Complete Generic-Level Phylogenetic Analyses of Palms (Arecaceae) with Comparisons of Supertree and Supermatrix Approaches

WILLIAM J. BAKER[1,*], VINCENT SAVOLAINEN[1,2], CONNY B. ASMUSSEN-LANGE[3], MARK W. CHASE[1],
JOHN DRANSFIELD[1], FÉLIX FOREST[1], MADELINE M. HARLEY[1], NATALIE W. UHL[4],
AND MARK WILKINSON[5]

[1]*Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK;*
[2]*Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK;*
[3]*Department of Ecology, University of Copenhagen, Rolighedsvej 21, DK-1958 Frederiksberg C, Denmark;*
[4]*Department of Plant Biology, Cornell University, 412 Mann Library Building, Ithaca, NY 14853, USA;*
[5]*Department of Zoology, Natural History Museum, Cromwell Road, London SW7 5BD, UK;*
*Correspondence to be sent to: Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK; E-mail: w.baker@kew.org.

Abstract.—Supertree and supermatrix methods have great potential in the quest to build the tree of life and yet they remain controversial, with most workers opting for one approach or the other, but rarely both. Here, we employed both methods to construct phylogenetic trees of all genera of palms (Arecaceae/Palmae), an iconic angiosperm family of great economic importance. We assembled a supermatrix consisting of 16 partitions, comprising DNA sequence data, plastid restriction fragment length polymorphism data, and morphological data for all genera, from which a highly resolved and well-supported phylogenetic tree was built despite abundant missing data. To construct supertrees, we used variants of matrix representation with parsimony (MRP) analysis based on input trees generated directly from subsamples of the supermatrix. All supertrees were highly resolved. Standard MRP with bootstrap-weighted matrix elements performed most effectively in this case, generating trees with the greatest congruence with the supermatrix tree and fewest clades unsupported by any input tree. Nonindependence due to input trees based on combinations of data partitions was an acceptable trade-off for improvements in supertree performance. Irreversible MRP and the use of strictly independent input trees only provided no obvious benefits. Contrary to previous claims, we found that unsupported clades are not infrequent under some MRP implementations, with up to 13% of clades lacking support from any input tree in some irreversible MRP supertrees. To build a formal synthesis, we assessed the cross-corroboration between supermatrix trees and the variant supertrees using semistrict consensus, enumerating shared clades and compatible clades. The semistrict consensus of the supermatrix tree and the most congruent supertree contained 160 clades (of a maximum of 204), 137 of which were present in both trees. The relationships recovered by these trees strongly support the current phylogenetic classification of palms. We evaluate 2 composite supertree support measures (rQS and V) and conclude that it is more informative to report numbers of input trees that support or conflict with a given supertree clade. This study demonstrates that supertree and supermatrix methods can provide effective, explicit, and complimentary mechanisms for synthesizing disjointed phylogenetic evidence while emphasizing the need for further refinement of supertree methods. [Arecaceae; congruence; matrix representation with parsimony (MRP); Palmae; phylogeny; supermatrix; supertree.]

The palms (Arecaceae or Palmae) are an iconic family of flowering plants comprising around 2400 species distributed throughout the tropical and subtropical regions of the world (Govaerts and Dransfield 2005). The family has a rich fossil record dating back to the Turonian (89–93.5 Ma; Harley 2006), although molecular dating analyses suggest that the lineage is substantially older (Bremer 2000; Bremer et al. 2004). Palms are important components of many habitats and can have major ecological impact (e.g., Peters et al. 2004). They are also of immense economic significance both in international trade (e.g., oil palm, date palm, coconut, rattan) and at the subsistence level in some of the poorest communities in the world.

Palms fall within the commelinid clade of the monocotyledons, and their monophyly is well established (Asmussen et al. 2006; Chase et al. 2006). The relationships within the family, however, are far from unambiguously resolved. More than 40 phylogenetic studies of the palm family have been published since the early 1990s (Dransfield, Uhl, et al. 2008). Although this impressive body of literature has facilitated a systematic reappraisal of the entire family (Dransfield et al. 2005; Dransfield, Uhl, et al. 2008), arbitrating among these works, which vary widely in their taxonomic inclusiveness and information content, has been problematic. A formal synthesis of this heterogeneous body of evidence is highly desirable for systematic reasons and to provide a comprehensive framework for comparative research. The construction of large, synthetic phylogenetic trees is the subject of a polarized methodological debate. Advocates of supertree methods have promoted the expediency of their methods for building very large trees that will minimize the detrimental impact of missing data on phylogenetic accuracy (Sanderson et al. 1998; Bininda-Emonds et al. 2002; Bininda-Emonds 2004a). In contrast, opponents have argued that supertrees are dislocated from real data and thus represent less valid phylogenetic hypotheses than trees inferred using conventional methods (Rodrigo 1993; Gatesy et al. 2002, 2004; Gatesy and Springer 2004). They highlight a range of negative features of supertree studies, such as nonindependence among input trees, the use of poor-quality input trees, and a lack of signal enhancement, while maintaining that missing data in supermatrices may not necessarily jeopardize phylogenetic accuracy after all (Wiens 1998, 2003, 2006; de Queiroz and Gatesy 2007). There have, however, been relatively few empirical comparisons of supermatrix and supertree approaches, and few practitioners have explored the potential merits

of using both approaches (Kennedy and Page 2002; Salamin et al. 2002; Fritsch et al. 2006; Fulton and Strobeck 2006; Hughes et al. 2006).

Here, we aim to build a synthesis of palm phylogenetic evidence by employing supermatrix and supertree methods to produce comprehensive phylogenetic trees of all palm genera (Fig. 1). From a supermatrix encompassing all relevant published data, as well as some new data, notably a morphological data set covering

FIGURE 1. A schematic representation of the use of supermatrix and supertree methods in this study. See text for full details.

all genera published here for the first time, we use maximum parsimony analyses to generate a "total evidence" (Kluge 1989) supermatrix tree. We also infer maximum parsimony trees from subsamples of the supermatrix and use these as input trees for supertree construction employing variants of matrix representation with parsimony (MRP). The use of different approaches permits cross-corroboration (Levasseur and Lapointe 2001, 2003). Thus, we compare palm phylogenies produced using supermatrix and supertree approaches to identify clades that are found in both, clades that occur in one and are compatible with the other, and clades that are incompatible. This provides a ranking of confidence in clades additional to that provided by support measures (e.g., bootstrap proportion). Although many supertree methods have been and continue to be proposed, MRP (Baum 1992; Ragan 1992a, 1992b; Baum and Ragan 1993, 2004) is the most widely used, but is controversial (e.g., Bryant 2004; Gatesy and Springer 2004; Wilkinson et al. 2004). MRP (and some other) methods may be biased by the size (Purvis 1995) and shape (Wilkinson, Cotton, et al. 2005) of input trees. MRP supertrees can include unsupported groups, that is, relationships that are not supported by any of the input trees (Bininda-Emonds and Bryant 1998; Wilkinson et al. 2004; Wilkinson, Pisani, et al. 2005), or groups that are contradicted by every input tree (Wilkinson et al. 2007). Nevertheless, empirical studies (e.g., Davies et al. 2004; Price et al. 2005; Beck et al. 2006) and limited simulations (Bininda-Emonds and Sanderson 2001) suggest that MRP can produce valuable syntheses of phylogenetic information and merits further study.

A range of modifications of standard MRP have been proposed to reduce biases and/or improve accuracy, including irreversible MRP, in which reversals are prohibited (Bininda-Emonds and Bryant 1998) and weighting of MRP matrix elements by support measures (Ronquist 1996; Sanderson et al. 1998; Bininda-Emonds and Sanderson 2001; Bryant 2004), but there has been little comparison of these approaches. Similarly, there has been much concern but little investigation of the effect of nonindependence in input trees (Gatesy et al. 2002, 2004; Gatesy and Springer 2004). We use our example to investigate the effects of various MRP implementations and nonindependence among input trees on supertree performance. We evaluate supertree performance in terms of congruence with the supermatrix tree and numbers of unsupported groups, and we use congruence with the supermatrix tree to identify a "most congruent supertree" (Fig. 1). We also provide the first empirical comparison of recently proposed measures of support for supertree clades (Price et al. 2005; Wilkinson, Pisani, et al. 2005; Beck et al. 2006).
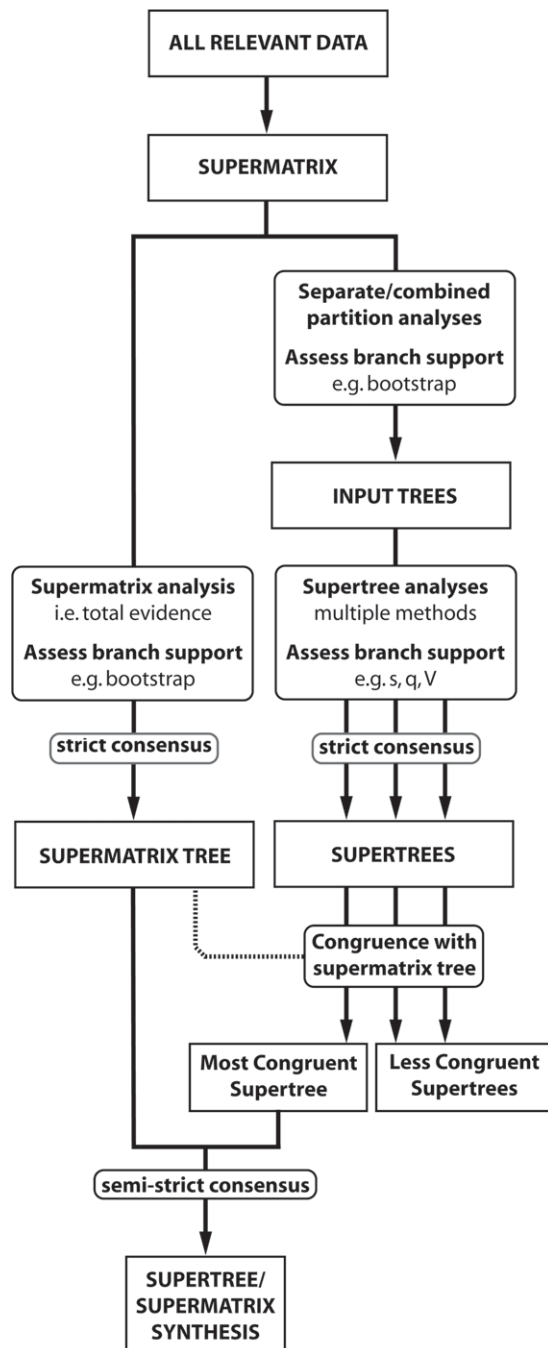
MATERIALS AND METHODS

Supermatrix Assembly

We built a supermatrix including all genera accepted in the *World Checklist of Palms* (Govaerts and Dransfield

2005) plus those subsequently recognized by Baker et al. (2006) and Lewis and Zona (2008). In total, 192 genera were included. Data from the recently discovered genus *Tahina* (Dransfield, Rakotoarinivo, et al. 2008) were not available in time for inclusion in this study. DNA sequence data sets were assembled from all relevant studies up to and including 2006 (Baker et al. 1999; Asmussen et al. 2000, 2006; Baker, Hedderson, and Dransfield 2000; Asmussen and Chase 2001; Lewis and Doyle 2001, 2002; Hahn 2002a, 2002b; Gunn 2004; Roncal et al. 2005; Loo et al. 2006; Norup et al. 2006; Savolainen et al. 2006; Thomas et al. 2006; online Supplementary Appendix 1, http://sysbio.oxfordjournals.org/). Plastid, nuclear ribosomal, and low-copy nuclear DNA data but not mitochondrial data were available. Terminal taxa were treated at the generic level, resulting in 42 palm (22%) genera being represented by sequences from more than one species, equivalent to assumptions of generic monophyly. Robust evidence for nonmonophyly is available for very few of these 42 genera (Dransfield, Uhl, et al. 2008), and where such data exist, we have endeavored to sample closely related species.

Where possible, original DNA sequence alignments were obtained into which further data downloaded from GenBank were integrated. Additional *rbc*L sequence data were generated following the protocol of Asmussen et al. (2006) to fill gaps in the existing data set, thereby ensuring that at least one molecular data set was complete for all palm genera and out-groups (Table 1).

In addition to the DNA sequence data, a plastid restriction fragment length polymorphism (RFLP) data set was included (Uhl et al. 1995) along with a previously unpublished morphological data set compiled by J.D., N.W.U., and M.M.H. covering all palm genera (online Supplementary Appendix 2). Morphological data were gathered from herbarium and living palm specimens at the Royal Botanic Gardens, Kew, the L.H. Bailey Hortorium (Cornell University), the Montgomery Botanical Centre, and Fairchild Tropical Botanic Garden.

Thirteen commelinid monocot out-groups that had previously been used in family-wide studies of palms (Lewis and Doyle 2001; Hahn 2002a; Asmussen et al. 2006) were included in the supermatrix. Additional sequences for these genera were incorporated in the supermatrix when available from GenBank/EBI.

Supermatrix Analysis

The supermatrix was subjected to a 2-stage maximum parsimony analysis to promote thoroughness and efficiency in searching tree space using PAUP* version 4.0b10 (Swofford 2002). First, a 5000-replicate heuristic search was conducted in which a maximum of 5 trees were saved per replicate. Starting trees were obtained using stepwise addition with random taxon addition sequences. Characters were unordered and equally weighted with gaps treated as missing data. Branches were collapsed if their maximum length equaled zero. Tree bisection–reconnection (TBR) branch swapping was used. The trees saved in the first stage were then used as starting trees for a second stage in which similar settings were used, but in this case the trees were swapped to completion with no limit placed on the number of trees saved. Branch support was assessed using PAUP* by bootstrapping the supermatrix 1000 times and conducting a single search for each bootstrap replicate using simple taxon addition, subtree pruning–regrafting (SPR) swapping, and saving no more than 5 trees per replicate (Salamin et al. 2003).

Supertree Analyses

Unlike the majority of supertree studies in which input trees have been harvested from the literature, we generated input trees de novo directly from subsamples of the supermatrix (Fig. 1), which permitted standardization of phylogenetic protocols and greater control of data independence. Input trees were inferred for each individual partition (e.g., morphology, a single molecular marker) and also from combinations of partitions. Bininda-Emonds et al. (2004) argue that trees inferred from combinations of partitions are independent from those based on the component partitions because they may yield different results due to data interactions and signal enhancement. However, whereas signal enhancement may be desirable, the repetition and potential increase in the influence of a partition render combinations dependent to some degree and are a cause for concern. To investigate this trade-off, we conducted analyses using only input trees generated from individual partitions (termed basic analyses) as well as analyses based on input trees generated from both individual partitions and combinations of partitions (termed expanded analyses). In the former, the input trees are independent, but there is no possibility of signal enhancement through interactions of secondary signals (Pisani and Wilkinson 2002) in the individual data partitions. In the latter, signal enhancement can occur, but at the expense of the independence of input trees, which may result in some data partitions having undue influence on results.

Combinations of partitions were selected using a strict protocol that aimed to 1) maximize taxon and data sampling and 2) minimize missing data, as follows. Taxon and data sampling were determined by building each combination around a core partition, with each individual data partition being used as the core partition in turn. Further partitions were added to a combination if they contained data for all taxa represented in the core partition. Because taxon sampling varied only slightly between some data sets, we excluded up to 10% of the taxa in the core partition if this permitted the incorporation of a further partition in the respective combination. However, for rooting purposes, we did not permit the complete elimination of all nonpalm out-group taxa, if any were present in the core partition.

Maximum parsimony analyses were conducted for each individual partition and combination under standardized conditions in PAUP*. Where possible, conventional 100-replicate heuristic searches were run to

TABLE 1. Data partitions and combinations used to build input trees

Data partition	Genome	Taxa	Informative characters	Trees	Tree length	CI	RI	Rooting
18S	Nuclear (ribosomal DNA)	73	243	5000	839	0.41	0.56	Nonpalms
*atp*B	Plastid	88	167	5000	392	0.53	0.62	Nonpalms
ITS	Nuclear (ribosomal DNA)	25	318	5	1423	0.39	0.33	Coryphoideae, Arecoideae, Ceroxy-loideae
*mat*K	Plastid	176	628	5000	1972	0.49	0.74	Nonpalms
ms (long)	Nuclear (low-copy region)	18	212	18	575	0.53	0.51	Nonpalms
ms (short)	Nuclear (low-copy region)	75	172	5000	843	0.36	0.51	Nonpalms
*ndh*F	Plastid	68	483	5000	1479	0.48	0.62	Nonpalms
prk	Nuclear (low-copy region)	119	308	5000	1167	0.46	0.74	Calamoideae
*rbc*L[a]	Plastid	205	239	5000	903	0.36	0.64	Nonpalms
*rpb*2	Nuclear (low-copy region)	105	380	5000	1257	0.51	0.66	Calamoideae
*rps*16 intron	Plastid	185	296	5000	844	0.56	0.75	Nonpalms
*trn*D–*trn*T	Plastid	58	70	5000	171	0.54	0.74	Calamoideae
*trn*L–*trn*F	Plastid	176	276	5000	777	0.57	0.70	Nonpalms
*trn*Q–*rps*16	Plastid	58	110	5000	226	0.61	0.75	Calamoideae
RFLP	Plastid	61	131	5000	259	0.51	0.80	Calamoideae
Morphology	—	192	102	5000	1277	0.12	0.68	Calamoideae
Combination 1: morphology/*rbc*L/*rps*16		174	328	5000	2081	0.21	0.68	Calamoideae
Combination 2: *atp*B/*rbc*L		88	331	5000	969	0.43	0.56	Nonpalms
Combination 3 : *ndh*F/*rbc*L		68	640	2250	2020	0.45	0.58	Nonpalms
Combination 4: morphology/*rbc*L		192	204	5000	1874	0.15	0.66	Calamoideae
Combination 5: *mat*K/*rbc*L/*rps*16/*trn*L–*trn*F		175	1392	5000	4450	0.48	0.69	Nonpalms
Combination 6: morphology/*ndh*F/*rbc*L/*trn*D–*trn*T/*trn*Q–*rps*16		56	485	24	1596	0.38	0.57	Calamoideae
Combination 7: *mat*K/morphology/RFLP/*rbc*L/*rps*16/*trn*L–*trn*F		60	526	92	1543	0.39	0.71	Calamoideae
Combination 8: 18S/*atp*B/*mat*K/*rbc*L/*rps*16/*trn*L–*trn*F		67	1240	5000	3665	0.47	0.55	Nonpalms
Combination 9: ITS/morphology/*rbc*L/*rps*16		25	438	3	1773	0.39	0.39	Coryphoideae, Arecoideae, Ceroxy-loideae
Combination 10: *mat*K/*ms* (short)/*rbc*L/*rps*16/*trn*L–*trn*F		70	793	5000	2420	0.46	0.56	Nonpalms
Combination 11: *mat*K/*ms*(long)/*rbc*L/*rps*16/*trn*L–*trn*F		18	481	5000	1066	0.58	0.58	Nonpalms
Combination 12: morphology/*prk*/*rbc*L		119	455	5000	2359	0.31	0.63	Calamoideae
Combination 13: *mat*K/morphology/*prk*/*rbc*L/*rpb*2/*rps*16/*trn*L–*trn*F		95	1025	900	3788	0.41	0.61	Calamoideae

Notes: Underlining indicates the core partitions in combinations. Features of the data sets and resultant trees are given. Note that *ms* (long) is a larger fragment of the malate synthase gene that is approximately twice as long as region *ms* (short) (Lewis and Doyle 2001). It includes *ms* (short) but is available for far fewer taxa than the short fragment. Abbreviations: ITS = Internal Transcribed Spacer region.

[a] New *rbc*L sequences were generated for the following taxa (voucher details and EBI accession numbers provided): *Adonidia merrillii* (Becc.) Becc., Forest 479 (K), AJ829848; *Calospatha scortechinii* Becc.,1990-2783 (K) AJ829855; *Ceratolobus pseudoconcolor* J.Dransf., 1975-3398 (K), AJ829860; *Daemonorops acamptostachys* Becc., Baker 703 (K), AJ829866; *Deckenia nobilis* H.Wendl. ex Seem., Lewis 98-031 (BH), AJ829867; *Eleiodoxa conferta* (Griff.) Burret, Dransfield 6514 (K), AJ829868; *Jubaea chilensis* (Molina) Baill., Forest 482 (K), AJ829875; *Lepidocaryum tenue* Mart., Dransfield 7012 (K), AJ829880; *Mauritiella aculeata* (Kunth) Burret, 1988-4331 (K), AJ829883; *Myrialepis paradoxa* (Kurz) J.Dransf., Baker 491 (KEP), AJ829887; *Nephrosperma van-houtteanum* (H.Wendl. ex Van Houtt.) Balf.f., Lewis 98-006 (BH), AJ829889; *Normanbya normanbyi* (F.Muell.) L.H. Bailey, Lewis 98-091 (BH), AJ829890; *Parajubaea torallyi* (Mart.) Burret, Vargas 3183 (NY), AJ829891; *Pinanga simplicifrons* (Miq.) Becc., Loo 314 (K), AJ829898; *Plectocomiopsis geminiflora* (Griff.) Becc., Baker 492 (KEP), AJ829900; *Pogonotium ursinum* (Becc.) J.Dransf., Baker 517 (K), AJ829901; *Polyandrococos caudescens* (Mart.) Barb.Rodr., 1997-122 (K), AJ829902; *Retispatha dumetosa* J.Dransf., Baker 530 (K), AJ829908; *Solfia samoensis* Rech., Tipama'a 001 (FTG), AJ829912; *Verschaffeltia splendida* H.Wendl., Lewis 98-039 (BH), AJ829916.

completion using TBR swapping and other settings as described above. However, most data sets could not be readily run to completion in this manner and were analyzed using a pragmatic 2-stage approach, the first stage comprising a 1000-replicate heuristic search to explore the tree space thoroughly, saving up to 5 trees per replicate and using SPR and other settings as described above for the supermatrix analysis. In the second stage, the trees from Stage 1 were used as starting trees, swapping them with TBR and saving a maximum of 5000 trees. Each data set was also bootstrapped as described for the supermatrix. Where possible, input trees were rooted on nonpalm out-group taxa. In their absence, trees were rooted according to the best available published evidence (Asmussen et al. 2006) and the outcome

of the supermatrix analysis. The trees were saved as rooted topologies with the root drawn as an internal polytomy.

Both standard and irreversible MRP were employed, using equally weighted "characters" as well as differential weight sets derived from bootstrap support values of the input trees (bootstrap weights). Counterintuitively, standard MRP allows supertree clades to be supported entirely or in part by reversals, and thus by evidence of nonmembership of some input tree clades. The inappropriateness of this led Bininda-Emonds and Bryant (1998) to suggest treating MRP matrix elements as irreversible (0 → 1) characters. Weighting MRP matrix elements by their corresponding bootstrap values has been shown to improve the

performance of MRP (Salamin et al. 2002; Moore et al. 2006) and may be applied to both standard and irreversible implementations.

The MRP matrix was built using Supertree 0.85b (Salamin et al. 2002) from strict consensus trees derived from each of the input tree analyses. Bootstrap weight sets were derived using a function of their bootstrap percentages (BPs) as described by Salamin et al. (2002) and implemented in Supertree 0.85b. The program requires a user-defined maximum weight for branches with 100% support. Because the impact of varying this maximum value has not been examined formally, we repeated our analyses with a range of maximum values (1, 2, 5, and 10) that would vary the influence of bootstrap-supported input tree clades from moderate to extreme. In the absence of an automated mechanism for accurately appending bootstrap values to the strict consensus input trees and to avoid introducing error by manually transcribing them, a pragmatic method for introducing bootstrap weights to the MRP analyses was used. For bootstrap-weighted MRP analyses, both a strict consensus tree and a bootstrap majority-rule consensus tree were submitted as input trees from each input tree analysis, based on the assumption, correct in this instance, that the latter was consistently congruent with the former. Supertree 0.85b automatically assigns a weight of 1 to any resolved input tree branch with a bootstrap value <50%. Thus in bootstrap-weighted analyses, all clades lacking bootstrap support <50% received a weight of 1, whereas branches with bootstrap values ≥50% received a weight of 1 plus a value proportionate to the bootstrap value up to the maximum base weight. The analyses were conducted using the 2-stage approach described above for input tree generation and the resultant trees summarized by strict consensus. Hereafter, the term supertree refers to the strict consensus of all optimal supertrees produced by an analysis.

Parsimony Ratchet Analyses

To assess confidence in the effectiveness of our search strategies, the parsimony ratchet (Nixon 1999) as implemented in PAUP* by PAUPRat (Sikes and Lewis 2001) was used to search for trees shorter than those recovered in the analyses described above. Where practicable, 10 ratchet searches were conducted on each of the various MRP matrix implementations and on the supermatrix. Individual ratchet searches consisted of 200 ratchet iterations, with 15% of characters perturbed in each iteration.

Comparisons of Trees

We used congruence with the phylogeny inferred from the supermatrix as a basis for selecting a "most congruent supertree" from among those produced using the various combinations of weighting, standard or irreversible parsimony, and basic or expanded input tree sets (Fig. 1). Thus, the most congruent of the supertrees is that which shares the highest number of clades with the supermatrix phylogeny. This does not assume that the latter is accurate, rather it rests on the central premise of taxonomic congruence that clades are more compelling and, in some sense, better supported if they are inferred using alternative approaches than if they are not. Operationally, the use of congruence as an objective function to select among supertrees is justified simply by our desire to produce a maximally informative synthesis of the 2 approaches. We also use the number of shared clades, the number of compatible clades, and the number of clades that are unsupported by the expanded input tree set to further investigate the effect of the methodological variants on inferred supertrees.

Supertree Clade Support

To assess clade support in the most congruent supertree, we used recently developed measures based on comparison of supertree clades to input trees using ST-support (Wilkinson, Pisani, et al. 2005) and QualiTree.pl version 1.2.1 (Price et al. 2005). Given a supertree clade C and an input tree T, then T supports C (if T includes a branch implied by C), is in conflict with C (if C and T are incompatible), permits C (if some resolution of a polytomy in T is a branch implied by C), or is irrelevant to C (Wilkinson, Pisani, et al. 2005). The relationship between each supertree clade and each tree in the expanded input tree set was characterized as one of support, conflict, permission, or irrelevance and the number of such relationships summed across all input trees for each supertree clade to give s, q, p, and r, respectively. From these measures, 2 additional indices, $V = (s - q)/(s + q)$ (Wilkinson, Pisani, et al. 2005) and $rQS = (s - q)/(s + q + p + r)$ (Price et al. 2005; Beck et al. 2006), were calculated for each supertree clade. Both these indices range from $+1$ to -1, with high scores reflecting strong support (see Discussion). Following Wilkinson, Pisani, et al. (2005), a supertree clade is considered unsupported if $s = 0$. The number of clades not supported by the expanded input tree set was determined for all supertrees.

Data and Trees

The supermatrix and MRP matrix (including bootstrap weight sets) can be downloaded from http://sysbio.oxfordjournals.org/ or http://eunops.org/en/content/palmsupertree and are also available from the first author on request. The input trees (with their taxon sets), supertrees, and the supermatrix tree are included in these files.

RESULTS

Supermatrix Analysis

In total, the supermatrix consisted of 16 data partitions representing 205 taxa and contained 20 355 characters, of which 4050 (20%) were parsimony informative. The completeness of taxonomic sampling

in each partition ranged from 12% to 100%, with an average of 51% (Table 1). Nine of the 16 partitions contained nonpalm out-groups. The previously unpublished morphological data set accounted for 105 of the characters. Twenty new *rbc*L sequences were generated to complete the generic sampling of palms for this gene.

Parsimony analysis of the supermatrix yielded 46 080 trees of 15 173 steps in length (consistency index [CI] = 0.41, retention index [RI] = 0.62). The parsimony ratchet did not improve upon this tree length. The strict consensus of these trees (hereafter termed the supermatrix tree) contains 176 resolved clades out of a maximum of 204 clades (i.e., 86%). Of these clades, 126 (72%) were supported by BPs greater than 50% (Fig. 2). Poor resolution is principally concentrated within the western Pacific clade of tribe Areceae (Norup et al. 2006), although some relationships within the Iriarteeae, Trachycarpeae, Cryosophileae, and Borasseae also remain ambiguous. All other groups showed no conflict among most parsimonious trees. Five weakly supported relationships (≤56%) in the bootstrap 50% majority-rule consensus tree are in conflict with or unresolved in the strict consensus tree of 46 080 trees. These branches resolve 1) *Sabal* sister to a clade of Corypheae, Chuniophoeniceae, Caryoteae, and Borasseae (53%), 2) *Howea* sister to *Laccospadix* and *Linospadix* (56%), 3) *Rhapidophyllum* sister to *Trachycarpus* (52%), 4) the monophyly of subtribe Lataniinae (53%), and 5), in the out-groups, *Calectasia* sister to *Dasypogon* (54%).

Supertree Analyses

In total, 29 analyses were conducted to produce input trees, 16 using individual partitions and 13 using unique combinations of partitions. Some core partitions shared identical taxon sampling and thus yielded identical combinations (*rps*16, *mat*K, and *trn*L–*trn*F; *trn*D–*trn*T and *trn*Q–*rps*16). Details of the input tree data sets and their tree statistics are presented in Table 1. All bootstrap 50% majority-rule consensus input trees were congruent with strict consensus input trees.

The MRP matrix consisted of 1561 parsimony-informative matrix elements. Basic analyses used only the 740 matrix elements derived from input trees inferred from individual partitions, whereas expanded analyses used the additional 821 matrix elements derived from input trees inferred from combinations of partitions. Of the 1561 elements, 970 (62%) were given additional weight in proportion to input tree bootstrap values in weighted analyses. All standard MRP analyses reached the preset tree limit of 5000 trees, whereas some irreversible MRP analyses yielded as few as 179 trees. Where tested, the shortest tree lengths recovered in these analyses were not bettered by the parsimony ratchet. Details of the degree of resolution, congruence with the supermatrix tree, and number of unsupported groups for each supertree are given in Table 2, and examples of supertrees are given in Figures 3 and 4.

Comparisons of Trees

The supertree produced using standard MRP with maximum bootstrap weights based on the expanded input tree set is most congruent with the supermatrix tree (Table 2) and is thus termed the "most congruent supertree" on this basis. The most congruent supertree also has the highest average (pairwise) congruence with the other supertrees (data not shown) and the lowest number of unsupported groups (Table 2). This tree is presented in Figure 3, and support values (*s*, *q*, *p*, *r*, rQS, V) for each of its clades are given in Appendix 1.

A conservative formal synthesis of the most congruent supertree and the supermatrix tree, both of which we take to be plausible alternative phylogenies of the palm family, can be provided in the form of their semistrict consensus tree (online Supplementary Appendix 3). This consensus is well resolved with 160 clades, of which 137 are present in both trees (Fig. 2), 19 are present in the most congruent supertree and compatible with the supermatrix tree (Fig. 3 and Appendix 1), and 4 are present in the supermatrix tree and compatible with the most congruent supertree (Fig. 2).

Standard MRP supertrees are always more congruent with the supermatrix tree and have fewer unsupported groups than the corresponding irreversible MRP supertree. Supertrees inferred using bootstrap-weighted MRP are always more congruent with the supermatrix tree than are supertrees inferred from equally weighted matrix elements, with more extreme weighting resulting in greater congruence. However, differential weighting has no consistent effect on the number of unsupported groups. Supertrees inferred using the expanded input tree set always have fewer unsupported clades and are always more congruent with the supermatrix tree than corresponding supertrees inferred from the basic input tree set.

Phylogenetic Relationships among Palm Genera

Supertrees derived from the basic input tree set fail to recover some phylogenetic relationships among palms that are well substantiated by previous studies. Examples of unorthodox relationships recovered in these supertrees include Coryphoideae being sister to Ceroxyloideae and Calamoideae being sister to Arecoideae. The monophyly of subfamilies Arecoideae, Ceroxyloideae, and Coryphoideae, tribes Areceae, Calameae, Euterpeae, and Lepidocaryeae, and subtribes Ancistrophyllinae, Hyphaeninae, Livistoninae, and Plectocomiinae are questioned by at least one, but usually several of these supertrees. Tribes Areceae and Euterpeae and subtribe Livistoninae are nonmonophyletic in all cases. These findings erode the credibility of basic input tree supertrees.

In contrast, the supermatrix tree, the most congruent supertree, and most supertrees based on the expanded input tree set are in substantial agreement with the classification of Dransfield et al. (2005; Dransfield, Uhl, et al. 2008), which was based on all available phylogenetic evidence from palms, including a preliminary version of
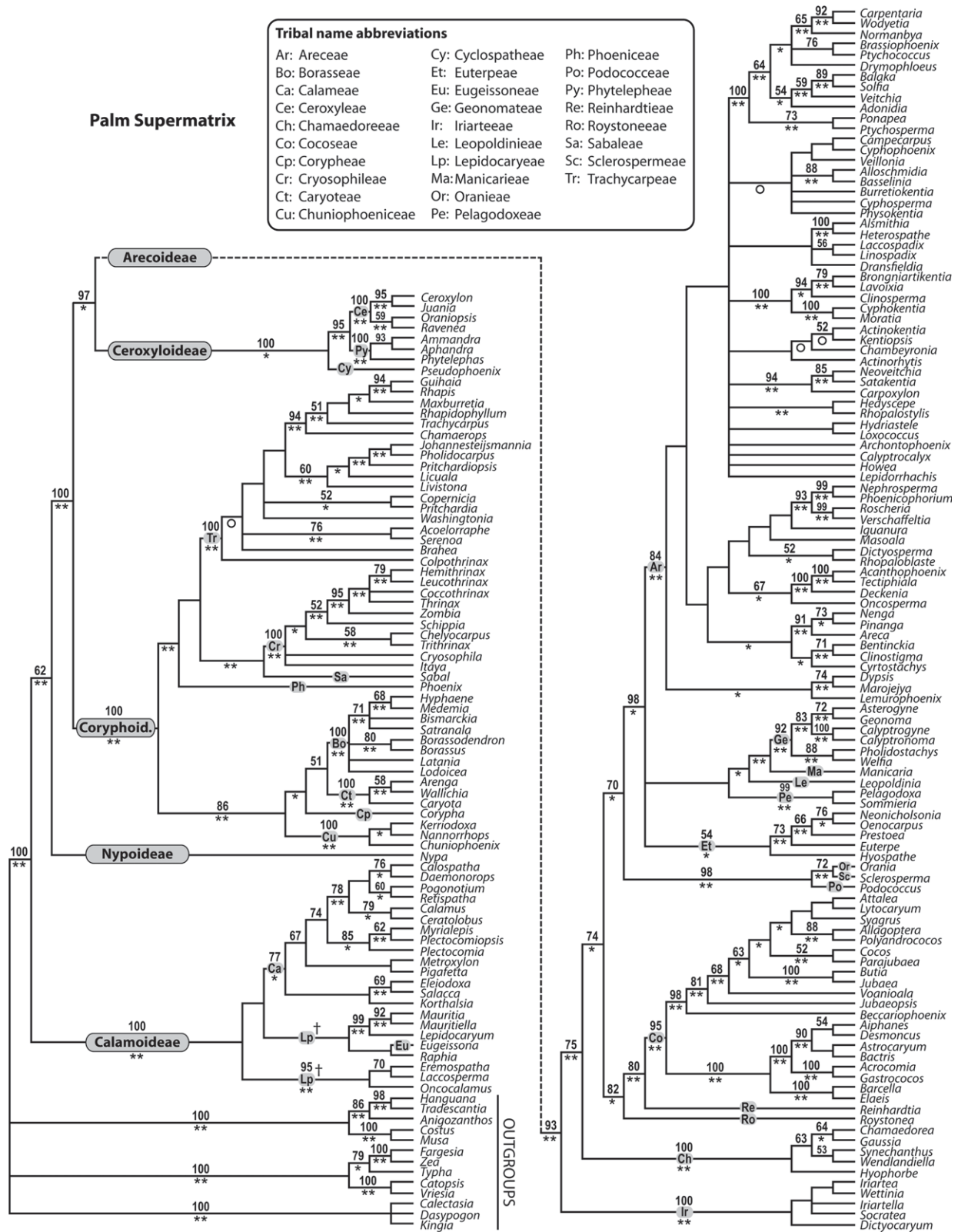
FIGURE 2. Strict consensus of the 46 080 trees from the supermatrix analysis of all palm genera (tree length = 15 173, CI = 0.41, RI = 0.62). Values above branches are BPs (>50%). Asterisks below branches indicate clades also recovered in supertrees; 1 asterisk denotes a clade recovered in the "most congruent supertree" (Fig. 2), 2 asterisks denote a clade recovered in all supertrees summarized in Figure 3 (including the most congruent supertree, see Fig. 3 legend). An open circle indicates a clade absent from, but compatible with, the most congruent supertree. Palm subfamilies and tribes are indicated. Nonmonophyletic higher taxa are marked with a dagger symbol.

TABLE 2. Numbers of clades in each supertree that are resolved, unsupported by the expanded input tree set (including the percentage of total resolved clades), shared with the supermatrix tree, and present in the semistrict consensus of the supertree and supermatrix tree (i.e., shared clades plus compatible clades)

Supertree analysis	Resolved clades	Input tree comparison Unsupported clades (%)	Supermatrix tree comparisons Shared clades	Shared + compatible clades
Basic input tree set				
Standard, equal weights	135	7 (5.19)	86	129
Standard, bootstrap weights (max = 1)	178	8 (4.49)	110	125
Standard, bootstrap weights (max = 2)	177	6 (3.39)	117	129
Standard, bootstrap weights (max = 5)	173	4 (2.31)	119	131
Standard, bootstrap weights (max = 10)	166	3 (1.81)	119	132
Irreversible, equal weight	140	15 (10.71)	83	123
Irreversible, bootstrap weights (max = 1)	182	24 (13.19)	101	115
Irreversible, bootstrap weights (max = 2)	183	18 (9.78)	106	117
Irreversible, bootstrap weights (max = 5)	177	14 (7.91)	111	124
Irreversible, bootstrap weights (max = 10)	172	12 (6.98)	111	125
Expanded input tree set				
Standard, equal weights	171	2 (1.17)	112	136
Standard, bootstrap weights (max = 1)	187	3 (1.60)	126	142
Standard, bootstrap weights (max = 2)	191	2 (1.05)	127	149
Standard, bootstrap weights (max = 5)	193	**1 (0.52)**	133	156
Standard, bootstrap weights (max = 10)	193	**1 (0.52)**	**137**	**160**
Irreversible, equal weight	176	9 (5.11)	111	129
Irreversible, bootstrap weights (max = 1)	**196**	8 (4.08)	123	137
Irreversible, bootstrap weights (max = 2)	195	9 (4.62)	127	143
Irreversible, bootstrap weights (max = 5)	193	13 (6.74)	127	145
Irreversible, bootstrap weights (max = 10)	193	11 (5.70)	130	148

Note: The best values are indicated in bold. Abbreviation: max = maximum.

this study. The monophyly of all subfamilies is widely supported (Arecoideae: BP = 93, s = 8; Calamoideae: BP = 100, s = 16; Ceroxyloideae: BP = 100, s = 6; Coryphoideae: BP = 100, s = 14; Nypoideae is monogeneric). The Ceroxyloideae clade collapses in the strict consensus of the equally weighted standard MRP supertree (although it is resolved in the majority of the most parsimonious trees).

Of the expanded input tree set supertrees, only the equally weighted irreversible MRP supertree is incongruent with palm subfamily delimitation, contradicting the monophyly of the Arecoideae and Ceroxyloideae. This supertree, which is the least congruent with the supermatrix tree of all expanded analysis supertrees, displays a number of unorthodox relationships among major lineages that are not supported by any input tree (s = 0). For these reasons, the remaining systematic interpretation focuses on comparisons between the supermatrix tree and all expanded input tree set supertrees except for the equally weighted irreversible MRP supertree. These supertrees are summarized in Figure 4, and key systematic features are described briefly here.

The monophyly of the 28 tribes of Dransfield et al. (2005; Dransfield, Uhl, et al. 2008) is almost universally supported. Tribe Euterpeae is not monophyletic in some standard MRP supertrees (equal and maximum weight = 1 and 2) due to the exclusion of *Hyospathe*. The calamoid tribes Calameae and Lepidocaryeae are also problematic in some cases. Tribe Lepidocaryeae is nonmonophyletic in the supermatrix tree, although the alternative relationships are poorly supported by the bootstrap. Similarly, the tribe is nonmonophyletic in 1 irreversible MRP supertree (maximum weight = 1). Tribe Calameae is not monophyletic in 2 irreversible

MRP supertrees (maximum weight = 1 and 2) as well as the equally weighted standard MRP supertree. At lower taxonomic levels, the monophyly of 7 subtribes was contradicted by at least one of the analyses summarized in Figure 4 (Archontophoenicinae, Basseliniinae, Dypsidinae, Lataniinae, Linospadicinae, Oncospermatinae, Plectocomiinae).

All analyses converge on the same relationships among subfamilies, that is, (Calamoideae (Nypoideae (Coryphoideae (Ceroxyloideae, Arecoideae)))). These relationships are for the most part strongly supported by bootstrap and s values (Fig. 4) and are consistent with the most extensive family-wide study published to date (Asmussen et al. 2006). Of these, the placement of Nypoideae as sister to all subfamilies except for the Calamoideae is the least well supported by the supermatrix (BP = 62) but reflects input tree topologies well (s = 15). There is, however, diversity in relationships at lower taxonomic levels (Fig. 4).

DISCUSSION

Supertrees and Supermatrices

Attitudes toward supertree and supermatrix approaches to phylogenetic inference have often been polarized (Gatesy et al. 2002, 2004; Bininda-Emonds 2003, 2004b). In contrast, we consider that there are no compelling reasons for dismissing either approach a priori and that comparisons of their results may be informative and helpful. Where the approaches support the same inferences, the cross-corroboration is a legitimate reason to have more confidence in those inferences than in inferences that are incompatible
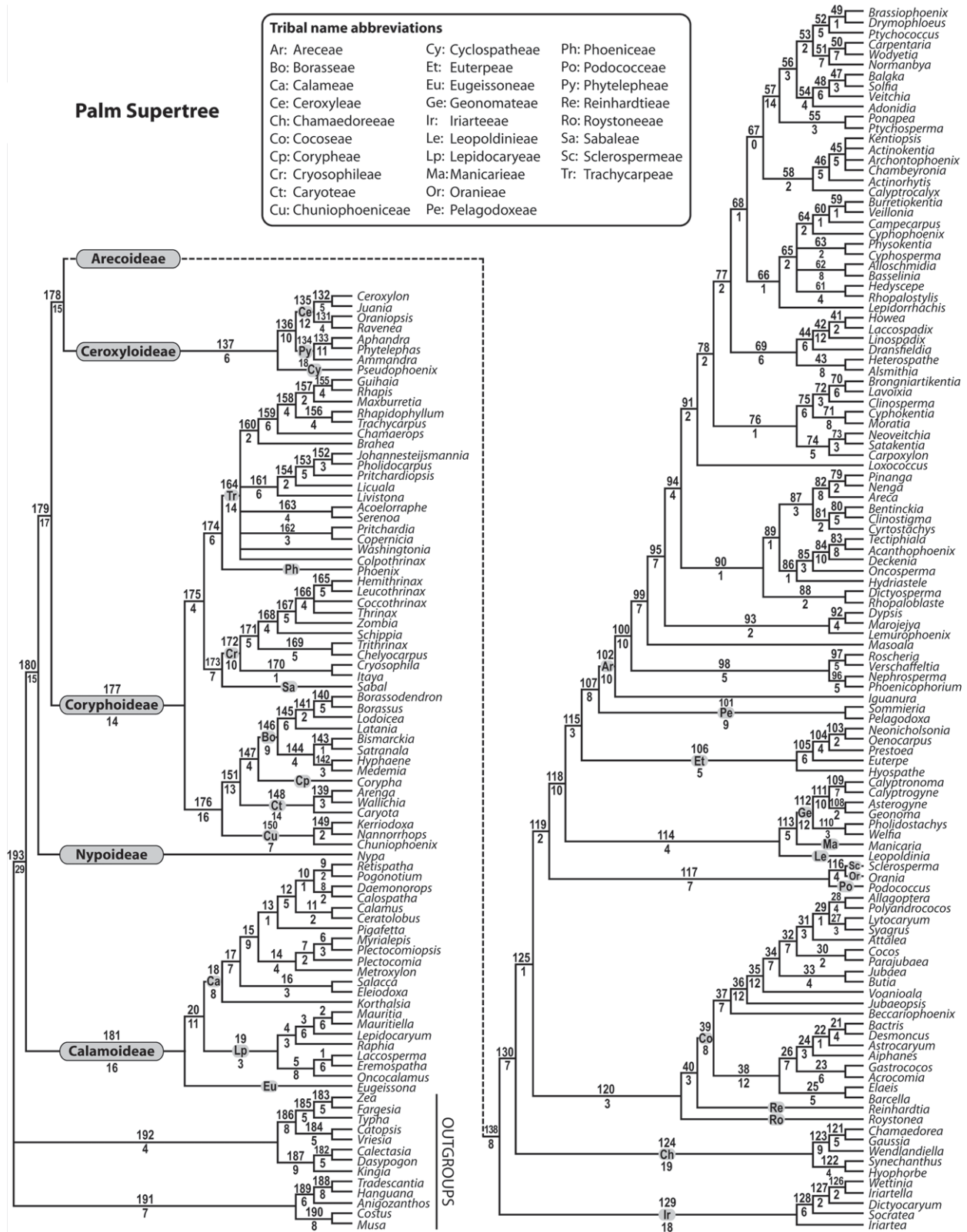
FIGURE 3.   The "most congruent supertree." Strict consensus of 5000 supertrees from standard MRP analysis based on the expanded input tree set with matrix elements weighted in proportion to bootstrap values of corresponding input tree clades (maximum weight = 10, tree length = 11 830.667, CI = 0.80, RI = 0.95). Numbers above branches are clade numbers (see online Supplementary Appendix 1 for support values). Those below branches indicate the number of input trees that support a given clade (*s*). Palm subfamilies and tribes are indicated. Only Node 67 is not supported by any input tree.

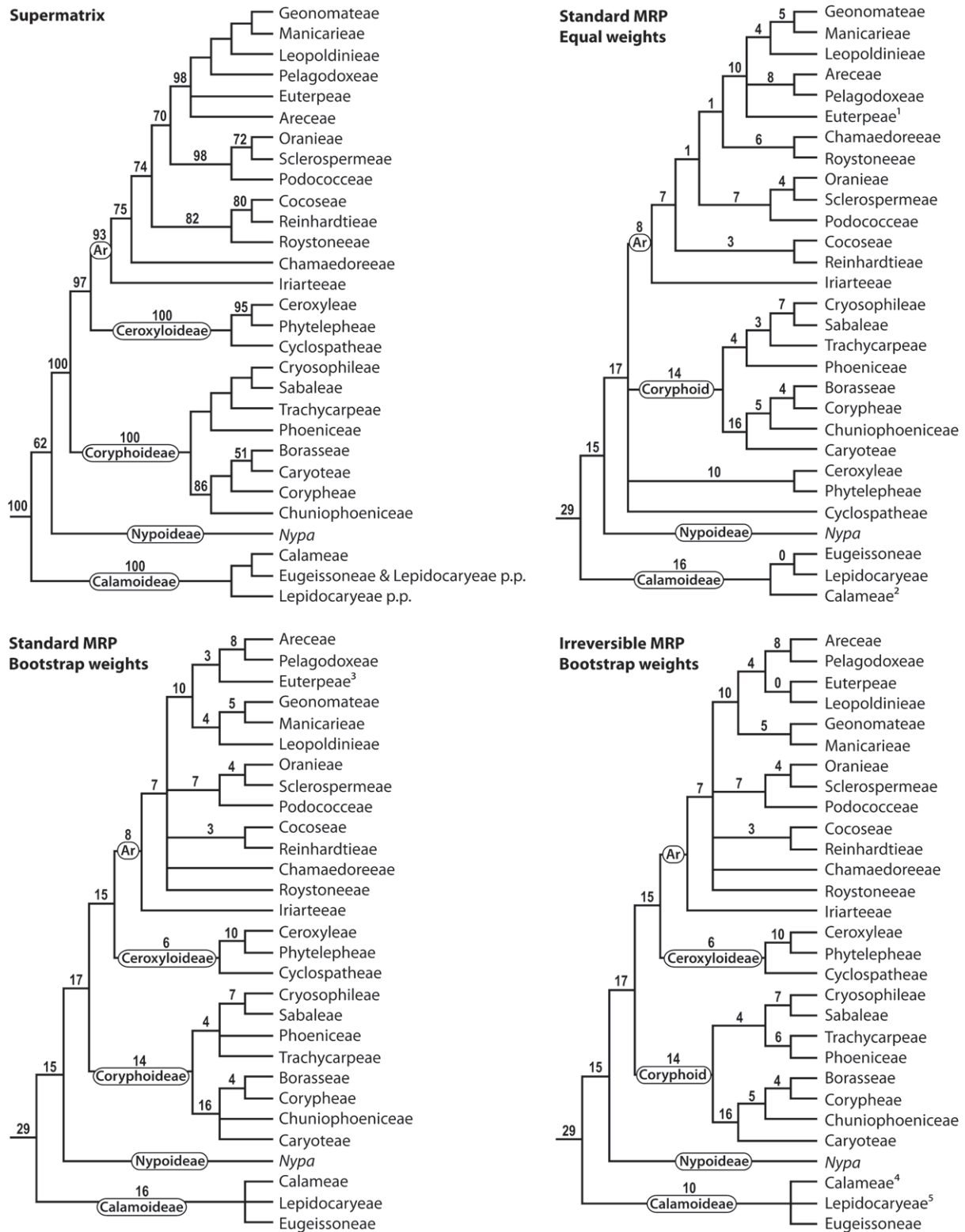FIGURE 4.   Summary trees depicting relationships between palm subfamilies and tribes that are resolved by the supermatrix and MRP analyses based on the expanded input tree set. The upper 2 trees show the relationships resolved by the supermatrix and equal weights standard MRP analyses. The lower 2 trees show relationships in common between all bootstrap-weighted standard MRP analyses and all bootstrap-weighted irreversible MRP analyses, respectively. The equal-weights irreversible MRP supertree is not shown here. Values above branches are BPs for the supermatrix tree and *s* for the supertrees. Palm subfamilies are indicated (Ar = Arecoideae). Notes: [1]*Hyospathe* resolves outside Euterpeae; [2]Calameae paraphyletic; [3]*Hyospathe* resolves outside Euterpeae in analyses with maximum weights = 1 and 2; [4]nonmonophyletic in analyses with maximum weights = 1 and 2; [5]nonmonophyletic in analyses with maximum weights = 1.

(Levasseur and Lapointe 2001, 2003). We think that the ranking of clades based on congruence and robustness to variation in methods of analysis is a useful adjunct to other measures of support. Thus, our comparisons of supertrees and the supermatrix tree help to highlight where uncertainty remains and where further work is most desirable.

In general, supertrees and the supermatrix are more congruent than incongruent, reflecting the experience of other workers who have employed both approaches (Kennedy and Page 2002; Salamin et al. 2002; Fritsch et al. 2006; Fulton and Strobeck 2006; Hughes et al. 2006). However, there are substantial differences in the performance of the variants of MRP, indicating that the choice of supertree method can be important in practice. In our example, standard MRP trees are always more congruent with the supermatrix tree, and irreversible MRP appears to promote the undesirable occurrence of unsupported clades, with between 4% and 13% of clades being unsupported by any input tree in irreversible MRP supertrees compared with 0.5–5% in standard MRP supertrees. These results further undermine the claim that unsupported clades are rare in MRP supertrees (Bininda-Emonds 2003; see also Wilkinson, Pisani, et al. 2005) and further support the expectation that differential weighting of matrix elements as a function of bootstrap values can improve the performance of MRP (Ronquist 1996; Sanderson et al. 1998; Bininda-Emonds and Sanderson 2001; Bryant 2004). In our example, the heaviest weighting schemes produced greatest congruence with the supermatrix tree.

The inclusion of input trees based on combinations of individual data partitions is controversial because of the lack of independence of these data partitions and the repetition of the most taxonomically complete partitions, which might lead to overweighting and potential biases. Thus, in the expanded tree set the *rbc*L partition featured in 14 of the 29 input tree analyses, whereas *rps*16 intron featured in 9, morphology in 8, and *mat*K and *trn*L–*trn*F in 7. In contrast, all remaining partitions were included in only 2 or 3 input tree analyses (Table 1). Nevertheless, supertrees based on the expanded input tree set have fewer unsupported clades and greater congruence with supermatrix tree than corresponding supertrees based on the basic input tree set. Whereas fewer unsupported groups (and greater resolution) might be expected simply from the increase in the number of input trees, the enhanced congruence with the supermatrix tree is less predictable a priori and appears to us to be a real benefit of this approach. In addition, these supertrees contain more credible phylogenetic results with respect to relationships that are well established in the literature. Although lack of independence is a reasonable concern, combining basic partitions allows some degree of signal enhancement (data interaction) that is not included if only single trees from independent partitions are used (Pisani and Wilkinson 2002). Future work should be directed at investigating the extent to which the introduction of such non-independence might bias results toward the signal in the largest or most overrepresented data partitions. An alternative approach to promoting signal enhancement while maintaining independence would be to employ sets of (e.g., bootstrap) trees rather than a single tree for each partition, so that secondary signals (sensu Pisani and Wilkinson 2002) are not entirely excluded from the analysis.

One advantage that has been claimed for supertree methods is that they avoid problems of missing data, which may be abundant in supermatrices (Sanderson et al. 1998; Bininda-Emonds et al. 2002; Bininda-Emonds 2004a). Despite the large amounts of missing data in our supermatrix, which, on average, lack data for around half of the taxa in each partition, it yields a well-resolved phylogeny with many well-supported clades. As others have suggested (de Queiroz and Gatesy 2007), the problem of missing data in supermatrix analyses may have been overstated. Indeed, if nonrandomly distributed missing data resulting from combining data partitions were a major problem for any analysis of a supermatrix, then such problems would be unlikely to be overcome by supertree methods, which simply replace the missing data with missing taxa and thereby transform the problem rather than solving it.

Achieving effective overlap (Wilkinson and Cotton 2007) among input trees is an important practical issue in supertree construction. The benefits of "seeding" both supermatrix and supertree studies with at least one completely sampled study have been demonstrated through simulation (Bininda-Emonds and Sanderson 2001). Many previous supertree studies have resorted to including a comprehensive taxonomy as one of the input trees in order to obtain effective overlap. In this study, the significant overlap in taxon sampling between data partitions undoubtedly contributes much to the success of the supertree and supermatrix analyses presented here and we have been able to avoid reliance upon taxonomies interpreted as phylogenies. At least one data partition (*rbc*L) is complete for all 205 taxa in the data set, whereas morphological data are present for all 192 palm genera. In addition, *mat*K, *rps*16 intron, and *trn*L–*trn*F data were available for more than 85% of taxa. These data sets are largely derived from a coordinated phylogenetic effort (Baker et al. 1999; Asmussen et al. 2000, 2006; Asmussen and Chase 2001) and provide a framework into which data sets from more sparsely sampled family-wide studies (e.g., Uhl et al. 1995; Lewis and Doyle 2001; Hahn 2002a) and less inclusive projects (e.g., Baker, Dransfield, Hedderson 2000; Hahn 2002b; Gunn 2004; Norup et al. 2006) can be readily integrated.

A major difference between our supertree analyses and many previous studies using MRP is that we generated input trees de novo from the supermatrix rather than using trees culled from the literature. Advantages of this approach are that dependencies in the underlying data highlighted by Gatesy et al. (2004) as well as variation in the analytical methods by which input trees are derived can be controlled. It also permits direct comparison of the supertrees with trees derived from conventional analysis of the supermatrix, providing

an objective function for the selection of the most congruent supertree from among variants. This objective function is entirely justified operationally, if our aim is to produce a well-resolved synthesis of supermatrix and supertree approaches, but it rests on the premise that congruent clades are generally more plausible than those that are not. Thus, it is worth noting that average V and rQS are much higher (0.205, 0.062) for the 137 clades of the most congruent supertree that are also in the supermatrix tree than they are for the 56 clades that are not (−0.193, −0.082). Similarly, of these 137 clades, 21 have bootstrap support of less than 50% and the remainder have an average of 84% in the supermatrix tree, whereas 30 of the 39 clades that are unique to the supermatrix tree have bootstrap support less than 50% and the average of the other 9 is just 60%. Judged in terms of supermatrix- and supertree-specific measures of support, clades that are present in both the supermatrix tree and the most congruent supertree are on average better supported than those that are not.

Support for Supertree Clades

The assessment of clade support is no less desirable for supertrees than for any other phylogenetic tree. However, the use of conventional measures with MRP, for example, BPs and decay indices, is complicated by the nonindependence of matrix elements derived from the same input tree and/or the potential for resampled data sets to lack particular taxa. Although these approaches have sometimes been used (e.g., Creevey et al. 2004) and some progress has been made in addressing these problems (Burleigh et al. 2006; Moore et al. 2006), the development of supertree-specific support indices has lagged behind the evolution of supertree construction methodology and has proven controversial. Bininda-Emonds (2003) developed a measure of quantitative support (QS) for supertree clades based on their relationship to individual input trees. Wilkinson, Pisani, et al. (2005) showed that QS can be difficult to interpret and may be counterintuitive and misleading. They defined the 4 mutually exclusive relationships of support, conflict, permission, and irrelevance that may exist between a supertree clade and an input tree and the corresponding measures s, q, p, and r that sum these relationships over all input trees, as employed here (Appendix 1). They also developed composite measures, the simplest of which, V, ranges from +1 to −1 like QS. In response, Price et al. (2005) introduced rQS that they claimed addressed the criticisms of QS. However, Price et al. provided no formal definition and did not explore the relationship between rQS and V. Given the importance of assessment of support, we provide here some comments on these measures and illustrate the differences in practice using the most congruent supertree presented in Figure 3. Appendix 1 summarizes s, q, p, r, V, and rQS for every clade in the supertree.

Both V and rQS are ratios that have the same numerator ($s − q$) and differ only in the denominator. With V, the denominator ($s + q$) is the sum of those trees that either support or conflict with the supertree clade. With rQS, the denominator is the total number of input trees ($= s+q+p+r$). Thus, the important difference is with respect to input trees that neither support nor conflict with the supertree clade. These trees are ignored in the calculation of V but do affect rQS. Consequently, if there are any irrelevant trees (i.e., trees that have no bearing on the support for the supertree clade because their taxon sampling is not informative with respect to the clade), then the absolute magnitude of rQS will be less than that of V, and the range of rQS will be less than that of V (Fig. 5). If rQS is +1 or −1, this tells us that all trees support or conflict with the supertree clade, respectively. However, neither is likely in most supertree analyses because of the limitations in taxon sampling of some input trees (Fig. 5). This is one of the original criticisms of QS that, *contra* Price et al. (2005), is not addressed by rQS. Different treatments of irrelevant trees also affect the ease of interpretation of rQS and V. Thus, if V is the same for 2 or more supertree clades, then we know that the ratio of support to conflict (though not their absolute values) is identical for these clades. In contrast, rQS values may be quite different when the ratio of support and conflict is identical, and identical rQS values may obtain under very different circumstances. For example, 4 clades (60, 94, 143, 154) of the most congruent supertree have V = −0.5 and a ratio of s/q of one-third, whereas rQS ranges from −0.07 to −0.27. Thus, we recommend V in preference to rQS if a composite measure combining support and conflict is desired. In general though, we think it more useful to report s and q uncombined because their meaning and relevance are clear, they are readily comparable across clades, and other measures are readily calculable from them. One important caveat here is that a single input tree clade may be taken to support multiple supertree clades as a result of differences in the included taxa. Although not employed here, Wilkinson, Pisani, et al. (2005) suggested reporting both s and a weighted support measure that takes this into account.
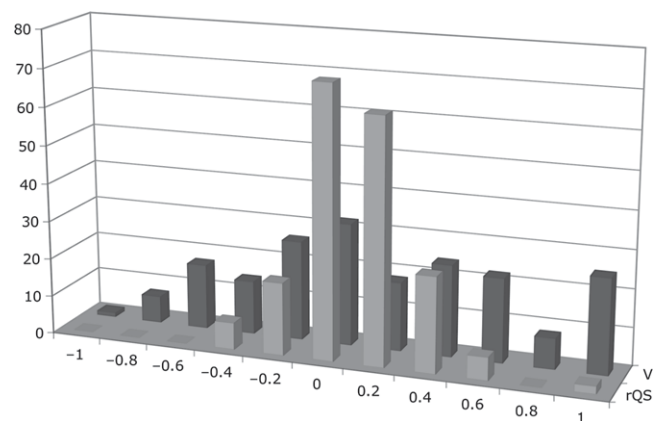

FIGURE 5.   Histogram showing distribution of values of V and rQS derived from the "most congruent supertree" presented in Figure 3.

Palm Phylogenetics

Our principle results are 2 synthetic, complete generic-level phylogenies of palms produced using supertree and supermatrix approaches. Neither the supermatrix tree nor the most congruent supertree is definitive, but each is well resolved and plausible. In addition, their semistrict consensus (online Supplementary Appendix 3) provides a more conservative synthesis of the 2 approaches, and comparisons with other suboptimal supertrees (Fig. 4) reveal some possible alternative relationships that might merit further investigation.

Our analyses agree strongly with the emerging consensus on the phylogenetics and classification of palms (Baker et al. 1999; Asmussen et al. 2000, 2006; Asmussen and Chase 2001; Hahn 2002a; Dransfield et al. 2005; Dransfield, Uhl, et al. 2008). By integrating data sets that are informative at different levels, we have succeeded where other studies have failed in achieving resolution across the family. For example, the most comprehensive phylogeny published to date (Asmussen et al. 2006) resolves many of the relationships that we too have recovered, but resolution and support are not evenly distributed across the family, with polytomies concentrated in subfamily Arecoideae in particular. In our study, relationships with Arecoideae are much more highly resolved with major lineages such as the core arecoid clade (Areceae, Euterpeae, Geonomateae, Leopoldinieae, Manicarieae, Pelagodoxeae) and the clade comprising Oranieae, Podococceae, and Sclerospermeae, as well as many lower-level relationships receiving good support in all analyses. However, ambiguities remain in some areas. As well as uncertainties in calamoid relationships, we have recovered a range of conflicting topologies among the tribes of the syncarpous clade of Coryphoideae (Borasseae, Caryoteae, Chuniophoeniceae, and Corypheae), among the Trachycarpeae, Phoeniceae, and the New World thatch palm clade (Cryosophileae and Sabaleae), and among major lineages of Arecoideae. Poor support and resolution also remain within 2 of the larger tribes, Areceae and Trachycarpeae. In all these areas, our study is limited by the data at hand, and improvements will only be obtained by targeted new data gathering for these groups.

Groups from the classification of Dransfield et al. (2005; Dransfield, Uhl, et al. 2008) that were not resolved as monophyletic highlight areas that would benefit in particular from closer investigation. For example, only 6 of the 16 data partitions contain a broad sampling of Calamoideae. The relationships that contradict the monophyly of tribes Lepidocaryeae and Calameae are not sufficiently well supported to undermine the current classification but should be prioritized for future phylogenetic research. In the case of tribe Euterpeae (Arecoideae), nonmonophyly in some supertrees is most likely caused by *Hyospathe*, which is represented in only 5 of the data partitions and 8 of the 29 input trees in the expanded set. Only 15% of genera appear in 8 input trees or less, with some genera featuring in as few as 5 input trees (e.g., *Jubaea*, *Parajubaea*, *Polyan-*

drococos), whereas some others are represented in as many as 27 (*Hyophorbe*, *Orania*, *Nypa*; online Supplementary Appendix 1). The relationships and classification within tribe Areceae are a particular cause for concern, with the monophyly of several of its subtribes (Archontophoenicinae, Basseliniinae, Dypsidinae, Linospadicinae, Oncospermatinae) called into question.

PROSPECTS

Using supermatrix and supertree methods, we have been able to construct the first phylogenetic trees for all palm genera, a valuable outcome not only for systematists but also for workers needing a comprehensive and highly resolved phylogenetic framework to pursue comparative research. However, it should be noted that our data sets are biased toward plastid DNA evidence. Although we have no reason to suspect that plastid data lead to erroneous phylogeny reconstructions in palms, it would be advantageous to sample the plant genomes more evenly, especially the nuclear genome. In addition, we have concentrated on parsimony-based analyses for this study, whereas model-based methods may provide alternative perspectives, especially those that can accommodate varied data types (DNA sequence, RFLP, morphology). With expanded taxon and data sampling, as well as methodological refinements, we can expect supertree and supermatrix approaches to be instrumental in achieving the ultimate goal of a species-level phylogeny of the palms. Although access to material and knowledge of informative nonplastid markers limit such ambitions currently, these obstacles are gradually being overcome by new molecular phylogenetic developments (e.g., Bacon et al. 2008) and targeted field collecting campaigns.

This study demonstrates empirically that supertree and supermatrix methods can provide effective, explicit, and complimentary mechanisms for synthesizing disjointed phylogenetic evidence. Relationships recovered by both approaches are more likely to be accurate representations of the underlying data than artifacts caused by the failure of just one of the approaches. Conversely, we may consider relationships that are not confirmed by both approaches less well supported and in need of further evaluation. It is not yet clear to what extent differences in the performance of MRP variants are general or context specific, and this merits further investigation. Nevertheless, we suggest that users of MRP should evaluate a range of implementations, where practical, rather than picking one approach a priori. The utility of MRP supertrees in the current case does not obviate the need for a critical attitude to supertrees, for assessments of the properties of MRP variants and other methods (e.g., Wilkinson, Cotton, et al. 2005) and their further development and refinement.

SUPPLEMENTARY MATERIAL

Supplementary material can be found at http:// sysbio.oxfordjournals.org/.

References

Asmussen C.B., Baker W.J., Dransfield J. 2000. Phylogeny of the palm family (Arecaceae) based on *rps*16 intron and *trn*L-*trn*F plastid DNA sequences. In: Wilson K.L., Morrison D.A., editors. Monocots: systematics and evolution. Melbourne (Australia): CSIRO. p. 525–535.

Asmussen C.B., Chase M.W. 2001. Coding and noncoding plastid DNA in palm systematics. Am. J. Bot. 88:1103–1117.

Asmussen C.B., Dransfield J., Deickmann V., Barfod A.S., Pintaud J.C., Baker W.J. 2006. A new subfamily classification of the palm family (Arecaceae): evidence from plastid DNA phylogeny. Bot. J. Linn. Soc. 151:15–38.

Bacon C.D., Feltus F.A., Paterson A.H., Bailey C.D. 2008. Novel nuclear intron-spanning primers for Arecaceae evolutionary biology. Mol. Ecol. Res. 8:211–214.

Baker W.J., Asmussen C.B., Barrow S.C., Dransfield J., Hedderson T.A. 1999. A phylogenetic study of the palm family (Palmae) based on chloroplast DNA sequences from the *trn*L-*trn*F region. Plant Syst. Evol. 219:111–126.

Baker W.J., Dransfield J., Hedderson T.A. 2000. Phylogeny, character evolution, and a new classification of the calamoid palms. Syst. Bot. 25:297–322.

Baker W.J., Hedderson T.A., Dransfield J. 2000. Molecular phylogenetics of subfamily Calamoideae (Palmae) based on nrDNA ITS and cpDNA *rps*16 intron sequence data. Mol. Phylogenet. Evol. 14:195–217.

Baker W.J., Zona S., Heatubun C.D., Lewis C.E., Maturbongs R.A., Norup M.V. 2006. *Dransfieldia* (Arecaceae)—a new palm genus from western New Guinea. Syst. Bot. 31:61–69.

Baum B.R. 1992. Combining trees as a way of combining datasets for phylogenetic inference, and the desirability of combining gene trees. Taxon. 41:3–10.

Baum B.R., Ragan M.A. 1993. A comment on Baum's method for combining phylogenetic trees—reply. Taxon. 42:637–640.

Baum B.R., Ragan M.A. 2004. The MRP method. In: Bininda-Emonds O.R.P., editor. Phylogenetic supertrees: combining information to reveal the tree of life. Dordrecht (the Netherlands): Kluwer. p. 17–34.

Beck R.M.D., Bininda-Emonds O.R.P., Cardillo M., Liu F.G.R., Purvis A. 2006. A higher-level MRP supertree of placental mammals. BMC Evol. Biol. 6:93.

Bininda-Emonds O.R.P. 2003. Novel versus unsupported clades: assessing the qualitative support for clades in MRP supertrees. Syst. Biol. 52:839–848.

Bininda-Emonds O.R.P. 2004a. The evolution of supertrees. Trends Ecol. Evol. 19:315–322.

Bininda-Emonds O.R.P. 2004b. Trees versus characters and the supertree/supermatrix "paradox." Syst. Biol. 53:356–359.

Bininda-Emonds O.R.P., Bryant H.N. 1998. Properties of matrix representation with parsimony analyses. Syst. Biol. 47:497–508.

Bininda-Emonds O.R.P., Gittleman J.L., Steel M.A. 2002. The (super)tree of life: procedures, problems, and prospects. Annu. Rev. Ecol. Syst. 33:265–289.

Bininda-Emonds O.R.P., Jones K.E., Cardillo S.A.M., Grenyer R., Purvis A. 2004. Garbage in, garbage out: data issues in supertree construction. In: Bininda-Emonds O.R.P., editor. Phylogenetic

supertrees: combining information to reveal the tree of life. Dordrecht (the Netherlands): Kluwer. p. 267–280.

Bininda-Emonds O.R.P., Sanderson M.J. 2001. Assessment of the accuracy of matrix representation with parsimony analysis supertree construction. Syst. Biol. 50:565–579.

Bremer K. 2000. Early Cretaceous lineages of monocot flowering plants. Proc. Natl. Acad. Sci. USA. 97:4707–4711.

Bremer K., Friis E.M., Bremer B. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. Syst. Biol. 53:496–505.

Bryant H.N. 2004. The cladistics of matrix representation with parsimony analysis. In: Bininda-Emonds O.R.P., editor. Phylogenetic supertrees: combining information to reveal the tree of life. Dordrecht (the Netherlands): Kluwer. p. 353–368.

Burleigh J.G., Driskell A.C., Sanderson M.J. 2006. Supertree bootstrapping methods for assessing phylogenetic variation among genes in genome-scale data sets. Syst. Biol. 55:426–440.

Chase M.W., Fay M.F., Devey D.S., Maurin O., Ronsted N., Davies T.J., Pillon Y., Peterson G., Seberg O., Tamura M.N., Asmussen C.B., Hilu K., Borsch T., Davis J.I., Stevenson D.W., Pires J.C., Givnish T.J., Sytsma K.J., McPherson M.A., Graham S.W., Rai H.S. 2006. Multigene analyses of monocot relationships: a summary. Aliso. 22:63–75.

Creevey C.J., Fitzpatrick D.A., Philip G.K., Kinsella R.J., O'Connell M.J., Pentony M.M., Travers S.A., Wilkinson M., McInerney J.O. 2004. Does a tree-like phylogeny only exist at the tips in the prokaryotes? Proc. R. Soc. Lond. Ser. B Biol. Sci. 271:2551–2558.

Davies T.J., Barraclough T.G., Chase M.W., Soltis P.S., Soltis D.E., Savolainen V. 2004. Darwin's abominable mystery: insights from a supertree of the angiosperms. Proc. Natl. Acad. Sci. USA. 101:1904–1909.

de Queiroz A., Gatesy J. 2007. The supermatrix approach to systematics. Trends Ecol. Evol. 22:34–41.

Dransfield J., Rakotoarinivo M., Baker W.J., Bayton R.P., Fisher J.B., Horn J.W., Leroy B., Metz X. 2008. A new coryphoid palm genus from Madagascar. Bot. J. Linn. Soc. 156:79–91.

Dransfield J., Uhl N.W., Asmussen C.B., Baker W.J., Harley M.M., Lewis C.E. 2005. A new phylogenetic classification of the palm family, Arecaceae. Kew Bull. 60:559–569.

Dransfield J., Uhl N.W., Asmussen C.B., Baker W.J., Harley M.M., Lewis C.E. 2008. Genera *Palmarum*—the evolution and classification of palms. Richmond (UK): Royal Botanic Gardens, Kew.

Fritsch P.W., Cruz B.C., Almeda F., Wang Y.G., Shi S.H. 2006. Phylogeny of *Symplocos* based on DNA sequences of the chloroplast *trn*C-*trn*D intergenic region. Syst. Bot. 31:181–192.

Fulton T.L., Strobeck C. 2006. Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. Mol. Phylogenet. Evol. 41:165–181.

Gatesy J., Baker R.H., Hayashi C. 2004. Inconsistencies in arguments for the supertree approach: supermatrices versus supertrees of Crocodylia. Syst. Biol. 53:342–355.

Gatesy J., Matthee C., DeSalle R., Hayashi C. 2002. Resolution of a supertree/supermatrix paradox. Syst. Biol. 51:652–664.

Gatesy J., Springer M.S. 2004. A critique of matrix representation with parsimony supertrees. In: Bininda-Emonds O.R.P., editor. Phylogenetic supertrees: combining information to reveal the tree of life. Dordrecht (the Netherlands): Kluwer. p. 369–388.

Govaerts R., Dransfield J. 2005. World checklist of palms. Richmond (UK): Royal Botanic Gardens, Kew.

Gunn B.F. 2004. The phylogeny of the Cocoeae (Arecaceae) with emphasis on *Cocos nucifera*. Ann. Mo. Bot. Gard. 91:505–522.

Hahn W.J. 2002a. A molecular phylogenetic study of the Palmae (Arecaceae) based on *atp*B, *rbc*L, and 18S nrDNA sequences. Syst. Biol. 51:92–112.

Hahn W.J. 2002b. A phylogenetic analysis of the Arecoid line of palms based on plastid DNA sequence data. Mol. Phylogenet. Evol. 23:189–204.

Harley M.M. 2006. A summary of fossil records for Arecaceae. Bot. J. Linn. Soc. 151:39–67.

Hughes J., Longhorn S.J., Papadopoulou A., Theodorides K., de Riva A., Mejia-Chang M., Foster P.G., Vogler A.P. 2006. Dense taxonomic

EST sampling and its applications for molecular systematics of the Coleoptera (beetles). Mol. Biol. Evol. 23:268–278.

Kennedy M., Page R.D.M. 2002. Seabird supertrees: combining partial estimates of procellariiform phylogeny. Auk. 119:88–108.

Kluge A.G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *epicrates* (Boidae, Serpentes). Syst. Zool. 38:7–25.

Levasseur C., Lapointe F.J. 2001. War and peace in phylogenetics: a rejoinder on total evidence and consensus. Syst. Biol. 50:881–891.

Levasseur C., Lapointe F.J. 2003. Increasing phylogenetic accuracy with global congruence. In: Janowitz M., Lapointe F.-J., McMorris F.R., Mirkin B., Roberts F.S., editors. Bioconsensus. Providence (RI): American Mathematical Society. p. 221–230.

Lewis C.E., Doyle J.J. 2001. Phylogenetic utility of the nuclear gene malate synthase in the palm family (Arecaceae). Mol. Phylogenet. Evol. 19:409–420.

Lewis C.E., Doyle J.J. 2002. A phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes. Plant Syst. Evol. 236:1–17.

Lewis C.E., Zona S. 2008. *Leucothrinax morrisii*, a new name for a familiar Caribbean palm. Palms. 52:84–88.

Loo A.H.B., Dransfield J., Chase M.W., Baker W.J. 2006. Low-copy nuclear DNA, phylogeny and the evolution of dichogamy in the betel nut palms and their relatives (Arecinae; Arecaceae). Mol. Phylogenet. Evol. 39:598–618.

Moore B.R., Smith S.A., Donoghue M.J. 2006. Increasing data transparency and estimating phylogenetic uncertainty in supertrees: approaches using nonparametric bootstrapping. Syst. Biol. 55:662–676.

Nixon K.C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics. 15:407–414.

Norup M.V., Dransfield J., Chase M.W., Barfod A.S., Fernando E.S., Baker W.J. 2006. Homoplasious character combinations and generic delimitation: a case study from the Indo-Pacific arecoid palms (Arecaceae: Areceae). Am. J. Bot. 93:1065–1080.

Peters H.A., Pauw A., Silman M.R., Terborgh J.W. 2004. Falling palm fronds structure Amazonian rainforest sapling communities. Proc. R. Soc. Lond. Ser. B Biol. Sci. 271:S367–S369.

Pisani D., Wilkinson M. 2002. MRP, taxonomic congruence and total evidence. Syst. Biol. 51:151–155.

Price S.A., Bininda-Emonds O.R.P., Gittleman A.L. 2005. A complete phylogeny of the whales, dolphins and even-toed hoofed mammals (Cetartiodactyla). Biol. Rev. 80:445–473.

Purvis A. 1995. A modification to Baum and Ragans method for combining phylogenetic trees. Syst. Biol. 44:251–255.

Ragan M.A. 1992a. Matrix representation in reconstructing phylogenetic relationships among the eukaryotes. Biosystems. 28:47–55.

Ragan M.A. 1992b. Phylogenetic inference based on matrix representation of trees. Mol. Phylogenet. Evol. 1:53–58.

Rodrigo A.G. 1993. A comment on Baum's method for combining phylogenetic trees. Taxon. 42:631–636.

Roncal J., Francisco-Ortega J., Asmussen C.B., Lewis C.E. 2005. Molecular phylogenetics of tribe Geonomeae (Arecaceae) using nuclear DNA sequences of phosphoribulokinase and RNA polymerase II. Syst. Bot. 30:275–283.

Ronquist F. 1996. Matrix representation of trees, redundancy, and weighting. Syst. Biol. 45:247–253.

Salamin N., Chase M.W., Hodkinson T.R., Savolainen V. 2003. Assessing internal support with large phylogenetic DNA matrices. Mol. Phylogenet. Evol. 27:528–539.

Salamin N., Hodkinson T.R., Savolainen V. 2002. Building supertrees: an empirical assessment using the grass family (Poaceae). Syst. Biol. 51:112–126.

Sanderson M.J., Purvis A., Henze C. 1998. Phylogenetic supertrees: assembling the trees of life. Trends Ecol. Evol. 13:105–109.

Savolainen V., Anstett M.C., Lexer C., Hutton I., Clarkson J.J., Norup M.V., Powell M.P., Springate D., Salamin N., Baker W.J. 2006. Sympatric speciation in palms on an oceanic island. Nature. 441:210–213.

Sikes D.S., Lewis P.O. 2001. PAUPRat: PAUP* implementation of the parsimony ratchet. Beta version 1. Available from: http://users.iab.uaf.edu/~derek_sikes/software2.htm.

Swofford D.L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland (MA): Sinauer.

Thomas M.M., Garwood N.C., Baker W.J., Henderson S.A., Russell S.J., Hodel D.R., Bateman R.M. 2006. Molecular phylogeny of the palm genus *Chamaedorea*, based on the low-copy nuclear genes PRK and RPB2. Mol. Phylogenet. Evol. 38:398–415.

Uhl N.W., Dransfield J., Davis J.I., Luckow M.A., Hansen K.S., Doyle J.J. 1995. Phylogenetic relationships among palms: cladistic analyses of morphological and chloroplast DNA restriction site variation. In: Rudall P.J., Cribb P.J., Cutler D.F., Humphries C.J., editors. Monocotyledons: systematics and evolution. Richmond (UK): Royal Botanic Gardens, Kew. p. 623–661.

Wiens J.J. 1998. Does adding characters with missing data increase or decrease phylogenetic accuracy? Syst. Biol. 47:625–640.

Wiens J.J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52:528–538.

Wiens J.J. 2006. Missing data and the design of phylogenetic analyses. J. Biomed. Inform. 39:34–42.

Wilkinson M., Cotton J.A. 2007. Supertree methods for building the tree of life: divide-and-conquer approaches to large phylogenetic problems. In: Hodkinson T.R., Parnell J.A.N., editors. Reconstructing the tree of life—taxonomy and systematics of species rich taxa. Boca Raton (FL): CRC Press. p. 61–75.

Wilkinson M., Cotton J.A., Creevey C., Eulenstein O., Harris S.R., Lapointe F.J., Levasseur C., McInerney J.O., Pisani D., Thorley J.L. 2005. The shape of supertrees to come: tree shape related properties of fourteen supertree methods. Syst. Biol. 54:419–431.

Wilkinson M., Cotton J.A., Lapointe F.J., Pisani D. 2007. Properties of supertree methods in the consensus setting. Syst. Biol. 56:330–337.

Wilkinson M., Pisani D., Cotton J.A., Corfe I. 2005. Measuring support and finding unsupported relationships in supertrees. Syst. Biol. 54:823–831.

Wilkinson M., Thorley J.L., Pisani D., Lapointe F.J., McInerney J.O. 2004. Some desiderata for liberal supertrees. In: Bininda-Emonds O.R.P., editor. Phylogenetic supertrees: combining information to reveal the tree of life. Dordrecht (the Netherlands): Kluwer. p. 227–246.

APPENDIX 1

Support measures for each clade in the most congruent supertree (Fig. 3). BPs from the supermatrix tree (Fig. 2) are also given for all shared clades. An asterisk indicates a clade shared by the 2 trees, but with BP <50%. An open circle indicates a clade that is absent from the supermatrix tree, but compatible with it, whereas a line indicates a clade that is incongruent with the supermatrix tree. Support measure definitions: s = number of input trees supporting a supertree clade, q = number of input trees conflicting with a supertree clade, p = number of input trees permitting a supertree clade, r = number of input trees irrelevant to a supertree clade, V = $(s - q)/(s + q)$ (Wilkinson, Pisani, et al. 2005), and rQS = $(s - q)/(s + q + p + r)$ (Price et al. 2005; Beck et al. 2006).

Clade	s	q	r	p	V	rQS	BP
1	6	3	19	1	0.333	0.103	70
2	6	0	22	1	1.000	0.207	92
3	6	1	22	0	0.714	0.172	99
4	3	5	19	2	−0.250	−0.069	—
5	8	2	19	0	0.600	0.207	95
6	3	3	22	1	0.000	0.000	62
7	2	3	22	2	−0.200	−0.034	85
8	2	3	22	2	−0.200	−0.034	76
9	2	4	22	1	−0.333	−0.069	60
10	1	5	22	1	−0.667	−0.138	*
11	2	4	22	1	−0.333	−0.069	79
12	5	2	22	0	0.429	0.103	78

Continued.

Clade	s	q	r	p	V	rQS	BP
13	1	4	19	5	−0.600	−0.103	—
14	4	6	15	4	−0.200	−0.069	—
15	9	5	11	4	0.286	0.138	67
16	3	3	22	1	0.000	0.000	69
17	7	8	11	3	−0.067	−0.034	—
18	8	7	11	3	0.067	0.034	77
19	3	5	19	2	−0.250	−0.069	—
20	11	6	4	8	0.294	0.172	—
21	4	5	12	8	−0.111	−0.034	—
22	1	7	12	9	−0.750	−0.207	—
23	6	8	12	3	−0.143	−0.069	100
24	3	5	12	9	−0.250	−0.069	90
25	5	2	19	3	0.429	0.103	100
26	7	5	12	5	0.167	0.069	100
27	3	3	17	6	0.000	0.000	—
28	4	1	24	0	0.600	0.103	88
29	1	10	12	6	−0.818	−0.310	—
30	2	1	24	2	0.333	0.034	52
31	3	10	10	6	−0.538	−0.241	*
32	7	9	8	5	−0.125	−0.069	*
33	4	1	24	0	0.600	0.103	100
34	7	8	7	7	−0.067	−0.034	63
35	12	6	7	4	0.333	0.207	68
36	12	6	7	4	0.333	0.207	81
37	7	6	5	11	0.077	0.034	98
38	12	5	9	3	0.412	0.241	100
39	8	8	5	8	0.000	0.000	95
40	3	13	5	8	−0.625	−0.345	80
41	2	2	23	2	0.000	0.000	—
42	12	1	10	6	0.846	0.379	—
43	8	4	17	0	0.333	0.138	100
44	6	4	10	9	0.200	0.069	—
45	5	3	8	13	0.250	0.069	—
46	5	3	6	15	0.250	0.069	○
47	3	0	23	3	1.000	0.103	89
48	6	2	13	8	0.500	0.138	59
49	1	5	23	0	−0.667	−0.138	—
50	6	0	17	6	1.000	0.207	92
51	7	2	17	3	0.556	0.172	65
52	5	5	17	2	0.000	0.000	○
53	2	5	17	5	−0.429	−0.103	*
54	4	5	13	7	−0.111	−0.034	54
55	3	4	17	5	−0.143	−0.034	73
56	3	7	12	7	−0.400	−0.138	64
57	14	2	6	7	0.750	0.414	100
58	2	7	6	14	−0.556	−0.172	○
59	1	0	23	5	1.000	0.034	—
60	1	3	17	8	−0.500	−0.069	—
61	4	1	17	7	0.600	0.103	*
62	8	1	17	3	0.778	0.241	88
63	2	3	15	9	−0.200	−0.034	○
64	2	3	14	10	−0.200	−0.034	○
65	2	4	12	11	−0.333	−0.069	○
66	1	4	12	12	−0.600	−0.103	○
67	0	8	6	15	−1.000	−0.276	○
68	1	10	6	12	−0.818	−0.310	○
69	6	6	8	9	0.000	0.000	○
70	6	1	17	5	0.714	0.172	79
71	8	0	17	4	1.000	0.276	100
72	3	2	17	7	0.200	0.034	94
73	3	3	17	6	0.000	0.000	85
74	5	4	15	5	0.111	0.034	94
75	6	4	17	2	0.200	0.069	100
76	1	5	15	8	−0.667	−0.138	○
77	2	11	6	10	−0.692	−0.310	○

Clade	s	q	r	p	V	rQS	BP
78	2	10	6	11	−0.667	−0.276	○
79	2	4	23	0	−0.333	−0.069	73
80	5	4	10	10	0.111	0.034	71
81	2	15	6	6	−0.765	−0.448	*
82	8	2	15	4	0.600	0.207	91
83	8	3	15	3	0.455	0.172	100
84	10	1	15	3	0.818	0.310	100
85	3	5	15	6	−0.250	−0.069	67
86	1	13	6	9	−0.857	−0.414	—
87	3	16	6	4	−0.684	−0.448	*
88	2	9	15	3	−0.636	−0.241	52
89	1	17	6	5	−0.889	−0.552	—
90	1	16	6	6	−0.882	−0.517	—
91	2	13	6	8	−0.733	−0.379	—
92	4	6	15	4	−0.200	−0.069	74
93	2	8	15	4	−0.600	−0.207	*
94	4	12	4	9	−0.500	−0.276	—
95	7	12	4	6	−0.263	−0.172	—
96	5	0	22	2	1.000	0.172	99
97	5	0	22	2	1.000	0.172	99
98	5	5	15	4	0.000	0.000	93
99	7	13	4	5	−0.300	−0.207	—
100	10	10	4	5	0.000	0.000	—
101	9	3	15	2	0.500	0.207	99
102	10	11	4	4	−0.048	−0.034	84
103	2	0	26	1	1.000	0.069	76
104	4	2	14	9	0.333	0.069	66
105	6	5	12	6	0.091	0.034	73
106	5	6	12	6	−0.091	−0.034	54
107	8	13	4	4	−0.238	−0.172	—
108	2	5	17	5	−0.429	−0.103	72
109	7	4	17	1	0.273	0.103	100
110	3	3	17	6	0.000	0.000	88
111	10	5	8	6	0.333	0.172	83
112	12	7	6	4	0.263	0.172	92
113	5	12	4	8	−0.412	−0.241	*
114	4	11	4	10	−0.467	−0.241	*
115	3	14	2	10	−0.647	−0.379	—
116	4	4	15	6	0.000	0.000	72
117	7	12	4	6	−0.263	−0.172	98
118	10	11	2	6	−0.048	−0.034	98
119	2	19	2	6	−0.810	−0.586	70
120	3	14	5	7	−0.647	−0.379	82
121	5	7	13	4	−0.167	−0.069	64
122	4	3	19	3	0.143	0.034	—
123	9	9	6	5	0.000	0.000	—
124	19	4	6	0	0.652	0.517	100
125	1	18	2	8	−0.895	−0.586	74
126	2	3	21	3	−0.200	−0.034	*
127	2	4	17	6	−0.333	−0.069	—
128	6	3	11	9	0.333	0.103	—
129	18	3	6	2	0.714	0.517	100
130	7	14	2	6	−0.333	−0.241	75
131	4	3	19	3	0.143	0.034	59
132	5	3	19	2	0.250	0.069	95
133	11	4	6	8	0.467	0.241	—
134	18	2	6	3	0.800	0.552	100
135	12	1	15	1	0.846	0.379	100
136	10	8	4	7	0.111	0.069	95
137	6	11	4	8	−0.294	−0.172	100
138	8	13	2	6	−0.238	−0.172	93
139	3	0	21	5	1.000	0.103	58
140	5	4	19	1	0.111	0.034	80

Continued.

Clade	s	q	r	p	V	rQS	BP
141	2	4	19	4	−0.333	−0.069	○
142	3	1	21	4	0.500	0.069	68
143	1	3	21	4	−0.500	−0.069	○
144	4	2	19	4	0.333	0.069	71
145	6	0	19	4	1.000	0.207	○
146	9	1	15	4	0.800	0.276	100
147	4	4	15	6	0.000	0.000	—
148	14	0	15	0	1.000	0.483	100
149	2	2	19	6	0.000	0.000	*
150	7	3	19	0	0.400	0.138	100
151	13	8	4	4	0.238	0.172	*
152	3	4	19	3	−0.143	−0.034	*
153	5	5	19	0	0.000	0.000	*
154	2	6	15	6	−0.500	−0.138	*
155	4	2	19	4	0.333	0.069	94
156	4	2	19	4	0.333	0.069	—
157	2	5	19	3	−0.429	−0.103	*
158	4	4	15	6	0.000	0.000	51
159	6	2	15	6	0.500	0.138	94
160	2	5	15	7	−0.429	−0.103	—
161	6	3	15	5	0.333	0.103	60
162	3	5	19	2	−0.250	−0.069	52
163	4	4	19	2	0.000	0.000	76
164	14	5	8	2	0.474	0.310	100
165	5	0	21	3	1.000	0.172	79
166	4	1	19	5	0.600	0.103	*
167	5	0	19	5	1.000	0.172	95
168	4	1	19	5	0.600	0.103	52
169	5	4	15	5	0.111	0.034	58
170	1	5	19	4	−0.667	−0.138	○
171	5	4	15	5	0.111	0.034	*
172	10	1	15	3	0.818	0.310	100
173	7	5	15	2	0.167	0.069	*
174	6	8	6	9	−0.143	−0.069	—
175	4	10	6	9	−0.429	−0.207	*
176	16	6	4	3	0.455	0.345	86
177	14	8	2	5	0.273	0.207	100
178	15	10	0	4	0.200	0.172	97
179	17	6	0	6	0.478	0.379	100
180	15	6	0	8	0.429	0.310	62
181	16	1	4	8	0.882	0.517	100
182	5	0	23	1	1.000	0.172	—
183	5	0	24	0	1.000	0.172	100
184	5	0	24	0	1.000	0.172	100
185	5	6	16	2	−0.091	−0.034	79
186	8	4	16	1	0.333	0.138	100
187	9	0	19	1	1.000	0.310	○
188	8	2	18	1	0.600	0.207	98
189	6	4	18	1	0.200	0.069	86
190	8	0	21	0	1.000	0.276	100
191	7	4	17	1	0.273	0.103	100
192	4	7	16	2	−0.273	−0.103	○
193	29	0	0	0	1.000	1.000	100