Are mitochondrial genes useful for the analysis of monocot relationships?

Jerrold I. Davis[1], Gitte Petersen[2], Ole Seberg[2], Dennis W. Stevenson[3], Christopher R. Hardy[4], Mark P. Simmons[5], Fabian A. Michelangeli[3], Douglas H. Goldman[6], Lisa M. Campbell[3], Chelsea D. Specht[7] & James I. Cohen[1]

[1] *L. H. Bailey Hortorium and Department of Plant Biology, Cornell University, Ithaca, New York 14853, U.S.A. jid1@cornell.edu (author for correspondence).*

[2] *Botanical Garden and Museum, Natural History Museum of Denmark, Sølvgade 83, Opg. S, DK-1307 Copenhagen K, Denmark.*

[3] *Institute of Systematic Botany, New York Botanical Garden, Bronx, New York 10458, U.S.A.*

[4] *Department of Biology, Millersville University, Millersville, Pennsylvania 17551, U.S.A.*

[5] *Department of Biology, Colorado State University, Fort Collins, Colorado 80523, U.S.A.*

[6] *Harvard University Herbaria, Harvard University, Cambridge, Massachusetts 02138, U.S.A.*

[7] *Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, U.S.A.*

A phylogenetic analysis of monocots and related dicots was conducted, using a four-gene matrix consisting of two genes from the plastid genome (*matK* and *rbcL*) and two from the mitochondrial genome (*atpA/atp1* and *cob*). The taxon sample includes 101 monocots and 36 dicots, and all four genes were sampled for all 137 taxa. Jackknife support was assessed for clades resolved by the four-gene analysis, and compared to support for the same clades by each of the four three-gene subset matrices, in order to quantify the degree to which each gene contributed to or detracted from support for each clade. Instances of positively and negatively correlated support for clades by genes of the same and different genomes were observed. In particular, the placement of *Acorus* within a clade that also includes Tofieldiaceae, Araceae, and Alismatales s.s., as opposed to its frequent placement as sister of all other monocots, is supported by *atpA* and *matK*. The results indicate that genes from the mitochondrial genome provide a unique test of relationships that have been inferred with plastid-encoded genes.

KEYWORDS: *atpA*, *cob*, mitochondrial genes, *matK*, monocots, phylogenetics, plastid genes, *rbcL*.

INTRODUCTION

The plastid genome, as sampled by nucleotide sequencing of genes and other regions, has been used more extensively than any other genome in the study of plant phylogenetic relationships (e.g., Palmer & al., 2004; and associated review articles). Although the major initial thrust of these studies was in the area of higher-level relationships (generally at the level of the genus and above), predominantly using *rbcL* and other protein-encoding genes of the plastid genome, many studies also have utilized noncoding regions of the plastid genome, usually for the analysis of lower-level relationships (e.g., Shaw & al., 2005; and citations within). More than a decade ago, Palmer (1992) discussed the potential utility of mitochondrial genes for plant phylogenetic purposes, but in the intervening period this genome has been used relatively infrequently, while the use of low-copy-number nuclear genes, which was rare at that time, has increased dramatically. This is not to say that mitochondrial genes have been ignored. The utility of mitochondrial genes for phylogenetic studies of achlorophyllous plants has been demonstrated in groups such as the Rafflesiales (Nickrent & al., 2004), and multigene studies often include one or more mitochondrial genes (e.g., Qiu & al., 2006). The literature on the use of mitochondrial DNA in plant systematics, and on challenges associated with its use (such as RNA editing), are reviewed elsewhere in this symposium (Petersen & al., 2006; Qiu & al., 2006), and readers are referred to those contributions for further information. In the present paper we examine the contribution of individual mitochondrial genes to the support of particular clades, in the context of a multi-gene analysis that also includes widely used genes of the plastid genome, with the goal of determining whether the mitochondrial genes contribute significantly to the results. Further, if they do contribute significantly, it is useful to determine whether this contribution takes the form of generalized support or opposition to the relationships supported by the plastid genome, or is a mixture of these two.

The power and utility of various genes and other character sets for phylogenetic analysis are difficult to assess, in part because true historical descent relationships among the terminals (i.e., the taxa) are not known for real data matrices of nontrivial size and complexity.

Discussion of these matters often is anecdotal, where it takes the form of ad hoc descriptions of "good" and "bad" groups supported by a particular character set (often a single gene). Incongruence among data sets sometimes is regarded as a nuisance, but it is precisely this quality that allows investigators to discover errors in earlier work, and to focus future studies on groups for which support exists for conflicting hypotheses. Indeed, progress in science is precisely a matter of disproving hypotheses, and it should be borne in mind that conflict between existing hypotheses and new evidence is the means by which this progress occurs (Popper, 1959).

The present authors, and associates, have conducted a series of studies of relationships among monocots, using the plastid-encoded gene *rbcL* and the mitochondrial-encoded gene *atpA*, with an array of magnoliid dicots and other taxa as outgroups (Davis & al., 1998, 2004; Stevenson & al., 2000; Michelangeli & al., 2003). Apart from the phylogenetic problems themselves, we have examined intrinsic features of the data sets such as data decisiveness (Goloboff, 1991; Davis & al., 1998) and conflicting patterns of support for alternative groupings, as assessed by bootstrap (Felsenstein, 1985) and jackknife (Farris & al., 1996) analyses of subsets of the data (Davis & al., 2004). In the present contribution we add one gene from each of these genomes (*cob* from the mitochondrial genome, and *matK* from the plastid genome), and examine patterns of congruence and conflict among the four genes in the context of an analysis of relationships among 137 taxa, including representatives of all major monocot lineages and a range of dicot outgroups. The previous analyses were conducted using parsimony as an optimality criterion (Farris, 1983), and patterns of agreement and disagreement between the present results and those of our earlier contributions are facilitated by continuing to use the same criterion. The present results indicate that patterns of incongruence among the four genes are complex. Two genes from the same genome often support similar relationships, but the same often is true for genes from different genomes, while genes from the same genome often conflict with each other. In light of these results, several aspects of higher-level relationship within the monocots remain controversial and deserving of continued study.

In the evaluation of monocot relationships, we focus on a set of 15 mutually exclusive lineages within the monocots, as discussed by Davis & al. (2004) in their review of monocot relationships; these groups are identified in Figs. 1–4 of the present contribution. All 15 of these groups have been resolved as monophyletic by most recent analyses in which they have been sampled in reasonable depth, and each of the groups therefore is regarded as likely to be monophyletic (though there can be no certainty on this matter). Controversy and considerable uncertainty exist with respect to relationships among these lineages, and the present focus is on these relationships.

MATERIALS AND METHODS

Taxon sampling. — The taxon set includes 137 terminals (101 monocots and 36 dicots; Appendix), representing a subset of 136 of the 218 terminals examined by Davis & al. (2004), plus *Chlorophytum*. All 15 major lineages of monocots identified by Davis & al. (2004) as groups of interest in the analysis of higher-level monocot relationships are represented, and the dicot sample includes representatives of early-diverging angiosperm lineages (e.g., *Amborella*, Nymphaeaceae, Illiciaceae), plus early-diverging lineages within the tricolpate dicots (e.g., *Platanus*, *Nelumbo*, Berberidaceae), and various additional "magnoliid" lineages for which there is evidence of a close affinity with monocots (e.g., Piperales [including Aristolochiales], Laurales, Magnoliales, and Canellales). To facilitate comparisons with previous classifications, provisional assignments of genera to families follow Kubitzki & al. (1993) and Kubitzki (1998a, b), except as noted, and provisional assignments of families to orders and other higher-level groupings follow the revised classification of the Angiosperm Phylogeny Group (i.e., the APG II system; Angiosperm Phylogeny Group, 2003). Kubitzki's taxonomic system is comprehensive for monocot genera (except for Poaceae and Orchidaceae). The assignment here of each genus to a family differs from the treatment by Kubitzki only with respect to the circumscription of Nartheciaceae, which, as treated in that work by Tamura (1998), includes six genera in the present analysis. Two of these six genera (*Aletris* and *Narthecium*) are recognized here as elements of Nartheciaceae, and the remaining four are assigned to two other families (*Pleea* and *Tofieldia* to Tofieldiaceae, *Japonolirion* and *Petrosavia* to Petrosaviaceae; cf. Chase & al., 2000; Fuse & Tamura, 2000; Soltis & al., 2000; Hilu & al., 2003; Davis & al., 2004). The data matrix includes a sequence of each of the four genes for every terminal. This was achieved by excluding achlorophyllous taxa, and in several cases by assembling "composite" taxa (Nixon & Davis, 1991) in which the available gene sequences for individual terminals in the matrix represent different species, usually from a single genus, but in a few cases from different genera (Appendix). In each of these cases, the species represented by each gene is believed to be more closely related to the species represented by the other three genes than to any other taxon in the matrix, though this cannot be guaranteed.

Molecular methods. — DNA sequence variation was examined at four protein-encoding loci, including

two from the plastid genome (*rbcL* and *matK*, encoding the large subunit of ribulose 1,5-bisphosphate carboxylase, and maturase K, respectively) and two from the mitochondrial genome (*atpA* and *cob*, encoding the alpha subunit of F-1-ATPase and apocytochrome b, respectively). The sequences were generated by the authors from total genomic DNA isolations, following standard PCR and automated cycle sequencing protocols, or obtained either from GenBank (where some *atpA* sequences are listed as *atp1*) or from other investigators. The set of *atpA* and *rbcL* sequences used in the present study is a subset of those analyzed by Davis & al. (2004), except for those of *Chlorophytum* (for which the *rbcL* sequence was taken from GenBank, and the *atpA* sequence was determined by the authors, using the same primers that were used by Davis & al., 2004). All of the *cob* sequences, except that of *Oryza*, which was obtained from GenBank, were generated by the authors for the present analysis, using the primers described by Petersen & al. (2006). Most of the *matK* sequences were gathered from GenBank, the exceptions being the three sequences from Dasypogonaceae, which were generated using the primers described by Molvray & al. (2000), Cuénoud & al. (2002), and Barfuss & al. (2005). In some cases (*cob* only), where PCR amplification yielded only faint bands, PCR-generated DNA fragments were cloned with an Invitrogen TOPO TA Cloning® Kit (Invitrogen Corporation, Carlsbad, California). Fragments were ligated into the pCR® 2.1-TOPO® vector and introduced into chemically competent *Escherichia coli* cells of strain DH5a T1®. Plasmid DNA then was extracted using a QIAprep® Spin Miniprep Kit (Qiagen Inc., Valencia, California) and sequenced, either with the original amplification primers or with the M13 plasmid primers supplied with the cloning kit.

The portion of each gene included in the matrix, and reference sequences in GenBank for the designation of nucleotide positions, are as follows: *rbcL* (1,371 nucleotides, sites 31–1,401 in *Nicotiana tabacum*, NC001879); *cob* (1,043 nucleotides, sites 41–1,083 in *Oryza sativa*, BA000029); *atpA* (1,259 nucleotides, sites 98–1,356 in *Oryza sativa*, AB076666); and *matK* (1,394 nucleotides, sites 120–1,513 in *Oryza nivara*, NC_005973). Manual alignment, by introducing inferred insertion/deletion regions, increased the number of recognized nucleotide sites in *atpA* and *matK* to 1,277 and 1,652 sites, respectively (Table 1). However, some regions in each of these genes were determined to be alignable only ambiguously, and therefore were excluded from analyses. One region in *atpA* was excluded (corresponding to sites 581–604 in the reference sequence), and seven regions in *matK* were excluded (corresponding to sites 201–246, 316–371, 424–469, 534–544, 649–706, 785–796, and 863–904 in the reference sequence). Also, two *atpA* sites (220 and 255 in the reference sequence) were excluded because sequencing of the two DNA strands often yielded conflicting results at these sites. Artifacts of this sort, reflecting imperfections in sequencing accuracy, can occur when certain combinations of nucleotides lie in close proximity to each other within a sequence (Parker & al., 1995). As a result of these exclusions, the matrix used in the analyses included 1,233 *atpA* sites and 1,206 *matK* sites (Table 1). In addition to the nucleotide characters, two parsimony-informative insertion/deletion (indel) characters in *atpA* were included in the matrix (Davis & al., 2004). The four-gene data matrix used in the analysis was deposited in TreeBASE (study accession number S1631).

Data analysis. — All characters, including the two *atpA* indels, were weighted equally and treated as nonadditive (i.e., the states unordered; Fitch, 1971) for tree searches. Nine subsets of the complete four-gene data matrix were analyzed cladistically, these being the combined four-gene matrix, the four one-gene matrices, and the four three-gene matrices that are obtained by removing each of the four genes from the four-gene matrix. The

Table 1. Characteristics of four genes as sampled across 101 monocot and 36 dicot taxa. "Length of region" is the no. of nucleotides in the sequenced region in the reference sequence for each gene (see text), prior to alignment. For *atpA* and the combined matrix that includes it, "+ 2" refers to the two informative insertion/deletion characters in the matrix. Tree lengths and consistency indices are calculated using matrices that include only parsimony informative characters.

Gene	*matK*	*rbcL*	*atpA*	*cob*	combined matrix
Length of region (nucleotides)	1,394	1,371	1,259	1,043	5,067
Number of aligned (and unambiguously aligned) nucleotides	1,652 (1,206)	1,371 (1,371)	1,277 (1,233)	1,043 (1,043)	5,343 (4,853)
Number (and %) of unambiguously aligned characters that are parsimony informative	823 (68%)	523 (38%)	359 (29%) + 2	211 (20%)	1,916 (40%) + 2
Number of steps (and CI and RI) on trees from single-gene and combined analyses	6,896 (0.24, 0.59)	4,177 (0.21, 0.56)	1,670 (0.36, 0.70)	929 (0.33, 0.66)	13,934 (0.25, 0.59)
Minimum number of steps, length increase in steps and % for single-gene matrices mapped on trees from combined analysis, relative to single-gene analyses	6,926, 30, 0.4%	4,227, 50, 1.2%	1,730, 60, 3.6%	1,030, 101, 10.9%	n.a., [1]262, 1.9%

[1]relative to sum of steps in four single-gene analyses (13,672).

simultaneous analysis of all four genes was conducted to provide our best estimate of phylogenetic relationships (Kluge, 1989; Nixon & Carpenter, 1996). The one-gene analyses were conducted to determine tree lengths and related characteristics of the individual genes. The three-gene analyses were conducted to determine the contribution of each gene to the overall level of support for clades resolved by the four-gene analysis. Support was measured by the parsimony jackknife (Farris & al., 1996), and the contribution of each gene to the support of each clade was determined as the degree to which support was greater or lesser in a matrix that differed from the complete matrix only by the exclusion of that gene. All analyses and all calculations of tree lengths, ensemble consistency and retention indices (Kluge & Farris, 1969; Farris, 1989), and clade support were conducted after removing parsimony uninformative characters from the data matrices. Parsimony searches of all nine data sets were conducted using conventional search strategies and the parsimony ratchet (Nixon, 1999). Searches were conducted with the multi-thread version of NONA vers. 1.6 (i.e., "PARANONA", compiled February 26, 1998; Goloboff, 1993), except as noted otherwise. Searches conducted with NONA used the default polytomy settings, which allow polytomies to occur (*poly*=), and which provide a resolution, rather than a polytomy, only when support for the resolution is unambiguous (*amb-*; i.e., ambiguous support is insufficient for resolution); the criterion of unambiguous support for a group is that the group's branch length is greater than zero under all possible character optimizations; conversely, support for a group is regarded as ambiguous when its minimum length is zero, and in these cases it is determined that the group is not resolved.

Conventional searches involved 1,000 individual subsearches, with each subsearch initiated by the construction of a Wagner tree, using a random-taxon-entry sequence, and with this tree then subjected to TBR swapping, with up to 20 shortest trees retained and subjected to additional branch swapping, using the command *mult**, preceded by *rs 0* and *hold/20*. All most-parsimonious trees accumulated by the 1,000 search initiations were pooled, and TBR swapping was conducted on these and all additional trees propagated during this phase of the search, with up to 100,000 trees retained and swapped, using the commands *hold 100000* and *max**.

Ratchet searches were conducted with WinClada vers. 1.00.08 (Nixon, 2002), with NONA invoked as a daughter process for cladistic analysis. Ten ratchet searches were conducted, each initiated with the generation of a Wagner tree, using a random-taxon-entry sequence, followed by TBR branch swapping with one tree retained (*rs 0*, *hold/1*, *mult*1*) and used as the starting point for 500 ratchet cycles. In the weighted/constrained half of each ratchet cycle a randomly selected set of 10% of the characters were resampled, and a randomly selected set of 10% of the resolved clades were constrained. All most-parsimonious trees accumulated by the 10 ratchet searches were pooled, and TBR swapping was conducted on these and all additional trees propagated during this phase of the search, with up to 100,000 trees retained and swapped.

Support for clades resolved by the four-gene analysis (see below) was assessed by strict consensus jackknife analyses (Davis & al., 2004) of the four-gene matrix and each of the three-gene matrices. Each of the 1,000 replicates within each jackknife analysis consisted of four subsearches, with up to 20 trees retained during TBR swapping after each search initiation (*hold/20*; *mult*4*), followed by additional TBR swapping of all shortest trees, including those generated during this phase of swapping, with up to 100 trees retained (*hold 100*; *max**). Character sampling for each replicate jackknife search was implemented in WinClada as described by Farris & al. (1996; also see Davis & al., 2004).

RESULTS

The number of nucleotides in the sequenced regions of the reference sequences ranged from 1,043 in *cob* to 1,394 in *matK* (Table 1). For *rbcL*, which had no inferred indels in the sequenced regions, and for *cob*, which had one inferred 3-nucleotide deletion in two taxa relative to the rest of the taxon set, the number of aligned sites was identical to the length of the sequenced region. For *atpA* and *matK*, which exhibit substantially more length variation, the number of aligned sites exceeded the number of nucleotides in the reference sequences, while the number of sites regarded as unambiguously alignable, and hence included in the analyses, was fewer. For *atpA*, 97% of the aligned sites were determined to be sufficiently unambiguously aligned to be included in the analysis (i.e., 1,233 of 1,277 sites; Table 1), while 73% of the aligned sites of *matK* were included (1,206 of 1,652 sites). As a consequence of these factors (length before and after alignment, and number of sites deemed unambiguously aligned), *rbcL* provided 31% more unambiguously aligned nucleotides than *cob* (1,371 vs. 1,043), with *matK* and *atpA* providing intermediate numbers. Of the available characters, 20% of those from *cob* and 68% of those from *matK* are parsimony informative, with intermediate percentages of the available characters informative for *atpA* and *rbcL* (Table 1). Thus, *matK* provides 823 unambiguously aligned and informative sites, *cob* provides about one-fourth of this number (211), and *atpA* and *rbcL* provide intermediate numbers, for a total of 1,918 informative characters (including the two *atpA*

indel characters) in the four-gene matrix. Tree lengths for the single-gene analyses range from 929 for *cob* to 6,896 for *matK* (Table 1).

The two mitochondrial genes, when analyzed individually, have CIs greater than 0.30 and RIs greater than 0.65, while the CIs and RIs of the two plastid genes are less than 0.25 and 0.60, respectively (Table 1). Analysis of the four-gene matrix yielded 48 equally parsimonious trees of length 13,934, CI 0.25, and RI 0.59. Each of the four genes has a range of lengths when mapped on these trees. Using the minimum length for each gene (the sum of which is 13,913 steps, 21 steps fewer than the actual number of steps in each of the 48 trees), the number of steps in *matK* is 0.4% greater on the four-gene trees than on the *matK*-only trees, and the corresponding increases for *rbcL*, *atpA*, and *cob* are 1.2%, 3.6%, and 10.9%, respectively (Table 1).

Phylogenetic results. — Fifteen major lineages of monocots were identified by Davis & al. (2004), and summary relationships among these groups are depicted in Fig. 1, as resolved in the consensus of the 48 most-parsimonious trees obtained by the present four-gene analysis, and by the two-gene 215-taxon analysis of Davis & al. As in the previous analysis, the deepest branch within the monocots, as resolved by the present analysis, is between a clade that consists of four of these lineages (Araceae, Tofieldiaceae, Acorales [i.e., *Acorus*] and Alismatales s.s.) and another that consists of the other 11 lineages. However, within the clade of four lineages Araceae is not placed as the sister of the other three (as in the two-gene analysis). Instead, there is a trichotomy from which Araceae, Tofieldiaceae, and *Acorus* + Alismatales s.s. diverge. This consensus reflects the occurrence of two basic structures among the 48 trees. In some of the trees Tofieldiaceae is sister of a clade that includes the other three groups, with Araceae sister of *Acorus* + Alismatales s.s., and in the other trees the positions of Tofieldiaceae and Araceae are reversed.

Within the group of 11 lineages, Petrosaviaceae is the sister of the remaining ten, as in the two-gene analysis (Fig. 1). Liliales diverge next, from the group that includes the remaining nine lineages, and in this respect the present analysis differs from the previous one, in which Dioscoreales were the next group to diverge, though jackknife support for this structure (in the form of support for the clade that included the remaining nine lineages) was minimal, at 3%. Also, among the remaining nine lineages, a clade consisting of Liliales, Nartheciaceae, and Pandanales was resolved by the two-gene analysis as the sister of Asparagales + commelinids s.l., with jackknife support of just 5%, while the present analysis places a group consisting of Pandanales, Nartheciaceae, and Dioscoreales s.l. as the sister of Asparagales + commelinids s.l. This group of three lineages has jackknife support of 83%, with 89% support for a nested clade consisting of Nartheciaceae and Dioscoreales s.l. The alternative groupings in the two-gene analysis had less support.

Among the six remaining groups, Asparagales are sister of the commelinids s.l. (which include the remaining five lineages), as in the two-gene analysis (Fig. 1), but support for the commelinids s.l. is 91%, as opposed to 36% in the previous analysis, and relationships within this group are substantially different. Among the five groups within the commelinids s.l., the only major relationship that remains unchanged is the placement of Zingiberales and Commelinales as sisters. In the two-gene analysis, a pectinate structure existed at the base of

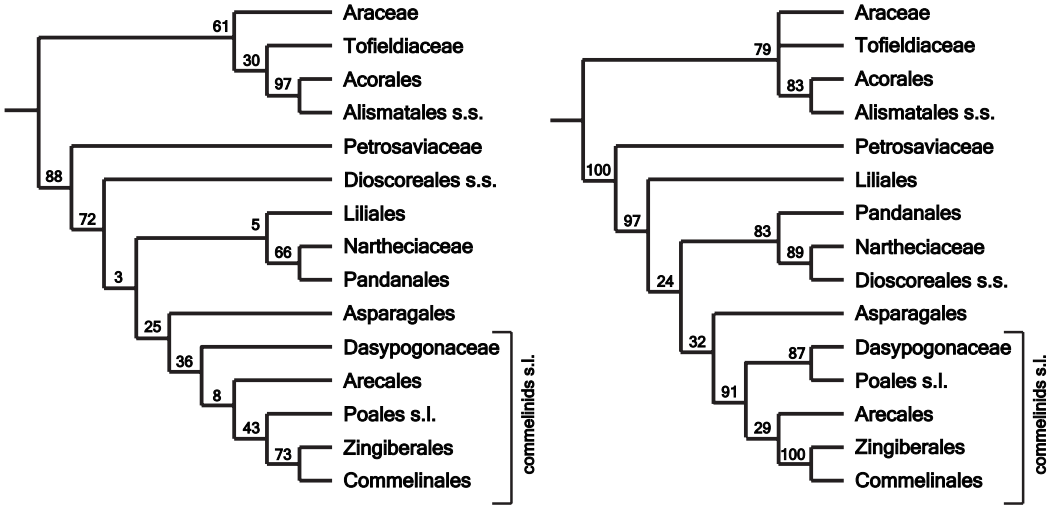

Fig. 1. Summary relationships among 15 major lineages of monocots, as resolved (left) by the principal two-gene analysis of Davis & al. (2004) and (right) by the present four-gene analysis. Both trees are summaries of consensus trees, with strict consensus jackknife scores (see text) indicated for relationships among the fifteen major lineages. Each tree is a portion of more inclusive tree that also includes dicot outgroups.

the commelinids, with Dasypogonaceae, Arecales, and Poales s.l. diverging in sequence from the line that includes Zingiberales and Commelinales. In the present analysis, Arecales are the sister of Zingiberales + Commelinales, and Dasypogonaceae are sister of Poales s.l.

A possible 3-nucleotide inversion of was identified in *cob*, involving sites 945–947 in the reference sequence. The sequence for these sites is AAG in most taxa in the study (AGA in *Oryza sativa* and several close relatives). An inversion is suggested by the occurrence of CTT (the reverse complement of AAG) in all representatives of Typhaceae, Bromeliaceae, Rapateaceae, and Eriocaulaceae. These four families are not resolved as a clade by the analysis, though they are in close proximity within the Poales s.l. Hence, if the observed nucleotide pattern does reflect the occurrence of one or more inversion events, either the inferred relationships are incorrect in this region, or multiple inversions have occurred (possibly including reversions). Analysis of the four-gene matrix, following the removal of these three sites, resulted in a set of 48 trees, though not the same 48 trees that were obtained when these sites were included. However, the two sets of trees yield the same consensus tree. All further analyses included these sites.

Patterns of support for groups by different genes. — A detailed consensus tree for the four-gene analysis is presented in Figs. 2–4. We do not discuss these relationships in depth, because a study based on a larger taxon sample is in preparation. However, we note areas in which the various genes have substantial individual impacts on the results obtained, as identified by jackknife percentages in the various three-gene analyses that differ from those of the four-gene analysis by 20 or more percentage points. There are 36 groups for which the exclusion of *matK* has an impact of this magnitude, 18 that are affected this much by *rbcL*, 15 by *atpA*, and three by *cob*.

Among the groups for which the various genes provide substantial amounts of support, there are instances in which genes from the same genome have substantial and positively correlated effects on the same lineages, others in which genes from the same genome have substantial negatively correlated effects, and corresponding instances in which genes from different genomes have substantial positively and negatively correlated effects. Using a criterion of at least 20% difference in jackknife support for a group between the four-gene analysis and any of the three-gene analyses, there are two clades for which genes from the same genome have positively correlated effects. One of these is the grouping of *Chlorophytum* with two representatives of Hemerocallidaceae (Fig. 3). This clade has 97% jackknife support in the four-gene analysis, but the removal of *matK* causes support to drop to 40%, and the removal of *rbcL* (with *matK* then included) causes support to drop to 63%. The other clade that exhibits this pattern is the one that consists of Typhaceae and Bromeliaceae (Fig. 4).

Positively correlated support by genes of different genomes is observed in nine cases. A striking example involves the clade that consists of Tofieldiaceae, Araceae, Acorales, and Alismatales s.s. (Fig. 2) Overall support for this group by the four-gene matrix is 79%, but it is 0% in the three-gene analysis that excludes *atpA*, and 27% in the three-gene analysis that excludes *matK*. A similar pattern is observed within Bromeliaceae (Fig. 4), where support for the clade that consists of all representatives of the family except *Brocchinia* is 88%, but it is 59% in the three-gene matrix that excludes *atpA*, and 32% in the three-gene matrix that excludes *matK*. Similarly, within the dicots (Fig. 2), support by the four-gene matrix is 89% for the clade that includes *Ceratophyllum* and the tricolpate dicots (Nelumbonaceae, Platanaceae, Lardizabalaceae, Berberidaceae), and it drops to 69% in the three-gene matrix that excludes *atpA*, and to 58% in the three-gene matrix that excludes *rbcL*.

There are four cases of negatively correlated support by genes from the same genome. One example is the clade that is the sister of *Flagellaria*, within Poales s.l., which consists of all representatives of Restionaceae, Joinvilleaceae, and Poaceae (Fig. 4). There is 63% jackknife support for this clade in the four-gene analysis, and support drops to 1% when *rbcL* is removed (demonstrating support for this group by *rbcL*), but rises to 88% when *matK* is removed (demonstrating conflict between *matK* and this grouping). A similar pattern is evident for the clade that consists of *Eichhornia* and *Monochoria* (Fig. 4). This group has 58% jackknife support by the four-gene matrix, and it drops to 5% when *matK* is removed, but rises to 90% when *rbcL* is removed. Similarly (but in this case involving the mitochondrial genome), the clade that consists of *Gyrocarpus* and *Neolitsea* has 51% jackknife support by the overall matrix, and support drops to 18% when *atpA* is removed, but rises to 83% when *cob* is removed.

Negatively correlated support by genes from different genomes is observed for just one clade, the grouping of Lowiaceae with Strelitziaceae (Fig. 4). Overall support for this group by the four-gene matrix is 29%, and it drops to 6% when *atpA* is removed, but rises to 50% when *rbcL* is removed.

DISCUSSION

The present analysis is based upon a data matrix that is a modified version of the two-gene (*rbcL* and *atpA*), 218-taxon data set analyzed by Davis & al. (2004). Two

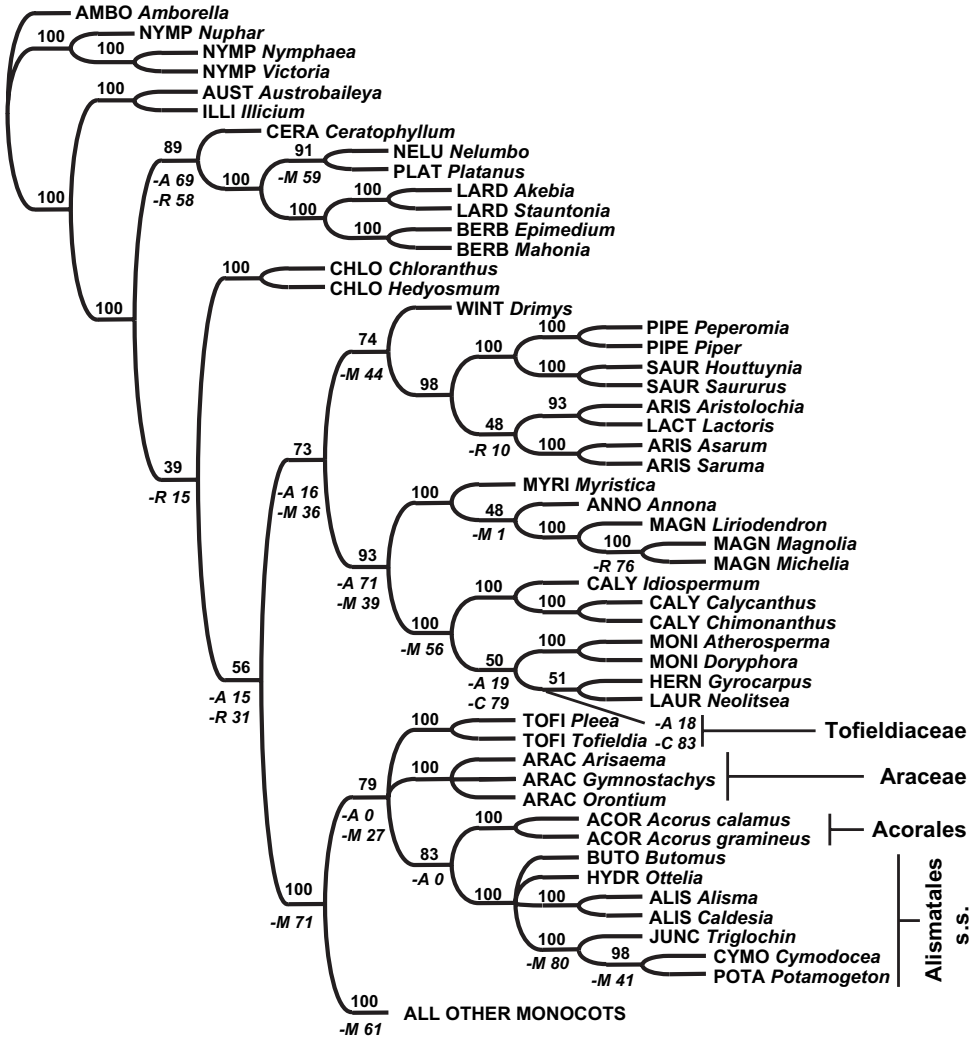

Fig. 2. Strict consensus of 48 most-parsimonious trees for 137 angiosperm terminals; basal region of consensus tree. Each terminal's name is preceded by a four-letter family code (cf. Appendix). Fifteen major mutually exclusive monocot lineages that encompass all monocots in the taxon sample (cf. Fig. 1) are labeled at right. The number above each branch is the strict consensus jackknife score for that clade (see text). Notations below branches identify one or more genes (A = *atpA*, C = cob, M = *matK*, R = *rbcL*) that, when individually removed from the four-gene matrix (yielding a three-gene matrix in each case), cause jackknife support to increase or decrease by >20 % relative to the level of support by the four-gene matrix; in each case the resulting jackknife percentage is specified.

additional genes (*matK* and *cob*) have been added for a subset of 136 of the taxa from the earlier analysis, and one new taxon (*Chlorophytum*) has been included, and scored for all four genes. Because the principal goal of the present study is to examine patterns of support by the four genes for different groups, taxa with data available for fewer than all four genes are not included. This taxon subset includes most of the dicot outgroups from the 218-taxon set, and among monocots it includes representatives of all 15 of the major lineages of interest discussed by Davis & al. (2004). Thus, the present analysis provides an initial perspective on patterns of variation for four genes across this taxon sample, and on the interacting patterns of support by these genes for relationships among these monocot lineages. Given this goal, achlorophyllous taxa (e.g., Triuridaceae) are excluded from the present study.

The four genes utilized in this analysis exhibit a range of properties that influence their utility for phylogenetic analysis. The number of nucleotides sequenced from each of the plastid-encoded genes exceeds the number from each of the mitochondrial genes (Table 1), but the addition of inferred indel regions, and the subsequent removal of those deemed ambiguously aligned, results in a net loss in sites for one gene from each genome (*atpA* and *matK*). This leaves *rbcL* as the gene providing the greatest number of unambiguously aligned sites (1,371), followed by *atpA* and *matK* (each with slightly more than

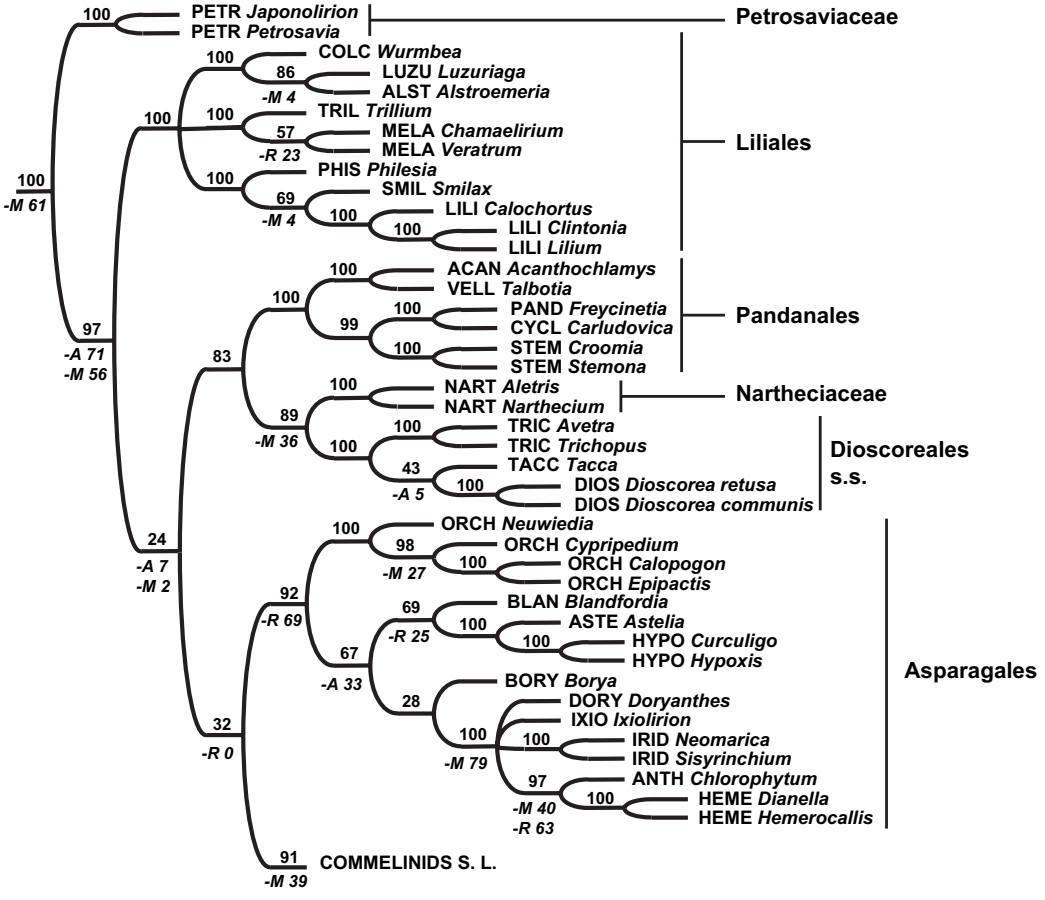

Fig. 3. Strict consensus of 48 most-parsimonious trees for 137 angiosperm terminals; structure of group labeled "all other monocots" in Fig. 2. Abbreviations as in Fig. 2.

1,200 sites), and finally by *cob* with 1,043 sites (Table 1). The percentage of unambiguously aligned sites that are parsimony informative is substantially greater for the two plastid-encoded genes than for the two mitochondrial genes, so 70% of the total number of informative characters are provided by the plastid-encoded genes. Thus, in terms of the criterion of numbers of characters, *rbcL* and *matK* contribute more to the overall matrix. By this criterion, *matK* contributes more than *rbcL* to the matrix, despite the exclusion of several regions of questionable alignment from the analysis. In both the plastid and mitochondrial genomes, the gene that presents the greatest challenge in terms of alignment also provides a greater percentage of parsimony-informative characters. The overall pattern that emerges, then, is one in which nucleotide substitution rates are greater in the two plastid genes than in the two mitochondrial genes, with a relatively fast-evolving and length-variable gene identifiable within each pair. Notably, it can never be known with certainty that all regions deemed unambiguously alignable actually are aligned correctly. A small number of misaligned sites can damage the results of an analysis dramatically, and if it is a general property of genes that those with higher rates of nucleotide transformation also have higher rates of indel activity, it may not always be appropriate to favor genes that provide the greatest numbers of (apparently) alignable nucleotide sites.

The two plastid-encoded genes account for 70% of the parsimony-informative characters in the four-gene matrix, and 82% of all steps in the trees obtained from this matrix (Table 1; the latter figure varies slightly among the 48 trees, and the value of 82% is based on the minimum for each gene across all trees). It should not be surprising, then, that the two mitochondrial genes both exhibit greater incongruence with the trees from the combined analysis (*atpA* having 3.6% more steps on these trees than on trees obtained from *atpA* alone, and *cob* having 10.9% more steps on the combined trees than on trees obtained from *cob*, while corresponding figures for *matK* and *rbcL* are 0.4% and 1.2%, respectively). As might be predicted, the two plastid-encoded genes also have substantial effects on a greater number of clades than do the mitochondrial genes (with "substantial effect" defined for this purpose as having a difference in jackknife support ≥ 20% when a gene is excluded from the four-gene matrix). This magnitude of influence by

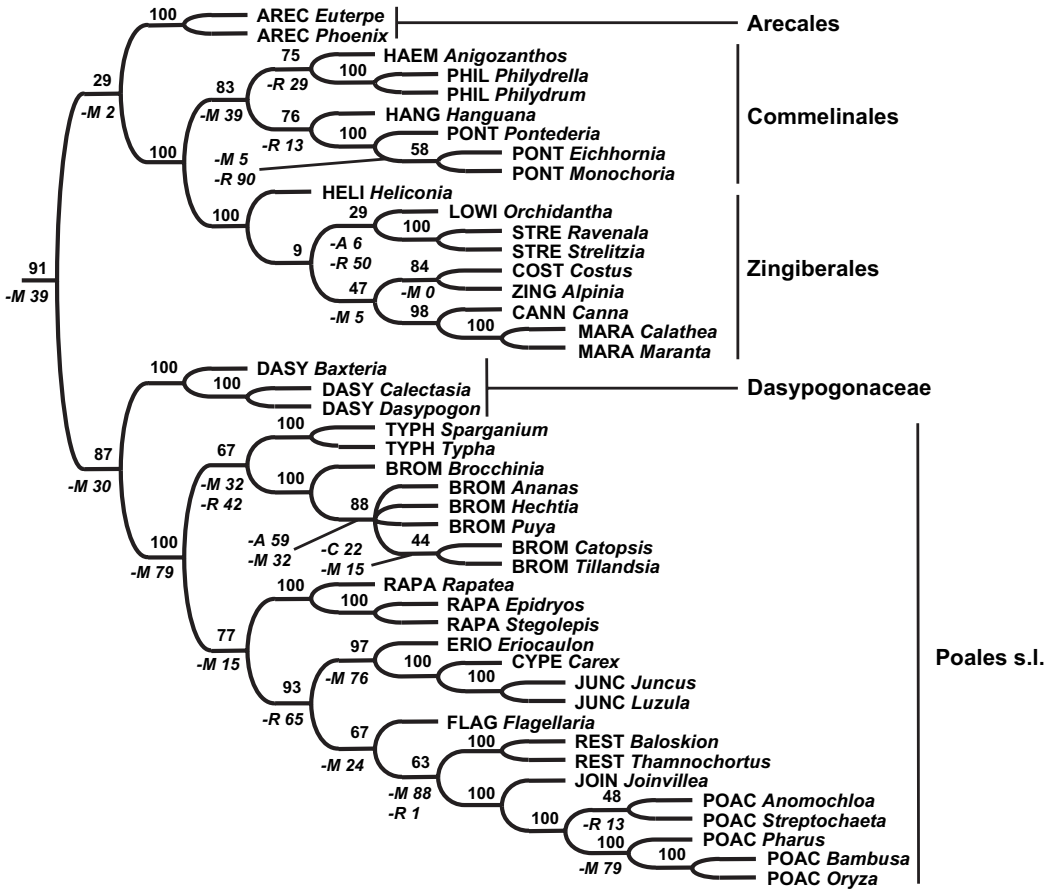

Fig. 4. Consensus of 48 most-parsimonious trees for 137 angiosperm terminals; structure of group labeled "commelinids s. l." in Fig. 3. Abbreviations as in Fig. 2.

matK is observed for a greater number of clades than for *rbcL* (36 vs. 18), and a similar pattern is observed for *atpA* vs. *cob* (15 clades vs. 3). Thus, if this measure of influence is adopted, each of the plastid-encoded genes has a greater influence on the results than either of the mitochondrial genes, and within each pair, the gene with the greatest substitution rate and the greatest amount of length variation has the greatest influence. The clades that are influenced in this manner are distributed throughout the overall tree (Figs. 2–4). However, some groups that are influenced in the same way by a particular gene are in close proximity (e.g., the strong influence of *atpA* on the clade consisting of Tofieldiaceae, Araceae, Acorales, and Alismatales s.s., and on the clade that consists of just the latter two of these taxa, Fig. 2). Thus, strong support or conflict between a gene and a particular group may be manifested multiple times, and it would be inadvisable to regard a simple count of the number of groups strongly influenced by a gene as more than an approximate indication of that gene's overall contribution to the four-gene matrix. With this caveat in mind, it is still true that a gene that influences many groups throughout the tree is playing a substantial role in the overall set of relationships, and *atpA* has an influence on support that is comparable to that of *rbcL*, if not equal to it.

Three of the genes influence small and large clades, but *cob*, which affects only three clades in this manner, has an effect only on groups of four terminals or fewer (with two of these groups involving Monimiaceae and related taxa, and the other involving relationships within Bromeliaceae). In this respect, the present results are consistent with those of Källersjö & al (1999), who observed that the most variable and homoplasious molecular characters (i.e., individual nucleotide sites) in their sample of 2,538 *rbcL* sequences made the greatest contributions to overall phylogenetic structure. The present study involves four genes which differ in terms of their average characteristics, while that of Källersjö & al. (1999) focused on differences among codon positions of nucleotides within a single gene. However, the results are similar in refuting prior notions that characters that are more variable and homoplasious are of lesser value than those that are less variable and less homoplasious, and that the least variable and least homoplasious characters

tend to influence deeper rather than shallower relationships in a tree.

In light of these results, which indicate a smaller overall influence on tree structure by the mitochondrial genes than by the plastid genes, should it be concluded that the effort and expense involved in the sequencing of mitochondrial genes is unnecessary? In reply to this question, we focus primarily on the influence of *atpA*. In the present analysis we have focused on jackknife support, rather than on resolution vs. lack of resolution, but the role of this gene in the resolution of many groups by the two-gene matrix was demonstrated by Davis & al. (2004), and need not be described again here. With the four-gene matrix, *atpA* has a substantial effect on support for 15 clades. In some cases this support is positively correlated with that of one or the other of the two plastid-encoded genes. For example, the placement of *Ceratophyllum* as sister of the tricolpate dicots has 89% jackknife support by the overall four-gene matrix, but this support drops to 69% in the three-gene matrix that excludes *atpA*, and to 58% in the three-gene matrix that excludes *rbcL* (Fig. 2). Thus, *atpA* reinforces the support that is provided for this group by *rbcL*, and in the absence of relevant data from other genes, *atpA* provides an essential element of the 89% support provided by the four-gene matrix. Additional instances of this sort are evident in the results (Figs. 2–4).

A more controversial role for *atpA* is evident in the support it provides for the grouping of Araceae, Tofieldiaceae, Acorales, and Alismatales s. s. (Figs. 1, 2). This group, which was resolved previously by the two-gene analysis of Davis & al. (2004), conflicts with the placement of *Acorus* as the sister of all other monocots, as resolved by several previous analyses (e.g., Duvall & al., 1993a, b; Chase & al., 2000). A recent nine-gene analysis also places *Acorus* among other monocots (Qiu & al., 2005), but in that case *Ceratophyllum* is placed as sister of the monocots (in contrast to the placement of *Ceratophyllum* by the present analysis). Thus, the position of *Acorus* as sister of all other monocots cannot be regarded as a settled matter, and *atpA* (which is one of the nine genes in the matrix analyzed by Qiu & al. [2005]) has played an important role in refuting such a placement. One notable result of the present analysis is the correlated role of *matK* in support of the placement of *Acorus*. The clade that consists of *Acorus* and three other monocot lineages, with 79% jackknife support, is strongly dependent on *atpA*, as evidenced by jackknife support of 0% for this group by the three-gene matrix that excludes *atpA* (Fig. 2). However, support for this group also is also low (27%) in the three-gene matrix that excludes *matK*. This evidence for an underlying commonality of support for the placement of *Acorus* by the mitochondrial gene *atpA* and the plastid gene *matK* should be sufficient by itself to refute arguments for the exclusion of certain genes from comprehensive analyses of plant phylogenetic relationships.

ACKNOWLEDGEMENTS

We thank various persons (J. Allen, S. Barrett, A. Bruneau, L. Caddick, K. Cameron, M. Chase, J. Conran, R. Dirig, A. Doust, J. Freudenstein, A. Gholson, Jr., S. Graham, J. Grimes, K. Hansen, Hedegaard, C. Kelloff, L. Kelly, J. Kress, P. Linder, J. Morris, O'Donohue, D. Potter, F. Rasmussen, H. Rasmussen, P. Rudall, O. Sønderhousen R. Soreng, T. Stuessy, N. Uhl, P. Weston) for providing plant materials, assistance during fieldwork, technical assistance, DNA isolations, and unpublished DNA sequences. We thank curators at the institutions listed in the caption of the Appendix for kindly providing access to their living materials. We acknowledge financial support from the U.S. National Science Foundation and the Danish National Sciences Research Council.

LITERATURE CITED

Angiosperm Phylogeny Group. 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.

Barfuss, M. H. J., Samuel, R., Till, W. & Stuessy, T. F. 2005. Phylogenetic relationships in subfamily Tillandsioideae (Bromeliaceae) based on DNA sequence data from seven plastid regions. *Amer. J. Bot.* 92: 337–351.

Chase, M. W., Soltis, D. E., Soltis, P. S., Rudall, P. J., Fay, M. F., Hahn, W. H., Sullivan, S., Joseph, J., Molvray, M., Kores, P. J., Givnish, T. J., Sytsma, K. J. & Pires, J. C. 2000. Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification. Pp. 3–16 in: Wilson, K. L. & Morrison, D. A. (eds.), *Monocots: Systematics and Evolution.* CSIRO, Collingwood.

Cuénoud, P., Savolainen, V., Chatrou, L. W., Powell, M., Grayer, R. J. & Chase, M. W. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *Amer. J. Bot.* 89: 132–144.

Davis, J. I., Simmons, M. P., Stevenson, D. W. & Wendel, J. F. 1998. Data decisiveness, data quality, and incongruence in phylogenetic analysis: an example from monocotyledons using mitochondrial *atpA* sequences. *Syst. Biol.* 47: 282–310.

Davis, J. I., Stevenson, D. W., Petersen, G., Seberg, O., Campbell, L. M., Freudenstein, J. V., Goldman, D. H., Hardy, C. R., Michelangeli, F. A., Simmons, M. P., Specht, C. D., Vergara-Silva, F. & Gandolfo, M. A. 2004. Phylogeny of the monocots, as inferred from *rbcL* and *atpA* sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. *Syst. Bot.* 29: 467–510.

Duvall, M. R., Clegg, M. T., Chase, M. W., Clark, W. D.,

Kress, W. J., Hills, H. G., Eguiarte, L. E., Smith, J. F., Gaut, B. S., Zimmer, E. A. & Learn, G. H., Jr. 1993a. Phylogenetic hypotheses for the monocotyledons constructed from *rbcL* sequence data. *Ann. Missouri Bot. Gard.* 80: 607–619.

Duvall, M. R., Learn, G. H., Jr., Eguiarte, L. E. & Clegg, M. T. 1993b. Phylogenetic analysis of *rbcL* sequences identifies *Acorus calamus* as the primal extant monocotyledon. *Proc. Nat. Acad. Sci. U.S.A.* 90: 4641–4644.

Farris, J. S. 1983. The logical basis of phylogenetic analysis. Pp. 1–36 in: Platnick, N. I. & Funk, V. A. (eds.), *Advances in Cladistics.* Columbia Univ. Press, New York.

Farris, J. S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D. & Kluge, A. G. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99–124.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20: 406–416.

Fuse, S. & Tamura, M. N. 2000. A phylogenetic analysis of the plastid *matK* gene with emphasis on Melanthiaceae sensu lato. *Plant Biol.* 2: 415–427.

Goloboff, P. A. 1991. Homoplasy and the choice among cladograms. *Cladistics* 7: 215–232.

Goloboff, P. A. 1993. *NONA vers. 1.6.* Tucuman, Argentina. Published by the author [distributed through http:\\www.cladistics.org].

Hilu, K. W., Borsch, T., Müller, K., Soltis, D. E., Soltis, P. S., Savolainen, V., Chase, M. W., Powell, M. P., Alice, L. A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T. A. B., Rohwer, J. G., Campbell, C. S. & Chatrou, L. W. 2003. Angiosperm phylogeny based on *matK* sequence information. *Amer. J. Bot.* 90: 1758–1776.

Källersjö, M., Albert, V. A. & Farris, J. S. 1999. Homoplasy increases phylogenetic structure. *Cladistics* 15: 91–93.

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

Kluge, A. G. & Farris, J. S. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18: 1–32.

Kubitzki, K. (ed.). 1998a. *The Families and Genera of Vascular Plants*, vol. 3. Springer-Verlag, Berlin.

Kubitzki, K. (ed.). 1998b. *The Families and Genera of Vascular Plants*, vol. 4. Springer-Verlag, Berlin.

Kubitzki, K., Rohwer, J. G. & Bittrich, V. (ed.). 1993. *The Families and Genera of Vascular Plants*, vol. 2. Springer-Verlag, Berlin.

Michelangeli, F. A., Davis, J. I. & Stevenson, D. W. 2003. Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *Amer. J. Bot.* 90: 93–106.

Molvray, M., Kores, P. J. & Chase, M. W. 2000. Polyphyly of mycoheterotrophic orchids and functional influences on floral and molecular characters. Pp. 441–448 in: Wilson, K. L. & Morrison, D. A. (ed.), *Monocots: Systematics and Evolution.* CSIRO, Collingwood.

Nickrent, D. L., Blarer, A., Qiu, Y.-L., Vidal-Russell, R. & Anderson, F. E. 2004. Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. *BMC Evol. Biol.* 4: 40: 1–17.

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

Nixon, K. C. 2002. *WinClada vers. 1.00.08.* Ithaca, New York. Published by the author [distributed through http:\\www.cladistics.org].

Nixon, K. C. & Carpenter, J. M. 1996. On simultaneous analysis. *Cladistics* 12: 221–241.

Nixon, K. C. & Davis, J. I. 1991. Polymorphic taxa, missing values, and cladistic analysis. *Cladistics* 7: 233–241.

Palmer, J. D. 1992. Mitochondrial DNA in plant systematics: applications and limitations. Pp. 36–49 in: Soltis, P. S., Soltis, D. E. & Doyle, J. J. (ed.), *Molecular Systematics of Plants.* Chapman and Hall, New York.

Palmer, J. D., Soltis, D. E. & Chase, M. W. 2004. The plant tree of life: an overview and some points of view. *Amer. J. Bot.* 91: 1437–1445.

Parker, L. T., Deng, Q., Zakeri, H., Carlson, C., Nickerson, D. A. & Kwok, P. Y. 1995. Peak height variations in automated sequencing of PCR products using Taq dye-terminator chemistry. *Biotechniques* 19: 116–121.

Petersen, G., Seberg, O., Davis, J. I. & Stevenson, D. W. 2006. RNA editing and phylogenetic reconstruction in two monocot mitochondrial genes. *Taxon* 55: 871–886.

Popper, K. 1959. *The Logic of Scientific Discovery.* Harper and Row, New York.

Qiu, Y.-L., Dombrovska, O., Lee, J., Li, L., Whitlock, B. A., Bernasconi-Quadroni, F., Rest, J. S., Davis C. C., Borsch, T., Hilu, K. W., Renner, S. S., Soltis, D. E., Soltis, P. S., Zanis, M. J., Cannone, J. J., Gutell, R. R., Powell, M., Savolainen, V., Chatrou, L. W. & Chase, M. W. 2005. Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Plant Sci.* 166: 815–842.

Qiu, Y.-L., Li, L., Hendry, T. A., Li, R., Taylor, D. W., Issa, M. J., Ronen, A. J., Vekaria, M. L. & White, A. M. 2006. Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes. *Taxon* 55: 837–856.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E. & Small, R. L. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Amer. J. Bot.* 92: 142–166.

Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savolainen, V., Hahn, W. H., Hoot, S. B., Fay, M. F., Axtell, M., Swensen, S. M., Prince, L. M., Kress, W. J., Nixon, K. C. & Farris, J. S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Bot. J. Linn. Soc.* 133: 381–461.

Stevenson, D. W., Davis, J. I., Freudenstein, J. V., Hardy, C. R., Simmons, M. P. & Specht, C. D. 2000. A phylogenetic analysis of monocotyledons based on morphological and molecular character sets, with comments on the placement of *Acorus* and Hydatellaceae. Pp. 17–24 in: Wilson, K. L. & Morrison, D. A. (eds.), *Monocots: Systematics and Evolution.* CSIRO, Collingwood.

Tamura, M. N. 1998. Nartheciaceae. Pp. 381–392 in: Kubitzki, K. (ed.), *The Families and Genera of Vascular Plants*, vol. 3. Springer-Verlag, Berlin.

Appendix. Taxa sampled for DNA sequences, accessions used for DNA isolations, and GenBank accession numbers of sequences. Genera are assigned to families according to Kubitzki (1998a, b) and Kubitzki & al. (1993), except as noted in the text. Families are assigned to orders according to the Angiosperm Phylogeny Group (2003), and are listed alphabetically within each order. Family names are accompanied by four-letter codes that also are used in the figures. For each sequence determined by the authors, accession information is provided for one or more DNA samples, including species name (or genus name if species is undetermined) and taxonomic authority, location of collection when known (for wild-collected plants), name of institution and accession number of living plant when known (for plants from curated living collections), and acronym of the herbarium in which the voucher specimen is deposited. GenBank accession numbers are provided without parentheses for sequences generated by the authors from the listed accessions, and within parentheses for sequences obtained from GenBank or directly from other persons. For sequences obtained from GenBank, the species name is provided as listed there (without taxonomic authority) if it differs from that of the DNA accession used by the authors, even if the difference is only in spelling. For unpublished sequences obtained directly from other persons, the donor's name and other available information are provided, in parentheses. Sources of plant accessions obtained from curated living collections are listed as follows (with accession numbers when known): Bailey Conserv. (L. H. Bailey Hortorium Conservatory, Cornell University, U.S.A.); Bogor Bot. Gard. (Bogor Botanic Garden, Indonesia); Copenhagen Bot. Gard. (University of Copenhagen Botanic Garden, Denmark); Cornell (Cornell University Campus, U.S.A.); Fairchild Trop. Gard. (Fairchild Tropical Garden, U.S.A.); Geneva Bot. Gard. (Geneva Botanical Gardens, Switzerland); Herrick Conserv. (Herrick Conservatory, Kent State University, USA); Lyon Arb. (Harold L. Lyon Arboretum, U.S.A.); Mo. Bot. Gard. (Missouri Botanical Garden, U.S.A.); NMNH (Smithsonian Institution, National Museum of Natural History, U.S.A.); NYBG (New York Botanical Garden, U.S.A.); Poring Centre (Poring Orchid Centre, Malaysia); RBG Kew (Royal Botanic Gardens, Kew, U.K.); RBG Melbourne (Royal Botanic Gardens Melbourne, Australia). Accessions from noncurated plantings are listed as "cult.".

Voucher information; *atpA*; *rbcL*; *cob*; *matK*.

DICOTYLEDONS: NO ORDINAL ASSIGNMENT: Amborellaceae (AMBO): *Amborella trichopoda* Baill., New Caledonia, *M. Simmons 1846* (BH); (AY009407); (L12628); DQ916607; (AJ506156). **Chloranthaceae (CHLO):** *Chloranthus spicatus* (Thunb.) Makino, NYBG 732/89 (NY); AY299746; (AY236835); DQ916608; *Chloranthus brachystachys* (AF543733). *Hedyosmum* sp., Puerto Rico, *D. Stevenson 1188* (NY); AY299777; *Hedyosmum orientale* (AY236848); DQ916609; *Hedyosmum arborescens* (AF465296). **Nymphaeaceae (NYMP):** *Nuphar variegata* Durand, U.S.A., Michigan, *J. Freudenstein 2564* (OS); *Nuphar* sp. (AF197638); (M77029); DQ916610; *Nuphar lutea* (AF543741). *Nymphaea odorata* Aiton, *K. Hansen s.n.*, June 1993 (BH); AY299814; (M77034); DQ916611; *Nymphaea alba* (AJ627251). *Victoria cruziana* Orb., Copenhagen Bot. Gard., not vouchered; AY299855; (M77036); DQ916612; *Victoria amazonica* (AF092991). **AUSTROBAILEYALES: Austrobaileyaceae (AUST):** *Austrobaileya scandens* C.T. White, NYBG 371/82A (NY); AY299723; (L12632); DQ916613; (AF465286). **Illiciaceae (ILLI):** *Illicium anisatum* L., NYBG 206/80A (NY); AY299786; *Illicium parviflorum* (L12652); DQ916614; *Illicium floridanum* (AF543738). **CERATOPHYLLALES: Ceratophyllaceae (CERA):** *Ceratophyllum demersum* L., Herrick Conserv., *J. Freudenstein 2555* (OS); AY299743; (D89473); DQ916615; (AF543732). **CANELLALES: Winteraceae (WINT):** *Drimys winteri* J.R. Forst. & G. Forst., RBG Melbourne, *A. Doust 1125* (MELU); AY299761; (AF093734); DQ916616; (AY437816). **LAURALES: Calycanthaceae (CALY):** *Calycanthus occidentalis* Hook. & Arn., Cornell, *M. Simmons 1899* (BH); AY299739; (AF022951); DQ916617; *Calycanthus fertilis* var *ferax* (AJ428413). *Chimonanthus praecox* (L.) Link, NYBG 78616 (NY); (AF197679); (L12639); DQ916618; (AY525340). *Idiospermum australiense* (Diels) S.T. Blake, NYBG 870/79A (NY); AY299785; (L12651); DQ916619; (AY525342). **Hernandiaceae (HERN):** *Gyrocarpus americanus* Jacq., NYBG 732/89, *L. Campbell s.n.* (NY); AY299773; *Gyrocarpus* sp. (L12647); DQ916620; (AF465295). **Lauraceae (LAUR):** *Neolitsea cassia* (L.) Kosterm., NYBG 245/85A (NY); AY299811; AY298841; DQ916621; *Neolitsea levinei* (AF244393). **Monimiaceae (MONI):** *Atherosperma moschatum* Labill., Australia, *J. Allen s.n.*, 4 Nov 1999 (NSW 433824); (AF197683); (AF121362); DQ916622; *Laurelia sempervirens* (AJ627928). *Doryphora sassafras* Endl., Australia, *J. Allen s.n.*, 4 Nov 1999 (NSW 433826); *Doryphora sassafras* (AF197688); *Doryphora aromatica* (L77211); DQ916623; *Doryphora sassafras* (AF542568). **MAGNOLIALES: Annonaceae (ANNO):** *Annona muricata* L., NYBG 921/92A, (NY); AY299712; (L12629); DQ916624; (AF543722). **Magnoliaceae (MAGN):** *Liriodendron tulipifera* L., NYBG 938/96 (NY); *Liriodendron chinense* (AF197690); *Liriodendron chinense* (L12654); DQ916625; (AF123480). *Magnolia grandiflora* L. cv. Edith Bogue, cult., *M. Simmons 1902* (BH); AY299800; AY298837; DQ916626; (AF548640). *Michelia figo* (Lour.) Spreng., Cornell, *M. Simmons 1898* (BH); AY299802; (L12659); DQ916627; (AF123467). **Myristicaceae (MYRI):** *Myristica fragrans* Houtt., NYBG 3/95B (NY); AY299808; AY298839; DQ916628; *Myristica maingayi* (AY220452). **PIPERALES: Aristolochiaceae (ARIS):** *Aristolochia gigantea* Mart. & Zucc., NYBG 1677/94 (NY); AY299718; *Aristolochia macrophylla* (L12630); DQ916629; (AB060794). *Asarum debile* Franch, *L. Kelly & Y. Tang 936* (BH, CDBI); *Asarum canadense* (AF197671); *Asarum canadense* (L14290); DQ916630; *Asarum caudigerum* (AY952420). *Saruma henryi* Oliver, RBG Kew 181-91-00951, *L. Kelly 688* (BH); (AF197672); (L12664); DQ916631; (AF543748). **Lactoridaceae (LACT):** *Lactoris fernandeziana* Phil., Chile, Juan Fernández Islands, *T. Stuessy et al. 15135* (OS); (AF197710); (L08763); DQ916632; (AF543739). **Piperaceae (PIPE):** *Peperomia polybotrya* Kunth, NYBG 380/49 (NY); AY299819; *Peperomia* sp. (L12661); DQ916633; *Peperomia graveolens* (AF542574). *Piper nigrum* L., Bailey Conserv. 68-334, *K. Hansen s.n.* (BH); *Piper betle* (AF197630); AY298847; DQ916634; (AB040153). **Saururaceae (SAUR):** *Houttuynia cordata* Thunb., Cornell, not vouchered; (AF197632); (L08762); DQ916635; (AF543737). *Saururus cernuus* L., *K. Hansen & J. Davis s.n.*, June 1994 (BH); AY299833; (L14294); DQ916636; (AF543749). **PROTEALES: Nelumbonaceae (NELU):** *Nelumbo lutea* Willd., NYBG 1686/95A (NY); (AY009420); (M77032); DQ916637; *Nelumbo nucifera* (AF543740). **Platanaceae (PLAT):** *Platanus occidentalis* L., NYBG 06793 (NY); (AF197655); (AF081073); DQ916638; (AF543747). **RANUNCULALES: Berberidaceae (BERB):** *Epimedium grandiflorum* Morr., Cornell, *J. Davis s.n.*, September 1999 (BH); AY299765; *Epimedium koreanum* (L75869); DQ916639; *Epimedium koreanum* (AB069837). *Mahonia aquifolium* (Pursh) Nutt., Cornell, *M. Simmons 1900* (BH); *Mahonia bealei* (AF197659); *Mahonia bealei* (L75871); DQ916640; *Mahonia japonica* (AB038184). **Lardizabalaceae (LARD):** *Akebia quinata* (Houtt.) Decne., cult., *J. Davis s.n.* (BH); AY299704; (L12627); DQ916641; (AB069851). *Stauntonia hexaphylla* Decne., NYBG 4225/95B (NY); AY299841; (L37922); DQ916642; *Lardizabala biter-*

Appendix. Continued.

Voucher information; *atpA*; *rbcL*; *cob*; *matK*.

nata (AY437809). **MONOCOTYLEDONS: NO ORDINAL ASSIGNMENT: Petrosaviaceae (PETR):** *Japonolirion osense* Nakai, Japan, *M. Chase 3000* (K); AY299790; (AF206784); DQ916643; (AB040161). *Petrosavia stellaris* Becc., Malaysia, *K. Cameron 2154* (K, NY); AY299821; (AF206806); DQ916644; *Petrosavia sakuraii* (AB040156). **ACORALES: Acoraceae (ACOR):** *Acorus calamus* L., U.S.A., New York, *R. Dirig 2990* (BH) for *atpA*; *Acorus calamus* L., Denmark, *G. Petersen & O. Seberg C-998* (C) for *cob*; AF039256; (M91625); DQ859124; (AB040154). *Acorus gramineus* Aiton, Herrick Conserv., *J. Freudenstein s.n.* (OS); AY299699; (D28866); DQ916645; (AB040155). **ALISMATALES: Alismataceae (ALIS):** *Alisma plantago-aquatica* L., Denmark, *O. Seberg C-472* (C); (AF197717); (L08759); DQ859125; *Alisma canaliculatum* (AB040179). *Caldesia oligococca* (F. Von Mueller) Buche, Australia, *F. Rasmussen et al. C-246* (C); AY277800; AY277799; DQ859129; (AY952427). **Araceae (ARAC):** *Arisaema triphyllum* (L.) Schott, U.S.A., New York, *N. Uhl 93-03* (BH); AY299717; AY298817; DQ916646; *Arisaema tortuosum* (AF387428). *Gymnostachys anceps* R. Br., Bailey Conserv. 95-101, *K. Hansen s.n.* (BH); AF039244; (M91629); DQ916647; (AB040177). *Orontium aquaticum* L., NYBG 49/80 (NY); AY299816; (AJ005632); DQ916648; (AF543744). **Butomaceae (BUTO):** *Butomus umbellatus* L., U.S.A., New York, *N. Uhl 92-05* (BH); AY299733; (U80685); DQ916649; (AY952416). **Cymodoceaceae (CYMO):** *Cymodocea serrulata* (R. Br.) Ascherson & Magh., *O'Donohue 21395* (BRN); AY277801; (U80687); DQ859131; *Halodule uninervis* (AY952424). **Hydrocharitaceae (HYDR):** *Ottelia ovalifolia* (R.Br.) Rich., Australia, *F. Rasmussen et al. C-245* (C); AY277802; *Ottelia alismoides* (U80707); DQ859146; *Ottelia acuminata* (AY952432). **Juncaginaceae (JUNG):** *Triglochin maritima* L., U.S.A., New York, *D. Goldman s.n.*, June 1993 (BH); AY299852; *Triglochin "maritimum"* (U80714); DQ916650; *Triglochin "maritimum"* (AB088782). **Potamogetonaceae (POTA):** *Potamogeton natans* L., *K. Hansen s.n.*, 1992 (BH); AY299829; *Potamogeton richardsonii* (U03730); DQ916651; *Potamogeton perfoliatus* (AY952425). **Tofieldiaceae (TOFI):** *Pleea tenuifolia* Michx., U.S.A., North Carolina, *M. Chase 152* (NCU); AY299827; (AJ131774); DQ916652; (AB183407). *Tofieldia calyculata* (L.) Wheldon, Geneva Bot. Gard., *M. Chase 1851* (K); AY299851; *Tofieldia pusilla* (AJ286562); DQ916653; (AB183403). **ASPARAGALES: Anthericaceae (ANTH):** *Chlorophytum nepalense* (Lindl.) Baker, Nepal, *Hedegaard 68, C-556* (C); DQ859074; *Chlorophytum orchidastrum* (*Z77257*); DQ859086; *Chlorophytum comosum* (AB017313). **Asteliaceae (ASTE):** *Astelia* sp., RBG Melbourne, *J. Grimes 3525* (MEL); AY299722; *Astelia pumila* (AF307906); DQ916654; *Astelia alpina* (AY368372). **Blandfordiaceae (BLAN):** *Blandfordia grandiflora* R. Br., RBG Melbourne 821598 Z2833, not vouchered; AY299727; *Blandfordia punicea* (Z73694); DQ916655; *Blandfordia punicea* (AY557206). **Boryaceae (BORY):** *Borya* aff. *sphaerocephala* R. Br., Australia, *J. Conran et al. 944* (ADU, PERTH); AY299728; *Borya septentrionalis* (Y14985); DQ916656; *Borya laciniata* (AY368373). **Doryanthaceae (DORY):** *Doryanthes excelsa* Corrêa, Australia, *M. Chase 188* (NCU); AY299760; (Z73697); DQ916657; (AB088785). **Hemerocallidaceae (HEME):** *Dianella caerulea* Sims, NYBG 88/3, *J. Davis s.n.* (BH); AY299756; *Dianella ensifolia* (M96960); DQ916658; *Dianella ensifolia* (AB088787). *Hemerocallis* sp. cv. Stella d'Oro, cult., *K. Hansen s.n.*, September 1992 (BH); AY299780; *Hemerocallis fulva* (L05036); DQ916659; *Hemerocallis fulva* (AB017318). **Hypoxidaceae (HYPO):** *Curculigo capitulata* (Lour.) Kuntze, Bailey Conserv. 95-103, *K. Hansen & J. Davis s.n.* (BH); AF039249; (Z73701); DQ916660; (AB088783). *Hypoxis occidentalis* Benth. var. *occidentalis*, Australia, *J. Conran et al. 919* (ADU, PERTH); AY299784; *Hypoxis glabella* (Y14989); DQ916661; *Hypoxis leptocarpa* (AY368375). **Iridaceae (IRID):** *Neomarica northiana* (Schneev.) Sprague, Bailey Conserv., *D. Goldman 1758* (BH); AY299812; AY298842; DQ916662; (AJ579972). *Sisyrinchium angustifolium* Mill., *K. Hansen 92-05* (BH); AY299837; *Sisyrinchium micranthum* (Z77290); DQ916663; *Sisyrinchium micranthum* (AJ579982). **Ixioliriaceae (IXIO):** *Ixiolirion tataricum* (Pall.) Herb., RBG Kew 1986-2579, *M. Chase 489* (K) for *atpA*; *Ixiolirion tataricum* (Pall.) Herb., Copenhagen Bot. Gard., *O. Sønderhousen 810* (C) for *cob*; AY299789; (Z73704); DQ916731; (AB017327). **Orchidaceae (ORCH):** *Calopogon tuberosus* (L.) Britton, Sterns & Poggenb., *D. Goldman 532* (TEX, BH); AY299738; (AF074119); DQ916664; (AF263635). *Cypripedium calceolus* L. var. *pubescens* (Willd.) Correll, *J. Morris 3A* (KE); AY299755; *Cypripedium passerinum* (AF074142); DQ916665; (AY557208). *Epipactis helleborine* (L.) Crantz, U.S.A., New York, *D. Potter s.n.* (OS); AY299766; (Z73707); DQ916666; (AF263659). *Neuwiedia veratrifolia* Blume, Poring Centre, *M. Chase O-883* (K); AY299813; (AF074200); DQ916667; (AY557211). **DIOSCOREALES: Dioscoreaceae (DIOS):** *Dioscorea retu.s.a.* Mast., Bailey Conserv. 91-058, *K. Hansen s.n.* (BH); AY299759; *Dioscorea polygonoides* (AJ235803); DQ916668; *Dioscorea alata* (AB040208). *Dioscorea communis* (L.) Caddick & Wilkin (formerly *Tamus communis* L.), Portugal, *F. & H. Rasmussen C-1170* (C); AY277804; (AF307474); DQ916732; (AF465303). **Nartheciaceae (NART):** *Aletris farinosa* L., U.S.A., North Carolina, *M. Chase 105* (NCU); AY299706; (*rbcL* sequence provided by M. Chase, from same DNA isolation as *atpA* and *cob*); DQ916669; *Aletris stenoloba* (AB040175). *Narthecium ossifragum* (L.) Huds., RBG Kew 1980-1915, *M. Chase 610* (K); AY299809; (AJ286560); DQ916670; *Narthecium asiaticum* (AB040162). **Taccaceae (TACC):** *Tacca parkeri* Seem., Colombia, *R. Schultes & F. López 9298b* (NY); AY299845; *Tacca chantrieri* (AJ235810); DQ916671; *Tacca* sp. (AB088792). **Trichopodaceae (TRIC):** *Avetra sempervirens* H. Perrier, Madagascar, *L. Caddick 304* (K); AY299724; AY298818; DQ916672; *Trichopus sempervirens* (AY973844). *Trichopus zeylanicus* Gaertn., Sri Lanka, *L. Caddick 346* (*MWC6634*) (K); AY277805; (AF307477); DQ916733; (AY973845). **LILIALES: Alstroemeriaceae (ALST):** *Alstroemeria caryophyllaea* Jacq., Fairchild Trop. Gard. 81-563 (FTG); AF039254; *Alstroemeria* sp. (Z77254); DQ916673; *Alstroemeria* sp. (AY624481). **Calochortaceae (CALO):** *Calochortus minimus* Ownbey, *Ness 606* (PUA); AY299737; (Z77263); DQ916674; *Calochortus uniflorus* (AY624478). **Colchicaceae (COLC):** *Wurmbea* sp., Australia, *J. Conran et al. 899* (ADU, PERTH); AY299856; AY298853; DQ916675; *Colchicum speciosum* (A040181). **Liliaceae (LILI):** *Clintonia borealis* (Aiton) Raf., RBG Kew 1981-6330, *M. Chase 498* (K); AY299748; (D17372); DQ916676; (AB024542). *Lilium superbum* L., cult., *M. Chase 112* (NCU); AY299797; (L12682); DQ916677; (AB024546). **Luzuriagaceae (LUZU):** *Luzuriaga radicans* Ruiz & Pav., RBG Kew 1961-64905, *M. Chase 499* (K); AY299798; (Z77300); DQ916678; *Drymophila moorei* (AB040180). **Melanthiaceae (MELA):** *Chamaelirium luteum* (L.) A. Gray, U.S.A., North Carolina, *M. Chase 224* (NCU); AY299745; (AJ276347); DQ916679; (AB040196). *Veratrum viride* Aiton, U.S.A., New York, *N. Uhl 92-02* (BH); AF039255; *Veratrum album* (D28168); DQ916680; *Veratrum maackii* (AB040183). **Philesiaceae (PHIS):** *Philesia buxifolia* Lam., RBG Kew 1965-68407, *M. Chase 545* (K) for *atpA*; *Philesia magellanica* J.F. Gmel., Argentina, *A. Rizzeto s.n.*, 31 March 1999, *C-520* (C) for *cob*; AY299822; *Philesia buxifolia* (Z77302); DQ916734; *Philesia buxifolia* (AY624479). **Smilacaceae (SMIL):** *Smilax rotundifolia* L., U.S.A., New York, *N. Uhl 92-07* (BH); AF039251; *Smilax glauca* (Z77310); DQ916681; *Smilax china* (AB040204). **Trilliaceae (TRIL):** *Trillium grandiflorum* (Michx.) Salisb., U.S.A., New York, *N. Uhl s.n.*, 1993 (BH); AF039253; (D28164); DQ916682; (AB017392). **PANDANALES: Acantho-**

Appendix. Continued.

Voucher information; *atpA*; *rbcL*; *cob*; *matK*.

chlamydaceae (ACAN): *Acanthochlamys bracteata* P.C. Kao, *P. Kao 1993* (K); AY299698; (*rbcL* sequence provided by M. Chase, from same DNA isolation as *atpA* and *cob*); DQ916683; (AY952413). **Cyclanthaceae (CYCL):** *Carludovica drudei* Mast., Bailey Conserv. 73-574 (BH); *Carludovica palmata* (AF197707); *Carludovica palmata* (AF197596); DQ916684; *Carludovica palmata* (AB088793). **Pandanaceae (PAND):** *Freycinetia multiflora* Merrill, Mo. Bot. Gard. 811323 (MO) for *atpA*; *Freycinetia excelsa* F. Muell., Australia, *F. Rasmussen et al. C-243* (C) for *cob*; AY299770; *Freycinetia scandens* (AF206770); DQ916735; *Freycinetia formosana* (AB040209). **Stemonaceae (STEM):** *Croomia pauciflora* (Nutt.) Torr., U.S.A., Florida, *A. Gholson, Jr. 10360* (FLAS); (AF197708); AY298827; DQ916685; (AY437815). *Stemona javanica* (Kunth) Engl., Bogor Bot. Gard. XV.B.5a, *M. Chase 2156* (K); AY299842; *Stemona japonica* (AJ131948); DQ916686; *Stemona japonica* (AB040210). **Velloziaceae (VELL):** *Talbotia elegans* Balf., Bailey Conserv. 91-069 (BH) for *atpA*; *Talbotia elegans* Balf., Copenhagen Bot. Gard. C-521 (C) for *cob*; AF039247; *Barbacenia elegans* (AJ131946); DQ916736; (AY491664). **NO ORDINAL ASSIGNMENT: Dasypogonaceae (DASY):** *Baxteria australis* R. Br. ex Hook., Australia, *J. Conran et al. 906* (ADU, PERTH); AY124504; AY123230; DQ916687; DQ888764. *Calectasia cyanea* R. Br., Australia, *J. Conran et al. 928* (ADU, PERTH); AY124505; AY123231; DQ916688; DQ888765. *Dasypogon hookeri* J.R. Drumm., Australia, *J. Conran et al. 917* (ADU, PERTH); AY124503; AY123229; DQ916689; DQ888766. **ARECALES: Arecaceae (AREC):** *Euterpe oleracea* Mart., Lyon Arb. L-70.0017 (BH); AY299769; AY298832; DQ916690; *Areca triandra* (AY952428). *Phoenix dactylifera* L., Bailey Conserv. 78-121 (BH); *Phoenix reclinata* (U58831); *Phoenix reclinata* (M81814); DQ916691; (AB040211). **COMMELINALES: Haemodoraceae (HAEM):** *Anigozanthos flavidus* DC. in Redouté, Bailey Conserv. 95-102, *K. Hansen & J. Davis s.n.* (BH); AF039246; (AJ404843); DQ916692; (AB088796). **Hanguanaceae (HANG):** *Hanguana malayana* Merr., RBG Kew, *P. Rudall s.n.* (K); AY299775; (AJ417896); DQ916693; (AB088800). **Philydraceae (PHIL):** *Philydrella pygmaea* (R. Br.) Caruel, Australia, *J. Conran et al. 915* (ADU, PERTH); AY299823; AY298845; DQ916694; (AF434870). *Philydrum lanuginosum* Banks & Sol. ex Gaertn., RBG Kew 1987-8002 (K); AY299824; (U41596); DQ916695; (AY952429). **Pontederiaceae (PONT):** *Eichhornia azurea* (Sw.) Kunth, RBG Kew 1991-1656 (K); AY299762; (U41573); DQ916696; *Eichhornia crassipes* (AB040212). *Monochoria korsakowii* Regel & Maack, *S. Barrett 1415* (TRT); AY299803; *Monochoria korsakovii* (U41590); DQ916697; (AB088795). *Pontederia cordata* L., NYBG 2844/95, *L. Campbell 755* (NY); AY299828; *Pontederia cordata* var. *cordata* (U41592); DQ916698; (AF434872). **POALES: Bromeliaceae (BROM):** *Ananas comosus* (L.) Merr., Bailey Conserv., not vouchered; AY299710; (L19977); DQ916699; *Ananas ananassoides* (AF162227). *Brocchinia reducta* Baker, Venezuela, *F. Michelangeli 525* (VEN); AY299729; AY298820; DQ916700; (AY614018). *Catopsis nutans* (Sw.) Griseb., Bailey Conserv. 72-783 (BH); AF039257; *Catopsis montana* (L19976); DQ916701; *Catopsis nutans* var. *nutans* (AY614026). *Hechtia texensis* S. Watson, NYBG 135/80 (NY); AY299776; *Hechtia montana* (L19974); DQ916702; *Hechtia carlsoniae* (AY614020). *Puya berteroniana* Mez, NYBG 30/77A (NY); AY124508; *Puya dyckioides* (L19973); DQ916703; *Puya laxa* (AY614022). *Tillandsia usneoides* (L.) L., Bailey Conserv., not vouchered; AY124507; *Tillandsia elizabethae* (L19971); DQ916704; (AY614122). **Cyperaceae (CYPE):** *Carex interior* L.H. Bailey, *K. Hansen 93-06* (BH); AY124514; *Carex monostachya* (Y12998); DQ916705; *Cyperus alternifolius* (AY952421). **Eriocaulaceae (ERIO):** *Eriocaulon humboldtii* Kunth, Venezuela, *F. Michelangeli 542* (VEN); AY124517; AY123236; DQ916706; *Eriocaulon septangulare* (AY952430). **Flagellariaceae (FLAG):** *Flagellaria indica* L., Bailey Conserv. 77-394, *K. Hansen s.n.*, May 1994 (BH); AF039248; (L12678); DQ916707; (AB040214). **Joinvilleaceae (JOIN):** *Joinvillea ascendens* Gaudich. ex Brongn. & Gris, U.S.A., Hawaii, *A. Bruneau s.n.*, August 1992 (BH); AY124519; *Joinvillea plicata* (L01471); DQ916708; (AF164380). **Juncaceae (JUNC):** *Juncus* sp., *K. Hansen s.n.* (BH); AY124520; *Juncus effusus* (L12681); DQ916709; *Juncus effusus* (AB088803). *Luzula acuminata* Raf., New York, *F. Michelangeli 543* (BH); AY124521; *Luzula multiflora* (AJ419945); DQ916710; *Luzula wahlenbergii* (AY973518). **Poaceae (POAC):** *Anomochloa marantoidea* Brongn., Bailey Conserv., *K. Hansen & J. Davis s.n.* (BH); AY124526; (AF021875); DQ916711; (AF164381). *Bambu.s.a. multiplex* (Lour.) Raeusch. ex Schult. & Schult. f., Bailey Conserv. 71-470, *R. Soreng s.n.* (BH); AY124525; (M91626); DQ916712; *Phyllostachys aurea* (AF164390). *Oryza* L. (all sequences taken from GenBank); *Oryza sativa* (X51422); *Oryza sativa* (D00207); *Oryza sativa* (BA000029.3); *Oryza nivara* (NC_005973). *Pharus latifolius* L., Bailey Conserv., *K. Hansen s.n.*, 29 July 92 (BH); AY124524; (AY357724); DQ916713; (AF164388). *Streptochaeta angustifolia* Soderstr., Bailey Conserv. (BH); AY124523; *Streptochaeta spicata* (AJ419949); DQ916714; (AF164382). **Rapateaceae (RAPA):** *Epidryos allenii* (Steyerm.) Maguire, Panama, *D. Stevenson 1210* (NY); AY299764; AY298830; DQ916715; (AF162225). *Rapatea xiphoides* Sandwith, Guyana, *C. Kelloff 975* (BH); AY124511; AF460969; DQ916716; *Rapatea* sp. (AF539958). *Stegolepis parvipetala* Steyerm., Venezuela, *F. Michelangeli 513* (VEN); AY124535; AY123242; DQ916717; (AY614014). **Restionaceae (REST):** *Baloskion tetraphyllum* (Labill.) B.G. Briggs & L.A.S. Johnson, RBG Kew 1977-6565, *M. Chase 560* (K); AY124529; (AF148761); DQ916718; *Restio tetraphyllus* (AF164379). *Thamnochortus cinereus* H.P. Linder, NYBG 227/86A (NY); AY124531; (*rbcL* sequence provided by H. Linder, from *Thamnochortus cinereus* H.P. Linder, *H. Linder et al. 7281*, Z); DQ916719; (AY690724). **Typhaceae (TYPH):** *Sparganium eurycarpum* Engelm., *K. Hansen s.n.*, June 1993 (BH); AY124509; *Sparganium americanum* (M91633); DQ916720; *Sparganium glomeratum* (AY952426). *Typha latifolia* L., U.S.A., New York, *N. Uhl 92-04* (BH); AY124510; (M91634); DQ916721; (AB088801). **ZINGIBERALES: Cannaceae (CANN):** *Canna indica* L., Bailey Conserv. 72-117 (BH); AY299741; (AF378763); DQ916722; *Canna flaccida* (AF478906). **Costaceae (COST):** *Costus lateriflorus* Baker, NMNH 98-224, *W.J. Kress 00-6599* (US) for *atpA & rbcL*; *Costus erythrophyllus* Loes., NYBG 4105/95 (NY) for *cob*; AY299753; AY298826; DQ916723; *Costus pulverulentus* (AF478907). **Heliconiaceae (HELI):** *Heliconia rostrata* Ruiz & Pav., NYBG 1380/91A (NY); AY299778; *Heliconia indica* (AF378765); DQ916724; *Heliconia irrasa* (AF478908). **Lowiaceae (LOWI):** *Orchidantha maxillarioides* K. Schum., NYBG 1639/91 (NY); AY299815; *Orchidantha fimbriata* (AF243841); DQ916725; *Orchidantha fimbriata* (AY952417). **Marantaceae (MARA):** *Calathea loeseneri* J.F. Macbr., NYBG 345/95A (NY); AY299735; (AF243842); DQ916726; (AY140273). *Maranta leuconeura* E. Morren, Bailey Conserv. 84-107, not vouchered; AY299801; *Maranta bicolor* (AF378768); DQ916727; (AY140303). **Strelitziaceae (STRE):** *Ravenala madagascariensis* Sonn., NYBG 331/99 (NY); AY299830; (L20138); DQ916728; (AF434873). *Strelitzia nicolai* Regel & Körn., NYBG 400/56 (NY); AY299843; (AF243846); DQ916729; *Strelitzia alba* (AF434874). **Zingiberaceae (ZING):** *Alpinia purpurata* (Vieill.) K. Schum., Bailey Conserv. 72-114 (BH); AY299708; AY298816; DQ916730; *Alpinia intermedia* (AB040213).