Robust Inference of Monocot Deep Phylogeny Using an Expanded Multigene Plastid Data Set

Sean W. Grahan  
*University of British Columbia*

Jessie M. Zgurski  
*University of Alberta*

Marc A. McPherson  
*University of Alberta*

Donna M. Cherniawsky  
*University of Alberta; Royal Alberta Museum*

Jeffery M. Saarela  
*University of British Columbia*

*See next page for additional authors*

Follow this and additional works at: [https://scholarship.claremont.edu/aliso](https://scholarship.claremont.edu/aliso)

Part of the [Botany Commons](https://scholarship.claremont.edu/aliso)

**Recommended Citation**

Grahan, Sean W.; Zgurski, Jessie M.; McPherson, Marc A.; Cherniawsky, Donna M.; Saarela, Jeffery M.; Horne, Elvira F. C.; Smith, Selena Y.; Young, Winson A.; O'Brien, Heath E.; Brown, Vincent L.; Pires, J. Chris; Olmstead, Richard G.; Chase, Mark W.; and Rai, Hardeep S. (2006) "Robust Inference of Monocot Deep Phylogeny Using an Expanded Multigene Plastid Data Set," *Aliso: A Journal of Systematic and Floristic Botany*: Vol. 22: Iss. 1, Article 2.  
Available at: [https://scholarship.claremont.edu/aliso/vol22/iss1/2](https://scholarship.claremont.edu/aliso/vol22/iss1/2)
Robust Inference of Monocot Deep Phylogeny Using an Expanded Multigene Plastid Data Set

Authors
Sean W. Grahan, Jessie M. Zgurski, Marc A. McPherson, Donna M. Cherniawsky, Jeffery M. Saarela, Elvira F. C. Horne, Selena Y. Smith, Winson A. Young, Heath E. O'Brien, Vincent L. Brown, J. Chris Pires, Richard G. Olmstead, Mark W. Chase, and Hardeep S. Rai
ROBUST INFERENCE OF MONOCOT DEEP PHYLOGENY USING AN EXPANDED MULTIGENE PLASTID DATA SET

SEAN W. GRAHAM,1,6 JESSIE M. ZGURSKI,2 MARC A. McPHERSON,2,7 DONNA M. CHERNIAWSKY,2,8 JEFFERY M. SAARELA,1 ELYVIRA F. C. HORNE,2 SELENA Y. SMITH,2 WINSON A. WONG,2 HEATH E. O'BRIEN,2,9 VINCENT L. BIRON,2,10 J. CHRIS PIRES,3,11 RICHARD G. OLMSTEAD,4 MARK W. CHASE,4 AND HARDEEP S. RAI1

1UBC Botanical Garden and Centre for Plant Research, University of British Columbia, 6804 SW Marine Drive, Vancouver, British Columbia V6T 1Z4, Canada; 2Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada; 3Deartment of Agronomy, University of Wisconsin, Madison, Wisconsin 53706, USA; 4Department of Biology, Box 351800, University of Washington, Seattle, Washington 98195, USA; 5Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK
*Corresponding author (swgraham@interchange.ubc.ca)

ABSTRACT

We use multiple photosynthetic, chlororespiratory, and plastid translation apparatus loci and their associated noncoding regions (ca. 16 kb per taxon, prior to alignment) to make strongly supported inferences of the deep internal branches of monocot phylogeny. Most monocot relationships are robust (an average of ca. 91% bootstrap support per branch examined), including those poorly supported or unresolved in other studies. Our data strongly support a sister-group relationship between Asparagales and the commelinid monocots, the inclusion of the orchids in Asparagales, and the status of Petrosaviaceae as the sister group of all monocots except Acorus and Alismatales. The latter finding supports recognition of the order Petrosaviaceae. Also strongly supported is a placement of Petermannia disjunct from Colchicaceae (Liliales) and a sister-group relationship between Commelinales and Zingiberales. We highlight the remaining weak areas of monocot phylogeny, including the positions of Dioscoreales, Liliales, and Pandanales. Despite substantial variation in the overall rate of molecular evolution among lineages, inferred amounts of change among codon-position data partitions are correlated with each other across the monocot tree, consistent with low incongruence between these partitions. Ceratophyllum and Chloranthaceae appear to have a destabilizing effect on the position of the monocots among other angiosperms; the issue of monocot placement in broader angiosperm phylogeny remains problematic.

Key words: deep phylgeny, heterotachy, Japonolirion, monocot backbone, monocot sister group, monocot tree of life, Petermanniaceae, plastid genome evolution.

INTRODUCTION

Available sources of phylogenetic data do not allow for well-supported inference of all of the deep branches of monocot phylogeny. The internal branches that have proven most resilient to well-supported resolution are often relatively short (e.g., Chase et al. 2000), including portions of the backbone linking those major clades defined as orders in the classification of the Angiosperm Phylogeny Group II [APG II] (2003), and a number of poorly resolved branches deep within each order. To address these and other problematic areas of higher-order monocot phylogeny, we collected new data for a large number of exemplar monocot taxa from an expanded range of regions in the plastid genome, including many that have not been examined intensively in the monocots. The plastid regions surveyed consist of portions of ten photosystem II genes, two NADH dehydrogenase subunit genes, three ribosomal protein genes, atpB, rbcL, and a diverse collection of noncoding regions that span these and other genes from this genome. Collectively, these represent ca. 16 kb of DNA sequence data per taxon, or about a ninth of the nonduplicated information in the plastid genome. Most of the regions were sequenced using primers developed for the inference of deep nodes of angiosperm phylogeny (Graham and Olmstead 2000).

Discussions of strategies for large-scale phylogenetic inference are usually framed in terms of a trade off between taxon sampling vs. amount of data collected per taxon. Aiming for a relatively dense taxon sampling is generally a beneficial strategy for accurate inference of phylogenetic relationships among and within large groups of organisms (e.g., Hillis 1998; Pollock et al. 2002; Zwickl and Hillis 2002), and examining more characters per taxon can also be very useful (e.g., Graham et al. 1998; Soltis et al. 1998; Poe and Swofford 1999; Hillis et al. 2003). Both strategies have demonstrated their effectiveness in ongoing studies of monocot phylogeny (e.g., Chase et al. 1993, 1995a, b, 2000, 2006; Stevenson et al. 2000; Givnish et al. 2006). While it is relatively clear which areas of monocot phylogeny are more (or less) sturdy, the level of data sampling appropriate for tackling the remaining areas of uncertainty is not clear in
advance (e.g., Hillis et al. 2003). Fortunately, the economic trade off between taxa and characters is becoming less limiting. With the development and application of methods for efficiently examining multiple genomic regions among distantly related taxa, we can collect and analyze substantially more data per taxon (and research dollar) than was possible at the time of the first two monocot conferences.

We therefore sampled a large number of characters for a broad sampling of monocot taxa. Our preliminary monocot study constitutes the largest to date, in terms of the amount of data examined per taxon. It also has a sufficient taxon density to represent all major branches of monocot phylogeny (excluding several achlorophyllous lineages), with several major clades examined at a taxon density comparable to other recent large studies. We report on the strong inferences that can be drawn from the current taxon sampling, highlight several remaining areas of uncertainty, and briefly review extensive among-lineage rate shifts in the context of the molecular evolutionary dynamics of different codon-position data partitions.

MATERIALS AND METHODS

Taxonomic and Genomic Sampling

The taxa surveyed are an expansion of the set examined in Graham et al. (2000) and McPherson et al. (submitted). Our sampling includes representatives of all ten clades recognized as monocot orders and 60 of 93 clades recognized at the rank of family in the classification scheme of APG II (2003), with Petermanniaceae accepted as a distinct family (discussed below). This sampling includes Petrosaviaceae and Dasyogonaceae, two families unplaced to order in the classification of APG II (2003) and in other large-scale studies. We use 25 “dicots” as outgroup taxa (24 from Graham et al. [2000] and Trimenia moorei (Oliv.) W. R. Philpoin [Austrobaileyales: Trimeniaceae; NSW 433770, GenBank nos. AY116652–AY116659]). The exemplars were chosen to represent a broad phylogenetic diversity of families, guided by the classification of APG II (2003) and other large-scale studies. Details of the species used, sample provenance, and GenBank accession information are presented in McPherson et al. (submitted) and in the Appendix. Ninety-four exemplar taxa are represented, 69 from the monocots. The latter number is an increase of 23 monocots from McPherson et al. (submitted).

Details of DNA extraction, amplification, and sequencing protocols are provided in Graham and Olmstead (2000) and McPherson et al. (submitted). The data set considered for phylogenetic analysis here spans 17 protein-coding plastid genes (atpB, ndhB, ndhF, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbL, psbN, psbT, rbcL, rpl2, rps7, 3'-rps12), and includes six intergenic regions in two of the plastid inverted regions (psbE-psbf-psbL-psbI and psbB-psbT-psbN-psbH), two intergenic regions in a cluster of genes spanning 3'-rps12, rps7 (two ribosomal small subunit genes) and ndhB, and three introns (one each in ndhB, rpl2 and 3'-rps12). Most monocots were sampled for a larger portion of ndhF (ca. 2.08 kb, representing most of the locus) than the outgroup taxa (ca. 1.29 kb from the relatively slowly evolving 5'-end of the gene). Three additional regions that are largely noncoding were also sampled for Asparagales and Liliales (and several other taxa; Appendix): an intergenic region between ndhB and trnL (CAA), an intergenic region between atpB and rbcL, and a contiguous region spanning trnL (UAA) and trnF (GAA) that includes an intron in the former gene and an intergenic spacer region. In a few cases closely related alternative taxa were sampled for the latter region (see Appendix). With these and a few other minor exceptions, all taxa were completely represented for all regions. Further details, including methods of data compilation, are provided in Graham and Olmstead (2000) and McPherson et al. (submitted).

The genes examined include a mix of those involved in photosynthesis (atpB, rbcL, and the photosystem II genes), chlororespiration (the ndh genes) and plastid translation (the trn, rps, and rpl loci). Those situated in the plastid inverted repeat regions (rpl2, rps7, 3'-rps12, ndhB and associated noncoding regions) are exceptionally slowly evolving (e.g., Graham et al. 2000), and the single-copy genes include slowly and rapidly evolving protein-coding and noncoding regions. The ndhF locus, for example, includes slowly and rapidly evolving portions (Olmsstead and Sweere 1994; Kim and Jansen 1995); the 3'-end of ndhF evolves substantially more rapidly (per nucleotide) than some of the single-copy noncoding regions examined here, such as the trnL–trnF region (Graham and Barrett 2004; Saarala et al. in press). The characters examined represent most of the spectrum of plastid evolutionary rates.

Alignment and Phylogenetic Analysis

We added taxa to previously published alignments (Rai et al. 2003) using criteria set out in Graham et al. (2000). The unaligned total DNA sequence length obtained in monocot exemplars ranges from ca. 12.7 kb in Burmannia L. to ca. 17.9 kb in Coelogyne Lindl., with a mean unaligned length in monocots of ca. 15.9 kb. Some monocots (such as Burmannia) lack sequence data for one or more major regions (see Appendix). The combined alignment is nearly 32 kb in length (31,900 base pairs [BP]), approximately twice the unaligned length of any individual taxon. This size is a result of fairly extensive gaps and/or unalignable regions, including some in nonangiosperm taxa that were part of the overall alignment, but that were not considered in the current study. Those noncoding regions that were too difficult to align were set aside in staggered gapped regions, with the staggered elements largely restricted to single taxa. Because they are unique, such single-taxon elements are parsimony uninformative and have no influence on the analysis. This allowed us to avoid defining character exclusion sets in hard to align regions, a substantial undertaking for a matrix of this size. However, in a subset of cases, we were able to align small blocks of taxa within otherwise unalignable regions. Variable characters within these elements can therefore contribute to tree searching (for a comparable example, see Steane et al. 1999). All gap cells are treated as missing data. We did not attempt to score insertion/deletion (indel) events. Graham et al. (2000) provide an overview of the (relatively limited) utility of indels inferred from several plastid inverted-repeat noncoding regions for inference of deep angiosperm phylogeny.

In total, 5617 aligned sites are potentially parsimony in-
formative in the full 94-taxon data set, or 4798 sites in the monocots alone. We analyzed the matrix using heuristic maximum-parsimony (MP) searches with PAUP* vers.

4.0b10 (Swofford 2002), using tree-bisection-reconnection (TBR) branch swapping, and 100 random-addition replicates. Branch support was estimated using bootstrapping (Felsenstein 1985) with 100 bootstrap replicates. No tree limits were set, and all other settings used were the default ones. Ceratophyllum L. and Chloranthaceae have variable and poorly supported placements in basal angiosperm phylogeny (e.g., Qiu et al. 1999; Graham and Olmstead 2000; Graham et al. 2000; Soltis et al. 2000; Hili et al. 2003; Davis et al. 2004), possibly because of a long, undivided subtending branch (Ceratophyllum) and short “basal” branches in angiosperm phylogeny (Chloranthaceae; e.g., Fig. 3 in Graham et al. 2000). We therefore performed a separate bootstrap analyses with these two taxa excluded, to examine their influence on phylogenetic analysis.

McPherson et al. (submitted) used simulated data sets based on several hypotheses of relationship for Aphyllanthes monspeliensis L., the sole member of Aphyllanthaceae (Asparagales), to demonstrate that there is substantial scope for systematic error in analyses that include this taxon, at least for the plastid regions examined here. Their analyses showed that a subset of hypotheses are recovered rarely when used as model trees, while others may be recovered frequently when they are not used to simulate data (i.e., high potential for type I and II errors). We therefore performed two basic sets of analyses with Aphyllanthes L. included or excluded from consideration (for more on this problematic taxon, see Fay et al. [2000]; Chase et al. 2006; Pires et al. 2006); Ceratophyllum and Chloranthaceae were included in both cases.

To examine whether the rate of molecular evolution shifts in parallel among different codon position classes across the monocot backbone, we set up character sets (CHARSETs) in PAUP* that classify each nucleotide in the protein-coding regions into one of the three codon positions (reading frames were defined with respect to Nicotiana tabacum L. and Ginkgo biloba L. sequences). The short overlap between psbD and psbc was ignored, as individual sites in this region have different codon positions in these two genes. Parsimony-inferred changes from each codon position were noted for each branch in monocot phylogeny, based on one of the most-parsimonious trees (see Fig. 2). We determined whether length estimates for the first two codon positions were correlated across branches with those for the third codon position, using JMP vers. 4.0.1 (SAS Institute 2000).

RESULTS

Most aspects of outgroup relationship are discussed elsewhere (Graham and Olmstead 2000; Graham et al. 2000). Amborella trichopoda Baill. (Amborellaceae) was considered to be the sister group of all remaining angiosperms, following the current consensus result of recent large-scale analyses that included seed-plant outgroups (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999; Soltis et al. 1999; Soltis et al. 2000; Barkman et al. 2000; Graham and Olmstead 2000; Graham et al. 2000; Zanis et al. 2002; Hili et al. 2003; Stefanovic et al. 2004). A recent result placing Amborella Baill. apart from the basal angio-

sperm split (Goremykin et al. 2003, 2004) is likely to be an artifact of extremely low taxon density, combined with exceptionally long branches in the exemplar taxa used by these authors for monocots, all of which are members of Poaceae (Soltis and Soltis 2004; Stefanovic et al. 2004; S. W. Graham unpubl. data).

We typically use family (or higher taxon) names to describe terminal taxa (for a rationale, see Chase et al. 2006). A detailed consideration of relationships recovered in Asparagales is provided in McPherson et al. (submitted); the results presented here for this order focus primarily on two taxa not considered previously for these plastid regions (Agapanthus africanus [Agapanthaceae] and Doryanthes palmeri [Doryanthaceae]). With Aphyllanthes excluded, most random-addition replicate searches found the same pair of trees (Fig. 1, 2) for 96 of 100 random-addition replicates, with a tree length of 35,826 steps, consistency index (CI), including all sites, of 0.362, and retention index (RI) of 0.474. With Aphyllanthes included, a single tree was found (a portion is shown in Fig. 3) in 95 of 100 random-addition replicates, with length 36,276 steps (CI = 0.360; RI = 0.473). The latter tree is topologically equivalent to one of the former trees when Aphyllanthes is pruned from it. We considered well-supported (“strongly supported” or “robust”) branches to have bootstrap support of ca. 95% and more, and poorly supported branches to have ca. 75% or less bootstrap support. By this criterion, most branches of monocot phylogeny are well supported by the current data (Fig. 1). To simplify the presentation of the results, we focus initially on presenting those branches that are poorly to moderately supported, and then address the well-supported relationships. All bootstrap values below refer to analyses with Aphyllanthes excluded, and Ceratophyllum and Chloranthaceae included, unless otherwise stated.

Poorly to Moderately Supported Relationships

We infer the eudicots plus Ceratophyllum to be the sister group of the monocots, and Chloranthaceae to be the sister group of the magnoliids (see also Graham et al. 2000). Both placements have weak bootstrap support (BP) from bootstrap analysis (67% and 60% BP, respectively; Fig. 1A). When Ceratophyllum and Chloranthaceae are removed from phylogenetic analysis, this appears to stabilize the inference of some of the remaining basal angiosperm relationships, with substantially improved bootstrap support for monocot monophyly (100% vs. 84% BP; Fig. 1A), and marginal improvement in support for a close relationship between the monocots and eudicots (73% vs. 67% BP; Fig. 1A). The bootstrap values shown in the monocots (Fig. 1B) are from analyses with Ceratophyllum and Chloranthaceae included. Excluding these two taxa has very little effect on support values within the monocots (data not shown).

A few deep branches of monocot phylogeny are not strongly supported. These include the precise order of splits at the base of Alismatales: Araceae are found to be the sister group to a clade of three sampled alismatid families (86% BP). The arrangement of Poales, Commelinales-Zingiberales, Araceales, and Dasyypogonaceae at the base of the com-

mellinid monocots is also unclear; the two branches supporting the relationship observed here each have <50% BP.
Fig. 1A.—Relationships among the major angiosperm lineages. One of two most-parsimonious trees inferred from a large plastid data set (atpB, ndhB, ndhF; ten photosystem genes, rbcL, rpl2, rps2, 3′-rps12, and various introns and noncoding regions; see text). Numbers above branches are results of bootstrap analyses using all taxa; those below branches are results when *Ceratophyllum* and Chloranthaceae are excluded from consideration. *Aphyllanthes monspeliensis* is excluded from both analyses (but see Fig. 3).

A sister-group relationship between Liliales and Asparagales–commelinids is only weakly supported (70% BP), as is that between Dioscoreales and Pandanales (63% BP).

Melanthiaceae are found to be the sister group to the other families of Liliales sampled here, but with only 39% BP. Of the two members of Asparagales new to this study, Agapanthaceae are clearly a member of a clade of three closely related families (Agapanthaceae, Alliaceae s.s., Amaryllidaceae). However, inferred interrelationships among these three taxa are poorly supported, with only 37% support for one of the arrangements shown here (Alliaceae–Amaryllidaceae; Fig. 1B). An alternative arrangement, Alliaceae–Agapanthaceae, is found on the other most parsimonious tree, with 56% BP (Fig. 3). We find Sparganiaceae–Typhaceae and Bromeliaceae to be successive sister groups of the remaining families of Poales, but with only 45% BP for this arrangement. However, these two lineages are strongly supported as emerging from the base of Poales (Fig. 1B). An arrangement of Flagellariaceae and Restionaceae as successive sister groups of the remaining sampled graminid taxa (Ecdelocoleaceae and Poaceae) has moderately strong support (87% BP).
Strongly Supported Relationships

The remaining relationships are nearly all strongly supported by bootstrap analysis (ca. 95% BP and higher; Fig. 1B). Ignoring the depression in support for monocot monophyly when Ceratophyllum and Chloranthaceae are included (84% vs. 100%; Fig. 1A), Acorus L. is robustly inferred to be the sister group of all other monocots. Alismatales and Petrosaviaceae are (respectively) the next successive sister groups of the remaining monocots. Asparagales and the commelinid monocots are well supported as sister taxa. All orders consisting of more than one family (sensu APG II 2003) are well supported at the taxon samplings used here. Several multiordinal clades are also strongly supported, including Commelinales–Zingiberales, and the commelinid monocots as a whole (Arecales, Commelinales–Zingiberales, Poales, Dioscoreales).

Most relationships within orders are also robustly supported. The (partly) mycoheterotrophic family Burmanniaceae is strongly supported as the sister group of Dioscoreaceae (only these two members of Dioscoreales were sam-
Fig. 2.—One of two most-parsimonious trees (Fig. 1), presented as a phylogram (ACCTRAN optimization). The scale bar indicates 100 changes.
| Chlorophytum | Agavaceae |
|------------|---------|
| Yucca | Aphyllanthaceae |
| Aphyllanthes | Agavaceae |
| Muilla | Themidaceae |
| Muscari | Hyacinthaceae |
| Asparagus | Asparagaceae s.s. |
| Lomandra | Laxmaniaceae |
| Maianthemum | Ruscaceae |
| Allium | Alliaceae |
| Agapanthus | Agapanthaceae |
| Narcissus | Amaryllidaceae |

**Fig. 3.—Region of disruption of bootstrap values in Asparagales when *Aphyllanthes monspeliensis* (Aphyllanthaceae) is included in analysis ("1" = Asparagaceae s.l.; "2" = Alliaceae s.l.). Numbers above branches are bootstrap values with *Aphyllanthes* excluded, those below branches are bootstrap values with it included. The tree shown is a portion of the most-parsimonious tree found with *Aphyllanthes* included (see text); this tree is otherwise identical to one of the two trees found with *Aphyllanthes* excluded (the other is shown in Figs. 1, 2), with comparable bootstrap values for the rest of the tree.

Apart from the basal split in Liliales, all relationships in this order are well supported. Petermanniaceae are distinct from Colchicaceae, whose sister group is Alstroemeriaceae of taxa sampled. Within Alismatales, the inferred relationships of Alismataceae, Butomaceae, and Scheuchzeriaceae are all well supported (the former two are sister groups with respect to the taxa sampled). Within Pandanales, both internal branches are well supported; Velloziaceae are the sister group of the remaining Pandanales, and Pandanaceae and Cyclanthaceae are sister taxa. Pontederiaceae are the sister group of Haemodoraceae of families sampled in Commelinaceae. Apart from the base of Poales, relationships among the remaining taxa of Poales are almost all well supported. Cyperaceae–Xyridaceae are the sister group of Mayaceae. This cyperid clade is the sister group of the graminid families. Ecdyiciocoeaceae and Poaceae are sister groups with respect to the taxa included.

The position of Doryanthaceae (a member of Asparagales not sampled by McPherson et al. submitted) is quite well supported (92% BP). It is inferred to be the sister group of a large clade consisting of Iridaceae, Xeronemataceae, Alliaceae s.l., Asparagaceae s.l., and Xanthorrhoeaceae s.l. The latter three taxa (Alliaceae s.l., Asparagaceae s.l., and Xanthorrhoeaceae s.l.) are indicated with numerals in Fig. 1B, and represent the more inclusive versions of these families in APG II (2003). Iridaceae are sister to a clade consisting of Xeronemataceae, Alliaceae s.l., Asparagaceae s.l. and Xanthorrhoeaceae s.l. A well-supported clade comprised of Ixioliriaceae and Tecophilaeaceae is the sister group to all of these taxa. Bootstrap support for this major part of the backbone of Asparagales [Asparagales base, (Ixioliriaceae–Tecophilaeaceae), (Doryanthaceae, (Iridaceae, (Xeronemataceae, (Xanthorrhoeaceae s.l., (Asparagaceae s.l.–Alliaceae s.l.))))])] is thus robustly supported by our data.

**Effect of Inclusion of Aphyllanthes monspeliensis (Aphyllanthaceae)**

When *Aphyllanthes* is included in the analysis, it is resolved as the sister group of Agavaceae, but with poor support (50% BP; Fig. 3; see also McPherson et al. submitted). Its inclusion does not impinge on other inferred relationships, but it does moderately depress bootstrap values for five of ten branches (by 10–25%) in a local cluster of families corresponding to Asparagaceae s.l., plus Alliaceae s.l. (Fig. 3). There is strong support (100% BP) for the inclusion of *Aphyllanthes* within the major clade consisting of Alliaceae s.l. and Asparagaceae s.l., and bootstrap support outside this local cluster does not appear to be adversely affected by its inclusion (data not shown).

**DISCUSSION**

The plastid gene set considered here has been used for phylogenetic inference of other deep and difficult phylogenetic problems (Graham and Olmstead 2000; Rai et al. 2003) and provides strong support for a number of deep monocot relationships that were poorly supported or unresolved in earlier studies. Average bootstrap support for the monocot portion of the tree from our data is ca. 91% per branch, based on 65 internal branches involving 68 monocot taxa (*Aphyllanthes* excluded). Three-quarters of the internal monocot branches have at least 95% bootstrap support, and 90% have at least 70% bootstrap support. Only six monocot branches on either shortest tree have less than 50% bootstrap
support (Fig. 1B). Some of these improvements in support compared to earlier studies are paralleled in other recent studies, including the two-gene sampling of Tamura et al. (2004), and the multigene studies of Chase et al. (2006) and Pires et al. (2006). A few clades are poorly or moderately well supported here, but are nonetheless congruent with other studies (e.g., the poorly supported position of Melanthiaceae relative to other sampled members of Liliales, which is consistent with that in Tamura et al. [2004], Givnish et al. [2006], and one of the analyses in Chase et al. [2006]).

Assessing the Strength of Support of Deep Monocot Relationships

Sampling error will tend to reduce estimated branch support when too few data are collected per taxon ("not enough characters examined"), or when too few characters define individual branches ("rapid radiations"). However, even when deep branches are short, simulation results using available models of DNA sequence evolution (e.g., Hillis 1998) indicate that maximum parsimony can yield very accurate reconstructions for relatively small amounts of data per taxon (a few thousand kb of DNA sequence data), provided a sufficiently dense taxon sampling is employed. Nonetheless, empirical analyses based on real DNA sequence data sets that are densely sampled and of this order of size (several kb long) are incompletely congruent with each other concerning major and minor points of relationship, and have numerous areas with weak statistical support (e.g., Soltis et al. 2000; Davis et al. 2004; Tamura et al. 2004; Chase et al. 2006; Pires et al. 2006), or exclude many problematic taxa (Tamura et al. 2004). The existence of these uncertain areas (e.g., the composition or relative arrangements of Asparagales, Dioscoreales, Liliales, Pandanales, and the major commelind lineages) provides a continuing impetus for expansion of the amount of data collected per taxon.

Empiricists do not have access to the correct tree of monocot phylogeny, but instead use a variety of statistical methods, such as bootstrap analysis, to assess the degree of confidence in phylogenetic inference (e.g., Hillis and Bull 1993; Felsenstein 2004). There is no clear consensus on what the cutoff should be for considering a clade to be strongly supported ("robust"), and there is some disagreement about why different methods provide different pictures of clade support. Bayesian phylogenetic inference, for example, is thought to suffer from inflated clade posterior probability estimates (e.g., Suzuki et al. 2002; Erixon et al. 2003; Lemmon and Moriarty 2004), and empirical and theoretical studies have demonstrated that bootstrap analysis can provide biased measures of support (e.g., Felsenstein and Kishino 1993; Hillis and Bull 1993). Small differences in clade support are also demonstrable for different resampling methods (jackknife, bootstrap), or different implementation strategies for these methods (Davis et al. 2004). However, the numerical value that we accepted to indicate well-supported clades in bootstrap analysis (ca. 95%) may often be on the conservative side for bootstrap analysis (Felsenstein and Kishino 1993; Hillis and Bull 1993).

A caveat for bootstrap analysis (and other methods for statistical inference of branch support) is that there are conditions, such as the oft-cited phenomenon of "long-branch attraction," under which phylogenetic inferences can be misleading (Felsenstein 1978; Hendy and Penny 1989). Long-branch attraction has been invoked for the placement of several problematic monocot taxa (e.g., the positions of Ixioliriaeae and Aphyllanthaceae in Asparagales; Fay et al. 2000). Parametric phylogenetic methods such as maximum likelihood and Bayesian analysis have been found to be less prone to the distorting effects of long-branch attraction on phylogenetic analysis (e.g., Chang 1996; Swofford et al. 2001), unless there are discordant changes in evolutionary rates among characters ("heterotachy")—in which case, maximum parsimony may be more reliable (Kolaczkowski and Thornton 2004).

The overall rate of molecular evolution in monocot plastid genomes observed here is quite variable among lineages (Fig. 2), in line with previous studies (Wilson et al. 1990; Bousquet et al. 1992; Gaut et al. 1992, 1996). The effect of this rate variation on phylogenetic analysis is unclear. However, we can at least rule out substantial heterotachy between two codon position partitions in our protein-coding regions—the first two vs. third codon positions. The former should predominantly reflect nonsynonymous substitutions, and the latter synonymous ones (e.g., Sanderson et al. 2000). We might therefore expect these to have different substitution dynamics among different lineages, based on theoretical considerations (Kimura 1968, 1983; Ohta 1992). In practice, however, the rate of synonymous and nonsynonymous substitutions are correlated in a broad variety of organisms (e.g., Sharp 1991; Wolfe and Sharp 1993; Akashi 1994; Gaut et al. 1996). We examined the amount of change in the first two vs. third codon positions, using this information as a rough proxy of nonsynonymous vs. synonymous changes. Despite inconstancy in the overall rate of molecular evolution (Fig. 2), changes in these two functionally defined codon-position classes are strongly correlated across the sampled branches of the monocot tree (r = 0.9036; P < 0.0001; Fig. 4).

We will not enter further into the debate about different measures of statistical support. However, a different framework for documenting the reliability of phylogenetic results is to demonstrate that analyses involving the same taxa for different genomic regions depict similar relationships (e.g., Penny et al. 1982). If there is instead well-supported incongruence among different genetic linkage groups, this may reflect deviations of particular gene trees from a consensus organismal pattern (e.g., Maddison 1997) or various phenomena that lead to strong systematic biases in the data (Naylor and Brown 1998). In general, any misleading effects of these on inferences of higher-order relationships using plastid data alone are expected to be rather small (Savolainen et al. 2002; see also Chase et al. 2006). However, determining whether this is indeed the case will at least require the availability of trees from multiple linkage groups (the plastid genome is a single linkage group), with each tree as well supported as possible.

The correlated change we observe among codon positions (Fig. 4) may explain why inferences made from subpartitions of our data (different codon position sets) are generally highly congruent with each other (e.g., McPherson et al. submitted). Our results are also generally highly congruent with other studies (e.g., Davis et al. 2004; Tamura et al. 2004;
Fig. 4.—Demonstration of parallel substitution rates in codon positions 1 and 2, vs. codon position 3 across monocot phylogeny. Data points for each codon set represent individual branch length estimates, summed across relevant sites in protein-coding regions. All terminal and internal branches in monocot phylogeny were considered, including the branch immediately subtending the monocots (Fig. 2). Changes along each branch were computed using ACCTRAN optimization for the relevant character set.

Chase et al. 2006; Pires et al. 2006). Where incongruent, the other studies generally have poor support for the conflicting relationship (see below for several exceptions). Using the phylogenetic data presented here, we have therefore come close to, but not yet fully attained, a phylogenetic backbone of monocot relationships that is well supported across all deep nodes. Other work (in progress) will address more of the details of this plastid framework by adding currently unsampled families.

**Remaining Problematic Relationships in Monocot Phylogeny**

A broad circumscription of the magnoliids ("eumagnoliids") in Soltis et al. (1999) and Solis et al. (2000) includes the monocots, based on a weakly supported clade found in their analysis. In contrast, APG II (2003) was more wary, accepting only Canellales, Laurales, Magnoliales, and Pinales as magnoliids. Although a sister-group relationship between the magnoliids (in a narrow sense) and monocots cannot be ruled out from our bootstrap analysis, classifications that depend on this or other placements should be treated very cautiously. Duvall et al. (2006) use various optimality criteria to analyze four concatenated genes and find several placements of the monocots in angiosperm phylogeny; for example, their maximum parsimony analysis depicts the eudicots as the sister group of monocots plus *Ceratophyllum* (with weak support for these relationships), while their Bayesian analysis places the monocots as the sister group of the magnoliids, with posterior probabilities of 0.97–1.00 for all relevant branches. Given the potential for Bayesian support values to be inflated or misleading (discussed above), the relatively uncertain position of the monocots observed in our analysis, and their variable position in other studies (e.g., Hilu et al. 2003; Davis et al. 2004; Chase et al. 2006), the jury is thus still out on where the monocots belong in flowering-plant phylogeny. We also observed a moderately depressive effect on the strength of support for monocot monophyly with *Ceratophyllaceae* and *Chloranthaceae* included in analysis. Although straightforward to demonstrate, the cause of this intriguing phenomenon is not clear, although it is conceivably associated with whatever is causing the uncertain placement of these two families in all current angiosperm-wide studies.

The monophyly of the monocots, and the position of *Acorus* as the sister group of all other extant monocots are both strongly supported here. *Acorus* has not been uniformly supported as the sister group of all monocots in all recent studies (e.g., Qiu et al. 1999), but when it is not, this generally is a function of poor statistical support in individual studies. However, Duvall and Ervin (2004) and Duvall et al. (2006), documented problems with the nuclear 18S rDNA locus for this taxon. A further exception to this uniform picture is a recent study by Davis et al. (2004) who examined monocot deep phylogeny using *rbcL* and the mitochondrial locus.
atpA. In combined analyses of these two genes, they found a strongly supported sister-group relationship between Acoraceae and a major clade of Alismatales consisting of all the taxa they sampled in this order, except Araceae and Tofieldiaceae. Their analysis of rbcL alone depicted Acorus as the sister group to all other monocots, but that of atpA alone depicted Acoraceae and associated alismatid taxa on very long branches, in a position nested deep in monocot phylogeny, as part of a small clade that included several taxa from Asparagales (Ixioliriaceae, Iridaceae, and a member of Agavaceae; see Fig. 4 in Davis et al. 2004, one of their most-parsimonious trees). This result, and the discordant placement of Acorus in their combined analyses, may be an artifact (see also Chase 2004).

Very short branches are more prone to the effects of sampling error, which may consequently contribute to poor support for some deep internal branches in monocot phylogeny (e.g., Chase et al. 2000, 2006). However, some of the “short” branches referred to in earlier studies still are weakly supported in our expanded data sampling and in comparable recent studies (e.g., Chase et al. 2006), despite not being clearly different in inferred length from neighboring branches that have strong support (compare, for example, the lengths and support for branches subtending Asparagales vs. Dioscoreales–Pandanales; Fig. 1B, 2). Relative unevenness of branch lengths can potentially contribute to erroneous or unstable phylogenetic inference due to long-branch attraction (Felsenstein 1978; Hendy and Penny 1989). However, if there are any problematic long branches remaining in the current data set, their effect here may be primarily to destabilize local estimates of bootstrap support, rather than to lead to erroneous placement of the affected clades. Although it is not always possible to do so, additional taxon sampling may often (although not always; e.g., Ramalha et al. 1998; Poe and Swofford 1999) help ameliorate the effects of long branches by breaking them up.

Nonetheless, disparity in rates of evolution may contribute to difficulties in inferring some relationships accurately. The elevated rate in the grasses commented on by other workers (e.g., Gaut et al. 1992) is evidently not unique to them within Poales (rate elevation compared to other monocots is evident from visual inspection of branches subtending Cyperaceae, Ecdiocoleaceae, Mayaceae, Poaceae, Restionaceae, and Xyridaceae; Fig. 2). However, the branches immediately subtending Bromeliaceae, Sparganiaceae–Typhaceae, and Flagellariaceae are short relative to other members of Poales (Fig. 2). This disparity might explain why the relative positions of Bromeliaceae and Sparganiaceae–Typhaceae are unclear with our current taxon sampling (Fig. 1B), and why the backbone relationships inferred with regards to Flagellariaceae, Restionaceae, and the other graminid families show moderately strong conflict with those inferred in Chase et al. (2006) (they see the reciprocal relative arrangement of Flagellariaceae and Restionaceae).

A few other higher-order groupings are in moderately strong disagreement with clades reported in Chase et al. (2006). These include the relationships among three major clades of Alismatales. We find Tofieldiaceae to be the sister group of a moderately well-supported clade consisting of Araceae and the remaining Alismatales. This result was also observed in some analyses of the three-gene data set of Chase et al. (2000), but conflicts with a well-supported linkage between Tofieldiaceae and alismatid families in Chase et al. (2006). In addition, we observe Mayaceae to be the sister group of Cyperaceae and Xyridaceae of taxa sampled in the cyperid clade, with strong support. Chase et al. (2006) instead find moderate support for a closer relationship between Mayaceae and Cyperaceae. Addressing such conflicts may require improved taxon sampling from the current data, work that is currently in progress.

Fay et al. (2000) posited that addition of more data per taxon would be required to solve the problematic positions of Aphyllantheaceae and Ixioliriaceae in Asparagales, as both had labile positions in their analyses and both are relatively isolated taxa. Despite the addition of four more genes to the complement employed by Fay et al. (2000), the position of Aphyllanthes in the analysis of Chase et al. (2006) and Pires et al. (2006) is still labile and weakly supported. McPherson et al. (submitted) used simulation studies to demonstrate that inference of the phylogenetic position of Aphyllanthes in Asparagales has a high error rate. Uncertainty in the placement of Aphyllanthes in Asparagales s.l. may be a function of it being on a long terminal branch (e.g., Fay et al. 2000; McPherson et al. submitted). Its inclusion in analysis does not appear to affect the underlying relationships inferred for other taxa in this family, although it depresses bootstrap support values in the local clade that includes it (Asparagaceae s.l.–Alliaceae s.l.; Fig. 3).

We inferred strong support for a sister-group relationship between Ixioliriaceae and Tecophilaceae, a result that was seen with poor support in some trees inferred by Fay et al. (2000), and with moderate support in the analysis of Pires et al. (2006). This clade is contradicted by the four- and seven-gene analyses of Chase et al. (2006), and the two-gene analysis by Davis et al. (2004), who instead find strong to moderate support for a sister-group relationship between Ixioliriaceae and Iridaceae. Our analyses robustly resolved this midpoint of the Asparagales backbone (Fig. 1B), but the strongly discordant positions of Ixioliriaceae among studies clearly require further attention. Further taxon sampling (to density levels comparable to Pires et al. 2006) among the relatives of Ixioliriaceae and Aphyllantheaceae may help clarify their phylogenetic status.

Unusual and disparate placements of Burmanniaceae were observed in early analyses of monocot rbcL data. Gaut et al. (1992) found Burmanniaceae nested in what is now referred to as the commelinid monocots, and Duvall et al. (1993) found it nested in Asparagales. Both placements may have been due to the long branch associated with this family, coupled with the relatively limited taxon sampling used in both studies. A relatively long branch also subtends the exemplar taxon from Burmanniaceae (Burmannia capitata; Fig. 2). We find strong bootstrap support for its position as the sister group of Dioscoreaceae, of taxa examined here (Fig. 1B). Some members of Burmanniaceae are completely mycoheterotrophic and achlorophyllous, including some taxa in Burmannia. Although photosynthesis has not been characterized physiologically in B. capitata, this species is chlorophyllous (Imhof 1999). It also has uninterrupted reading frames for all 16 protein-coding regions examined here (ndhF was not examined for this species), including all ten...
photosystem II genes, apB and rbcL. This suggests that these loci produce gene products that are functional in photosynthesis. Nonetheless, partial heterotrophy (if present in this taxon) may contribute to its relatively long subtending branch. Whatever the cause, this long branch could also result in misleadingly high bootstrap support. However, the position we inferred for Burmanniaceae (Fig. 1B) is congruent with other recent phylogenetic studies based on molecular and morphological data, which lends more credence to the idea that this taxon has been correctly placed among the deep branches of monocot phylogeny. Chase et al. (1995b, 2000, 2006) and Caddick et al. (2002) found Burmanniaceae to be nested in a redefined Dioscoreales (APG 1998, APG II 2003), although in contrast to our study they found only poor support for this clade as a whole. Improved taxon sampling in Dioscoreales using the plastid regions sampled here should be valuable for further clarifying relationships among the constituent taxa in this order.

**Contributions to Our Knowledge of Monocot Higher-Order Relationships**

One of our most significant findings is the well-supported placement of Petrosaviaceae in monocot phylogeny. This family (represented here by *Japonolirion osense*; see Cameron et al. 2003) is strongly supported as the sister group of all monocots except *Acorus* and Alismatales, and is thus (apparently) the sole extant descendant of a very early split in monocot phylogeny. This supports the idea that the family should be recognized in its own order (Petrosaviales Takht.) in rank-based classifications (see also Chase 2004). The family’s position was only partly resolved in the plastid-based study of Chase et al. (2000), but the moderate to strong support for the relationship observed here is also seen in the seven-gene sampling of Chase et al. (2006), and the two-gene sampling of Tamura et al. (2004). However, Tamura et al. (2004) used only two outgroup taxa in total, and this sparse sampling may have inflated support levels for their basal monocot inferences.

Additional major findings include a well-supported relationship between Commelinales and Zingiberales (see also Tamura et al. 2004; Chase et al. 2006; Pires et al. 2006), and strong support for a sister-group relationship between Asparagales and the commelinid monocots. The latter relationship has been seen in several other studies with comparable taxon samplings, but with only poor support (e.g., Tamura et al. 2004; Chase et al. 2006). Our most densely sampled major clade is Asparagales. All “unbracketed” families of Asparagales (in the sense of APG II 2003) have now been sampled for the regions considered here, and most of their interrelationships are inferred with strong support (Fig. 1B). The spine of inferred relationships in Asparagales largely parallels other recent studies (Fay et al. 2000; Chase et al. 2006; Givnish et al. 2006; Pires et al. 2006), but is generally better supported here. The orchids are well supported here as a member of Asparagales (see also Tamura et al. 2004; Chase et al. 2006), and their inclusion helps define the deepest nodes in that order (McPherson et al. submitted).

The position of Doryanthaceae (as the sister group of a large group of Asparagales that includes Alliaceae s.l., Asparagaceae s.l., Iridaceae, Xanthorrhoeaceae s.l., and Xeromormataceae; Fig. 1B) is strongly supported in our analysis, and is also found with weak support in Fay et al. (2000) and Pires et al. (2006). Although their precise interrelationships are unclear, Agapanthaceae, Alliaceae s.s., and Amaryllidaceae are linked in a strongly supported clade (Alliaceae s.l.), a relationship seen by Meerow et al. (1999), but without strong support. Fay et al. (2000) found this relationship with strong support, although a reanalysis using standard parsimony by McPherson et al. (2004) found only moderate support using the same data. Pires et al. (2006) also find this clade with strong support, and find good support for Agapanthaceae as the sister group of a clade consisting of Alliaceae s.s. and Amaryllidaceae s.s. The larger clade (Alliaceae s.l.) is unperturbed by the inclusion of Aphyllanthes in our analyses, although its level of support is somewhat adversely affected (Fig. 3). Other aspects of Asparagales phylogeny inferred from our data are discussed in more detail in McPherson et al. (submitted).

Most other relationships that are strongly supported are also largely or completely congruent with the comparable taxonomic sampling by Chase et al. (2006). These include most relationships within Liliales, those among the four sampled families of Pandanales (Cyclanthaceae, Pandanaceae, Stemonaceae, and Velloziaceae) and most of the relationships within the commelinids, including relationships within Commelinales, and most relationships within Poales. Our sequences of *Petermannia* F. Muell. (Petermanniaceae) were not derived from the misidentified sample included in Rudall et al. (2000); see Chase et al. (2006). The distinct position of *Petermannia* in Liliales could be dealt with in the APG system by recognition of Petermanniaceae (Fig. 1B), or perhaps by a substantial expansion in the circumscription of Colchicaceae. Further inference of phylogenetic relationships within these orders will be addressed in more detail elsewhere using improved taxon sampling.

The phylogenetic status of Dasypogonaceae is still unclear, although our data and others (e.g., Chase et al. 2000, 2006; Davis et al. 2004; Givnish et al. 2006) indicate that it is the sole extant representative of a lineage that diverged very early in the history of the commelinid monocots. As in other recent studies, we did not find strong support for any particular arrangement of this family and Arecales, Commelinales–Zingiberales, and Poales at the base of the commelinid clade (Fig. 1B). Various arrangements of these four lineages have been observed (e.g., Davis et al. 2004; Chase et al. 2006; Givnish et al. 2006). Tamura et al. (2004) found moderate support for a sister-group relationship between Poales and Commelinales–Zingiberales, but they did not sample Dasypogonaceae. The relationships observed here, with Dasypogonaceae as the sister group of Commelinales–Zingiberales, and Arecales as the sister group of Poales, have very poor support (38% and 33% MP, respectively). However, if correct, this scenario would either require elevation of Dasypogonaceae to ordinal status, or a substantial reworking of current ordinal boundaries in the commelinid monocots. A sister-group relationship between Commelinales and Zingiberales has strong support, so it would not be acceptable to sink Dasypogonaceae in either order. If Dasypogonaceae, Commelinales, and Zingiberales were instead
combined in a single order, Commelinales is the name at the rank of order with the earliest use (see APG II 2003). Most other arrangements at the base of the commelinid monocots would also require recognition of Dasypogonaceae as a distinct order in the APG system of classification. However, sister-group relationships between Dasypogonaceae and Poales or Arecales have minor bootstrap support (36% BP for the former, 11% BP for the latter), and if either relationship is eventually shown to be correct, Dasypogonaceae could be included in the respective order. While Dasypogonaceae are confirmed here to be part of a deep-diverging split in the commelinid monocots, it is clear that more data are needed to satisfactorily resolve the phylogenetic and taxonomic status of this problematic family.

**Looking Forward**

The data presented here should contribute to an ongoing renaissance in our understanding of monocot systematics and evolution, which was sparked by the morphological work of Rolf Dahlgren and colleagues (e.g., Dahlgren et al. 1985) and further promoted by the first large-scale molecular studies (e.g., Chase et al. 1993). Our data provide a more robust framework for making inferences about molecular and morphological evolution, and should help with the fine-tuning of higher-order monocot classification. The largely robust framework of monocot deep phylogeny presented here, which includes exemplar taxa from all major chlorophyllous clades and many of the most problematic taxa, demonstrates clearly the value of an expanded plastid genomic sampling, by yielding results that are both well supported and congruent with those inferred in other recent studies. We are continuing taxon sampling using the current genomic set in undersampled groups (such as Alismatales, Liliales, and Poales). We need to improve our understanding of relationships within each of these major clades, and to address the remaining weak nodes along the spine of monocot phylogeny (particularly the relative positions of Asparagales–commelinids, Dioscoreales, Liliales, and Pandanales). Improved taxon sampling outside the monocots may also help address the position of the monocots in the angiosperms. We expect that at least some of these problems will require collection of further data per taxon (from the plastid and/or other genomes), and likely substantially more data in a subset of cases (perhaps of the order of whole plastid genome sampling).

Our study demonstrates the benefits of continued expansion in plastid genome sampling for addressing unresolved problems in monocot deep phylogenetics. While there are obvious benefits to examining other genomic regions (e.g., for evidence of interfamilial incongruence, or for finding strongly supported placements of achlorophyllous mycoheterotrophic monocot groups such as Triuridaceae) and morphological characters (e.g., for finding characters that support new groupings implied by molecular data), the time has not yet come to decelerate sampling in the plastid genome for monocot systematics. Indeed, the rate of increase in plastid-based studies at all levels of plant phylogeny has not yet leveled off (e.g., Shaw et al. 2005). This genome has proven to be a workhorse of modern monocot systematics, and we predict that it will likely remain so for the foreseeable future.

**ACKNOWLEDGMENTS**

The authors would like to thank two anonymous reviewers, Lorna Rollins, the staff of the Molecular Biology Service Unit in the Department of Biological Sciences at the University of Alberta, and the many people who generously assisted with tissue collections or provided DNA samples. Jerrold Davis (Cornell University) provided substantial assistance with material. Two anonymous reviewers made useful suggestions for improving this article. This research was funded by NSERC Discovery grants to SWG and by the Universities of Alberta and British Columbia; an NSF grant to RGO (DEB 9727025) also provided support for the early phases of this study. JCP acknowledges support from an NSF-NATO postdoctoral fellowship (DGE 0000658). NSERC and the Alberta Ingenuity Fund provided studentships to JMS, JMZ, and SYS.

**LITERATURE CITED**

AKASHI, H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translation accuracy. Genetics 136: 927–935.

**ANGIOSPERM PHYLOGENY GROUP** [APG]. 1998. An ordinal classification for the families of flowering plants. *Ann. Missouri Bot. Gard.* 85: 531–553.

**ANGIOSPERM PHYLOGENY GROUP II** [APG II]. 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.

BARKMAN, T. J., G. CHENERY, J. R. McNEAL, J. LYONS-WEILER, W. J. ELLISENS, G. MOORE, A. D. WOLFE, AND C. W. DEPAMPHILIS. 2000. Independent and combined analysis of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. *Proc. Natl. Acad. Sci. U.S.A.* 97: 13166–13171.

BOUSQUET, J., S. H. STRAUSS, A. H. DOERKEN, AND R. A. PRICE. 1992. Extensive variation in the evolutionary rate of rbcL genes sequences among seed plants. *Proc. Natl. Acad. Sci. U.S.A.* 89: 7844–7848.

CADDICK, L. R., P. J. RUDALL, P. WILKIN, T. A. HEDDERSON, AND M. W. CHASE. 2002. Phylogenetics of Dioscoreales based on combined analyses of morphological and molecular data. *Bot. J. Linn. Soc.* 138: 123–144.

CAMERON, K. M., M. W. CHASE, AND P. J. RUDALL. 2003. Re-circumscription of the monocotyledonous family Petrosaviaceae to include Japonolirion. *Brittonia* 55: 214–225.

CHANG, J. T. 1996. Full reconstruction of Markov models on evolutionary trees: identifiability and consistency. *Math. Biosci.* 137: 51–73.

CHASE, M. W. 2004. Monocot relationships: an overview. *Amer. J. Bot.* 91: 1645–1655.

———. M. R. DUVALL, H. G. HILLS, J. G. CONRAN, A. V. COX, L. E. EGUHARTE, J. HARTWELL, M. F. FAY, L. R. CADDICK, K. M. CAMERON, AND S. H. STEVENSON. 1995a. Molecular phylogenetics of Lilianae. *Bot. J. Linn. Soc.* 118 (2): 139–170.

———. M. R. DUVALL, H. G. HILLS, J. G. CONRAN, A. V. COX, L. E. EGUHARTE, J. HARTWELL, M. F. FAY, L. R. CADDICK, K. M. CAMERON, AND S. H. STEVENSON. 1995b. Molecular phylogenetics of Lilianae. *Bot. J. Linn. Soc.* 118 (2): 139–170.

CHASE, M. W. 2004. Monocot relationships: an overview. *Amer. J. Bot.* 91: 1645–1655.

———. M. R. DUVALL, H. G. HILLS, J. G. CONRAN, A. V. COX, L. E. EGUHARTE, J. HARTWELL, M. F. FAY, L. R. CADDICK, K. M. CAMERON, AND S. H. STEVENSON. 1995a. Molecular phylogenetics of Lilianae. *Bot. J. Linn. Soc.* 118 (2): 139–170.

———. M. R. DUVALL, H. G. HILLS, J. G. CONRAN, A. V. COX, L. E. EGUHARTE, J. HARTWELL, M. F. FAY, L. R. CADDICK, K. M. CAMERON, AND S. H. STEVENSON. 1995b. Molecular phylogenetics of Lilianae. *Bot. J. Linn. Soc.* 118 (2): 139–170.

FAY, J. Cribb, D. J. RONSTED, T. J. DAVIES, Y. PILLOW, G. PEETERS, O. SEBERG, M. N. TAMURA, C. B. ASMUSSSEN, K. HILL, T. BORSCH, J. I. DAVIS, D. W. STEVENSON, J. C. PIERIS, T. J. GIVNISI, K. J. SYTSMA, M. A. McPHERSON, S. W. GRAHAM, AND H. S. RAI. 2006. Multigene analyses of monocot relationships: a summary, pp. 63–75. In J. T. Columbus, E. A. FRIAR, J. M. PORTER, L. M. PRINCE, AND M. G. SIMPSON [eds.]. Monocots: comparative biology and evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California, USA.
In 1983, The neutral theory of molecular evolution. Cambridge University Press, New York, USA, 353 p.

Kolaczkowski, B., and J. W. Thornton. 2004. Performance of maximum parsimony and likelihood phylogenetic when evolution is heterogeneous. Nature 431: 980–984.

Lemmon, M. R., and E. C. Moriarty. 2004. The importance of proper model assumption in Bayesian phylogenetics. Syst. Biol. 53: 265–277.

Maddison, W. 1997. Gene trees in species trees. Syst. Biol. 46: 523–536.

Mathews, S., and M. J. Donoghue. 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. Science 286: 947–950.

McPherson, A. M., M. F. Fay, M. W. Chase, and S. W. Graham. 2004. Parallel loss of a slowly evolving intron from two closely related families in Asparagales. Syst. Biol. 53: 296–307.

Meierow, A. W., M. F. Fay, C. I. Guy, Q.-B. Li, F. Q. Zaman, and M. W. Chase. 1999. Systematics of Amaryllidaceae based on cladistic analysis of plastid sequence data. Amer. J. Bot. 86: 1325–1345.

Naylor, G. P., and W. M. Brown. 1998. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inferences based on comparisons of sequences. Syst. Biol. 47: 61–76.

Ohta, T. 1992. The nearly neutral theory of molecular evolution. Ann. Rev. Ecol. Syst. 23: 263–286.

Olmscheid, R. G., and J. A. Sweere. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43: 467–481.

Parrish, C. L., K. L. Adams, and J. D. Palmer. 1999. Multigene analyses identify the three earliest lineages of extant flowering plants. Curr. Biol. 9: 1485–1488.

Penne, D. L. R. Foulks, and M. D. Hendy. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. Nature 297: 197–200.

Pires, J. C., J. J. Maureira, T. J. Givnish, K. J. Systma, O. Seberg, G. Petersen, J. I. Davis, D. W. Stevenis, P. J. Rudall, M. F. Fay, and M. W. Chase. 2006. Phylogeny, genome size, and chromosome evolution of Asparagales, pp. 287–304. In J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson [eds.], Monocots: comparative biology and evolution (excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California, USA.

Poe, S., and D. L. Swoford. 1999. Taxon sampling revisited. Nature 398: 299–300.

Polloock, D. D., D. J. Zwickl, J. A. McGuire, and D. M. Hillis. 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. 51: 664–671.

Qu, Y.-L., J. Lee, F. Bernasconi-Quadroni, D. E. Soltis, P. S. Soltis, M. Zanis, E. A. Zimmer, Z. Chen, V. Savolainen, and M. W. Chase. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature 402: 404–407.

Rai, H. S., H. E. O’Brien, P. A. Reeves, R. G. Olmscheid, and S. W. Graham. 2003. Inference of higher-order relationships in the cycads from a large chloroplast data set. Molec. Phylog. Evol. 29: 350–359.

Rannala, B., J. P. Hulsenbeek, Z. Yang, and R. Nielsen. 1998. Taxon sampling and the accuracy of large phylogenies. Syst. Biol. 47: 702–710.

Rudall, P. J., K. L. Stobart, W.-P. Hong, C. A. Furness, G. C. Kiet, and M. W. Chase. 2000. Consider the lilies: systematics of Liliales, pp. 347–357. In K. L. Wilson and D. A. Morrison, [eds.], Monocots: systematics and evolution. CSIRO Publishing, Collingwood, Victoria, Australia.

Saarelka, J. M., P. M. Peterson, R. M. Keane, J. Cayoulette, and S. W. Graham. In press. Molecular phylogenetics of Bromus (Poaceae: Pooidae) based on chloroplast and nuclear DNA sequence data. In J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson [eds.], Monocots: comparative biology and evolution—Poales. Rancho Santa Ana Botanic Garden, Claremont, California, USA.

Sanderson, M. J., M. F. Woiciechowski, J.-M. Hu, T. Sher Khan, and S. G. Brady. 2000. Error, bias and long-branch attraction in data for two chloroplast genes in seed plants. Molec. Biol. Evol. 17: 783–797.

SAS Institute. 2000. JMP vers. 4.0.1. Statistical Discovery Software. Cary, North Carolina, USA.

Savolainen, V., M. W. Chase, N. Salamin, D. E. Soltis, P. S. Soltis, A. J. López, O. Fedirko, and G. J. P. Naylor. 2002. Phylogeny reconstruction and functional constraints in organelar genomes: plastid atpB and rbcL sequences versus animal mitochondria. Syst. Biol. 51: 638–647.

Sharp, P. M. 1991. Determinants of DNA sequence divergence between Escherichia coli and Salmonella typhimurium: codon usage, map position, and concerted evolution. J. Molec. Evol. 33: 23–33.

Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Shiripun, C. T. Winder, E. E. Schilling, and R. L. Small. 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., and P. S. Soltis. 2004. Amborella not a “basal angiosperm”? Not so fast. Amer. J. Bot. 91: 997–1001.

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

—, —, M. E. Mort, M. W. Chase, V. Savolainen, S. B. Hoot, and C. M. Morton. 1998. Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. Syst. Biol. 47: 32–42.

Soltis, P. S., D. E. Soltis, and M. W. Chase. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402: 402–403.

Steane, D. A., R. W. Scotland, D. J. Mabberley, and R. G. Olmscheid. 1999. Molecular systematics of Clerodendrum (Lamiaceae): ITS sequences and total evidence. Amer. J. Bot. 86: 98–107.

Stefanovic, S., D. W. Rice, and J. D. Palmer. 2004. Long branch attraction, taxon sampling, and the earliest angiosperms: Amborella or monocots? B. M. C. Evol. Biol. 4: 35.

Stevenson, D. W., J. I Davis, J. V. Freudenstein, C. R. Hardy, M. P. Simmons, and C. D. Specht. 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 K. L. Wilson and D. A. Morrison, [eds.], Monocots: systematics and evolution. CSIRO Publishing, Collingwood, Victoria, Australia.

Suzuki, Y., G. V. Glazko, and M. Nei. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA. 99: 16138–16143.

Swoford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (and other methods), vers. 4.0b10. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

—, —, P. J. Wadell, J. P. Hulsenbeek, P. G. Foster, P. O. Lewis, and J. S. Rogers. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Syst. Biol. 50: 525–539.

Tamura, M. N., J. Yamashita, S. Fuse, and M. Haraguchi. 2004. Molecular phylogeny of the monocotyledons inferred from combined analysis of plastid matK and rbcL gene sequences. J. Pl. Res. 117: 109–120.

Wilson, M. A., B. Gaut, and M. T. Clegg. 1990. Chloroplast DNA
evolves slowly in the palm family (Arecaceae). *Molec. Biol. Evol.* 7: 303–314.

WOLFE, K. H., AND P. M. SHARP. 1993. Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. *J. Molec. Evol.* 37:441–456.

ZANIS, M. J., D. E. SOLTIS, P. S. SOLTIS, S. MATHEWS, AND M. J. DONOGHUE. 2002. The root of angiosperms revisited. *Proc. Natl. Acad. Sci. U.S.A.* 99: 6848–6853.

ZWICKL, D. J., AND D. M. HILLIS. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* 51: 588–598.
Appendix 1. GenBank accession numbers and vouchers for exemplar monocot taxa. Alternative source details (e.g., living collection accession numbers) are given in place of voucher information in several cases. GenBank numbers of sequences generated by workers outside the Graham laboratory are underlined; several of these were generated from the same or closely related species (the latter are noted; see relevant GenBank accessions for further details). The three grasses (Poaceae) included here (Oryza sativa L., Triticum aestivum L. and Zea mays L.) are from whole plastid genomes deposited on GenBank (accessions NC001320, NC002762 and NC001666.2, respectively). Further details on the regions used are presented in Graham and Olmstead (2000) and McPherson et al. (submitted).

| Taxon | Gene or region | apB | ndhF | psbB, T, N, & pubH |
|-------|----------------|-----|------|-------------------|
|       |                |     |      |                   |
| Arecales |                |     |      |                   |
| Acorales |                |     |      |                   |
| Acorus calamus L. (Denver Botanic Gardens, Colorado; Olmstead 97-149 [DNA]) | AY147593 | AY007657 | AY147499 |
| Alismatales |                |     |      |                   |
| Butomus umbellatus L. (Chase 6414 [K]) | AY147594 | AY147500 | AY147500 |
| Sagittaria latifolia Willd. (Barrett s. n. [TRT]) | AY147606 | AY147788 | AY007471 |
| Scheuchzeria palustris L. (Waterway & Graham 97-60 [ALTA]) | AY147605 | AY147786 | AY007471 |
| Spathiphyllum wallisii Regel (Chase 201 [NCU]) | AY147595 | AY147785 | AY147501 |
| Tofieldia glutinosa (Michx.) Pers. (Stefanovic s. n. [UBC]) | AY147595 | AY147785 | AY147501 |
| Arecales |                |     |      |                   |
| Roystonea princeps (Becc.) Barret (Santiago J-4 [UPR]) | AY147608 | AY147772 | AY147515 |
| Asparagales |                |     |      |                   |
| Agapanthus africanus (L.) Hoffmanns. (McPherson 020420-1 [ALTA]) | AY147608 | AY147772 | AY147515 |
| Alania endlicheri Kunst (Vitt 27706 [ALTA]) | AY147610 | AY147773 | AY147519 |
| Allium textile A. Nelson & J. F. MacBr. (McPherson 990704-79 [ALTA]) | AY147628 | AY147774 | AY147536 |
| and A. senescens Pall. | | | |
| Amerorhizis rotundifolia (Banks ex Pursh) Hultén (McPherson 010610-1 [ALTA]) | AY147623 | AY147783 | AY147531 |
| Aphyllanthes monspeliensis L. (Graham 00-04-08 [ALTA]) | AY147629 | AY147787 | AY147537 |
| Asparagus officinalis L. (McPherson 010819-2 [ALTA]) | AY147630 | AY147788 | AY147538 |
| Asphodelus albus Mill. (Harder 1-000430 [ALTA]) | AY147613 | AY147774 | AY147520 |
| and A. aestivus Brot. | | | |
| Astelia alpina Banks & Sol. ex R. Br. (Chase 1103 [K]) | AY147614 | AY147775 | AY147521 |
| Blandfordia punicea (Labill.) Sweet (Chase 519 [K]) | AY147615 | AY147776 | AY147522 |
| and B. nobilis Sm. | | | |
| Chlorophytum comosum (Thumb.) Jacques (McPherson 000321-1 [ALTA]) | AY147631 | AY147789 | AY147539 |
| and Anthericum ramosum L. | | | |
| Coelogyne cristata Lindl. (McPherson 010921-1 [ALTA]) | AY147616 | AY147777 | AY147523 |
| and C. macdonalldii F. Muell. & Kraenzl. | | | |
| Curculigo capitulata (Lour.) Kuntze (Chase 205 [NCU]) | AY147617 | AY147778 | AY147524 |
| and Hypoxis villosa L. f. | | | |
| Cynaestrum cordifolium Oliv. (Graham & Barrett 2 [TRT]) | AF168902 | U79228 | AY147525 |
| and Kabuyea hostifolia (Engl.) Brummit | | | |
| Cyripedium passerinum Richardson (McPherson 010722-6 [ALTA]) | AY147618 | AY147779 | AY147526 |
| and C. irapeanum La Llave & Lex. | | | |
| Doryanthus palmeri W. Hill ex Benth. (Chase 2837 [K]) | AY146543 | AY465648 | AY465569 |
| Hemerocallis littorea Makino (Chase 3833 [K]) | AY147619 | AY147780 | AY147527 |
| Iris missouriensis Nutt. (McPherson 000707-5a-7 [ALTA]) | AY147620 | AY547003 | AY147528 |
| and I. uguicularis Poir. | | | |
| Isolirion tataricum (Pall.) Herb. (Chase 489 [K]) | AY147621 | AY147781 | AY147529 |
| Lanaria lanata (L.) Druce (Goldblatt 9410 [MO]) | AY147622 | AY147782 | AY147530 |
| Lomandra longifolia Labill. (Vitt 27411 [ALTA]) | AY147632 | AY547004 | AY147540 |
| and Thyssanotus spiniger Brittian | | | |
| Maianthemum racemosum (L.) Link (McPherson 990704-97 [ALTA]) | AY147633 | AY547005 | AY147541 |
| and M. bifolium (L.) F. W. Schmidt | | | |
| Muilla maritima (Torr.) S. Watson (Pires 98-028 [WIS]) | AY147634 | AY147790 | AY147542 |
| Muscari comosum (L.) Mill. (Harder 000419-1 [ALTA]) | AY147635 | AY547006 | AY147543 |
| Narcissus elegans (Haw.) Spach (Barrett 1434 [TRT]) | AY147636 | U79216 | AY147544 |
| Phormium tenax J. R. Forst. & G. Forst. (McPherson 000612-3 [ALTA]) | AY147624 | AY147784 | AY147532 |
| Styrrinchium montanae Greene (McPherson 990704-71 [ALTA]) | AY147625 | AY547008 | AY147533 |
| and S. micranthum Cav. | | | |
| Xanthonrhoa resinosa Pers. (Chase 192 [NCU]) | AY147626 | AY147785 | AY147534 |
| Xerokemia callistemum W. R. B. Oli (Chase 653 [K]) | AY147627 | AY147786 | AY147535 |
| Yucca glauca Nutt. (Provided by D. Hurlburt, University of Alberta; field collection at Onefour, Alberta; Graham 00121 [DNA]) | AY147637 | AY547014 | AY147545 |

and Agave celsii Hook,
Appendix 1. Extended.

| Gene or region (continued) | Length (unaligned) |
|----------------------------|--------------------|
|                           | All nucleotides    |
|                            | Protein-coding     |

| psbD & psbC | psbE, F, L & psbI | rbcL | rpl2 | 3′-rpsL2, rpsL7, ndhB, trnL (CAA) spacer region | trnL (UAA)-trnF (GAA) |
|-------------|-------------------|------|------|-----------------------------------------------|----------------------|
| AF123813    | AF123828          | D28865 | AF123785 | AF123771 | 13.8 kb | 11.0 kb |
| AF147565    | AF149357          | AF147701 | AF147466 | AF147735 | 16.4 kb | 11.5 kb |
| AF147662    | AF147566          | Y14982 | AF147705 | AF147471 | AF147737 | 16.6 kb | 11.3 kb |
| AF147675    | AF147583          | AF149372 | AF147721 | AF147489 | AF147753 | 17.0 kb | 11.3 kb |
| AF147670    | AF147578          | AF149368 | AF147716 | AF147484 | AF147748 | 16.0 kb | 11.2 kb |
| AF147676    | AF147584          | AF149373 | AF147722 | AF147490 | AF147754 | 17.2 kb | 11.3 kb |
| AF147677    | AF147585          | AF149374 | AF147723 | AF147491 | AF147755 | 17.2 kb | 11.3 kb |
| AF147659    | AF147567          | AF149360 | AF147706 | AF147472 | AF147738 | 16.8 kb | 11.3 kb |
| AF147660    | AF147568          | Z77261 | AF147707 | AF147473 | AF147739 | 16.7 kb | 10.7 kb |
| AF147661    | AF147569          | Z73694 | AF147708 | AF147474 | AF147740 | 17.3 kb | 11.5 kb |
| AF147678    | AF147586          | AF149375 | AF147724 | AF147492 | AF147756 | 16.8 kb | 11.3 kb |
| AF147662    | AF147570          | AF149361 | AF147709 | AF147475 | AF147741 | 17.9 kb | 11.3 kb |
| AF147663    | AF147571          | AF149362 | AF147710 | AF147476 | AF147742 | 16.7 kb | 11.3 kb |
| AF147664    | AF147572          | U41572 | AF147711 | AF147477 | AF147743 | 17.1 kb | 11.3 kb |
| AF147665    | AF147573          | AF149363 | AF147712 | AF147478/ | AF147479 | 14.7 kb | 10.6 kb |
| AF1465673   | AF1465596         | AF1465700 | AF1465725 | AF1465624 | AF699128 | 16.3 kb | 11.3 kb |
| AF147666    | AF147574          | AF149364 | AF147713 | AF147480 | AF147744 | 16.1 kb | 11.3 kb |
| AF147667    | AF147575          | AF149365 | AF147714 | AF147481 | AF147745 | 16.8 kb | 11.3 kb |
| AF147668    | AF147576          | AF149366 | AF147482 | AF147746 | AF147747 | 15.0 kb | 9.8 kb |
| AF147669    | AF147577          | AF149367 | AF147715 | AF147483 | AF147747 | 16.2 kb | 11.3 kb |
| AF147679    | AF147587          | L05039 | AF147725 | AF147493 | AF147757 | 17.2 kb | 11.4 kb |
| AF147680    | AF147588          | AF149376 | AF147726 | AF147494 | AF147758 | 17.2 kb | 11.3 kb |
| AF147681    | AF147589          | AF149377 | AF147727 | AF147495 | AF147759 | 17.2 kb | 11.3 kb |
| AF147682    | AF147590          | AF149378 | AF147728 | AF147496 | AF147760 | 17.1 kb | 11.3 kb |
| AF147683    | AF147591          | AF149379 | AF147729 | AF147497 | AF147761 | 17.3 kb | 11.3 kb |
| AF147671    | AF147579          | Z09232 | AF147717 | AF147485 | AF147749 | 17.2 kb | 11.5 kb |
| AF147672    | AF147580          | AF149369 | AF147718 | AF147486 | AF147750 | 16.7 kb | 11.3 kb |
| AF147673    | AF147581          | AF149370 | AF147719 | AF147487 | AF147751 | 17.3 kb | 11.3 kb |
| AF147674    | AF147582          | AF149371 | AF147720 | AF147488 | AF147752 | 17.2 kb | 11.3 kb |
| AF147684    | AF147592          | AF149380 | AF147730 | AF147498 | AF147762 | 17.3 kb | 11.3 kb |
|              |                   |        |        |                   | AF508509 |

References:

- Fay et al. (2000)
Appendix 1. Continued.

| Taxon | Voucher [herbarium] | Gene or region |
|-------|---------------------|----------------|
|       |                     | apB, ndhF, psbB, T, N, & psbH |

**Commelinales**

Hydrothrix gardneri Hook f. (Barrett 1414 [TRT])

Philydrum lanuginosum Banks & Sol. ex Gaertn. (Graham & Barrett 1 [TRT])

Xiphidium caeruleum Aubl. (Provided by M. Simpson, California State University at San Diego; Graham 5.7.94 [DNA])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY147604, U41606, AY147511 |
|       |                     | AY147607, U41622, AY147514 |
|       |                     | AY147611, AF547013, AY147518 |

**Dioscoreales**

Burmannia capitata (Walter ex J. E Gmel.) Mart. (Neyland 958 [MCN])

Dioscorea bulbifera L. (E.P.O Biology, University of Colorado, Boulder; Olmstead 97-151 [DNA])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY147596, —, AY147502 |
|       |                     | AF187059, AY007652, AF123849 |

**Liliales**

Alstroemeria aurea Graham (Sheahan s. n. [K])

Aanticlea elegans (Pursh) Rydb. (McPherson 990704-68 [ALTA])

Lilium superbum L. (Chase 112 [NCU])

Petermannia cirrosa E Muell. (Frederiksen et al. s. n. [C])

Smilax rotundifolia L. (Uhl 92-07 [BH])

Trillium grandiflorum (Michx.) Salisb. (Eades s. n. [ALTA])

Wurmbea pygmaea (Endl.) Benth. (unpubl. data.) (Case 77 [PERTH])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY465546, AY465651, AY465572 |
|       |                     | AY147600, AY147765, AY147506 |
|       |                     | AY465558, AY465585, AY465743 |
|       |                     | AY147599, AF547009, AY147505 |
|       |                     | AY147609, AF547011, AY147516 |

**Pandanales**

Carludovica drudei Mast. (Bailey Hortorium 73:574 [BH])

Pandanus copelandii Merr. (Sherman & Bynum 303 [MO])

Stemona tuberosa Lour. (Rothwell & Stockey 46 [ALTA])

Talbotia elegans Balf. (Rothwell & Stockey 48 [ALTA])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY465545, AY465650, AY465571 |
|       |                     | AY465544, AY465649, AY465570 |
|       |                     | AY465557, AY465561, AY465583 |
|       |                     | AY147609, AF547011, AY147516 |

**Poales**

Ananas comosus (L.) Merr. (Rai 1003 [ALTA])

Cyperus papyrus L. (Alan Yen 174 [private collection])

Ecleiocolea monostachya F. Muell. (Hopper 8531 [KPBGI])

Elegia fenestrata Pillans [New York Botanical Garden accession 1697/95]

Flagellaria indica L. (Bailey Hortorium 77:394 [BH])

Mayaca fiuviatilis Aubl. (Abbott 10526 [FLAS])

Sparganium eurycarpum Engelm. (Hansen s. n. [BH])

Typha angustifolia L. (Graham 1040 [ALTA])

Typha latifolia L. (McPherson 010819-3 [ALTA])

Xyris jupicai Rich. (Goldman 1766 [BH])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY465545, AY465650, AY465571 |
|       |                     | AY465554, AY465649, AY465570 |
|       |                     | AY465534, AY465642, AY465560 |
|       |                     | AY465535, AY438617, AY465561 |
|       |                     | AY465539, AY465645, AY465565 |
|       |                     | AY147605, AY147770, AY147512 |
|       |                     | AY147610, —, AY147517 |
|       |                     | AY465541, AY465657, AY465567 |

**Zingiberales**

Ensete ventricosum (Welw.) Cheesman (Kress 96-5372 [US])

Strelitzia reginae Aiton (O'Brien s. n. [ALTA])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AF168910, AY147769, AY147510 |
|       |                     | AF165540, AY465646, AY465566 |

**Uncertain Ordinal Placement**

Dasypogonaceae

Dasypogon hooker J. R. Drumm. (Chase 430 [NCU] & Conran et al. 917 [PERTH])

Kingia australis R. Br. (Conran et al. 922 [PERTH])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY147603, AY147768, AY147509 |
|       |                     | AY465538, AY465644, AY465564 |

**Petrosaviaceae**

Japonolirion osense Nakai (Chase 3000 [K])

|       |                     |                       |
|-------|---------------------|-----------------------|
|       |                     | AY147598, AY147764, AY147504 |

1 We generated an ndhF sequence here using different source material: Tofieldia glutinosa (Morton & Venn 9282 [ALTA]).

2 Sequences presented in Fay et al. (2000) but not yet available on GenBank.
### Appendix 1. Continued. Extended.

| Gene or region (continued) | Length (unaligned) |
|---------------------------|-------------------|
|                          | All nucleotides   |
|                          | Protein-coding    |
|                          | nucleotides      |

| pubD & psbC | psbE, F, L & psbI | rbcL | rpl2 | 3'-rps12, rps7, ndhB, trnL, (CAA) | apB-rbcL spacer region | trnL (UAA)-trnF (GAA) | | |
|-------------|------------------|------|------|---------------------------------|------------------------|---------------------|---|---|
| AY147650    | AY147558         | U41582 | AY147697    | AY147462                       | —                      | —                   | 14.8 kb 11.2 kb |
| AY147653    | AY147561         | AY147596 | AY147700    | AY147465 AY147734              | —                      | —                   | 16.3 kb 11.3 kb |
| AY147657    | AY147565         | AY149359 | AY147704    | AY147470                       | —                      | —                   | 14.8 kb 11.1 kb |
| AY147641    | AY147549         | AY149347 | AY147688    | AY147453                       | —                      | —                   | 12.7 kb 9.9 kb  |
| AF123819    | AF123834         | D28327 | AF123791    | AF123777                       | —                      | —                   | 13.9 kb 11.0 kb |
| AY465676    | AY465599         | AY465703 | AY465728    | AY465627 AY699131              | AY699225               |                    | 16.8 kb 11.4 kb |
| AY147645    | AY147553         | AY149351 | AY147692    | AY147457 AY699130              | AY699168/AY699169      |                    | 16.3 kb 11.2 kb |
| AF239783    | Y007480          | L12682 | AY007493    | AF238070 AY699129             | Fay et al. (2000)²     |                    | 15.6 kb 11.4 kb |
| AY465689    | AY465612         | AY465714 | AY465741    | AY465640 AY699144             | —                      |                    | 13.2 kb 9.4 kb  |
| AY465685    | AY465608         | AY465710 | AY465737    | AY699140                       | AY699170/AY699171      |                    | 16.6 kb 11.2 kb |
| AY465687    | AY465610         | AY465712 | AY465739    | AY699142                       | AY699166/AY699167      |                    | 16.5 kb 11.3 kb |
| AY465688    | AY465611         | AY465713 | AY465740    | AY699143                       | A. L. Case, Duke U.    |                    | 16.6 kb 11.3 kb |
|             |                  |        |             |                                 |                       |                    |                |
| AY465675    | AY465598         | AY465702 | AY465727    | AY465626                       | —                      | —                   | 15.5 kb 11.3 kb |
| AY465674    | AY465597         | AY465701 | AY465726    | AY465625                       | —                      | —                   | 15.0 kb 11.4 kb |
| AY147644    | AY147552         | AY149350 | AY147691    | AY147456                       | —                      | —                   | 15.4 kb 11.3 kb |
| AY147655    | AY147563         | AY149358 | AY147702    | AY147467                       | —                      | —                   | 14.8 kb 11.1 kb |
| AY147646    | AY147554         | L19977 | AY147693    | AY147458 AY147731             | —                      | —                   | 16.5 kb 11.5 kb |
| AY465664    | AY465587         | Y12966 | AY465717    | AY465615                       | —                      | —                   | 14.8 kb 11.2 kb |
| AY465665    | AY465588         | AY465692 | AY465718    | AY465616                       | —                      | —                   | 15.4 kb 11.1 kb |
| AY465666    | AY465589         | AY465693 | AY465719    | AY465617                       | —                      | —                   | 15.4 kb 11.4 kb |
| AY465667    | AY465590         | AY465694 | AY465720    | AY465618                       | —                      | —                   | 15.4 kb 11.2 kb |
| AY147651    | AY147559         | AY149355 | AY147698    | AY147463                       | —                      | —                   | 13.8 kb 10.5 kb |
| AY465669    | AY465592         | AY465696 | AY465721    | AY465620                       | —                      | —                   | 15.6 kb 11.4 kb |
| AY147656    | AY147564         | —      | —           | AY147468                       | —                      | —                   | 16.4 kb 11.4 kb |
|             |                  |        |             | M91634                         | AY147703              | AY147469 AY147736  | 15.3 kb 11.4 kb |
|             |                  |        |             | —                               | —                      | —                   |                |
|             |                  |        |             | —                               | AY465671              | AY465694 AY465723   | 15.3 kb 11.2 kb |
|             |                  |        |             | AY465698 | AY465622                       | —                      | —                   | 15.3 kb 11.2 kb |
|             |                  |        |             | —                               | AY465672              | —                   |                |
|             |                  |        |             | —                               | AY465621              | —                   |                |
|             |                  |        |             | —                               | AY465673              | —                   |                |
|             |                  |        |             | —                               | AY465674              | —                   |                |
|             |                  |        |             | AY465675 | AY147557 AY149354 | AY147696 AY147461 | AY147733           | 16.4 kb 11.3 kb |
|             |                  |        |             | AY465670 | AY465593 | AY465697 | AY465722 | AY465621 | — | — | — | 13.9 kb 11.1 kb |
|             |                  |        |             | — | AY465668 | AY465591 | AY465695 | — | AY465619 | — | — | — | 13.9 kb 10.4 kb |
|             |                  |        |             | — | AY465643 | AY147551 | AY149349 | AY147690 | AY147455 | — | — | — | 14.1 kb 11.1 kb |