# Overcoming Deep Roots, Fast Rates, and Short Internodes to Resolve the Ancient Rapid Radiation of Eupolypod II Ferns 

Carl J. Rothfels ${ }^{1, *}$, Anders Larsson ${ }^{2}$, Li-Yaung $\mathrm{KuO}^{3}$, Petra Korall ${ }^{2}$, Wen-Liang Chiou ${ }^{4}$, and Kathleen M. Pryer ${ }^{1}$<br>${ }^{1}$ Department of Biology, Duke University, Box 90338, Durham, NC 27708, USA; ${ }^{2}$ Systematic Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden; ${ }^{3}$ Institute of Ecology and Evolutionary Biology, National Taizan University, No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan; and ${ }^{4}$ Botanical Garden Division, Taiwan Forestry Research Institute, 53 Nan-hai Road, Taipei 10066, Taiwan;<br>*Correspondence to be sent to: Department of Biology, Duke University, Box 90338, Durham, NC 27708, USA; E-mail: carl.rothfels@duke.edu.<br>Carl J. Rothfels, Anders Larsson and Li-Yaung Kuo have contributed equally to this work.

Received 29 December 2010; reviews returned 7 April 2011; accepted 23 August 2011
Associate Editor: Roberta J. Mason-Gamer


#### Abstract

Backbone relationships within the large eupolypod II clade, which includes nearly a third of extant fern species, have resisted elucidation by both molecular and morphological data. Earlier studies suggest that much of the phylogenetic intractability of this group is due to three factors: (i) a long root that reduces apparent levels of support in the ingroup; (ii) long ingroup branches subtended by a series of very short backbone internodes (the "ancient rapid radiation" model); and (iii) significantly heterogeneous lineage-specific rates of substitution. To resolve the eupolypod II phylogeny, with a particular emphasis on the backbone internodes, we assembled a data set of five plastid loci (atpA, atpB, matK, rbcL, and trnG-R) from a sample of 81 accessions selected to capture the deepest divergences in the clade. We then evaluated our phylogenetic hypothesis against potential confounding factors, including those induced by rooting, ancient rapid radiation, rate heterogeneity, and the Bayesian star-tree paradox artifact. While the strong support we inferred for the backbone relationships proved robust to these potential problems, their investigation revealed unexpected model-mediated impacts of outgroup composition, divergent effects of methods for countering the star-tree paradox artifact, and gave no support to concerns about the applicability of the unrooted model to data sets with heterogeneous lineage-specific rates of substitution. This study is among few to investigate these factors with empirical data, and the first to compare the performance of the two primary methods for overcoming the Bayesian star-tree paradox artifact. Among the significant phylogenetic results is the near-complete support along the eupolypod II backbone, the demonstrated paraphyly of Woodsiaceae as currently circumscribed, and the well-supported placement of the enigmatic genera Homalosorus, Diplaziopsis, and Woodsia. [Moderate data; outgroup rooting; Phycas; phylogeny evaluation; rate heterogeneity; reduced consensus; star-tree paradox; Woodsiaceae.]


A classic problem in phylogenetics is the reconstruction of "ancient rapid radiations," broadly defined as evolutionary histories where long branches are intercalated among a series of short backbone internodes (see Fig. 1; Whitfield and Lockhart 2007; Jian et al. 2008). Accurately resolving such topologies is a welldocumented challenge for phylogenetic inference (Gaut and Lewis 1995; Huelsenbeck 1995; Jackman et al. 1999; Anderson and Swofford 2004; Wang et al. 2009) and is also of considerable practical importance-this ancient rapid radiation model is a prominent feature of many phylogenetic problems (Whitfield and Lockhart 2007). Furthermore, the ancient rapid radiation pattern rarely exists unaccompanied; rather, it tends to coincide with other well-recognized analytical challenges. First, the phylogenetic root is often long with respect to ingroup branches (Fig. 1; Bergsten 2005; Schuettpelz and Hoot 2006). Because signal deteriorates along phylogenetic branches (in a likelihood framework), long branches are less likely than short ones to strongly affix to any single point in the topology (Wheeler 1990; Swofford et al. 1996; Huelsenbeck et al. 2002). Furthermore, although the monophyly of the ingroup and of all ingroup relationships may be fully supported, uncertainty in the placement of the root may nonetheless reduce apparent support for relationships among
ingroup clades when one uses consensus-based measures to assess support (Wilkinson 1996; Roberts et al. 2009). Second, lineage-specific heterogeneity in rates of substitution is common, making "fast" taxa particularly difficult to place (Fig. 1; Felsenstein 1978; Hillis and Bull 1993; Soltis et al. 1999; Takezaki and Gojobori 1999; but see Ho and Jermiin 2004; Nickrent et al. 2004; Drummond et al. 2006). Finally, the presence of both very short and very long branches-regardless of their topological arrangement-poses additional challenges. While long-branch attraction has been well characterized (Felsenstein 1978; Anderson and Swofford 2004; Bergsten 2005), other branch-length related inconsistencies are just beginning to attract attention (Lewis et al. 2005; Yang and Rannala 2005; Yang 2008; Marshall 2009; Roberts et al. 2009; Brown et al. 2010).

One option for tackling problems associated with reconstructing ancient rapid radiations is to amass char-acter-rich (often genome-scale) data sets (e.g., Pereira and Baker 2006; Hallstrom et al. 2007; Jian et al. 2008; Wang et al. 2009; Regier et al. 2010). However, the specific challenges inherent to this sort of phylogenetic problem are not necessarily amenable to resolution by greatly expanded character data (Philippe et al. 2011). Rather, increasing character data can yield increasingly strong support for erroneous relationships, especially


Figure 1. Challenges inherent in resolving the eupolypod II phylogeny. Eupolypods II phylogram modified from Schuettpelz and Pryer (2007), in (a) unrooted and (b) rooted form. (i) Outgroup taxa are on long branches. (ii) Backbone internodes are very short, suggesting an "ancient rapid radiation." (iii) The ingroup is marked by significant heterogeneity in rates of evolution, with the members of Aspleniaceae on much longer branches than other eupolypod II taxa.
in cases of branch-length variation such as is inherent in the long-root, short internode, and rate heterogeneity features common under the ancient rapid radiation model (Gaut and Lewis 1995; Soltis et al. 2004; Bergsten 2005; Philippe et al. 2005; Steel and Matsen 2007; Whitfield and Lockhart 2007; Rannala and Yang 2008; Susko 2008; Yang 2008). Here, we focus on resolving an ancient rapid radiation, that of the eupolypods II clade, using moderate amounts of character data but a strongly expanded taxon sample (for a recent application of this "moderate data/many taxa" approach, see Parfrey et al. 2010). This fern clade has resisted elucidation by both morphological and molecular data (Ching 1964a; Sledge 1973; Smith 1995; Sano et al. 2000a; Smith et al. 2006; Schuettpelz and Pryer 2007; Wei et al. 2010; Kuo et al. 2011), and previous molecular studies indicate that it exhibits all of the analytical challenges outlined above (see Fig. 1).

The Eupolypods II, together with its sister group, Eupolypods I, comprise the large eupolypod clade, which encompasses two-thirds of living fern species (Fig. 2; Pryer et al. 2004; Schneider et al. 2004b; Smith et al. 2006). The ancestors of Eupolypods I and II diverged from each other in the Early Cretaceous (Pryer et al. 2004; Schneider et al. 2004b; Schuettpelz and Pryer 2009). The eupolypod II clade started to diversify shortly thereafter (its crown group is approximately 100 million years old; Schuettpelz and Pryer 2009) and has subsequently grown into a lineage-rich clade comprising nearly $30 \%$ of extant fern diversity. Eupolypods II includes some of the most familiar groups of ferns (the lady ferns, ostrich fern, sensitive fern, marsh ferns, and spleenworts), as well as some of the most speciesrich genera: Thelypteris s. lat. ( $\sim 950$ species); Asplenium ( $\sim 700$ species); Diplazium ( $\sim 350$ species); Athyrium ( $\sim 220$ species); and Blechnum s. lat. ( $\sim 150$ species).

The eupolypod II clade is cosmopolitan in distribution, with the subgroups primarily temperate to tropical, and the larger subclades each well represented in both areas. However, many of the phylogenetically enigmatic taxa in this clade are limited to the Himalayas or Southeast Asia, and critical members of several genera are rare and/or infrequently collected. This pattern of rarity and endemism, in conjunction with the richness and geographical breadth of the clade as a whole, is undoubtedly a contributing factor to the incomplete sampling of these ferns in previous phylogenetic studies.

Not surprisingly, given the clade's size and age, eupolypod II taxa are morphologically disparate and seemingly incohesive. However, early workers did tend to recognize the close affinities among many of the taxa in this clade, although frequently with members of Eupolypods I interdigitated among them (Holttum 1947; Sledge 1973; Mickel 1974; Tryon and Tryon 1982). The cohesiveness of the Eupolypods II started to become apparent with the earliest applications of molecular phylogenetic techniques to ferns (Wolf et al. 1994; Hasebe et al. 1995; Sano et al. 2000a) and has been strongly supported in recent broad studies (Schneider et al. 2004b; Schuettpelz and Pryer 2007; Kuo et al. 2011). None of these studies, however, found support for the backbone relationships within Eupolypods II, and only Kuo et al. (2011) attempted to sample its major lineages. It remains one of the few areas of the fern tree-of-life where the backbone relationships remain elusive (Smith et al. 2006; Schuettpelz and Pryer 2007).

Our approach to resolving the eupolypod II phylogeny couples a considerably expanded taxon sample with moderate character sampling. Our objectives include identifying well-supported major (approximately "family-level") clades and determining the backbone relationships among these clades. Given the anticipated


Figure 2. Broad phylogeny of ferns. Approximate number of species per clade is given in parentheses. Modified from Smith et al. (2006).
phylogenetic challenges and potential for artifacts in our data, we explicitly evaluate our phylogenetic hypothesis against these analytical pitfalls, placing strong emphasis on the use of the reduced consensus technique (Wilkinson 1996) to isolate the effects of signal weakness from those of signal conflict (e.g., Wiens 2003; Cobbett et al. 2007). Our study aims for a comprehensive and well-supported phylogeny of this important group of ferns, and for novel inferences about the behavior of our choice of methods, gleaned from their performance on this data set.

## Materials and Methods

## Taxon Sampling

We selected an ingroup of 67 species, intended to maximize our capture of the deep divergences (Zwickl and Hillis 2002) within Eupolypods II. Decisions for inclusion were based on data from previous molecular (Gastony and Ungerer 1997; Murakami et al. 1999; Sano et al. 2000a, 2000b; Smith and Cranfill 2002; Tzeng 2002; Cranfill and Kato 2003; Wang et al. 2003; Schneider et al. 2004a; Schuettpelz and Pryer 2007; Kuo et al. 2011) and morphological studies (Brown 1964; Kato 1975a, 1975b, 1977, 1979, 1984; Kato and Darnaedi 1988; Sano et al. 2000b; Wang et al. 2004; Wang 2008). While 67 species is sparse coverage of the approximately 2600 species in the clade, our utilization of past results (both molecular and morphological/taxonomic) in selecting our taxon sample allows us a high degree of confidence that we have captured a great majority of the deepest branches, if not all of them. Most unsampled taxa are known to be deeply nested in crown clades, especially in the large genera Asplenium, Athyrium, Blechnum, Diplazium, and Thelypteris (sensu lato).

Wherever possible, we included generic and familial types, to facilitate future taxonomic revisions. Based on data from Schuettpelz and Pryer (2007) and Liu et al. (2007), our broad outgroup sample included 10 representatives from the sister group to the Eupoly-
pods II (Eupolypods I, see Fig. 2). To better evaluate the effect of uncertainty in outgroup placement on the ingroup topology and to better understand the divergence between Eupolypods I and II, we also included two representatives from each of the two potentially successive sister groups to the Eupolypods (Notholaena and Cryptogramma from Pteridaceae; Dennstaedtia and Pteridium from Dennstaedtiaceae; see Fig. 2 and Schuettpelz and Pryer 2007). Our total sample has 81 terminal taxa (Appendix 1).

## Amplification and Sequencing

DNA was extracted from silica-dried or herbarium material, using either (i) a modified Carlson-Yoon protocol ( $<0.01 \mathrm{~g}$ dried plant material, silica beads, $750 \mu \mathrm{~L}$ Carlson buffer, and $20 \mu \mathrm{~L}$ mercaptoethanol added to a 2-mL tube and ground for 45 s using a Mini-Beadbeater (BioSpec Products), followed by incubation at $65^{\circ} \mathrm{C}$ for 45 min ; Yoon et al. 1991) or (ii) the protocol of Pryer et al. (2004) or (iii) the protocol of Kuo et al. (2011). For material extracted under the Carlson-Yoon protocol, the extracted DNA was purified by Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare).

Five plastid loci were selected for analysis: $\operatorname{atp} A, \operatorname{atp} B$, matK, $r b c L$, and the $\operatorname{trnG}-\operatorname{trn} R$ intergenic spacer (henceforth " $\operatorname{trn} G-R$ "). All loci, except for matK, were amplified according to either the "standard" or "difficult" reaction protocols (below) depending on the source of the material (standard for most extractions; difficult for those from herbarium specimens greater than 10 years old), using the primers listed in Table 1. The standard amplification reaction used standard Taq polymerase with the following cycle: a 3 min initial denaturation at $95^{\circ} \mathrm{C}$, followed by 35 cycles of 30 s denaturation at $95^{\circ} \mathrm{C}, 1$ min annealing at $54^{\circ} \mathrm{C}$, and 2 min elongation at $72^{\circ} \mathrm{C}$, followed by a final elongation of 10 min at $72^{\circ} \mathrm{C}$. The difficult amplification reaction, using Phusion High Fidelity DNA Polymerase (Finnzymes), was 1 min initial denaturation at $98^{\circ} \mathrm{C}$, followed by 35 cycles of 10 s denaturation at $98^{\circ} \mathrm{C}, 30 \mathrm{~s}$ annealing at $58^{\circ} \mathrm{C}$, and 1 min elongation at $72^{\circ} \mathrm{C}$, followed by a final elongation of 8 min at $72^{\circ} \mathrm{C}$. Amplification of all matK sequences followed the protocol of Kuo et al. (2011).

PCR products from Carlson-Yoon extractions were purified using MultiScreen Plates in a vacuum manifold (Millipore) and sequenced by Macrogen Inc. (South Korea). For material extracted under the protocol of Pryer et al. (2004), each PCR product was cleaned using $0.5 \mu \mathrm{~L}$ of exonuclease I and $1 \mu \mathrm{~L}$ of Shrimp Alkaline Phosphatase (USB, Cleveland, OH); reaction tubes were incubated at $37^{\circ} \mathrm{C}$ for 15 min and then heated to $80^{\circ} \mathrm{C}$ for 15 min to inactivate the enzymes, prior to sequencing on a Applied Biosystems 3730 xl at the Duke IGSP Sequencing Facility (Duke University). Material extracted under the protocol of Kuo et al. (2011) was sequenced at Genomics (Taipei, Taiwan). We completed our sampling with an additional 100 previously published sequences from GenBank (Appendix 1).

TABLE 1. Primers used in amplification and sequencing

| Locus | Primer | Usage |  | Sequence (5'-3') | Reference |
| :---: | :---: | :---: | :---: | :---: | :---: |
| atp $A$ | ESATPA535F | F | S | ACAGCAGTAGCTACAGATAC | Schuettpelz et al. (2006) |
| $\operatorname{atp} A$ | ESATPA557R | R | S | ATTGTATCTGTAGCTACTGC | Schuettpelz et al. (2006) |
| atpA | ESATPA856F | F | S | CGAGAAGCATATCCGGGAGATG | Schuettpelz et al. (2006) |
| $a t p A$ | ESATPA877R | R | S | CATCTCCCGGATATGCTTCTCG | Schuettpelz et al. (2006) |
| $a t p A$ | ESATPF412F | F | A, S | GARCARGTTCGACAGCAAGT | Schuettpelz et al. (2006) |
| $a t p A$ | ESTRNR46F | R | A, S | GTATAGGTTCRARTCCTATTGGACG | Schuettpelz et al. (2006) |
| $\operatorname{atpB}$ | ATPB910R | R | S | TTCCTGYARAGANCCCATTTCTGT | Pryer et al. (2004) |
| $\operatorname{atpB}$ | ESATPB172F | F | A, S | AATGTTACTTGTGAAGTWCAACAAT | Schuettpelz and Pryer (2007) |
| $\operatorname{atpB}$ | ESATPB701F | F | S | TATGGTCAGATGAATGAACC | Schuettpelz and Pryer (2007) |
| $\operatorname{atp} B$ | ESATPE45R | R | A, S | ATTCCAAACWATTCGATTWGGAG | Schuettpelz and Pryer (2007) |
| matK | ASPmatKrLVV ${ }^{\text {a }}$ | R | A, S | TTCGTGTCCRTAAAACAACCAA | This study |
| matK | Athymat rHTY ${ }^{\text {b }}$ | R | A, S | CACACRAAGTTTYGTAYGTGTGAA | This study |
| matK | BLEmatKrDVP ${ }^{\text {c }}$ | R | A, S | AATAGATGTRRAAATGGCACATC | This study |
| matK | CYSmatKrCGK ${ }^{\text {d }}$ | R | A, S | AACTGAGTRACCTTTCCACACG | This study |
| matK | DematKrHTY ${ }^{\text {e }}$ | R | A, S | ACGAAGTTTTGTACGTGTGAA | This study |
| matK | DimatKrTYK ${ }^{\text {f }}$ | R | A, S | CCACACRAAGTTTTGTACGTGT | This study |
| matK | DIPZmatKrDSIg | R | A, S | GTCCATAAAACTACCAATATCGAATC | This study |
| matK | EuIImatKrHLL ${ }^{\text {h }}$ | R | A, S | GTGARAAACYATCCTTARTAGATG | Kuo et al. (2011) |
| matK | EuIImatKfSIH | F | A, S | TCRAAAATBTCRCAGTCTATTCATTC | This study |
| matK | FERmatKfEDR | F | A, S | ATTCATTCRATRTTTTTATTTHTGGARGAYAGATT | Kuo et al. (2011) |
| matK | FERmatKrAGK | R | A, S | CGTRTTGTACTYYTRTGTTTRCVAGC | Kuo et al. (2011) |
| matK | FERNchlBfYAA | F | A | GATGTRAYGTATGCRGCYAAAGA | Kuo et al. (2011) |
| matK | FERNrps16fQCGR | F | A | CRMTRTGGTAGRAAGCAAC | Kuo et al. (2011) |
| matK | FERNrps16fSRQE | F | A | CCCGRMRAGAAGGGARAG | Kuo et al. (2011) |
| matK | ONOmatKrIRD ${ }^{\text {i }}$ | R | A, S | GTRGAAATGGCACATCCCTAAT | This study |
| matK | PtmatKrIHY ${ }^{\text {j }}$ | R | A, S | TTTCTMYATCTTSCRTARTGAAT | Kuo et al. (2011) |
| matK | THEmatKrVRL ${ }^{\text {k }}$ | R | A, S | TCGACGAAACAAGCGAAC | This study |
| matK | WOOmatKrVRL ${ }^{1}$ | R | A, S | TCKACGAAACAGGCGAAC | This study |
| $r b c L$ | ESRBCL1361R | R | A, S | TCAGGACTCCACTTACTAGCTTCACG | Schuettpelz and Pryer (2007) |
| $r b c L$ | ESRBCL1F | F | A, S | ATGTCACCACAAACGGAGACTAAAGC | Schuettpelz and Pryer (2007) |
| $r b c L$ | ESRBCL663R | R | S | TACRAATARGAAACGRTCTCTCCAACG | Schuettpelz and Pryer (2007) |
| $r b c L$ | PKRBCL556F | F | S | GGTAGRGCYGTCTAYGAATGYC | This study |
| $t r n G-R$ | trnG1F | F | A, S | GCGGGTATAGTTTAGTGGTAA | Nagalingum et al. (2007) |
| $\operatorname{trnG-R}$ | trnG353F | F | S | TTGCTTMTAYGACTCGGTG | Korall et al. (2007) |
| $t r n G-R$ | trnG63R | R | S | GCGGGAATCGAACCCGCATCA | Nagalingum et al. (2007) |
| $t r n G-R$ | trnGR43F1 | F | S | TGATGCGGGTTCGATTCCCG | Nagalingum et al. (2007) |
| $t r n G-R$ | trnR22R | R | A, S | CTATCCATTAGACGATGGACG | Nagalingum et al. (2007) |

Notes: $\mathrm{F}=$ forward; $\mathrm{R}=$ reverse; $\mathrm{A}=$ used in amplifications; $\mathrm{S}=$ used in sequencing reactions. While most primers were applied across the phylogeny, superscripts following primer names indicate lineage specificity: ${ }^{\mathrm{a}}$ Aspleniaceae; ${ }^{\mathrm{b}}$ athyriids; ${ }^{\mathrm{C}}$ Blechnaceae; ${ }^{\mathrm{d}}$ Cystopteris/Gymnocarpium; ${ }^{\mathrm{e}}{ }_{\text {depariids; }}{ }^{\mathrm{f}}$ diplaziids; $\mathrm{g}_{\text {Diplaziopsis/Homalosorus; }}{ }^{\mathrm{h}}$ Rhachidosorus; ${ }^{\mathrm{i}}$ Onocleaceae; ${ }^{\mathrm{j}}{ }^{\mathrm{F}}$ teridaceae; ${ }^{\mathrm{k}}$ Thelypteridaceae; ${ }^{1}$ Woodsia and allies.

## Alignment and Tree Search

Sequences were manually aligned in Mesquite 2.72 (Maddison and Maddison 2009). Indels (limited to matK, $\operatorname{trn} G-R$, and the ends of the atpA alignment) were assessed by eye, and ambiguously aligned areas were excluded prior to phylogenetic analysis. Any gaps associated with unambiguous indel regions were treated as missing data. In one rapidly evolving region of the $\operatorname{trn} G-R$ alignment, we were unable to confidently align the Pteridaceae sequences to those of the other taxa. In order to retain this otherwise unambiguous region, we excised those portions of the Pteridaceae sequences, replacing them with question marks.

To evaluate congruence among our loci, we performed maximum likelihood (ML) tree searches on 1000 bootstrap data sets for each locus individually, under a GTR $+\mathrm{I}+\mathrm{G}$ model using the default settings in Garli v1.0.695 (Zwickl 2006, see SI Table 1, available from http://www.sysbio.oxfordjournals.org). The majority rule bootstrap consensus trees from each locus were manually compared and examined for strongly supported conflicts (Mason-Gamer and Kellogg 1996),
after which we concatenated the full data with abioscript (Larsson 2010), producing a single annotated fivelocus data set, with excluded regions removed. This alignment is largely complete ( 361 of the possible 405 sequences are present, for an average of 4.5 loci per terminal taxon) and contains $13.3 \%$ missing data and 6595 characters, of which 3641 are variable (Table 2). Our alignment is available on TreeBase (accession number S11464); the full length of all newly generated sequences (including any portions excluded prior to analysis) are deposited in GenBank (see Appendix 1).

To obtain a point estimate of the phylogeny, we performed 10 tree bisection-and-reconnection heuristic searches of the concatenated (unpartitioned) data, each from a different random-addition-sequence starting tree, under ML using a GTR $+\mathrm{I}+\mathrm{G}$ model as implemented in PAUP* 4.0610 for Unix (Swofford 2002). The values for the exchangeability parameters, base frequencies, gamma shape parameter, and proportion of invariant sites were fixed at their ML values as optimized under a Garli 0.951 (Zwickl 2006) tree search, using default genetic algorithm and termination settings.

TABLE 2. Data set statistics

| $\underline{\text { Data set }}$ | Taxa | Sites | Variable sites | Missing data (\%) | $\begin{gathered} \text { Mean } \\ \text { MLBS (\%) } \\ \hline \end{gathered}$ | Bipartitions with |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  | $\begin{gathered} \geqslant 70 \% \\ \text { MLBS (\%) } \\ \hline \end{gathered}$ | $\begin{gathered} \geqslant 0.95 \\ \mathrm{PP}(\%) \\ \hline \end{gathered}$ |
| $\operatorname{atp} A$ | 74 | 1706 | 809 | 3.1 | 79 | 69 | 69 |
| atpB | 69 | 1278 | 507 | 0.8 | 74 | 58 | 62 |
| matK | 75 | 1377 | 1142 | 9.0 | 84 | 75 | 74 |
| $r b c L$ | 78 | 1308 | 417 | 1.0 | 76 | 68 | 68 |
| $\operatorname{trnG-R}$ | 65 | 926 | 706 | 3.4 | 81 | 71 | 71 |
| Combined | 81 | 6595 | 3641 | 13.3 | 92 | 91 | 90 |

Notes: Missing data include both uncertain bases (e.g., ?, R, Y) and gaps (-). Support values are listed as MLBS or Bayesian PPs.

We assessed support using ML bootstrapping and Bayesian inference. For the ML bootstrapping, we performed 5000 replicates in each of PAUP* 4.0 b10 for Unix (Swofford 2002), Garli 1.0.695 (MPI parallel version; Zwickl 2006), and RAxML v7.2.6 (Stamatakis 2006). The PAUP* runs were performed with the parameters optimized as above, reconnection limit set to eight ("reconlim $=8^{\prime \prime}$ ), and with only a single random-additionsequence per bootstrap replicate. In Garli, we ran 5000 bootstrap replicates on the concatenated data, under a GTR+I+G model using the default genetic algorithm and termination settings. In RAxML, we ran 5000 bootstrap replicates on the data partitioned by locus, with each locus assigned a GTR+G model. For Bayesian inference, we ran four runs of four chains each (one cold; three heated), for 15 million generations, under a partitioned GTR+I+G model in the parallel version of MrBayes v3.1.1 (Ronquist and Huelsenbeck 2003; Altekar et al. 2004). Each of the five loci was assigned its own partition, with substitution parameters unlinked among partitions, and branch lengths linked (with a proportionality parameter to account for rate heterogeneity among partitions); the posterior was sampled every 1000 generations. Visual inspection in AWTY (Wilgenbusch et al. 2004; Nylander et al. 2008) revealed that the runs converged within the first 500,000 generations. To be conservative, we excluded the first 2 million generations of each run as burn-in prior to summarizing the posterior. The posterior thus comprised 52,000 samples ( 13,000 post-burn-in samples from each of the four runs).

## Phylogeny Evaluation

As stated above, earlier studies (e.g., Schuettpelz and Pryer 2007) indicate that the eupolypod II phylogeny is likely to include several key challenges for phylogenetic inference, specifically a series of long branches among very short backbone internodes (an ancient rapid radiation), lineage-specific rate heterogeneity, and a distantly related outgroup. Given these concerns, we sought to explicitly evaluate our topology and support values against these potential artifacts, with particular emphasis on the support values along the backbone of the tree.

Our approach to phylogeny evaluation involved permutations of both the models and the implementation of those models (i.e., programs). The models were deliberately selected according to their varying degrees
of susceptibility for each of the risk factors in question, in an attempt to isolate potential model-based biases. The choice to additionally vary the programs used was in part due to constraints of implementation-no single program offered all the models we wished to compare. This approach has the added benefit of demonstrating a further level of robustness: if our phylogenetic results are insensitive to both the differing models and the myriad of incompletely quantified differences among programs, we can be all the more confident in our conclusions. Additionally, varying both the models and their implementation more closely matches the options available to empirical phylogeneticists seeking to resolve ancient rapid radiations. This approach suffers a clear liability, however, in that the effects of model differences and implementation differences are conflated. In the event of differing results, we may not be able to isolate the effects of one from the other; therefore, the added value to empirical phylogeneticists comes at the cost of reduced utility of our results to program developers and theorists.

The specific evaluations performed are described more thoroughly in the Results section. Computationintensive analyses were run on either the Duke Shared Cluster Resource (https://wiki.duke.edu/display/SC SC/DSCR) or the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX). When appropriate, multiple tree files were summarized onto a target phylogram with SumTrees 2.0.2 (Sukumaran and Holder 2010) for subsequent inspection or manipulation in FigTree 1.3.1 (Rambaut 2006).

## Results

## Data and Topology Point Estimate

Tree-wide mean ML bootstrap support (MLBS) values (summed bootstrap values on the ML tree divided by the number of internodes in that tree) for the individual loci ranged from $74 \%$ (atpB) to $84 \%$ ( $\operatorname{matK}$ ). The concatenated data have a mean MLBS of $92 \%$ and strongly support (i.e., have $\geqslant 70 \%$ MLBS and $\geqslant 0.95$ posterior probability [PP]) $90 \%$ of the partitions (Table 2). Across data sets, ML and Bayesian inference consistently inferred strong support for a comparable number of bipartitions (Table 2), a result that offers further empirical corroboration for the approximate equivalence of 70\% MLBS and 0.95 PP (Hillis and Bull 1993; Alfaro et al. 2003).


FigURE 3. Phylogeny evaluation: rate heterogeneity and the Bayesian star-tree paradox artifact. a) Unrooted ML phylogram of the concatenated data, with the backbone internodes highlighted and labeled. b) Accounting for the impact of rate heterogeneity on backbone support values. The four values listed for each backbone internode are: MLBS on full data, MLBS with Aspleniaceae (Asplenium and Hymenasplenium) pruned from trees, MLBS with Aspleniaceae removed from analysis, and posterior support from BEAST. c) Controlling for the Bayesian star-tree paradox artifact using the polytomy prior in Phycas. The four PPs listed for each internode are from: MrBayes 3.1.1 (susceptible to the star-tree artifact), Phycas with $C=1$, Phycas with $C=e$, and Phycas with $C=10$. d) Controlling for the Bayesian star-tree paradox artifact using the Yang branch-length prior. The four PPs listed for each internode are from: MrBayes 3.1.1 (susceptible to the star-tree artifact), MrBayes 3.1.2 with branch-length prior $\mathrm{mu} / \mathrm{mu} 0=0.01, \mathrm{Mr}$ Bayes 3.1 .2 with branch-length prior mu1/mu0 $=0.001$, and MrBayes 3.1 .2 with branch-length prior $\mathrm{mu} 1 / \mathrm{mu} 0=0.0001$.

There are two well-supported conflicts among the individual-locus ML trees. The first involves a tip relationship (matK unites Deparia pterorachis with D. unifurcata, with $72 \%$ MLBS, whereas rbcL places D. pterorachis as sister to the rest of the genus, with $75 \%$ support) that is peripheral to the focus of this study. The second is deeper in the tree: matK has $80 \%$ MLBS for a clade uniting Thelypteridaceae with the athyrioids, Blechnaceae, and Onocleaceae, to the exclusion of Woodsia and allies, whereas both $\operatorname{atp} A$ and $\operatorname{atpB}$ place Woodsia and its allies within that clade ( $92 \%$ in $\operatorname{atpA} ; 71 \%$ in $\operatorname{atpB}$ ). Given that we confirmed this conflict to not be attributable to laboratory or identification errors, and because the loci involved are linked and the taxa are long-diverged, we do not ascribe this conflict to differences in evolutionary history and proceeded to concatenate all the data for subsequent analyses.

Each of our ten ML best-tree searches of the concatenated data in PAUP* (from different random-additionsequence starting trees) inferred the same tree (Fig. 3a),
suggesting that tree space is unimodal for our data set. Most partitions in this topology point estimate were strongly supported by both ML bootstrapping and Bayesian PPs (Table 2); the different ML programs (PAUP*, Garli, and RAxML) inferred very similar MLBS levels (data not shown).

## Bayesian Star-Tree Paradox Artifact

For certain branches, we observed very high Bayesian PPs from the MrBayes analysis, but much lower levels of support from the ML bootstrapping (Fig. 4a); these support discrepancies are disproportionately represented among short branches (Fig. 4b). This pattern is consistent with artificially high Bayesian support due to the star-tree paradox artifact-most implementations of Bayesian phylogenetic inference do not consider polytomies among the option set for the MCMC sampler and thus can return high PPs for branches that are unsupported by the data (Lewis et al. 2005; Yang and


Figure 4. Discrepancies between ML and Bayesian support values. a) ML phylogram of the concatenated data, internal branches only (all tip branches have been deleted). Branches are colored according to the magnitude of the difference between their PP (from MrBayes 3.1 .1 ) and their percent MLBS (from 5000 pseudoreplicates in $P A U P^{*}$ ). b) Internal branches rotated to be vertical and sorted by length. Colors follow Figure 4a.

Rannala 2005; Yang 2008; see early hints in Cummings et al. 2003).

To ensure that this "star-tree paradox" artifact was not influencing our assessment of support, we compared the results from our original MrBayes 3.1.1 analysis (Ronquist and Huelsenbeck 2003; MrBayes 3.1.1 is potentially vulnerable to the star-tree paradox artifact) with those of a non-Bayesian analysis (ML bootstrapping in PAUP* from our initial assessment of support), as well as with two implementations of Bayesian inference that use different approaches to reduce their vulnerability to the star-tree paradox artifact.

First, we analyzed our data with Phycas 1.1.2-r (Lewis et al. 2010). Phycas uses reversible-jump MCMC to allow the sampling of incompletely resolved topologies, controlled via the incorporation of a polytomy prior, " $C$." A value of $C=1$ means that unresolved and fully resolved topologies are weighted equally; under a value of $C=10$ a trichotomy is 10 times more likely, a priori, than either of its fully resolved resolutions (Lewis et al. 2005). We performed three runs of our full concatenated data, under a GTR $+\mathrm{I}+\mathrm{G}$ model, using a branchlength hyperprior (default values), for 200,000 cycles (sampling from the posterior every 10 cycles; note: Phycas makes proposals to each free parameter in each cycle, and thus Phycas cycles are not comparable to MrBayes generations, which include a proposal to only a single parameter), with $C=1, C=e(2.718)$, and $C=$ 10 , respectively. Inspection of the $A W T Y$-type plots (see Nylander et al. 2008) and sojourn plots (see Lewis and Lewis 2005) revealed that the runs converged before 40,000 cycles; to be conservative, we excluded the first 50,000 cycles ( 5000 samples) as burn-in. These trees are available as SI Figures 1-3.

We then reanalyzed our data with a modified version of MrBayes 3.1.2 that incorporates exponential priors on
internal and external branch lengths (Yang 2007, 2008). These "Yang branch-length priors" allow the concentration of the prior mass on topologies with very short internal branches, for an intended effect similar to that of the polytomy prior (above); only if the data strongly support a branch can the short internal branch prior be overcome. We performed three runs of four chains each under the settings used in the initial MrBayes analysis but with the addition of the branch-length priors. The mean of the external branch-length prior (mu1) was set to 0.1, and the mean of the internal branch-length prior (mu0) to $0.00001,0.0001$, and 0.001 , successively. As in the initial analyses, 2 million generations from each run were discarded as burn-in (trees available as SI Figs. 4-6).

The results of these analyses show that four of the backbone internodes (A, B, C, and D; Fig. 3a) were largely insensitive to either the polytomy prior (Fig. 3b) or the Yang branch-length prior (Fig. 3c); their PPs stayed at or within three percentage points of 1.0 for all seven analyses (the original MrBayes 3.1.1 analysis, three Phycas runs with increasingly strong polytomy prior values, and three MrBayes 3.1.2 runs with increasingly strong Yang branch-length prior values). Interestingly, the two approaches (polytomy prior vs. branch-length prior) had very different effects on the other backbone internodes, despite the approaches being designed to overcome the same shortcoming in Bayesian phylogenetic inference. For example, the only backbone internode that was not well supported by the original ML and Bayesian analyses (internode F; Fig. 3a) exhibited increased support under weak versions of either the polytomy prior or the branch-length prior. However, the norms of reaction for the two priors were opposed: as the polytomy prior increased in strength ( $C=1, e$, and 10), the posterior support for internode $F$
decreased ( $\mathrm{PP}=0.93,0.89,0.83$; Fig. $3 b$ ), whereas as the branch-length prior strength increased (mu0/mu1 = $0.01,0.001,0.0001$ ), the PP on internode F also increased ( $\mathrm{PP}=0.97,1.0,1.0$; Fig. 3c).

The remaining three internodes (E, G, and H; Fig. 3a) were well supported by the original ML and Bayesian analyses but showed some sensitivity to either the polytomy or the branch-length prior, again in opposing ways. Internode E was largely insensitive to the branchlength prior (Fig. 3c) but was strongly weakened by the polytomy prior (Fig. 3b), whereas internodes G and H were largely unaffected by the polytomy prior but were unsupported under strong values of the branch-length prior (Fig. 3b,c).

## Lineage-Specific Rate Heterogeneity

To investigate whether the rapid rate of evolution for the Aspleniaceae (Figs. 1 and 3) was biasing tree reconstruction, we attempted to isolate the effects of this rate heterogeneity in three ways. First, we pruned the Aspleniaceae from 1000 full-data Garli ML bootstrap trees prior to building the consensus tree and evaluating support. This "reduced consensus" approach (Wilkinson 1996; Burleigh et al. 2009) removes any effects due solely to uncertainty in the placement of these long-branch taxa. If the remaining relationships are well supported, then overall support values will appear low in the standard consensus but will be restored under the reduced consensus. Second, we reran the Garli ML bootstrap analysis on a data set where the Aspleniaceae had been removed prior to analysis, to eliminate any effect that these taxa might have on the optimization of model parameters, and to allow the model to better fit the remaining data (the "reduced data" approach). Third, we ran the full data set in BEAST 1.5.4 (Drummond and Rambaut 2007), incorporating a relaxed-clock model that explicitly models lineage-specific rate variation (Drummond et al. 2006) and thus should be less sensitive to any artifacts induced by the strongly heterogeneous rates in our data. We ran three independent runs on the full concatenated data set, each for 20 million generations (sampling the posterior every 1000 generations), with the following settings: birth-death tree prior; lognormal uncorrelated relaxed clock; and GTR+I+G substitution model. Priors were left at their default values, with the exception of those for six time-to-most-recent-common-ancestor (TMRCA) age parameters, which were each given normal distributions with a mean equal to the inferred age estimated for that clade by Schuettpelz and Pryer (2009) and a standard deviation equal to $10 \%$ of that mean. The relevant taxon sets, and their TMRCA prior means, are: tree root (165.6 MA), Dennstaedtiaceae (119.3 MA), Eupolypods (116.7 MA), Pteridaceae (110.8 MA), Eupolypods II (103.1 MA), and Eupolypods I ( 98.9 MA ). None of the taxon sets was constrained to be monophyletic. The use of secondary constraints such as these is clearly inferior to the use of fossil data for divergence time dating (Magallón 2004), but as no such data are available, and our interest is more in the relative than absolute divergence times,
this approach seemed best. Visual inspection in Tracer (Rambaut and Drummond 2007) demonstrated that the runs converged before 1 million generations; to be conservative, we excluded the first 3 million generations of each run prior to analyzing the pooled posterior of 51,000 samples ( 17,000 from each run; SI Fig. 7). For this sample, the effective sample size for each parameter was above 300 .

None of these attempts to mitigate potential effects of the increased rates of molecular evolution associated with Aspleniaceae strongly affected support values along the backbone. Support values from the full taxonsample consensus data (Fig. 3d, first values) differed from those from the reduced consensus (Fig. 3d, second values) by at most one percentage point. Removing Aspleniaceae from the data set prior to the bootstrap tree searches had a larger effect (up to a five percentage point change in support; Fig. 3d, third values), but in no case resulted in an internode moving from well supported ( $\geqslant 70 \%$ MLBS) to poorly supported or vice versa. The support values from the BEAST analysis (Fig. 3d, fourth values) were concordant with those of the ML runs, especially in that the internode uniting Rhachidosorus with the diplazioids, Hemidictyum, and Aspleniaceae (internode F; Fig. 3d) was the only one without strong support (it had a PP of 0.90).

## Rooting Uncertainty

To evaluate any effects that uncertainty in root-branch placement might have on apparent levels of support within the ingroup, we compared ingroup backbone MLBS values from the analysis of our complete data (full outgroup) with those from each of six different variations in outgroup composition: (i) ingroup only; (ii) ingroup + Dryopteris; (iii) ingroup + Dryopteris and Didymochlaena; (iv) ingroup + Dryopteris and Notholaena; (v) ingroup + Dryopteris and Notholaena and Pteridium; and (vi) ingroup + our full eupolypod I sample (Fig. 5). This outgroup sampling regime was selected to successively bisect the longest outgroup branches, with a particular emphasis on breaking the proximate root branch (the branch connecting the ingroup to the first outgroup node).

We evaluated support for each of the six outgroup sampling regimes using both the reduced consensus approach (full data included in the analysis, but outgroup pruned down to the desired sample prior to forming the consensus tree; Fig. 5, first values; Wilkinson 1996; Burleigh et al. 2009) and a reduced data approach (outgroup reduced to the desired sample prior to the analyses; Fig. 5, second values). The former approach controls for uncertainty in outgroup placement alone (i.e., it offers a metric of the signal strength), whereas the latter approach additionally accounts for model fit. All analyses were based on 1000 ML bootstrap replicates of the concatenated data in Garli 1.0.695 (MPI parallel version; Zwickl 2006), under a GTR+I+G substitution model, using the default genetic algorithm and termination settings.


FIGURE 5. Effects of outgroup composition on ingroup backbone support values. Values on each internode indicate the percentage point difference in MLBS between the focal analysis and the analysis with the full taxon sample. The first values are those from the reduced consensus approach; second values are from the reduced data approach. NA indicates an internode not present with that outgroup sample. Excluded taxon branches are in gray; included outgroup branches are black. This figure shows only outgroup and ingroup backbone branches; most ingroup branches have been deleted (but were included in the analyses). Labeling of the backbone internodes follows Figure 3a.

The results of these rooting comparisons demonstrate that our initial concerns-that the outgroup would wander and thus reduce support measures within the ingroup-were largely unfounded. The reduced consensus support values were minimally different from those with the full outgroup (Fig. 5, first values). When a "consensus interference" effect did appear (first values $>0$ in Fig. $5 \mathrm{~b}, \mathrm{c}, \mathrm{f})$, it was correlated with the maximum root length rather than with proximate root length, i.e., it is the long Notholaena branch that wanders rather than the outgroup as a whole.

In stark contrast, outgroup composition had a strong effect on backbone support if the outgroup was changed prior to the tree searching steps. When we reduced our outgroup sample and reran the ML bootstrapping (the reduced data approach), backbone internode support values changed from their full-outgroup values by up to 32 percentage points (Fig. 5, second values). The largest of these changes are reductions in support for branches proximate to the root (internodes E, F, and G; Fig. 3a) and are due to uncertainty in the ancestral state
of the smaller outgroup sample (as demonstrated by the reduced consensus values from each of the smaller data sets; data not shown).

## Eupolypod II Phylogeny

Our results demonstrate that the vast majority of internodes in the ML tree are strongly supported by both ML bootstrapping and Bayesian PPs (Fig. 6 and Table 2), and these support values proved robust to our phylogeny evaluations. In particular, the ML tree has 10 highly supported major (approximately family unit) ingroup clades: Cystopteris/Gymпосarpium; Rhachidosorus; Diplaziopsis/Homalosorus; Hemidictyum; Aspleniaceae; Thelypteridaceae; Woodsia and allies; Onocleaceae; Blechnaceae; and the athyrioids (Fig. 6b). Of these, Cystopteris/Gymnocarpium is sister to the remaining eupolypod II taxa, followed by the loosely supported assemblage of Rhachidosorus with Diplaziopsis/Homalosorus + Hemidictyum + Aspleniaceae. Blechnaceae is sister to Onocleaceae, and they together are


FIGURE 6. ML phylogram of the concatenated data. Support values indicated are MLBS followed by Bayesian PPs. Bold branches have $\geqslant 70 \%$ MLBS and $\geqslant 0.95$ PP. Support values of $100 \%$ or 1.0 are indicated with an asterisk ( ${ }^{*}$ ). a) Most recent family designations (Smith et al. 2006), with the paraphyletic Woodsiaceae highlighted. b) Major clade names used in this study: Cyst/Gymno = Cystopteris s. lat. and Gymnocarpium; Rha = Rhachidosorus; Dipls = Diplaziopsis and Homalosorus; $H=$ Hemidictyum; Aspleniaceae; Thelypteridaceae; Woods and allies $=$ Woodsia and allies; Onocleac = Onocleaceae; Blechnaceae; athyrioids. Backbone internodes labeled A through H following Figure 3a.
successively sister, in a pectinate pattern, to the athyrioids, then to Woodsia and allies, and finally to the Thelypteridaceae (Fig. 6). This broad phylogeny is in general agreement with earlier molecular phylogenetic studies that included members of the Eupolypods II (Gastony and Ungerer 1997; Murakami et al. 1999; Sano et al. 2000a, 2000b; Smith and Cranfill 2002; Tzeng 2002; Cranfill and Kato 2003; Wang et al. 2003; Schneider et al. 2004a; Wei et al. 2010; see particularly Schuettpelz and Pryer 2007; Kuo et al. 2011). However, the backbone of the phylogeny is strongly supported for the first time; the only backbone internode lacking such support is the one attaching the Rhachidosorus branch to the rest of the tree (Fig. 6, internode F). Additionally, we are finally able to confidently place the enigmatic genera Cheilanthopsis, Diplaziopsis, Homalosorus, Protowoodsia, and Woodsia.

## DISCUSSION

## Bayesian Star-Tree Paradox Artifact

Although one would anticipate that internodes across a topology would differ in their sensitivity to the startree paradox approaches (not all short branches are inferred equal), it is unclear what is driving the different responses in our data-neither the original Bayesian posteriors nor the MLBS levels correlate with the behavior of a given internode under the additional priors (Fig. 3). This study is the first to examine the performance of these star-tree paradox methods on empirical data; their nonparallel effects were perhaps the most surprising result of this portion of the analyses. However, while they were developed for the same function, the methods differ strongly in their approach and should not be expected to result in similar behavior. The branch-length prior, in effect, flattens the posterior for topologies. As the $\mathrm{mu} 0 / \mathrm{mu} 1$ ratio decreases, the relative influence of any data supporting an internal branch is reduced, and the external branches come closer to being randomly arranged. However, each topology sampled from the posterior must still be fully resolved and thus any reduction of support for a particular topology must be accompanied by increased support for some other one. In this sense, the branch-length prior is less a measure of intrinsic support for a given internode than it is a measure of whether that node is better supported than all alternative resolutions. Under circumstances of low support for an entire set of relationships, the branchlength prior favors the best of a bad lot. The polytomy prior, on the other hand, allows the direct comparison between a given resolution and a polytomy. Strong values of the branch-length prior lead to many trees, each with a low posterior, whereas strong values of the polytomy prior lead to a star tree, with a high posterior.

In interpreting the performance of these two methods on our data, it is important to stress that we did not attempt to tightly isolate the effects of the Yang branch-length prior and those of the polytomy prior. Rather, each was bundled with other elements of its host
program (MrBayes 3.1.2 and Phycas 1.1.2-r, respectively), which differ from each other in both their models, and their implementation of those models. In particular, important model differences include data partitioning in MrBayes (the Phycas runs were on unpartitioned data) and the incorporation of a branch-length hyperprior in Phycas (there is no such hyperprior in MrBayes); important implementation differences include the limitation of Phycas to Larget-Simon moves (Larget and Simon 1999), whereas MrBayes utilizes a broader suite of topology proposals.

Regardless of the different performance of the two methods, the backbone support levels in our data were generally robust to the star-tree paradox artifact approaches (Fig. 3b,c), suggesting that the high Bayesian support values for these internodes are valid. Even under extreme values of the polytomy prior, for example ( $C=10$, or trichotomies 10 times more likely, a priori, than their fully resolved alternatives), the posterior consensus tree still resolved each of the eight critical backbone internodes, and only one fell below 0.95 PP (internode E; Fig. 3b). The differences between Bayesian PP and MLBS values in our data (Fig. 4), therefore, reflect something other than the failure of the original Bayesian analyses to include polytomies in the option set; these differences may simply be due to Bayesian inference being more sensitive to small amounts of data than is bootstrapping and thus more likely to support short internodes (Alfaro et al. 2003).

## Lineage-Specific Rate Heterogeneity

The absence of any effect of lineage-specific rate heterogeneity on our topology estimation or support levels is particularly interesting in light of recent questions (Drummond et al. 2006) about the general applicability of the unrooted model (aka "no clock" model; Yang and Rannala 2006; Wertheim et al. 2010) in phylogenetic inference. Given the dramatic lineage-specific rate heterogeneity that is present in our data set, one might expect the unrooted and relaxed-clock models to fit very differently, and given that the fast lineages in our data are intercalated among short internodes, our topology would be expected to be sensitive to such model differences. However, no effects are seen; our data, at least, do not support concerns about the application of the unrooted model in phylogenetic inference, a result that provides empirical support to the simulation results of Wertheim et al. (2010).

## Rooting Uncertainty

The effects of differing outgroup compositions on support levels for branches phylogenetically distant from the root were unexpected and may reflect a combination of both stochastic variation in ML bootstrapping and of factors of model optimization on the different data sets. Neither of these explanations is heartening.

The latter-the "model-mediated" effect-requires that changes in outgroup composition have strong and somewhat idiosyncratic effects on support levels on parts of the tree phylogenetically distant from the root itself (for a similar case, see Roberts et al. 2009). These effects are not due to the outgroup itself changing position (that possibility is eliminated by comparison with the reduced consensus values) and must instead be mediated through nontopological factors. The former explanation-stochastic variation in bootstrap support values-would suggest that 1000 pseudoreplicates are insufficient to get accurate support estimates for these data. Regardless of the precise mechanism by which the outgroup affects support values, these results emphasize the wisdom of including a broad outgroup sample, particularly when the outgroup is distantly related to the taxa of interest (Swofford et al. 1996; Graham and Iles 2009).

## Eupolypod II Phylogeny: Major Clades

The affinities of Cystopteris s. lat. (including Acystopteris, e.g., Blasdell 1963) and Gymnocarpium have been the object of considerable taxonomic disagreement. Both genera, individually or in tandem, have been thought to be allied with the Dryopteridaceae (in Eupolypods I) or the Athyriaceae; in either position they were inevitably highlighted as being anomalous (see Sledge 1973). Early molecular studies supported Cystopteris and Gymnocarpium as sister species and demonstrated their lack of close affinity to either Dryopteridaceae or Athyrium but were unable to pinpoint their phylogenetic position (Wolf et al. 1994; Hasebe et al. 1995). Recent studies (Schuettpelz and Pryer 2007; Kuo et al. 2011) were the first to support a sister group relationship between a Cystopteris/Gymnocarpium clade and the rest of Eupolypods II, a result corroborated and strengthened by our data (Fig. 6, internode E), the first to include multiple accessions of Acystopteris and Cystopteris s. str.

Historically, arguments about Rhachidosorus focused on its validity as a genus, distinct from either Athyrium or Diplazium (Ching 1964a; Kato 1975b). Early molecular phylogenies (Sano et al. 2000a; Tzeng 2002; Wang et al. 2003) provided the first evidence that Rhachidosorus might not be closely related to either, a result further emphasized by the three-gene results of Kuo et al. (2011). In our study, the two included Rhachidosorus species form a tight clade phylogenetically distant from any other taxon; their closest relatives appear to be Diplaziopsis, Homalosorus, Hemidictyum, and Aspleniaceae. While our data do not strongly support a precise position for Rhachidosorus (Fig. 6, internode F), of note is the $100 \%$ MLBS and 1.0 PP for internode D (Fig. 6), which separates Rhachidosorus from the athyrioids. Thus, our data very strongly reject a close relationship between Rhachidosorus and its presumed allies, the athyrioids, an unanticipated conclusion based on morphological data (Kato 1975b). Indeed, our data suggest that the two groups last shared a common ancestor nearly 100 MA (SI Fig. 7).

As with Rhachidosorus, Homalosorus and Diplaziopsis were long thought to be allied with the athyrioids, where they are typically treated as members of Diplazium (Ching 1964b; Kato 1975a, 1977; Kato and Darnaedi 1988; Wang et al. 2004). The first indication that this placement might be inaccurate came from the study of Sano et al. (2000a), which strongly supported Homalosorus (a monotypic genus) as sister to their lone Diplaziopsis accession and placed the two genera distant from Diplazium, a result corroborated by Wei et al. (2010) and Kuo et al. (2011). Our study includes two Diplaziopsis species, which are strongly supported as sister to each other, and together are sister to Homalosorus. These two genera form a clade that is strongly supported, for the first time, as sister to Hemidictyum + Aspleniaceae (Fig. 6, internode G).

Woodsia has been underrepresented in molecular phylogenetic studies to date; no study has included more than one species, and none has been able to strongly infer the position of that species, either. Here, we establish that Woodsia s. lat. is likely to be monophyletic (Fig. 6, seven species included in our analysis) and we demonstrate that two of the three segregate genera (Cheilanthopsis and Protowoodsia) recognized by Shmakov (2003) are nested within Woodsia s. str.; only Hymenocystis is as-yet unsampled. Additionally, our study finds strong support for the position of Woodsia s. lat. to be far from the other Woodsiaceae genera (Athyrium, Acystopteris, Cornopteris, Cystopteris, Deparia, Diplazium, Gymnocarpium, Rhachidosorus, Diplaziopsis, Homalosorus, Hemidictyum), as circumscribed in the most recent family level fern classification (Smith et al. 2006); compare Figure 6a with Figure 6b.

The "athyrioids" have been a source of great disagreement in fern systematics (e.g., Ching 1940; Alston 1956; Ching 1964a; Sledge 1973; Tryon and Tryon 1982). Molecular data confirmed their distant relationship to the dryopteroid ferns (Dryopteridaceae, in Eupolypods I), but uncertainty regarding their delimitation and affinities has persisted until very recently. Sano et al. 2000a were the first to extensively sample the athyrioids, and they provided the initial evidence that the group, as then understood, was strongly heterogeneous. Our data corroborate the results of earlier studies (Sano et al. 2000a; Wang et al. 2003; Schuettpelz and Pryer 2007) in revealing three major clades within the athyrioids s. str.: one containing Athyrium and close allies ("athyriids"); one containing Diplazium s. lat. ("diplaziids"); and one containing Deparia s. lat. ("depariids"). Our novel finding is the well supported, early diverging position of Athyrium skinneri with respect to the other athyriids included in our sample. This species belongs to a small group of predominantly Mexican taxa, none of which had been included in previous phylogenetic studies. Its position as sister to the rest of the included athyriids (including Cornopteris and Pseudocystopteris) emphasizes the paraphyly of Athyrium as currently circumscribed and has important implications for our understanding of the evolution of both the athyriids and the diplaziids. Our study provides additional novel
support for the placement of the athyrioids as phylogenetically distant from Rhachidosorus, Cystopteris, Gymnocarpium, Woodsia, Diplaziopsis, Hemidictyum, and Homalosorus, a topology that is in conflict with the recent classifications of the group (Wang et al. 2004; Smith et al. 2006); both Athyriaceae sensu Wang et al. (2004) and Woodsiaceae sensu Smith et al. (2006) are shown here to be strongly paraphyletic (Fig. 6).

Our results for the remaining five major cladesAspleniaceae, Thelypteridaceae, Hemidictyum, Blechnaceae, and Onocleaceae-agree in all important respects with earlier studies of these groups (Gastony and Ungerer 1997; Murakami et al. 1999; Cranfill 2001; Smith and Cranfill 2002; Schneider et al. 2004a; Schuettpelz and Pryer 2007; Kuo et al. 2011), albeit with generally increased support.

## Eupolypod II Phylogeny: Morphological Stasis and Disparity

A striking pattern in our phylogeny is its incongruence with previous morphology-based hypotheses of relationship, particularly with respect to the position of the genera of Woodsiaceae sensu Smith et al. (2006): Acystopteris, Cystopteris, Diplaziopsis, Gymnocarpium, Hemidictyum, Homalosorus, Rhachidosorus, Woodsia and allies, as well as the athyrioids (Fig. 6). Some of these groups have been historically difficult to place and thus their isolation from Woodsia or the athyrioids (the bulk of Woodsiaceae sensu Smith et al. (2006) is in the athyrioids) is not particularly surprising. Smith et al. (2006) themselves noted that their Woodsiaceae might prove to be not monophyletic. The placement of three genera, however, was utterly unanticipated by morphological data: Diplaziopsis, Homalosorus, and Rhachidosorus. These taxa have not only been considered closely related to the athyrioids, they have been nearly universally considered to be members of the large genera Diplazium (first two) or Athyrium (Rhachidosorus). Their phylogenetic position, deeply isolated from their presumed relatives, underscores the complex patterns of morphological evolution in Eupolypods II; further morphological investigations are necessary to determine whether the apparent similarities between these three genera and the athyrioids are due to convergence or symplesiomorphy.

This trend of shared morphological syndromes across very deep splits in the tree by some members of the "Woodsiaceae" is in contrast to the interdigitation, among those same taxa, of a series of distinct morphologically unique groups, including the Aspleniaceae, Blechnaceae, Onocleaceae, and Thelypteridaceae. The coarse picture of eupolypod II morphological evolution, then, is marked by two seemingly opposing patterns. On the one hand are the autapomorphy-rich clades, whose individual phylogenetic coherence is strong, but whose deep relationships were obscure based on morphological data. And, on the other, the morphologically consistent yet phylogenetically incoherent members of the "Woodsiaceae".

Although not the focus of this study, our phylogeny contains rich information on relationships closer to the
tips of the tree, within the approximately family unit clades. For example, within the athyrioids and Blechnaceae, morphological evolution is complex, and nonmonophyletic generic concepts are common. Generic delimitation within these families is in need of much further study. In addition, a cursory comparison between the Onocleaceae and their sister group, the Blechnaceae, is revealing. Both clades have approximately the same crown ages (SI Fig. 7) yet exhibit strikingly different patterns of diversification. The Onocleaceae branch is marked by few well-spaced divergences leading to the five extant species. Conversely, the Blechnaceae branch features multiple, very short internodes; this family includes approximately 200 extant species.

## Phylogeny Evaluation

Despite the presence in our data set of each of the anticipated challenges to robust phylogenetic inference (long outgroup branch; strong lineage-specific rate heterogeneity; ancient rapid radiation model; Figs. 1 and 3a), we were able to infer a phylogeny with strong backbone support (Fig. 6), and our various evaluations gave no indication that the support for the internodes in our ML tree is due to artifacts. However, different approaches to controlling for the Bayesian star-tree paradox artifact, and different outgroup sampling regimes all influenced support levels; only lineage-specific rate heterogeneity had a negligible effect.

These effects give further weight to arguments for rigorously evaluating phylogenies against potential artifacts. While specific vulnerabilities may be data set dependent, the core elements of our analysis regime are broadly applicable, including the inspection of preliminary phylogenetic hypotheses for potential confounding factors, the investigation of those factors through scrutinizing the performance of multiple models and multiple implementations of those models, and the utilization of the reduced consensus approach to isolate topological effects of signal weakness from those of signal conflict. Although this study is focused on the post-data set steps, preanalysis components (taxon sampling, character sampling, character evaluation) are also vital. In particular, in our case, the use of a broad taxon sample with moderate character data proved effective.

## Supplementary Material

Supplementary material, including data files, can be found at http:/ /www.sysbio.oxfordjournals.org/.

## Funding

This work was supported in part by the National Science Foundation [DEB-0347840 to K.M.P.]; an NSF DDIG award to K.M.P. and C.J.R (DEB-1110767); the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning ("Formas"; 2006-429 to P.K.); the National Science Council of the Republic
of China (NSC 97-2313-B-054-006-MY3 to W.-L.C.); the Duke University Department of Biology (Grant-in-Aid to C.J.R.); the Society of Systematic Biologists (MiniPEET Award to C.J.R.); and the Canadian National Science and Engineering Research Council (PGS-D to C.J.R.).

## ACKNOWLEDGMENTS

We thank the many collectors who have donated material to the Pryer Lab DNA Database (www.pryerlab. net) and thus make research like this possible; Eric Schuettpelz, Masahiro Kato, Fay-Wei Li, and George Yatskievych made particularly important contributions to this study. In addition, we thank the staff of B, COLO, DUKE, E, GOET, MO, NY, P, TAIF, TNS, UC, UPS, and VT for curating this material and allowing access to their collections, and the Dr Cecilia Koo Botanic Conservation Center in Taiwan for providing living materials. Yea-Chen Liu, Atsushi Ebihara, and Chun-Neng Wang provided vital materials and access to facilities. David Swofford assisted with the installation and operation of Phycas, Nahid Heidari was invaluable in the lab, and this manuscript was improved immensely through comments from James Beck, Amanda Grusz, Layne Huiet, Fay-Wei Li, Eric Schuettpelz, Erin Sigel, Alan Smith, David Swofford, Michael Windham, Editor-in-Chief Ronald W. DeBry, Associate Editor Roberta J. Mason-Gamer, and two anonymous reviewers. Portions of the analyses were performed on resources provided by the Swedish National Infrastructure for Computing through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project p2009050.

## REFERENCES

Alfaro M.E., Zoller S., Lutzoni F. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 20:255-266.
Alston A. 1956. The subdivision of the Polypodiaceae. Taxon. 5:23-25.
Altekar G., Dwarkadas S., Huelsenbeck J.P., Ronquist F. 2004. Parallel Metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics. 20:407-415.
Anderson F., Swofford D. 2004. Should we be worried about longbranch attraction in real data sets? Investigations using metazoan 18S rDNA. Mol. Phylogenet. Evol. 33:440-451.
Bergsten J. 2005. A review of long-branch attraction. Cladistics. 21:163193.

Blasdell R.F. 1963. A monographic study of the fern genus Cystopteris. Memoirs of the Torrey Botanical Club. Vol. 21. Durham (NC): The Seeman Printery. p. 1-102.
Brown D.F.M. 1964. A monographic study of the fern genus Woodsia Beih. Nova Hedwigia. 16:1-156.
Brown J., Hedtke S., Lemmon A., Lemmon E. 2010. When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. Syst. Biol. 59:145-161.
Burleigh J., Hilu K., Soltis D.E. 2009. Inferring phylogenies with incomplete data sets: a 5-gene, 567-taxon analysis of angiosperms. BMC Evol. Biol. 9:61-72.
Ching R.-C. 1940. On natural classification of the family "Polypodiaceae". Sunyatsenia. 5:1-37.
Ching R.-C. 1964a. On some confused genera of the family Athyriaceae. Acta Phytotax. Sin. 9:41-84.

Ching R.-C. 1964b. On the genus Diplaziopsis C. Chr. Acta Phytotax. Sin. 9:31-38.
Cobbett A., Wilkinson M., Wills M. 2007. Fossils impact as hard as living taxa in parsimony analyses of morphology. Syst. Biol. 56: 753-766.
Cranfill R. 2001. Phylogenetic studies in the Polypodiales (Pteridophyta) with an emphasis on the family Blechnaceae [PhD thesis]. Berkeley (CA): University of California.
Cranfill R., Kato M. 2003. Phylogenetics, biogeography, and classification of the woodwardioid ferns (Blechnaceae). In: Chandra S., Srivastava M., editors. Pteridology in the new millennium. Dordrecht (the Netherlands): Kluwer Academic Publishers. p. 25-47.
Cummings M., Handley S., Myers D., Reed D., Rokas A., Winka K. 2003. Comparing bootstrap and posterior probability values in the four-taxon case. Syst. Biol. 52:477-487.
Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:699-710.
Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.
Duffy A.M., Kelchner S.A., Wolf P.G. 2009. Conservation of selection on matK following an ancient loss of its flanking intron. Gene. 438:17-25.
Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401-410.
Gastony G.J., Ungerer M.C. 1997. Molecular systematics and a revised taxonomy of the onocleoid ferns (Dryopteridaceae: Onocleeae). Am. J. Bot. 84:840-849.
Gaut B.S., Lewis P.O. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. Mol. Biol. Evol. 12:152-162.
Graham S., Iles W. 2009. Different gymnosperm outgroups have (mostly) congruent signal regarding the root of flowering plant phylogeny. Am. J. Bot. 96:216-227.
Hallstrom B., Kullberg M., Nilsson M. 2007. Phylogenomic data analyses provide evidence that Xenarthra and Afrotheria are sister groups. Mol. Biol. Evol. 24:2059-2068.
Hasebe M., Wolf P.G., Pryer K.M., Ueda K., Ito M., Sano R., Gastony G.J., Yokoyama J., Manhart J.R., Murakami N., Crane E.H., Haufler C.H., Hauk W.D. 1995. Fern phylogeny based on $r b c L$ nucleotide sequences. Am. Fern J. 85:134-181.
Hillis D., Bull J. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42: 182-192.
Ho S., Jermiin L. 2004. Tracing the decay of the historical signal in biological sequence data. Syst. Biol. 53:623-637.
Holttum R.E. 1947. A revised classification of leptosporangiate ferns. J. Linn. Soc. Lond. Bot. 53:123-155.

Huelsenbeck J. 1995. Performance of phylogenetic methods in simulation. Syst. Biol. 44:17-48.
Huelsenbeck J., Bollback J., Levine A. 2002. Inferring the root of a phylogenetic tree. Syst. Biol. 51:32-43.
Jackman T., Larson A., de Queiroz K., Losos J. 1999. Phylogenetic relationships and tempo of early diversification in Anolis lizards. Syst. Biol. 48:254-285.
Jian S., Soltis P.S., Gitzendanner M.A., Moore M.J., Li R., Hendry T.A., Qiu Y.-L., Dhingra A., Bell C.D., Soltis D.E. 2008. Resolving an ancient, rapid radiation in Saxifragales. Syst. Biol. 57:38-57.
Kato M. 1975a. Reticulate venation of Diplazium heterophlebium. Acta Phytotax. Geobot. 26:160-163.
Kato M. 1975b. On the systematic position of Athyrium mesosorum. Acta Phytotax. Geobot. 27:56-60.
Kato M. 1977. Classification of Athyrium and allied genera of Japan. Bot. Mag. Tokyo. 90:23-40.
Kato M. 1979. Taxonomic study of the genus Cornopteris (Athyriaceae). Acta Phytotax. Geobot. 30:101-118.
Kato M. 1984. A taxonomic study of the athyrioid fern genus Deparia with main reference to the Pacific species. J. Fac. Sci. Univ. Tokyo Bot. 13:375-430.
Kato M., Darnaedi D. 1988. Taxonomic and phytogeographic relationships of Diplazium flavoviride, D. pycnocarpon, and Diplaziopsis. Am. Fern J. 78:77-85.
Korall P., Conant D.S., Metzgar J.S., Schneider H., Pryer K.M. 2007. A molecular phylogeny of scaly tree ferns (Cyatheaceae). Am. J. Bot. 94:873-886.

Kuo L.-Y., Li F.-W., Chiou W.-L., Wang C.-N. 2011. First insights into fern matK phylogeny. Mol. Phylogenet. Evol. 59:556-566.
Larget B., Simon D. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Mol. Biol. Evol. 16: 750-759.
Larsson A. 2010. abioscript [Internet]. Available from: http:/ /botanik. dnsalias.com/phylogeny/abioscripts/. Accessed on November 1, 2010.

Lewis L., Lewis P.O. 2005. Unearthing the molecular phylodiversity of desert soil green algae (Chlorophyta). Syst. Biol. 54:936-947.
Lewis P.O., Holder M.T., Holsinger K.E. 2005. Polytomies and Bayesian phylogenetic inference. Syst. Biol. 54:241-253.
Lewis P.O., Holder M.T., Swofford D.L. 2010. Phycas [Internet]. Available from: http://hydrodictyon.eeb.uconn.edu/projects/phycas/ index.php/Phycas_Home. Accessed on November 1, 2010.
Liu H.-M., Zhang X.-C., Wang W., Qiu Y.-L., Chen Z.-D. 2007. Molecular phylogeny of the fern family Dryopteridaceae inferred from chloroplast $r b c L$ and atpB genes. Int. J. Plant Sci. 168:1311-1323.
Maddison W.P., Maddison D.R. 2009. Mesquite: a modular system for evolutionary analysis [Internet]. Available from: http://mesquite project.org. Accessed on July 1, 2010.
Magallón S. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. Int. J. Plant Sci. 165:7-21.
Marshall D. 2009. Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. Syst. Biol. 59:108-117.
Mason-Gamer R.J., Kellogg E.A. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst. Biol. 45:524-545.
Mickel J.T. 1974. Phyletic lines in the modern ferns. Ann. Mo. Bot. Gard. 61:474-482.
Murakami N., Nogami S., Watanabe M., Iwatsuki K. 1999. Phylogeny of Aspleniaceae inferred from $r b c L$ nucleotide sequences. Am. Fern J. 89:232-243.

Nagalingum N.S., Schneider H., Pryer K.M. 2007. Molecular phylogenetic relationships and morphological evolution in the heterosporous fern genus Marsilea. Syst. Bot. 32:16-25.
Nakahira Y. 2000. A molecular phylogenetic analysis of the family Blechnaceae, using the chloroplast gene rbcL [MSc thesis]. Tokyo (Japan). Graduate School of Science, University of Tokyo.
Nickrent D., Blarer A., Qiu Y., Vidal-Russell R., Anderson F. 2004. Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol. Biol. 4:40.
Nylander J., Wilgenbusch J., Warren D., Swofford D.L. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics. 24:581-583.
Parfrey L., Grant J., Tekle Y., Lasek-Nesselquist E., Morrison H., Sogin M., Patterson D., Katz L. 2010. Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst. Biol. 59:518-533.
Pereira S., Baker A. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol. Biol. Evol. 9:1731-1740.
Philippe H., Brinkmann H., Lavrov D.V., Littlewood D.T.J., Manuel M., Wörheide G., Baurain D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9:e1000602.
Philippe H., Yan Z., Henner B., Nicolas R., Delsuc F. 2005. Heterotachy and long-branch attraction in phylogenetics. BMC Evol. Biol. 5: 50-58.
Pinter I., Bakker F., Barrett J., Cox C., Gibby M., Henderson S., MorganRichards M., Rumsey F., Russell S., Trewick S., Schneider H., Vogel J. 2002. Phylogenetic and biosystematic relationships in four highly disjunct polyploid complexes in the subgenera Ceterach and Phyllitis in Asplenium (Aspleniaceae). Org. Divers. Evol. 2:299-311.
Pryer K.M., Schuettpelz E., Wolf P.G., Schneider H., Smith A.R., Cranfill R. 2004. Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. Am. J. Bot. 91:1582-1598.
Rambaut A. 2006. FigTree [Internet]. Available from: http:/ /tree.bio. ed.ac.uk/software/figtree/. Accessed on November 1, 2010.
Rambaut A., Drummond A.J. 2007. Tracer [Internet]. Available from: http://tree.bio.ed.ac.uk/software/tracer/. Accessed on November 1, 2010.

Rannala B., Yang Z. 2008. Phylogenetic inference using whole genomes. Annu. Rev. Genomics Hum. Genet. 9:217-231.
Regier J.C., Shultz J.W., Zwick A., Hussey A., Ball B., Wetzer R., Martin J.W., Cunningham C.W. 2010. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature. 463:1079-1083.
Roberts T.E., Sargis E.J., Olson L.E. 2009. Networks, trees, and treeshrews: assessing support and identifying conflict with multiple loci and a problematic root. Syst. Biol. 58:257-270.
Ronquist F., Huelsenbeck J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19: 1572-1574.
Rothfels C.J., Windham M.D., Grusz A.L., Gastony G.J., Pryer K.M. 2008. Toward a monophyletic Notholaena (Pteridaceae): resolving patterns of evolutionary convergence in xeric-adapted ferns. Taxon. 57:712-724.
Sano R., Takamiya M., Ito M., Kurita S., Hasebe M. 2000a. Phylogeny of the lady fern group, tribe Physematieae (Dryopteridaceae), based on chloroplast $r b c \mathrm{~L}$ gene sequences. Mol. Phylogenet. Evol. 15: 403-413.
Sano R., Takamiya M., Kurita S., Ito M., Hasebe M. 2000b. Diplazium subsinuatum and Di. tomitaroanum should be moved to Deparia according to molecular, morphological, and cytological characters. J. Plant Res. 113:157-163.

Schneider H., Russell S., Cox C., Bakker F., Henderson S., Rumsey F., Barrett J., Gibby M., Vogel J. 2004a. Chloroplast phylogeny of asplenioid ferns based on $r b c L$ and $\operatorname{trnL}-F$ spacer sequences (Polypodiidae, Aspleniaceae) and its implications for biogeography. Syst. Bot. 29:260-274.
Schneider H., Schuettpelz E., Pryer K.M., Cranfill R., Magallón S., Lupia R. 2004b. Ferns diversified in the shadow of angiosperms. Nature. 428:553-557.
Schuettpelz E., Hoot S.B. 2006. Inferring the root of Isoetes: exploring alternatives in the absence of an acceptable outgroup. Syst. Bot. 31:258-270.
Schuettpelz E., Korall P., Pryer K.M. 2006. Plastid atpA data provide improved support for deep relationships among ferns. Taxon. 55:897-906.
Schuettpelz E., Pryer K.M. 2007. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. Taxon. 56:1037-1050.
Schuettpelz E., Pryer K.M. 2009. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. Proc. Natl. Acad. Sci. U.S.A. 106:11200-11205.

Schuettpelz E., Schneider H., Huiet L., Windham M.D., Pryer K.M. 2007. A molecular phylogeny of the fern family Pteridaceae: assessing overall relationships and the affinities of previously unsampled genera. Mol. Phylogenet. Evol. 44:1172-1185.
Shmakov A. 2003. Review of the family Woodsiaceae (Diels) Herter of Eurasia. In: Chandra S., Srivastava M., editors. Pteridology in the new millennium. Dordrecht (the Netherlands): Kluwer Academic Publishers. p. 49-64.
Sledge W. 1973. Generic and family boundaries in the Aspidiaceae and Athyriaceae. Bot. J. Linn. Soc. 67:203-210.
Smith A.R. 1995. Non-molecular phylogenetic hypotheses for ferns. Am. Fern J. 85:104-122.
Smith A.R., Cranfill R.B. 2002. Intrafamilial relationships of the thelypteroid ferns (Thelypteridaceae). Am. Fern J. 92:131-149.
Smith A.R., Pryer K.M., Schuettpelz E., Korall P., Schneider H., Wolf P.G. 2006. A classification for extant ferns. Taxon. 55:705-731.

Soltis D., Albert V., Savolainen V., Hilu K., Qiu Y. 2004. Genome-scale data, angiosperm relationships, and 'ending incongruence': a cautionary tale in phylogenetics. Trends Plant Sci. 9:1360-1385.
Soltis P.S., Soltis D.E., Wolf P., Nickrent D., Chaw S., Chapman R. 1999. The phylogeny of land plants inferred from 18 S rDNA sequences: pushing the limits of rDNA signal? Mol. Biol. Evol. 16: 1774-1784.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688-2690.
Steel M., Matsen F. 2007. The Bayesian "Star Paradox" persists for long finite sequences. Mol. Biol. Evol. 24:1075-1079.
Sukumaran J., Holder M.T. 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics. 26:1569-1571.

Susko E. 2008. On the distributions of bootstrap support and posterior distributions for a star tree. Syst. Biol. 57:602-612.
Swofford D.L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sunderland (MA): Sinauer Associates.
Swofford D.L., Olsen G., Waddell P., Hillis D.M. 1996. Phylogenetic inference. In: Hillis D.M., Moritz C., Mable B., editors. Molecular systematics. Sunderland (MA): Sinauer Associates. p. 407-511.
Takezaki N., Gojobori T. 1999. Correct and incorrect vertebrate phylogenies obtained by the entire mitochondrial DNA sequences. Mol. Biol. Evol. 16:590-601.
Tryon R.M., Tryon A.F. 1982. Ferns and allied plants with special reference to Tropical America. New York: Springer.
Tzeng Y.-H. 2002. Phylogenetic relationships of athyrioid ferns inferred from chloroplast DNA sequences [MSc thesis]. Kaohsiung (Taiwan): National Sun Yat-sen University.
Wang H., Moore M.J., Soltis P.S., Bell C.D., Brockington S.F., Alexandre R., Davis C.C., Latvis M., Manchester S.R., Soltis D.E. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc. Natl. Acad. Sci. U.S.A. 106:3853-3858.
Wang M.-L., Chen Z.D., Zhang X.-C., Lu S.-G., Zhao G.-F. 2003. Phylogeny of the Athyriaceae: evidence from chloroplast $t r n \mathrm{~L}-\mathrm{F}$ region sequences. Acta Phytotax. Sin. 41:416-426.
Wang M.-L., Hsieh Y.-T., Zhao G.-F. 2004. A revised subdivision of the Athyriaceae. Acta Phytotax. Sin. 42:524-527.
Wang Z.-R. 2008. Athyriaceae (part 1) [First Draft] [Internet]. Available from: http:/ / flora.huh.harvard.edu/china/mss/volume02/ Athyriaceae-MO-part1_original.htm. Accessed on October 12, 2010.
Wei R., Zhang X.-C., Qi X.-P. 2010. Phylogeny of Diplaziopsis and Homalosorus based on two chloroplast DNA sequences: $r b c L$ and rps4+ rps4-trnS IGS. Acta Bot. Yunn. 32(Suppl):46-54.
Wertheim J.O., Sanderson M.J., Worobey M., Bjork A. 2010. Relaxed molecular clocks, the bias-variance trade-off, and the quality of phylogenetic inference. Syst. Biol. 59:1-8.

Wheeler W.C. 1990. Nucleic acid sequence phylogeny and random outgroups. Cladistics. 6:363-368.
Whitfield J.B., Lockhart P.J. 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22:258-265.
Wiens J.J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52:528-538.
Wilgenbusch J.C., Warren D.L., Swofford D.L. 2004. AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference [Internet]. Available from: http://ceb.csit.fsu. edu/awty. Accessed on November 1, 2010.
Wilkinson M. 1996. Majority-rule reduced consensus trees and their use in bootstrapping. Mol. Biol. Evol. 13:437-444.
Wolf P.G., Soltis P.S., Soltis D.E. 1994. Phylogenetic relationships of dennstaedtioid ferns: evidence from rbcL sequences. Mol. Phylogenet. Evol. 3:383-392.
Yang Z. 2007. Fair-balance paradox, star-tree paradox and Bayesian phylogenetics. Mol. Biol. Evol. 24:1639-1655.
Yang Z. 2008. Empirical evaluation of a prior for Bayesian phylogenetic inference. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363: 4031-4039.
Yang Z., Rannala B. 2005. Branch-length prior influences Bayesian posterior probability of phylogeny. Syst. Biol. 54:455-470.
Yang Z., Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23:212-226.
Yoon C.S., Glawe A., Shaw P.D. 1991. A method for rapid small-scale preparation of fungal DNA. Mycologia. 83:835-838.
Zwickl D.J. 2006. GARLI. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion [PhD thesis]. Austin (TX): The University of Texas.
Zwickl D.J., Hillis D. 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51:588-598.
APPENDIX 1. Voucher table and Genbank accession numbers

|  | DB\# | Voucher | Provenance | $r b c L$ | atpB | $\operatorname{atp} A$ | trnG-R | matK |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Outgroup Dennstaedtiaceae |  |  |  |  |  |  |  |  |
| Dennstaedtia dissecta (Sw.) T. Moore | 2465 | Schuettpelz 9 (DUKE) | Costa Rica: San Jose. | EF463166 ${ }^{\text {j }}$ | EF463373 ${ }^{\text {j }}$ | EF463652 ${ }^{\text {j }}$ | - | - |
| D. scabra(Wall. ex Hook.) T. <br> Moore | - | Li 608 (TAIF) | Taiwan | - | - | - | - | JF303916 ${ }^{\text {c }}$ |
| Pteridium aquilinum (L.) Kuhn | - | Kuo 825 (TAIF) | Taiwan | - | - | - | JF832237 |  |
| P. aquilinum (L.) Kuhn | - | Der 67 (UTC) | USA: California | - | - |  | - | EU223824 ${ }^{\text {a }}$ |
| P. esculentum (G. Forst.) Cockayne | 125 | Smith s.n. (UC) | Cult. (origin: Australia) | U05940 ${ }^{\text {m }}$ | - | EF452115 ${ }^{\text {k }}$ | - | - |
| Dryopteridaceae |  |  |  |  |  |  |  |  |
| Ctenitis eatonii (Baker) Ching | - | Kuo 441 (TAIF) | Taiwan | - | - | - | - | JF303947 ${ }^{\text {C }}$ |
| C. sloanei (Poepp. ex Spreng.) C.V. Morton | 3607 | Schuettpelz 552 et al. (DUKE) | Cult. (Botanischer Garten München-Nymphenburg) | EF463172 ${ }^{\text {j }}$ | EF463383 ${ }^{\text {j }}$ | JF832106 | JF832202 | - |
| Didymochlaena truncatula (Sw.) J. Sm. | - | RBGE 19933685 | Unknown | DQ508769 ${ }^{\text {d }}$ | - | - | - | - |
| D. truncatula (Sw.) J. Sm | 2435 | Schuettpelz 267 (DUKE) | Ecuador: Zamora-Chinchipe | - | EF452030 ${ }^{\text {j }}$ | JF832112 | JF832210 |  |
| D. truncatula (Sw.) J. Sm | - | No voucher taken | Cult. (Dr Cecilia Koo Bot. Cons. Center K017011) | - | - | - | - | JF303942 ${ }^{\text {c }}$ |
| Dryopteris filix-mas (L.) Schott | 7057 | Larsson 8 (UPS) | Sweden: Uppsala | JF832067 | JF832164 | JF832119 | JF832217 | - 3 c |
| D. sparsa (D. Don) Kuntze | - | Liu 481 (TAIF) | Taiwan |  | - |  |  | JF303945 ${ }^{\text {c }}$ |
| Hypodematium crenatum (Forssk.) Kuhn | 6981 | $\begin{aligned} & \text { Yatskievych 09-072 } \\ & \text { et al. (MO) } \end{aligned}$ | Nepal: Kaski | JF832072 | JF832169 | JF832128 | JF832225 | - |
| H. crenatum (Forssk.) Kuhn | - | No voucher taken. | Cult. (Dr Cecilia Koo Bot. Cons. Center 013155; origin: Taiwan) | - | - | - | - | JF303944 ${ }^{\text {c }}$ |
| Leucostegia immersa Wall. ex C. Presl | - | Kuo 170 (TAIF) | Taiwan | - | - | - | - | JF303943 ${ }^{\text {c }}$ |
| L. pallida (Mett.) Copel. | 3652 | Schuettpelz 605 et al. (B) | Cult. (Botanischer Garten Berlin-Dahlem) | - | EF463425 ${ }^{\text {j }}$ | EF463709 | - | - |
| Lomariopsidaceae |  |  |  |  |  |  |  |  |
| Cyclopeltis crenata (Fée) C. Chr. | - | No voucher taken | Cult. (Dr Cecilia Koo Bot. Cons. Center 0169403; origin: Thailand) | - | - | - | - | JF303954 ${ }^{\text {c }}$ |
| C. semicordata (Sw.) J. Sm. | 3501 | Barrington 2129 (VT) | Costa Rica: Punta Arenas | EF463234 ${ }^{\text {j }}$ | EF463480 ${ }^{\text {j }}$ | JF832107 | JF832203 |  |
| Nephrolepis cordifolia (L.) C. Presl | - | Li 335 (TAIF) | Taiwan | - | - | - | - | JF303953 ${ }^{\text {c }}$ |
| $N$. davallioides (Sw.) Kunze | 4131 | Schuettpelz 716 et al. (DUKE) | Malaysia: Pahang | JF832075 | JF832172 | JF832131 | JF832228 | - |
| Polypodiaceae |  |  |  |  |  |  |  |  |
| Polypodium vulgare L. | 7056 | Larsson 13 (UPS) | Sweden: Uppsala | JF832081 | JF832178 | JF832137 | JF832234 | - |
| Pteridaceae |  |  |  |  |  |  |  |  |
| Cryptogramma crispa (L.) R. Br. ex Hook. | 2949 | Christenhusz and Katzer 3871 (DUKE) | UK: Scotland | EF452148 ${ }^{\text {k }}$ | EF452027 ${ }^{\text {k }}$ | EU268740 ${ }^{\text {h }}$ | EU268687 ${ }^{\text {h }}$ | JF832265 |
| Notholaena grayi Davenport | 3187 | Schuettpelz et al. 480 (DUKE) | USA: Arizona | EU268794 ${ }^{\text {h }}$ | JF832173 | EU268749 ${ }^{\text {h }}$ | EU268697 ${ }^{\text {h }}$ | JF832280 |
| Tectariaceae |  |  |  |  |  |  |  |  |
| Tectaria zeylanica (Hout.) Sledge | 3569 | Schuettpelz 514 et al. (GOET) | Cult. (Alter Botanischer Garten Göttingen | EF463275 ${ }^{\text {j }}$ | EF463531 ${ }^{\text {j }}$ | JF832143 | JF832242 | - |
| T. zeylanica (Hout.) Sledge | - | No voucher taken. | Cult. (Dr Cecilia Koo Bot. Cons. Center 017295; origin: Taiwan) | - | - | - | - | JF303951 ${ }^{\text {c }}$ |
| Ingroup Aspleniaceae |  |  |  |  |  |  |  |  |
| Asplenium auritum Sw. | 3482 | Schneider s.n. (GOET) | Cult. (Goett. Bot. Gard.) | EF463146 | EF463327) | EF463591 | - | - |

APPENDIX 1. Continued

|  | DB\# | Voucher | Provenance | $r b c L$ | atpB | atpA | $\operatorname{trnG-R}$ | matK |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| A. juglandifolium Lam. | 3465 | Schneider s.n. (GOET) | Cult. (Goett. Bot. Gard.) | EF463151 ${ }^{\text {j }}$ | EF463333 ${ }^{\text {j }}$ | JF832092 | JF832190 | JF832252 |
| A. ritoense Hayata | - | Nogami and K.O. 27 (KYO?) | Japan | AB014692 ${ }^{\text {e }}$ | - |  | - |  |
| A. ritoense Hayata | 3479 | Ranker 2063 (COLO) | Taiwan | - | EF463343 | EF463607 | - |  |
| A. ruta-muraria L. | - | K018516 | Unknown | - | - | - | - | JF832253 |
| A. ruta-muraria L . | 2947 | Christenhusz and Katzer 3869 (DUKE) | UK: Scotland | - | EF463344 | EF463608 | - | - |
| A. ruta-muraria L. | - | Unknown | Austria | AF525273g | - 82152 | - 28093 | - 2101 | - 3 |
| A. septentrionale (L.) Hoffm. | 7058 | Larsson 11 (UPS) | Sweden: Uppsala | JF832054 | JF832152 | JF832093 | JF832191 | JF832254 |
| A. theciferum (Kunth) Mett. | 2426 | Schuettpelz 258 (DUKE) | Ecuador: Zamora-Chinchipe | JF832055 | JF832153 | JF832094 | JF832192 | JF832255 |
| A. trichomanes L. | 3129 | Schuettpelz 422 et al. (DUKE) (DUKE) | USA: Arizona | EF463157 ${ }^{\text {j }}$ | EF463349 ${ }^{\text {j }}$ | JF832095 | - | JF832256 |
| Hymenasplenium cheilosorum Kunze ex Mett. | 3529 | Schäfer 55 (GOET) | China: Yunnan | JF832071 | EF463350 | JF832126 | JF832223 | - |
| H. unilaterale (Kunth) Mett. | 3470 | Ranker 2072 (COLO) | Taiwan | EF452140 ${ }^{\text {k }}$ | EF452020 ${ }^{\text {k }}$ | JF832127 | JF832224 |  |
| H. unilaterale (Kunth) Mett. | - | Li 619 (TAIF) | Taiwan | - | - | - | - | JF303924 ${ }^{\text {c }}$ |
| Athyrioids |  |  |  |  |  |  |  |  |
| Athyrium filix-femina (L.) Roth | 7054 | Larsson 7 (UPS) | Sweden: Uppsala | JF832056 | JF832154 | JF832096 | JF832193 | - |
| A. filix-femina (L.) Roth | - | Kuo 117 (TAIF) | USA: Massachusetts | - | - | - | - | JF303941 ${ }^{\text {C }}$ |
| A. niponicum (Mett.) Hance | 2852 | Unknown | Japan | JF832057 | JF832155 | JF832097 | JF832194 | J |
| A. niponicum (Mett.) Hance | - | No voucher taken. | Cult. (Dr Cecilia Koo Bot. Cons. Center 016544) | - | - | - | - | JF832257 |
| A. otophorum (Miq.) Koidz. | 3744 | Smith s.n. (UC) | Cult. | EF463305 ${ }^{\text {j }}$ | EF463563 ${ }^{\text {j }}$ | JF832098 | JF832195 | - |
| A. otophorum (Miq.) Koidz. | - | Ebihara et al. 070210 02 (TNS) | Japan: Shizuoka | - | - | - | - | JF832258 |
| A. skinneri (Baker) Diels | 6584 | Rothfels 3155 et al. (DUKE) | Mexico: Nayarit | JF832058 | JF832156 | JF832099 | JF832196 | JF832259 |
| Cornopteris decurrenti-alata (Hook.) Nakai | 4962 | Schuettpelz 1215A et al. (DUKE) | Taiwan: Ilan | JF832061 | JF832159 | JF832104 | JF832200 | JF832263 |
| Deparia lancea (Thunb.) R. Sano | 2558 | Schuettpelz 298 (DUKE) | Cult. (Duke U. Greenh.) | EF463306 ${ }^{\text {j }}$ | EF463567 | JF832109 | JF832207 |  |
| D. lancea (Thunb.) R. Sano | 28 | Kuo 112 (TAIF) | Taiwan | JF832065 | JF832161 | JF832110 | - 82208 | JF303940 ${ }^{\text {c }}$ |
| D. petersenii (Kunze) M. Kato | 2864 | Kato s.n. (?) | Cult. (UTBG 90-554) | JF832065 | JF832161 | JF832110 | JF832208 |  |
| D. pterorachis (H. Christ) M. Kato | - | Fujimoto SF070906 (TNS) | Japan: Nagano | - ${ }^{\text {a }}$ | - | - | - | JF832269 |
| D. pterorachis (H. Christ) M. Kato | - | Unknown | Unknown | AB021716 ${ }^{\text {i }}$ | - | - | - | - |
| D. pterorachis (H. Christ) M. Kato | - | Unknown | Unknown | D43907 ${ }^{\text {i }}$ | - | - | - | - |
| D. pycnosora (H. Christ) M. Kato | - | Ebihara and Kadota HK2007-794 (TNS) | Japan: Hokkaido | - | - | - | - | JF832270 |
| D. unifurcata (Baker) M. Kato | 2865 | Unknown | Cult. (UTBG 90-660) | EF463307 ${ }^{\text {j }}$ | EF463569 ${ }^{\text {j }}$ | JF832111 | JF832209 | JF832271 |
| Diplazium bombonasae Rosenst. | 3764 | Moran 7493 (NY) | Ecuador: Pastaza | EF463308 ${ }^{\text {j }}$ | EF463570 | JF832115 | JF832213 | JF832273 |
| D. dilatatum Blume | 3638 | Schuettpelz 588 et al. (GOET) | Cult. (Botanischer Garten München-Nymphenburg) | EF463311 ${ }^{\text {j }}$ | EF463573 ${ }^{\text {j }}$ | JF832116 | JF832214 | - |
| D. dilatatum Blume | - | Kuo 987 (TAIF) | Taiwan: Taipei | - | - | - | - | JF832274 |
| D. proliferum (Lam.) Thouars | - | No voucher taken. | Cult. (Dr Cecilia Koo Bot. Cons. Center K014215) | - | - | - | - | JF303939 ${ }^{\text {C }}$ |
| D. proliferum (Lam.) Thouars | 3639 | Schuettpelz 590 and Schneider (GOET) | Cult. (Botanischer Garten München-Nymphenburg) | EF463315 ${ }^{\text {j }}$ | EF463577 | EF463918 | - | - |
| D. wichurae (Mett.) Diels | 2874 | Unknown | Cult. (UTBG 88-244) | - | EF463579 ${ }^{\text {j }}$ | JF832117 | JF832215 | - |
| D. wichurae (Mett.) Diels | - | Kuo 986 (TAIF) | Taiwan | - 2083 | - | - | - | JF832275 |
| Pseudocystopteris atkinsonii (Bedd.) Ching | 4837 | Schuettpelz 1094A et al. (DUKE) | Taiwan: Nantou | JF832083 | - | - | JF832235 | JF832285 |
| P. atkinsonii (Bedd.) Ching | - | Kuo 477 (TAIF) | Taiwan: Nantou | - | - | - | - | JF832285 |

APPENDIX 1. Continued

|  | DB\# | Voucher | Provenance | $r b c L$ | atpB | $\operatorname{atp} A$ | $\operatorname{trnG-R}$ | matK |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Blechnaceae |  |  |  |  |  |  |  |  |
| Blechnum orientale L. | - | Kuo 827 (TAIF) | Taiwan: Taipei |  | - | JF304007 ${ }^{\text {C }}$ | - | JF303938 ${ }^{\text {c }}$ |
| B. orientale L. | - | Unknown | Unknown | AB040567 ${ }^{\text {f }}$ | - | - | - | - |
| B. polypodioides Raddi | 2554 | Schuettpelz 294 (DUKE) | Cult. (Duke U. Greenh.; origin: Costa Rica) | EF463159 | EF463352 | JF832100 | JF832197 | JF832260 |
| B. schomburgkii (Klotzsch) C. Chr. | 2410 | Schuettpelz 242 (DUKE) | Ecuador: ZamoraChinchipe | EF463160 ${ }^{\text {j }}$ | EF463353 ${ }^{\text {j }}$ | JF832101 | JF832198 | JF832261 |
| B. spicant (L.) J. Sm. | 6943 | Windham 3395 (DUKE) | Canada: British Columbia | JF832059 | JF832157 | JF832102 | JF832199 | - 8326 |
| B. spicant (L.) J. Sm. | - | Kuo 191 (TAIF) | Canada: British Columbia | - | - | - | - | JF832262 |
| Doodia media R. Br. | 2555 | Schuettpelz 295 (DUKE) | Cult. (Duke U. Greenhouse 86-119) | - | EF463355 ${ }^{\text {j }}$ | JF832118 | JF832216 | JF832276 |
| Sadleria cyatheoides Kaulf. | 3432 | Schuettpelz 507 (DUKE) | Cult. (Duke U. Greenhouse 87-166) | EF463161 ${ }^{\text {j }}$ | EF463356 ${ }^{\text { }}$ | JF832141 | JF832240 | JF832288 |
| Stenochlaena tenuifolia (Desv.) T. Moore | 3429 | Schuettpelz 504 (DUKE) | Cult. (Duke U. Greenhouse) | EF463163 ${ }^{\text {j }}$ | EF463358 | JF832142 | JF832241 | JF832289 |
| Woodwardia areolata (L.) T. Moore | - | Unknown | Unknown | - | - | - | - | JF832296 |
| W. areolata (L.) T. Moore | - | Cranfill s.n. (UC) | Cultivation (U. C. Bot. Gard. 82.2087) | AF425102 ${ }^{1}$ | - | - | - | - |
| W. unigemmata (Makino) Nakai | - | Kuo 493 (TAIF) | Taiwan |  | - 2187 | - 2151 | - | JF832297 |
| W. unigemmata (Makino) Nakai | 4796 | Schuettpelz 1053A et al. (DUKE) | Taiwan: Nantou | JF832090 | JF832187 | JF832151 | JF832250 | - |
| Cyst./Gymno. |  |  |  |  |  |  |  |  |
| Acystopteris taiwaniana (Tagawa) Love \& Love | 4870 | Schuettpelz 1127A et al. (DUKE) | Taiwan: Nantou | JF832052 | - | - | JF832188 | - |
| A. taiwaniana (Tagawa) Love \& Love | - | Kuo 175 (TAIF) | Taiwan | - | - | JF832091 | - | JF303925 ${ }^{\text {c }}$ |
| A tenuisecta (Bl.) Tagawa | - | Kuo 474 | Taiwan: Nantou |  |  |  |  | JF832251 |
| A tenuisecta (Bl.) Tagawa | 4831 | Schuettpelz 1088A et al. (DUKE) | Taiwan: Nantou | JF832053 | - | - | JF832189 |  |
| Cystopteris fragilis (L.) Berhn. | 7103 | Larsson 21 (UPS) | Sweden: Uppsala | JF832062 | JF832160 | JF832108 | JF832204 | JF832266 |
| C. montana (Lam.) Bernh. ex Desv. | 5839 | LeBlond 6448 (DUKE) | Canada: Newfoundland | JF832063 | - | - | JF832205 | - |
| C. montana (Lam.) Bernh. ex Desv. | 6969 | Harris 09-073 (DUKE) | Canada: Ontario | - | - | - | - | JF832267 |
| C. moupinensis Franch. | 4861 | Schuettpelz 1118A et al. (DUKE) | Taiwan: Nantou | JF832064 | - | - | JF832206 | - |
| C. moupinensis Franch. | - | Kuo 460 | Taiwan: Nantou | - | - | - 832120 | - | JF832268 |
| Gymnocarpium dryopteris (L.) Newman | 7059 | Larsson 6 (UPS) | Sweden: Uppsala | JF832068 | JF832165 | JF832120 | JF832218 | JF832277 |
| G. oyamense (Baker) Ching | - | Liu 9688 (TAIF) | China: Sichuan | - | - | - |  | JF832278 |
| G. oyamense (Baker) Ching | 6399 | Nakato s.n. (DUKE) | Cult. (origin: Japan) | JF832069 | JF832166 | JF832121 | JF832219 | - |
| G. remotepinnatum (Hayata) Ching | 3066 | Yatskievych 02-31 (MO) | China: Yunnan | EF463317 | EF463580 | JF832122 | JF832220 |  |
| G. remotepinnatum (Hayata) Ching | - | Kuo 149 (TAIF) | Taiwan | - | - | - | - | JF303926 ${ }^{\text {c }}$ |
| Dipls/Homal. |  |  |  |  |  |  |  |  |
| Diplaziopsis cavaleriana (Christ) C. Chr. | - | Unknown | Unknown | D43909 ${ }^{\text {i }}$ | JF | JF80 | JF | - 282272 |
| D. cavaleriana (Christ) C. Chr. | - | Liu 9717 (TAIF) | China: Sichuan | - | JF832162 | JF832113 | JF832211 | JF832272 |
| D. javanica (Blume) C. Chr. | 4967 | Schuettpelz 1220A (DUKE) | Taiwan: Ilan | JF832066 | JF832163 | JF832114 | JF832212 | - |
| D. javanica (Blume) C. Chr. | - | Kuo 138 (TAIF) | Taiwan | JF8 | JF | JF | JF | JF303928 ${ }^{\text {C }}$ |
| Homalosorus pycnocarpos (Spreng.) Pic. Serm. | 1226 | Cranfill s.n. (UC) | Cult. (J.Mickel pers. garden) | JF832070 | JF832168 | JF832125 | JF832222 |  |
| H. pycnocarpos (Spreng.) Pic. Serm. | - | Kuo 122 (TAIF) | USA: Massachusetts | - | - | - | - | JF303929 ${ }^{\text {c }}$ |
| Hemidictyum |  |  |  |  |  |  |  |  |
| Hemidictyum marginatum (L.) C. Presl | 3054 | Christenhusz 2476 | French Guiana: <br> Montagnes Tortue | EF463318 | EF463581 | JF832124 | JF832221 | JF303927 ${ }^{\text {c }}$ |
| Onocleaceae |  |  |  |  |  |  |  |  |
| Matteuccia struthiopteris (L.) Tod. | 7053 | Larsson 5 (UPS) | Cult. (Uppsala Bot. Gard.) | JF832074 | JF832171 | JF832130 | JF832227 | - |
| M. struthiopteris (L.) Tod. | - | No voucher taken. | Cult. (Dr Cecilia Koo Bot. Cons. Center \#019687) | - | - | - | - | JF303936 ${ }^{\text {c }}$ |

Appendix 1. Continued

|  | DB\# | Voucher | Provenance | $r b c L$ | $\operatorname{atp} B$ | atp $A$ | trnG-R | matK |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Onoclea sensibilis L. | 2998 | Schuettpelz 353 (DUKE) | Cult. (Juniper Level Bot. Gard.; origin: USA) | - | EF463488 ${ }^{\text {j }}$ | JF832132 | JF832229 | - |
| O. sensibilis L. | - | Kuo 115 (TAIF) | USA: Massachusetts | - | - | - | - | JF303935 ${ }^{\text {c }}$ |
| O. sensibilis L. | - | Unknown | Unknown | JF832076 ${ }^{\text {b }}$ | - | - | - |  |
| Onocleopsis hintonii Ballard | 6729 | Rothfels 3360 et al. (DUKE) | Mexico: Oaxaca | JF832077 | JF832174 | JF832133 | JF832230 | JF832281 |
| Pentarhizidium intermedium (C. Chr.) Hayata |  | Liu 9689 (TAIF) | China: Sichuan |  |  |  |  | JF832282 |
| P. intermedium (C. Chr.) Hayata | 6991 | Dickoré 11519 (UC) | China: Tibet | JF832078 | JF832175 | JF832134 | JF832231 |  |
| P. orientale (Hook.) Hayata | 6992 | 1980 Sino-Amer. Exped. \#2050 (UC) | China: Hubei | JF832079 | JF832176 | JF832135 | JF832232 |  |
| P. orientale (Hook.) Hayata | - | No voucher taken | Cult. (Dr Cecilia Koo Bot. Cons. Center K032501; origin: Taiwan) | - | - | - | - | JF832283 |
| Rhachidosorus |  |  |  |  |  |  |  |  |
| Rhachidosorus mesosorus (Makino) Ching |  | Fujimoto s.n. (TNS) | Cult. (TBG; origin: Japan) | - | - |  |  | JF832287 |
| R. mesosorus (Makino) Ching | 7075 | Matsumoto s.n. (DUKE) | Cult. (TBG; origin: Japan) | JF832084 | JF832180 | JF832140 | JF832238 |  |
| R. pulcher (Tagawa) Ching | - | Chen 441 (TNU) | Taiwan | JF303971 ${ }^{\text {c }}$ | JF832181 | JF303998 ${ }^{\text {c }}$ | JF832239 | JF303962 ${ }^{\text {c }}$ |
| Thelypteridaceae |  |  |  |  |  |  |  |  |
| Coryphopteris seemannii Holttum | 3774 | Game 95/147 (UC) | Fiji: VitiLevu | EF463300 ${ }^{\text {j }}$ | EF463556 ${ }^{\text {j }}$ | JF832105 | JF832201 | JF832264 |
| Macrothelypteris torresiana (Gaudich.) Ching | 6502 | Rothfels 3050 et al. (DUKE) | Mexico: San Luis Potosi | JF832073 | JF832170 | JF832129 | JF832226 |  |
| M. torresiana (Gaudich.) Ching |  | Kuo 826 (TAIF) | Taiwan |  |  |  |  | JF303931 ${ }^{\text {c }}$ |
| Phegopteris connectilis (Michx.) Watt | 7060 | Larsson 17 (UPS) | Sweden: Uppsala | JF832080 | JF832177 | JF832136 | JF832233 |  |
| P. connectilis (Michx.) Watt |  | Kuo 151 (TAIF) | Taiwan |  |  |  |  | JF303932 ${ }^{\text {c }}$ |
| Pseudophegopteris cruciata (Willd.) Holttum | 3559 | Janssen 2724 (P) | France: Ile de la Reunion | EF463279 ${ }^{\text {j }}$ | EF463535 ${ }^{\text {j }}$ | JF832139 | JF832236 | JF832286 |
| Thelypteris dentata (Forssk.) E.P. St. John | 3654 | Schuettpelz 607 et al. (B) | Cult. (Botanischer Garten Berlin-Dahlem) | EF463284 ${ }^{\text {j }}$ | EF463540 ${ }^{\text {j }}$ | JF832144 | JF832243 | JF832290 |
| T. gemmulifera (Hieron.) A.R. Sm. | 3747 | Huiet s.n. et al. (UC) | Cult. (UCBG 83.0943; origin: Venezuala) | EF463285 ${ }^{\text {j }}$ | EF463541 ${ }^{\text {j }}$ | JF832145 | JF832244 | JF832291 |
| T. palustris Schott | 7055 | Larsson 16 (UPS) | Sweden: Uppsala | JF832085 | JF832182 | JF832146 | JF832245 | JF832292 |
| T. pozoi (Lag.) C.V. Morton | - | Kuo 110 (TAIF) | Taiwan: Orchid Is. |  | - | JF304004 ${ }^{\text {c }}$ | - | JF303933 ${ }^{\text {c }}$ |
| T. pozoi (Lag.) C.V. Morton | - | Unknown | Unknown | AB059579 ${ }^{\text {n }}$ | - |  | - |  |
| T. uraiensis (Rosenst.) Ching | - | Kuo 139 (TAIF) | Taiwan: Taipei | JF303972 ${ }^{\text {c }}$ | - | JF304002 ${ }^{\text {C }}$ | - | JF303934 ${ }^{\text {c }}$ |
| Woodsia and allies |  |  |  |  |  |  |  |  |
| Cheilanthopsis elongata (Hook.) Copel. | 7169 | Polunin, Sykes and Williams (UPS) | Nepal: Maharigaon | JF832060 | JF832158 | JF832103 | - | - |
| Protowoodsia manchuriensis (Hook.) Ching | 7101 | Nakaike, Sakakibara and Ishizuka s.n. (E) | Japan: Koisawa | JF832082 | JF832179 | JF832138 | - | - |
| P. manchuriensis (Hook.) Ching | - | Fujimoto s.n. (TNS) | Cult. (TBG; origin: Japan) | - 83326 | - 232167 | - 832123 | - | JF832284 |
| Woodsia andersonii Beddome | 7102 | Ho et al. 2601 (E) | China: Qinghai | JF833268 | JF832167 | JF832123 | - | JF832279 |
| W. ilvensis (L.) R. Br. | 7050 | Karis s.n. (UPS) | Sweden: Runmarö | JF832086 | JF832183 | JF832147 | JF832246 | JF832293 |
| W. mollis (Kaulf.) J. Sm. | 7013 | Larsson 103 (UPS) | Mexico: Hidalgo | JF832087 | JF832184 | JF832148 | JF832247 | JF832294 |
| W. plummerae Lemmon | 5872 | Schuettpelz 1235A (DUKE) | USA: Arizona | JF832088 | JF832185 | JF832149 | JF832248 | JF832295 |
| W. polystichoides D.C. Eaton | 7015 | Larsson 4 (UPS) | Cult. (Uppsala Bot. Gard.) | JF832089 | JF832186 | JF832150 | JF832249 |  |
| W. polystichoides D.C. Eaton | - | Li 420 (TAIF) | Taiwan | - | - | - | - | JF303930 ${ }^{\text {c }}$ |

[^0]
[^0]:    Notes: DB\#: Pryer Lab DNA Database number (www.pryerlab.net). Shaded headings mark each "major clade," following Figure 6. Superscripts following GenBank numbers indicate previously published sequences from: ${ }^{a}$ Duffy et al. (2009), ${ }^{b}$ Gastony and Ungerer (1997), ${ }^{\mathrm{c}}$ Kuo et al. (2011), ${ }^{\mathrm{d}} \mathrm{J}$.-M. Lu and D.-Z Li. (unpublished data), ${ }^{\mathrm{e}} \mathrm{N}$. Murikami, S. Nogami, M. Watanabe, and K. Iwatsuki. (unpublished data), ${ }^{\mathrm{f}}$ Nakahira (2000), ${ }^{8}$ Pinter et al. (2002), ${ }^{\text {h }}$ Rothfels et al. (2008), ${ }^{\text {is }}$ Sano et al. (2000a), ischuettpelz and Pryer (2007), ${ }^{\mathrm{k}}$ Schuettpelz et al. (2007), ${ }^{1}$ Smith and Cranfill (2002), ${ }^{\text {m }}$ Wolf et al. (1994), " $Y$. Yatebe and N. Murikami (unpublished data). Data missing or not applicable are indicated with a "-".

