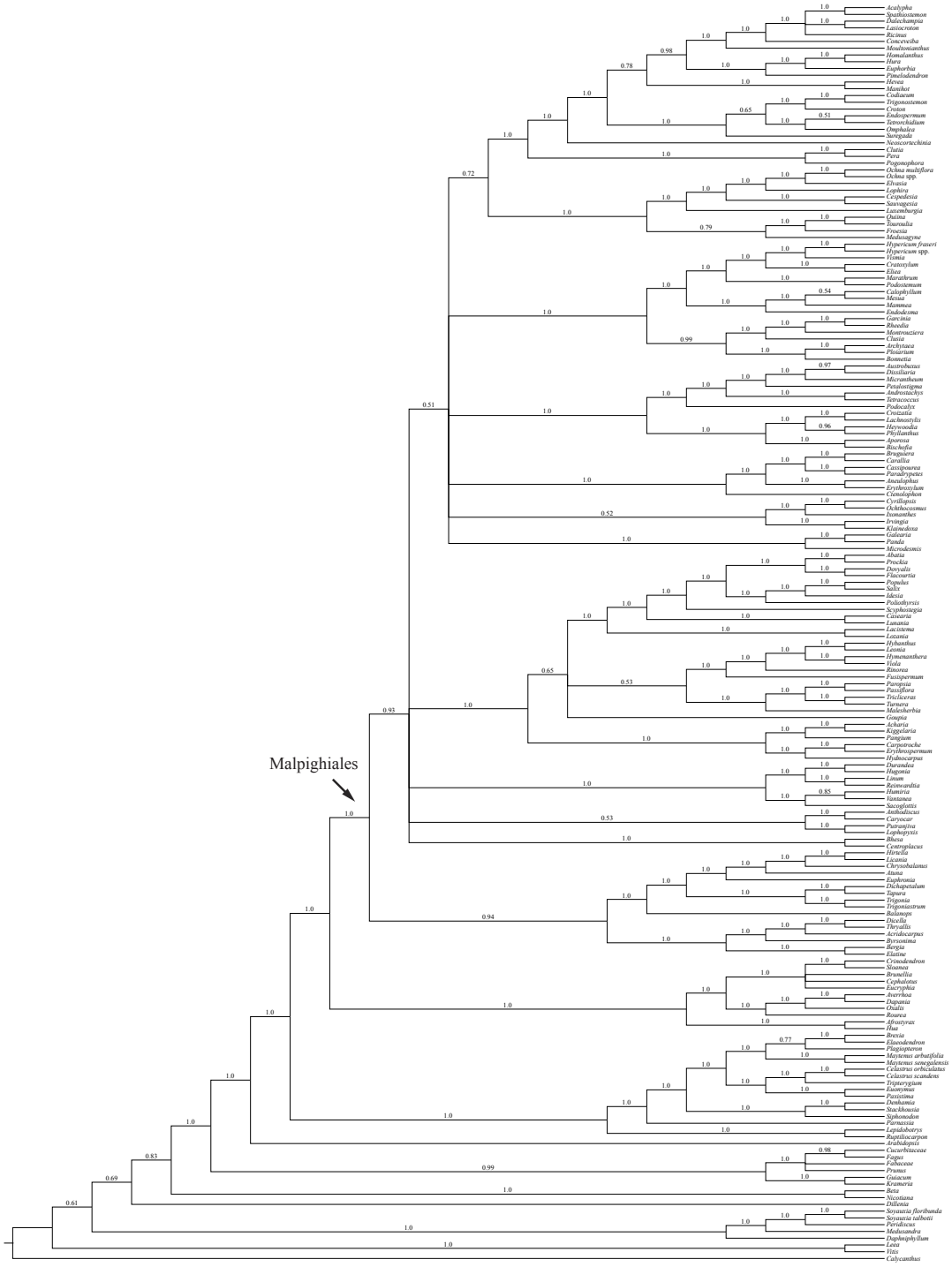

Fig. S28. Bayesian 50% majority-rule consensus tree inferred from the ***13-gene*** matrix. Bayesian posterior probabilities are indicated above each branch; spp. = composite terminals compiled from multiple closely related species.

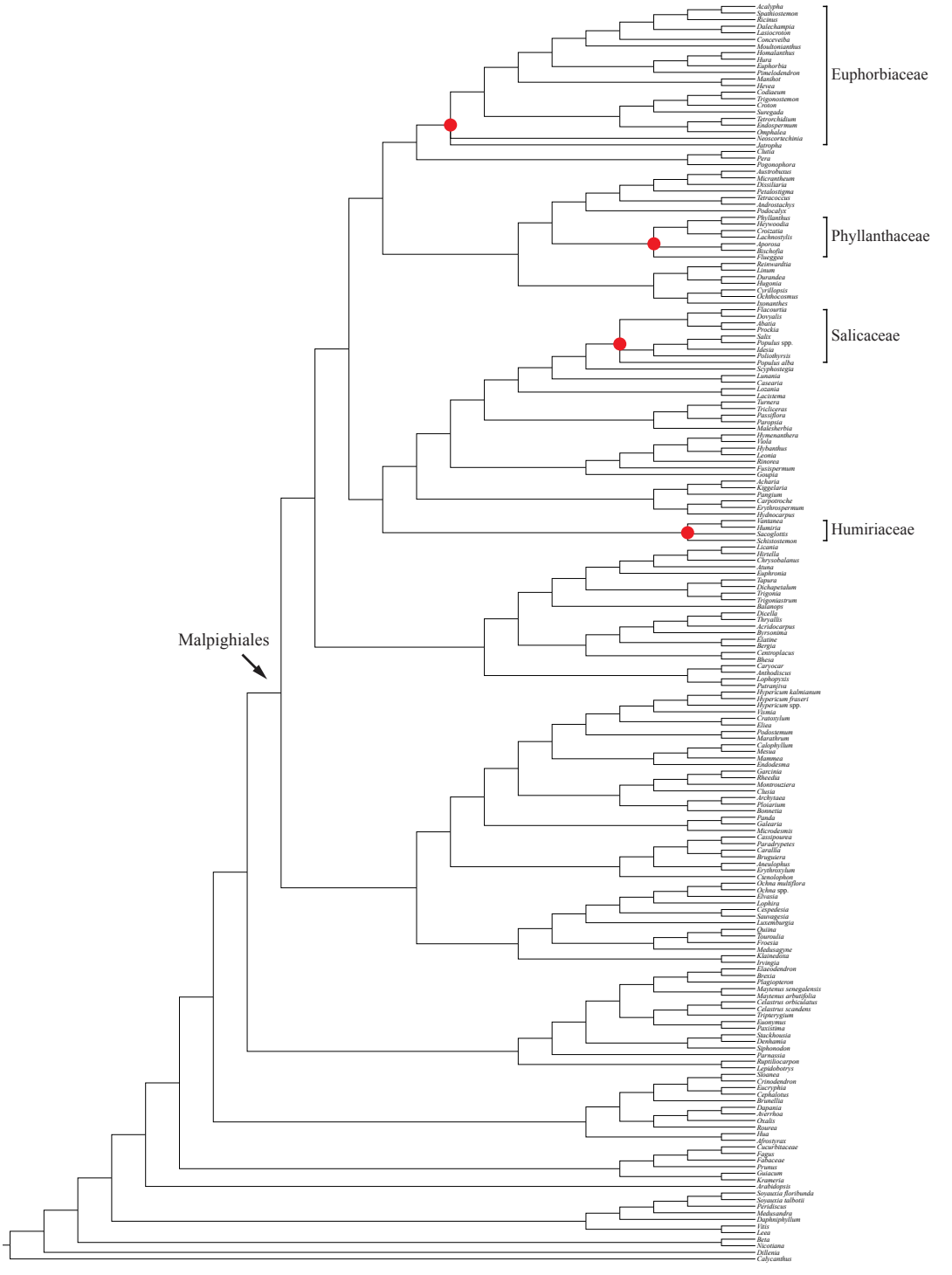

Fig. S29. The BUILD tree (i.e., the Adams consensus tree) of the 14,025 trees from a single terrace (Table 1). This terrace contains the best-scoring maximum likelihood tree inferred from the ***combined-incomplete*** matrix using the GenePart partitioning strategy. Polytomies are highlighted with red dots; spp. = composite terminals compiled from multiple closely related species.

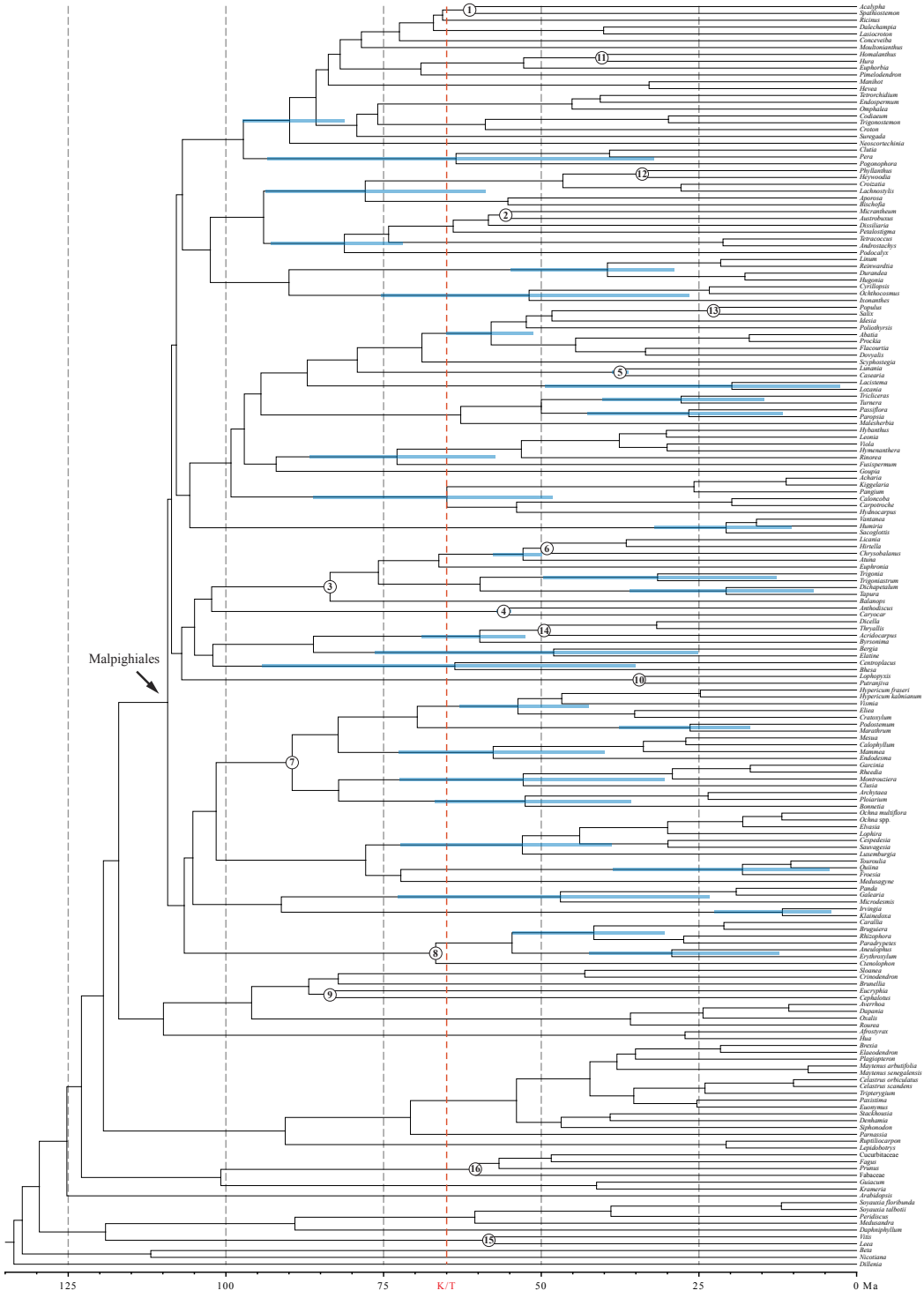

Fig. S30. Chronogram of Malpighiales inferred from BEAST. The node age error bar [95% highest posterior density interval (HPD)] for each Malpighiales family is shown in blue. Minimum age fossil constraints are numbered according to the summary Table S6. The K/T boundary (~65 Ma) is marked with a red line.

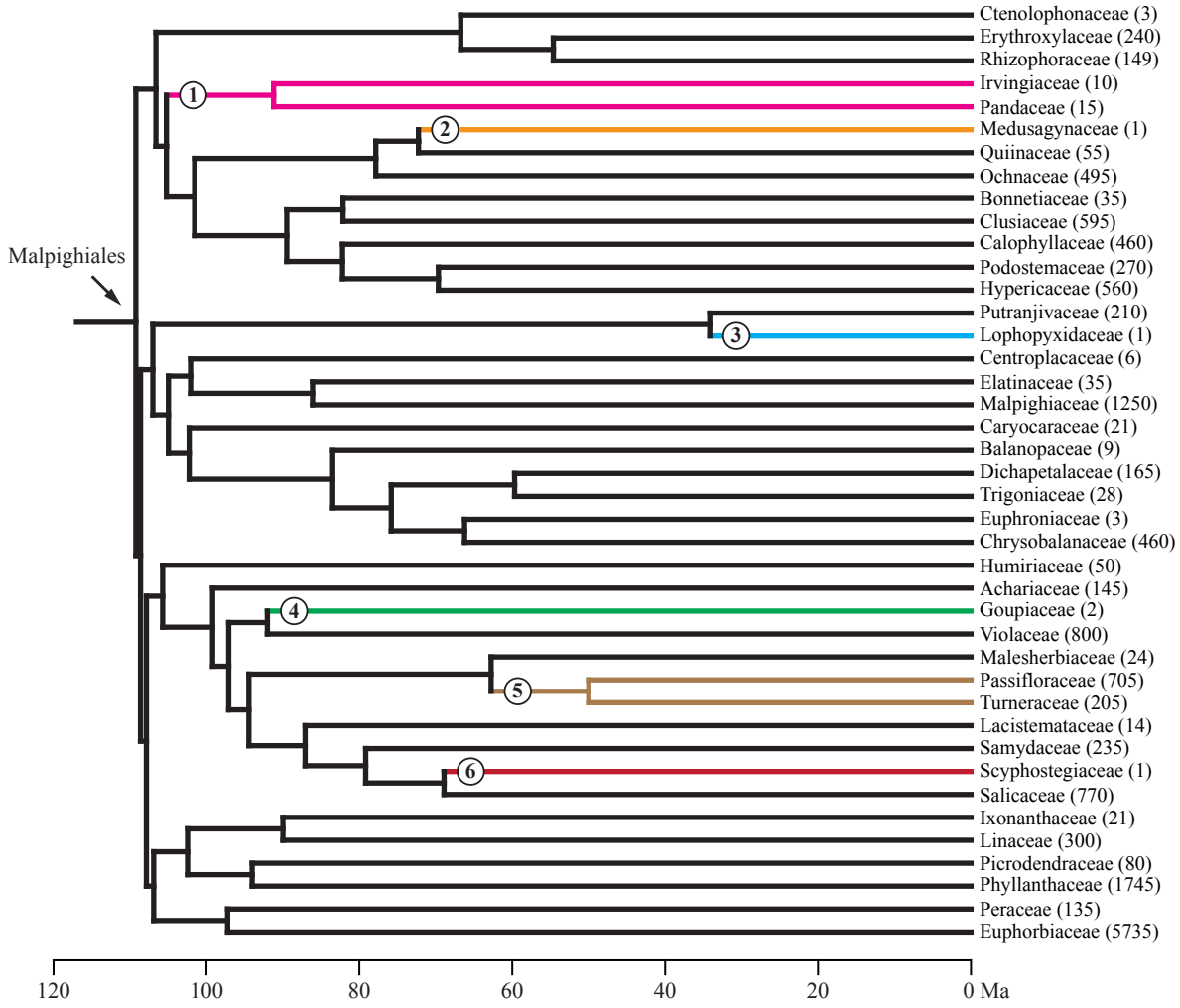

Fig. S31. Malpighiales family level chronogram derived from a pruned tree of the larger 191-taxon analysis (Fig. S29). Taxa were pruned from this dated tree such that each monophyletic family is represented by a single terminal node. The approximate number of accepted species for each family is given in parentheses to the right. Significant shifts in speciation/extinction rates identified by our MEDUSA analysis, and summarized in the main text, are marked by numbered circles.

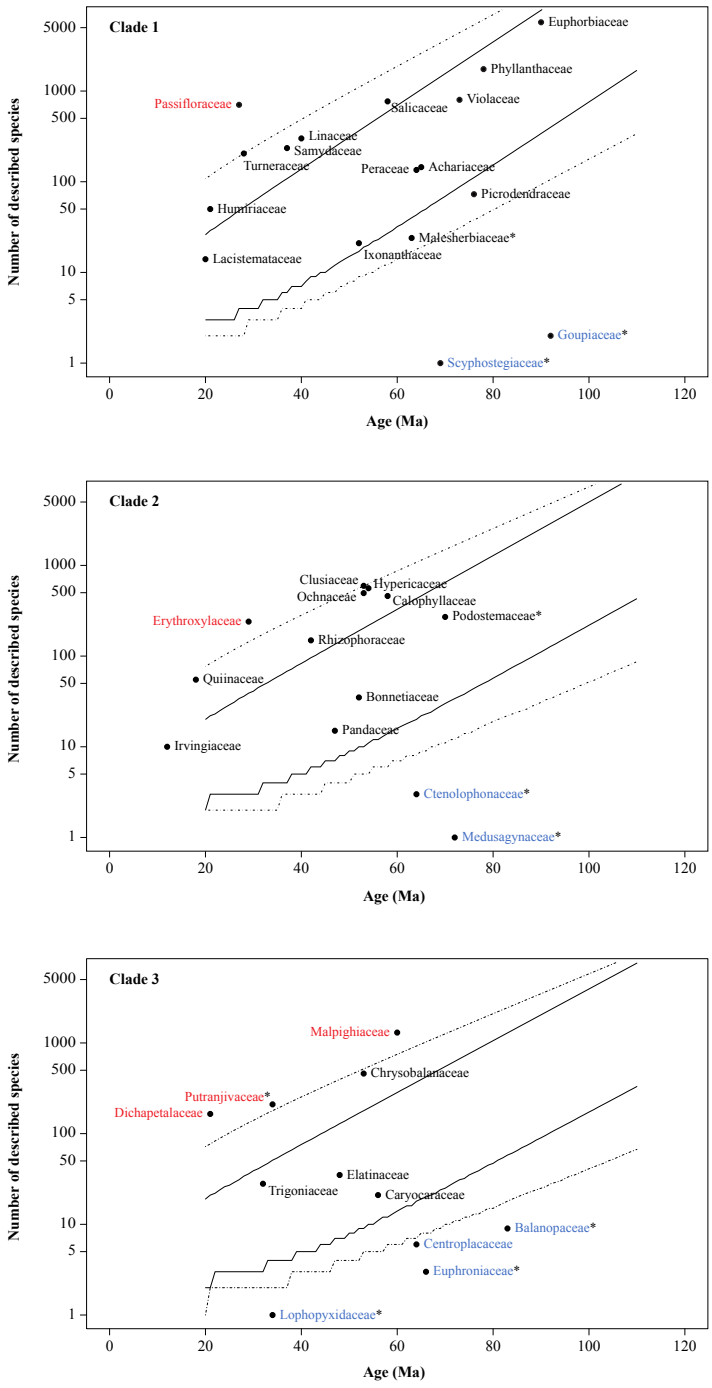

Fig. S32. Plot of species diversity versus age for Malpighiales families. The 95% confidence interval of expected species diversity through time is shown with solid lines in the absence of extinction (ε=0.0) and dashed lines under a high relative extinction rate (ε=0.9). When the crown group age is unavailable, stem group age is used (those families marked with asterisks). Families exhibiting a significant shift in net species diversification are highlighted in red (acceleration) and blue (deceleration).