A Plastid Phylogeny of the Cosmopolitan Fern Family Cystopteridaceae (Polypodiophyta)

Carl J. Rothfels, Michael D. Windham, and Kathleen M. Pryer

1Department of Biology, Duke University, Post Office Box 90338, Durham, North Carolina 27708, U.S.A.
2Author for correspondence (crothfels@yahoo.ca)

Abstract—Among the novel results of recent molecular phylogenetic analyses are the unexpectedly close evolutionary relationships of the genera Acystopteris, Cystopteris, and Gymnocarpium, and the phylogenetic isolation of these genera from Woodsia. As a consequence, these three genera have been removed from Woodsiaceae and placed into their own family, the Cystopteridaceae. Despite the ubiquity of this family in rocky habitats across the northern hemisphere, and its cosmopolitan distribution (occurring on every continent except Antarctica), sampling of the Cystopteridaceae in phylogenetic studies to date has been sparse. Here we assemble a three-locus plastid dataset (matK, rbcL, trnG-R) that includes most recognized species in the family and multiple accessions of widespread taxa from across their geographic ranges. All three sampled genera are robustly supported as monophyletic, Cystopteris is strongly supported as sister to Acystopteris, and those two genera together are sister to Gymnocarpium. The Gymnocarpium phylogeny is deeply divided into three major clades, which we label the disjunctum clade, the robertianum clade, and core Gymnocarpium. The Cystopteris phylogeny, similarly, features four deeply diverged clades: C. montana, the sudetica clade, the bulbifera clade, and the fragilis complex. Acystopteris includes only three species, each of which is supported as monophyletic, with A. taiwaniana sister to the japonica/tenuisecta clade. Our results yield the first species-level phylogeny of the Cystopteridaceae and the first molecular phylogenetic evidence for species boundaries. These data provide an essential foundation for further investigations of complex patterns of geographic diversification, speciation, and reticulation in this family.

Keywords—Cosmopolitan species, Cystopteris, fern phylogeny, Gymnocarpium, intralinkage incongruence, species complex.

Cystopteris Bernh. and Gymnocarpium Newman—including the bulblet fern, bladder ferns, fragile ferns, and oak ferns (see Fig. 1A–C, E–H)—are among the most frequently encountered and familiar ferns in the northern hemisphere, occurring in most forested and rocky habitats in North America, Europe, and Asia. However, despite their familiarity, their phylogenetic relationships have been contentious. Cystopteris Nakai (Fig. 1D) and Cystopteris have long been considered close relatives, with many authors historically treating them together under a broad concept of Cystopteris (e.g. Tagawa 1935; Blasdell 1963). Prior to the proliferation of molecular evidence, however, most taxonomists did not consider Cystopteris s.l. and Gymnocarpium to be closely allied. Indeed, these two taxa usually were assigned to the dryopteroid and aethrioid fern lineages, respectively, with the caveat that they were each morphologically anomalous within those lineages and that their phylogenetic positions were thus uncertain (e.g. Sledge 1973). This confusion regarding the affinities of Cystopteris and Gymnocarpium to other polypod ferns has continued until very recently. For example, when naming the family Cystopteridaceae, Schmakov (2001) included Pseudacystopteris Ching (which belongs in the Aethriaceae; Kato 1977; Sano et al. 2000; Fraser-Jenkins 2008; Liu 2008; Rothfels et al. 2012b). Quite recently, Z. R. Wang (1997), M. L. Wang et al. (2004) and Smith et al. (2006) each advocated familial concepts that grouped Cystopteris and Gymnocarpium with very distantly related taxa (Rothfels et al. 2012b).

Wolf et al. (1994) and Hasebe et al. (1995) provided the first molecular evidence that Cystopteris and Gymnocarpium are, indeed, closely allied to one another, and that they are evolutionarily distinct from both the dryopteroid (Dryopteridaceae sensu Smith et al. (2006)) and aethrioid ferns (Aethriaceae sensu Rothschild et al. (2012b)). Subsequent molecular phylogenetic studies—those with greater taxon and character sampling—have yielded an increasingly clear understanding of this group, culminating in the resurrection and recircumscription of the Cystopteridaceae (Rothfels et al. 2012b). This family occupies a critical position sister to the rest of the eupolypod II clade (Sano et al. 2000; Schuettpelz and Pryer 2007; Kuo et al. 2011; Rothfels et al. 2012a) and represents a deeply isolated lineage within eupolypod ferns—it last shared a common ancestor with other extant fern lineages approximately 100 million years ago (Schuettpelz and Pryer 2009; Rothfels et al. 2012a).

The Cystopteridaceae consists of at least three genera—Gymnocarpium, Acystopteris, and Cystopteris—and possibly Cystoathyrium Ching (Rothfels et al. 2012a, 2012b). The latter includes a single species from China that is known only from the type specimen at PE and may be extinct (Rothfels et al. 2012b). Although Cystoathyrium has not been included in any molecular phylogenetic study to date, the limited morphological information available suggests it may be a member of the Cystopteridaceae (Rothfels et al. 2012b; Sundue and Rothfels, unpubl.). The other two genera comprise approximately 37 species: seven in Gymnocarpium (Sarvela 1978; Pryer 1993), three in Acystopteris, and about 27 in Cystopteris (Blasdell 1963; Rothfels 2012). Although the circumscription of the Cystopteridaceae and its placement within the fern tree of life are now relatively well established, species boundaries and relationships among the included taxa remain unclear. Cystopteris s.l. was last monographed 50 yr ago (Blasdell 1963; he included Acystopteris in his generic concept) and the closest we have to a global monograph of Gymnocarpium is a six-page synopsis (Sarvela 1978). Both Cystopteris and Gymnocarpium have extremely broad geographical distributions (Fig. 2). Populations of G. dryopteris, for example, are scattered across the northern half of both North America and Eurasia (Fig. 2A), and Cystopteris fragilis s.l. is perhaps the most widely distributed fern in the world, ranging from the high arctic to Tierra del Fuego in the Americas, blanketing most of Eurasia, occurring in eastern and southern Africa, eastern Australia, New Zealand, Hawai‘i, and many isolated rocky oceanic islands (Fig. 2C, D; the “C. fragilis complex”). The few biosystematic investigations undertaken to date have revealed that the
widespread species of both Cystopteris and Gymnocarpium include multiple independent lineages; these often differ in ploidy level and typically involve reticulate evolutionary histories (Vida 1972, 1974; Sarvela 1978; Pryer et al. 1983; Haufler et al. 1990; Haufler and Windham 1991; Pryer and Haufler 1993).

In order to explore the intriguing patterns of diversification and phylogeography within these species complexes, we first need to understand their basic phylogenetic relationships. Here, we assemble a three-locus plastid dataset that includes most recognized species of Cystopteridaceae, including multiple accessions from across the geographic ranges of many taxa (especially the polyploids). Our primary goal is to establish the first robust phylogeny for the family—focusing on the branching (divergent) relationships as a necessary prerequisite for investigating species boundaries and reticulate evolution within Cystopteridaceae (Rothfels et al., unpubl.).

**Materials and Methods**

**Taxonomic Sampling**—We analyzed an ingroup sample of 75 accessions of Cystopteridaceae, selected to maximize the inclusion of named taxa appearing in published taxonomic works (Tagawa 1935; Blandell 1963; Mickel 1972; Bir and Trikha 1974; Sarvela 1978; Sarvela et al. 1981; Moran 1983; Pryer et al. 1983; Haufler et al. 1990; Haufler and Windham 1991; Pryer and Haufler 1993).

In order to explore the intriguing patterns of diversification and phylogeography within these species complexes, we first need to understand their basic phylogenetic relationships. Here, we assemble a three-locus plastid dataset that includes most recognized species of Cystopteridaceae, including multiple accessions from across the geographic ranges of many taxa (especially the polyploids). Our primary goal is to establish the first robust phylogeny for the family—focusing on the branching (divergent) relationships as a necessary prerequisite for investigating species boundaries and reticulate evolution within Cystopteridaceae (Rothfels et al., unpubl.).

**Materials and Methods**

**Taxonomic Sampling**—We analyzed an ingroup sample of 75 accessions of Cystopteridaceae, selected to maximize the inclusion of named taxa appearing in published taxonomic works (Tagawa 1935; Blandell 1963; Mickel 1972; Bir and Trikha 1974; Sarvela 1978; Sarvela et al. 1981; Moran 1983; Pryer et al. 1983; Haufler et al. 1990; Haufler and Windham 1991; Pryer and Haufler 1993).
Fig. 2. Geographic ranges of sampled taxa of Cystopteridaceae. Ranges are approximated by the colored polygons, with dotted lines indicating uncertain range limits. Colored circles indicate the collection location of the vouchers used in this study; a numeral inside the colored circle indicates the number of accessions from that location. A: Gymnocarpium. B: Acystopteris and Cystopteris montana. C: The C. bulbifera clade, the C. sudetica clade, and C. laurentiana and C. reevesiana. D: The C. fragilis complex (excluding C. laurentiana and C. reevesiana, which are presented in panel C). Range boundaries determined from published floristic works (Tagawa 1935; Mickel 1972; Bir and Trikha 1974; Sarvela 1978; Sarvela et al. 1981; Moran 1983; Britton et al. 1984; Fraser-Jenkins 1986; Lobin 1986; Prada 1986; Salvo and Otermin 1986; Breckle 1987; Haufler et al. 1990; Haufler and Windham 1991; Haufler et al. 1993; Pryer 1993; Denk 1998; Latorre 2000; Velayos et al. 2001; Mickel and Tejero-Diez 2004; Murphy and Rumsey 2005; Fraser-Jenkins 2008; Wang 2008; Crouch et al. 2011; Japanese Society for Plant Systematics 2012).
Table 1. Primers used in amplification and sequencing (F = forward; R = reverse).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer</th>
<th>Sequence (5'–3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>matK</td>
<td>AJmatKf1</td>
<td>GTATTACAKAAAAAGTGRAGCCTTAG</td>
<td>This study</td>
</tr>
<tr>
<td>matK</td>
<td>AJmatKf3</td>
<td>TGGAAGGTAYTCAGTTCCGGTCTG</td>
<td>This study</td>
</tr>
<tr>
<td>matK</td>
<td>AJmatKr1</td>
<td>ATYCTCAATCTACGCAATTCAT</td>
<td>This study</td>
</tr>
<tr>
<td>matK</td>
<td>AJmatKr3B</td>
<td>CGATTGIAAGTARAAATTTCC</td>
<td>This study</td>
</tr>
<tr>
<td>rbcL</td>
<td>ESRBCL1F</td>
<td>TCAGGACTCCACTTACTAGCCTCAG</td>
<td>(Schuettpelz and Pryer 2007)</td>
</tr>
<tr>
<td>rbcL</td>
<td>ESRBCL663R</td>
<td>ATGTCACGCAAAAGGCACTAAGC</td>
<td>(Schuettpelz and Pryer 2007)</td>
</tr>
<tr>
<td>rbcL</td>
<td>ESRBCL1361R</td>
<td>TACRAAATARGAAACRCTCCTCCACAG</td>
<td>(Schuettpelz and Pryer 2007)</td>
</tr>
<tr>
<td>trnG-R</td>
<td>trnGf1</td>
<td>GGCGGTATTAGTTAGTTAGTAA</td>
<td>(Nagalingum et al. 2007)</td>
</tr>
<tr>
<td>trnG-R</td>
<td>CRcsySRNGf1</td>
<td>GTGATACGCAACAGACCAGA</td>
<td>This study</td>
</tr>
<tr>
<td>trnG-R</td>
<td>CRcsySRNGr1</td>
<td>GTGCATCCTATAAAATCAGTACG</td>
<td>This study</td>
</tr>
<tr>
<td>trnG-R</td>
<td>trnR22R</td>
<td>CTATCATTAGACGATGAC</td>
<td>(Nagalingum et al. 2007)</td>
</tr>
</tbody>
</table>

Pryer 2007; Rothfels et al. 2012a). Our total sample includes 84 accessions (Appendix 1).

DNA Isolation, Amplification, and Sequencing—DNA was extracted from herbarium specimens or silica-dried material in the Fern Lab Silica Archive (http://fernlab.biology.duke.edu/) using a 96-well modification (Beck et al. 2011; doi:10.5061/dryad.11p75m70) of a standard CTAB protocol (Doley and Dickson 1987), or using DNeasy kits (Qiagen, Valencia, California, USA). Three plastid loci were selected for analysis: matK, rbcL, and the trnG-trnR intergenic spacer (hereafter "trnG-R"). Primer sequences and associated data are provided in Table 1. The full lengths of matK and trnG-R were amplified and sequenced in two overlapping fragments, using the primers AJmatKf1+AJmatKr3B and AJmatKf3+AJmatKr1 for matK and trnGf1+CRcsySRNGf1 and CRcsySRNGr1+trnR22R for trnG-R. Most rbcL sequences were amplified in one piece using the primer pair ESRBCL1F+ES1361R. For some herbarium specimens with degraded DNA, rbcL was amplified in two pieces, using the primer pairs ESRBCL1F+ES633R and ES645F+ES1361R. Loci were amplified in 21 μl reactions consisting of 2 μl Denville buffer (10x), 2 μl dNTPs (each 2 mM), 0.2 μl Denville Choice Taq (5 U/μl), 1 μl of each primer (10 μM), 1 μl of DNA, and 13.6 μl of water. Our thermal cycling protocol for matK consisted of an initial denaturation step (94˚C for 3 min), 35 denaturation, annealing, and elongation cycles (94˚C for 45 sec, 50˚C for 30 sec, 72˚C for 1.5 min), and a final elongation step (72˚C for 10 min). For rbcL and trnG-R we used the same program, except that the annealing temperature was 45˚C for rbcL and 35˚C for trnG-R. PCR products were purified using Shrimp Alkaline Phosphatase (USB, Cleveland, Ohio) following established protocols (Rothfels et al. 2012a) and sequenced on an ABI Prism 3700 DNA Analyzer (Applied Biosystems, Carlsbad, California) at the Duke University Genome Sequencing and Analysis Core Resource, again using established protocols (Schuettpelz and Pryer 2007). Chromatograms were assembled and edited in Sequencher v4.5 (Gene Codes Corporation, Ann Arbor, Michigan), and the resulting 169 newly generated sequences are deposited in GenBank (Appendix 1).

Sequence Alignment and Phylogenetic Analysis—Sequences for each locus were manually aligned in Mesquite v2.72 (Maddison and Maddison 2009). Ambiguously aligned regions (limited to trnG-R) were excluded prior to analysis. A total of four datasets were analyzed: the three single-locus datasets (to test for incongruence) and a combined three-locus dataset (Table 2). The single-locus datasets were analyzed under maximum likelihood (ML) in Garli v2.0 (Zwickl 2006), using the best model as determined by the small-sample correction for the Akaike Information Criterion (AICc; Akaile 1974; Burnham and Anderson 2004) in jModeltest v0.1.1 (Posada 2008; see Table 2). For each locus, ML tree searches were performed on 500 bootstrap pseudoreplicate datasets, each searched from two different random-addition starting trees; other settings were left at their default values. The majority-rule consensus trees from each pool of bootstrap trees were compared for highly supported (≥70% bootstrap support) incompatible splits (Mason-Gamer and Kellogg 1996).

The combined three-locus dataset was analyzed under both ML and Bayesian frameworks. The ML analyses used the same settings as the single-locus analyses (above), in a single Garli (Zwickl 2006) run with the data partitioned by locus, substitution parameters unlinked among partitions, and each partition permitted its own average rate (subset-specific rates = 1). The search for the ML best tree started from each of 20 different random-addition starting trees, and support was assessed with 1,000 bootstrap pseudoreplicates, each searched from two random-addition starting trees. The Bayesian analyses were performed in the parallel version of MrBayes v3.1.2 (Ronquist and Huelsenbeck 2003; Altekar et al. 2004), with parameters unlinked among the three partitions. The best-fitting models for rbcL and trnG-R (see Table 2) were implemented with the nst=mixed, statefreqpr=dirichlet(1,1,1,1) and rates=invgamma settings; rbcL differed in that base frequencies were fixed at 0.25 each (statefreqpr=fixed(equal)). The model settings for matK were more straightforward (nst=6 rates=gamma statefreqpr=dirichlet(1,1,1,1)). The average rates for each partition were allowed to be different (ratepr=variable); other priors were left at their default values. Four independent runs, each with four chains (one cold, three heated), were run for 50 million generations with a sample taken every 7,500 generations. The resulting sample parameter traces were visualized in Tracer v1.5 (Rambaut and Drummond 2007). The runs each converged (and to the same area of parameter space) well before 500,000 generations; to be very conservative, we excluded the first 5 million generations of each run as burn-in, prior to summarizing the posterior. Our final pool included 2,800,000 samples; effective sample sizes for all parameters were greater than 1,000. Our final three-locus dataset is available from TreeBASE (accession S13449).

Results

Phylogenetic Analyses—Four datasets were analyzed in this study: one for each locus individually, and one of the three loci combined. While the differing numbers of sequences among the datasets make precise comparisons difficult, matK appears to be more informative (i.e. it has a higher percentage of strongly supported bipartitions) than trnG-R, which in turn outperforms rbcL (Table 2).

There are two well-supported conflicts among the loci. The first of these conflicts is relatively minor: trnG-R supports a

Table 2. Statistics for the datasets analyzed in this study. Missing data includes both uncertain bases (?), N, R, Y, etc.) and gaps (-). MLBS: maximum likelihood bootstrap support. *The combined dataset was analyzed under a partitioned model, with each locus given its own best-fitting model.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Taxa</th>
<th>Included sites</th>
<th>Variable sites</th>
<th>Missing data (%)</th>
<th>Best-fitting model</th>
<th>Mean</th>
<th>MLBS (≥70%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>matK</td>
<td>74</td>
<td>1211</td>
<td>559</td>
<td>1.6</td>
<td>GTR+G</td>
<td>90%</td>
<td>85%</td>
</tr>
<tr>
<td>rbcL</td>
<td>61</td>
<td>1309</td>
<td>256</td>
<td>6.2</td>
<td>TrNef+I+G</td>
<td>81%</td>
<td>67%</td>
</tr>
<tr>
<td>trnG-R</td>
<td>84</td>
<td>1119</td>
<td>538</td>
<td>4.9</td>
<td>TVM+I+G</td>
<td>87%</td>
<td>80%</td>
</tr>
<tr>
<td>combined</td>
<td>84</td>
<td>3639</td>
<td>1353</td>
<td>17.4</td>
<td>*</td>
<td>90%</td>
<td>85%</td>
</tr>
</tbody>
</table>
The Gymnogarpini clade includes all our accessions except for those in the disjunctum clade, whereas rbcL supports a clade of all our Gymnogarpini accessions except for those in the robertianum clade (Fig. 3A). In this case, the conflicting bootstrap support is only marginally greater than our 70% cut-off (70.6% for trnG-R and 72% for rbcL). The second case of intralinkage incongruence is more substantial. Here, the conflicting relationships are near the base of the Cystopteris crown group (following the divergence of C. montana) and the discordance is much stronger. The matK locus supports a sister relationship between the bulbifera clade and the fragilis complex with 82.4% bootstrap support, whereas trnG-R supports a sister relationship between the bulbifera and sudetica clades with 92.6% bootstrap support (rbcL is equivocal; Fig. 3B). This incongruence is not the result of misidentification or lab error (the same extractions were used for all loci, and multiple accessions were involved in each case). Errors of alignment inference remain a possibility, but only one locus (trnG-R) has areas of ambiguous alignment, and they were excluded prior to analysis. Furthermore, careful review of the alignments failed to uncover any regions that could possibly be contributing to this result. The combined dataset includes 3639 sites for 84 taxa; additional dataset characteristics are provided in Table 2.

Cystopteridaceae Phylogeny—Analysis of the combined dataset inferred strong support for the vast majority of internodes across the tree, under both ML bootstrapping and Bayesian analysis (Fig. 4; Table 2). The monophyly of the family as a whole is maximally supported (1.0 posterior probability and 100% ML bootstrap support), as is the monophyly of the three genera. Cystopteris and Acystopteris are maximally supported as sister genera, and they, together, are sister to Gymnogarpini (Fig. 4).

The Gymnogarpini phylogeny features three deeply diverged "major clades": the disjunctum clade, the robertianum clade, and core Gymnogarpini. Though each of these is strongly supported, relationships among them are uncertain. The disjunctum clade includes the diploid G. disjunctum, the widespread allotetraploid G. dryopteris, and the triploid hybrid between them, G. × brittonianum. The G. robertianum clade includes all accessions of that species, as well as a Japanese accession of uncertain identity. The remainder of the genus comprises the core Gymnogarpini clade, including the morphologically anomalous G. oyamense, the eastern North American G. appalachianum, two taxa currently treated as subspecies of G. jessoense (subsp. jessoense and subsp. parvulum), and the east Asian G. remotepinnatum.

Within the Acystopteris clade, each of the three recognized species is maximally supported as monophyletic. Acystopteris japonica and A. tenuisecta are sister (1.0 posterior probability and 100% ML bootstrap support), and they, together, are sister to A. taiwaniensia.

The first split within Cystopteris is maximally supported, separating C. montana from the rest of the genus. The next branch is not well supported—similar to the situation at the base of Gymnogarpini. Consequently, relationships among three highly supported clades (sudetica, bulbifera, and the fragilis complex) remain uncertain. The sudetica clade is predominantly Asian; of its three recognized species, C. pellucida and C. moupinensis are exclusively East Asian, while C. sudetica extends west across Eurasia. In contrast, the bulbifera clade, which includes diploid C. bulbifera and related allopolyploids, is limited to North America. The bulk of Cystopteris species belong to the remaining clade (informally known as the fragilis complex), which is maximally supported as monophyletic. The first divergence within the C. fragilis clade robustly separates C. protrusa from the rest of the complex. Relationships among the remaining C. fragilis complex species are convoluted, with many taxa (including C. fragilis s.s.) appearing in multiple subclades.
Fig. 4. Maximum likelihood phylogram of the concatenated plastid data. Support values (maximum likelihood bootstrap support/Bayesian posterior probability) appear below the branches; values of 100% and 1.0, respectively, are indicated by asterisks. Bold branches are highly supported (>70% maximum likelihood bootstrap support and >0.95 posterior probability). Numbers in brackets following the species’ names are accession numbers from the Fern Lab database (http://fernlab.biology.duke.edu/); see Appendix 1 for further voucher information.
Discussion

Phylogenetic Analyses and Intralinkage Incongruence—The relative informativeness of the loci examined in our dataset (\textit{matK} outperforming \textit{trnG-R}, which was superior to \textit{rbcL}) is consistent with other molecular phylogenetic analyses of ferns across a variety of evolutionary depths (Kuo et al. 2011; F.-W. Li et al. 2011; Rothfels et al. 2012a). An unexpected result, however, was the appearance of well-supported conflicts among the individual loci (Fig. 3). Because all three loci are in the plastid genome, they should constitute a single non-recombining linkage group with a common evolutionary history. To the contrary, our data reveal two cases of intralinkage incongruence where at least one locus has ≥70% bootstrap support for a relationship that is incompatible with a relationship well-supported by another locus (again, with at least 70% support; Fig. 3; Mason-Gamer and Kellogg 1996). We suspect that this incongruence is due to the failure of our phylogenetic inference methods to fully capture the idiosyncrasies of molecular evolution in these taxa, as has been seen in other multi-region studies of single linkage groups (Rothfels et al. 2012a; Weisrock 2012). As with the mitogenomic data of Weisrock (2012), the conflicts in our data could be due to different patterns of selection operating on the individual loci, biases in base composition across loci and taxa, slight changes in taxon representation, or unusual variance in the stochastic substitution process.

Cystopteridaceae Phylogeny—Our broad results, demonstrating that both the family and its constituent genera are monophyletic and that \textit{Cystopteris} + \textit{Acystopteris} is sister to \textit{Gymnocarpium}, are consistent with the high support inferred for these relationships in earlier molecular phylogenies (Sano et al. 2000; Liu 2008; C. Li et al. 2011; Rothfels et al. 2012a). The monophyly of the family, and its deep divergence from its closest relatives (Rothfels et al. 2012a), further emphasizes the need to recognize these three genera as comprising their own family, rather than including them in a broad Athyriaceae sensu Wang et al. (2004) or Woodsiaceae sensu Smith et al. (2006).

Gymnocarpium Phylogeny—While \textit{Gymnocarpium} is maximally supported as monophyletic, the deepest divergence within the genus is uncertain (Fig. 4). Data from \textit{trnG-R} support the \textit{G. disjunctum} clade as sister to the remainder of the genus (71% bootstrap support), whereas \textit{rbcL} places \textit{G. robertianum} as the earliest diverging lineage (with 72% bootstrap support; Fig. 3A). On its own, \textit{matK} resolves the same relationship as \textit{trnG-R}, but without support, which is the same result obtained from the combined data (Fig. 4). Despite the addition of substantial amounts of data relative to previous studies (the largest previous sample of \textit{Gymnocarpium} was the three accessions included in Rothfels et al. (2012a)), the early evolutionary history of \textit{Gymnocarpium} remains enigmatic.

Of the three major groups that comprise \textit{Gymnocarpium}, the \textit{disjunctum} clade is the most straightforward to describe. In our sampling, it includes both accessions of the diploid species \textit{G. disjunctum}, all samples of the cosmopolitan tetraploid \textit{G. dryopteris}, and the only accession of triploid \textit{G. × brittonianum}, which is hypothesized to be a hybrid between the other two taxa (Pryer and Haufler 1993; Fig. 4). The grouping of \textit{G. disjunctum} and \textit{G. dryopteris} is expected, given that isozyme analyses suggest that \textit{G. dryopteris} is an allopolyploid between \textit{G. disjunctum} and \textit{G. appalachianum} (Pryer and Haufler 1993). Our data further support this hypothesis by indicating that \textit{G. disjunctum} is the maternal parent of all \textit{G. dryopteris} populations sampled. Another striking feature of this clade is the genetic uniformity observed across all loci—the sequences of \textit{G. dryopteris} are nearly identical regardless of whether they are from Alaska, Scandinavia, or Japan (Fig. 4). This genetic similarity of \textit{G. dryopteris} accessions across loci extends to the lone accession of \textit{G. dryopteris} var. aokigaharaense, providing no support to its recognition at the varietal level. However, in naming the variety, Nakaike (1969) pointed out that it was somewhat intermediate between \textit{G. dryopteris} and \textit{G. jessoense}. Thus, nuclear data will be necessary to corroborate or refute a possible hybrid origin of this taxon.

Our results for the \textit{robertianum} clade have three noteworthy elements. First, our single accession of \textit{G. robertianum} from the southern part of its North American range (5852: USA, Minnesota; see Appendix 1; Fig. 2A) is somewhat divergent from the other accessions. Second, as in tetraploid \textit{G. dryopteris}, the remaining accessions in this clade have near-identical sequences across loci, whether they are from North America, Scandinavia, or Japan (Fig. 4). And finally, one of these accessions is from far beyond the recognized range of \textit{G. robertianum} (7979: Japan, Iwate; see Fig. 2A). Plants of this morphology in that region are typically treated as \textit{G. jessoense} subsp. \textit{jessoense}. Our two samples of the latter from China and Pakistan are well-supported members of the core \textit{Gymnocarpium} clade and thus quite divergent from the \textit{robertianum} clade. Nuclear data will be needed to determine whether this Japanese accession represents an unrecognized range extension of \textit{G. robertianum} or is, instead, an unrecognized allopolyploid with \textit{G. robertianum} as the maternal parent. Future conclusions regarding the identity of these plants could have important consequences for \textit{Gymnocarpium} nomenclature, because \textit{G. jessoense} is typified on Japanese material (Koidzumi 1936).

The third major clade of \textit{Gymnocarpium} contains most of the named taxa, including \textit{G. oyamense}, which is sometimes recognized as the segregate genus \textit{Currania} Copel. (e.g. Copeland 1909; Lloyd and Klekowski 1970). The inclusion of this species within the core \textit{Gymnocarpium} is somewhat surprising given its anomalous morphology (including elongate sori and pinnatifid leaf dissection; Fig. 1C). It provides yet another example of a pattern seen elsewhere in Eupolypods II (e.g. Aspleniaceae, Onocleaceae, Blechnaceae; Rothfels et al. 2012a), where certain strongly apomorphic taxa are embedded within a larger group that has an apparently conserved, pleisiomorphic morphology. \textit{Gymnocarpium oyamense} also is noteworthy because it is on a much longer branch, suggesting a strongly elevated rate of molecular evolution. By comparison, all other species in the genus display remarkably clock-like rates of evolution. Such elevated rates of substitution have been seen in other groups of ferns (Des Marais et al. 2003; Schuettpelz and Pryer 2006; F.-W. Li et al. 2011; Rothfels et al. 2012a; Rothfels and Schuettpelz, unpubl.). In these studies, however, the elevated rates characterized significant portions of the phylogeny, not single, isolated species.

Within the core \textit{Gymnocarpium} clade, \textit{G. oyamense} is sister to the remaining species. The southeastern Appalachian endemic diploid \textit{G. appalachianum} is sister to a well-supported clade of \textit{G. jessoense} subsp. \textit{parvulum} accessions from North America and East Asia (Fig. 4). These two taxa are, in turn, sister to accessions of \textit{G. jessoense} subsp. \textit{jessoense} from
mainland Asia + the East Asian G. remotepinnatum (Fig. 4). The placement of G. jessoense subsp. jessoense (diploid) and G. jessoense subsp. parvulum (tetraploid) in different well-supported clades suggests that these taxa should be treated as distinct species. The taxonomy and nomenclature of Asian Gymnocalcarion species is particularly complex, and one problem that our dataset does not address is the uncertain identity of G. fedtschenkoanum Pojark. (Pojarkova 1950). Fraser-Jenkins (2008) assigns the majority of Himalayan Gymnocalcarion populations to G. fedtschenkoanum, restricting the name G. jessoense to more eastern populations. Future taxonomic work on Gymnocalcarion will need to emphasize a global approach, with a concerted focus on Asian taxa.

Acystopteris Phylogeny—Acystopteris has long been recognized as a distinctive element, either as a subgenus of Cystopteris (Blasdell 1963; Kato 1977) or as a separate genus (Nakai 1933; Smith et al. 2006; Wang 2008; Rothfels et al. 2012b; Japanese Society for Plant Systematics 2012). It shares with Cystopteris a base chromosome number of $x = 42$ (Mitui 1975) and a distinctive, hood-like indusium, but differs in having catenate scales and tuberculate light-tan spores (Blasdell 1963; Rothfels et al. 2012; Sundue and Rothfels, unpubl.), as well as an unusual, low-elevation, tropical distribution (Fig. 2B; Wang 2008). Our Acystopteris sampling includes two accessions from each of the three named species; each is monophyletic, as is the genus as a whole (Fig. 4). The well-supported position of Acystopteris as sister to Cystopteris permits either its continued recognition as a separate entity, or its merger into a broader concept of Cystopteris. Given that Acystopteris is easily diagnosed and has unique ecological and biogeographical features, as well as a relatively deep divergence from Cystopteris, we favor recognizing it at the generic level.

Within Acystopteris, A. tenuisecta is strongly supported as sister to A. japonica, and this clade is then sister to the Taiwan endemic A. taiwaniana. This result is surprising, given that A. taiwaniana is frequently treated as a variety of A. japonica and these two taxa are rarely, if ever, confused with A. tenuisecta. Wang (2008) suggested that A. taiwaniana might be an allopolyploid formed by hybridization between A. japonica and A. tenuisecta; however, this hypothesis is not supported by our data.

Cystopteris Phylogeny—Our phylogenetic hypothesis for Cystopteris includes four highly supported “major” clades (Fig. 4): montana, sudetica, bulbifera, and the C. fragilis complex (including C. prostrata). A clade comprising all C. montana accessions, including samples from Canada, Norway, and China, is strongly supported as sister to the remainder of the genus. This topology would permit the recognition of the genus Rhizomatopteris A.P. Khokhr. (typefied on C. montana; Khokhrjakov 1985), while still retaining a monophyletic Cystopteris. However, in naming Rhizomatopteris, Khokhrjakov (1985) included C. sudetica (and by extension, C. moupinensis and C. pellucida) in his generic concept. Rhizomatopteris sensu Khokhrjakov, then, includes all taxa with broadly deltate-to-pentagonal leaves and widely spaced internodes on long-creeping rhizomes (see Fig. 1E, H). This assemblage is not monophyletic, and there seems little value in recognizing a monotypic Rhizomatopteris (containing only C. montana).

Blasdell’s (1963: 80) “evolutionary tendencies” diagram (an early-tree-like visualization inferred using elements of Wagner’s (1980) “groundplan divergence scheme”) is remarkably similar to our molecular phylogeny. His diagram, based entirely on morphological characters, includes a basal division between Acystopteris and Cystopteris, followed by a subsequent split with one branch largely corresponding to our C. fragilis complex clade (but excluding C. diaphana), and another branch with C. montana at its base, C. bulbifera next, and C. sudetica and C. pellucida grouped together. Indeed, if the incongruent position of C. diaphana is ignored, and the extant species that he included on internal branches are moved to branch tips, then his diagram perfectly anticipates our four-clade (plus Acystopteris) result. On the other hand, Blasdell’s (1963) proposed classification of Cystopteris s.s. into two sections (Emarginatae and Cystopteris) is incompatible with certain aspects of our phylogeny. His Cystopteris section Emarginatae constitutes a paraphyletic grade comprised of C. montana, the sudetica and bulbifera clades, plus C. diaphana. In our phylogeny, the latter is deeply nested within the C. fragilis complex (Fig. 4), rendering Blasdell’s (1963) Cystopteris section Cystopteris paraphyletic as well.

Cystopteris montana, and the sudetica and bulbifera clades—The position of C. montana as sister to the rest of the genus is very strongly supported. Within the C. montana lineage, there is a shallow but highly supported split separating the two high-elevation Tibetan accessions from the North American and Scandinavian samples (Fig. 4). Although Blasdell (1963) reports both diploid and tetraploid cytotypes for C. montana based on spore size differences, the species is cytologically documented only as a tetraploid (Hauffler et al. 1993). Further investigations are necessary to determine if cryptic taxa exist within C. montana, and whether the Chinese taxon C. modesta Ching (not included in this study) is distinct from C. montana (see Fraser-Jenkins 2008).

Relationships among the three other major clades of Cystopteris—the sudetica clade, the bulbifera clade, and the C. fragilis complex—are not highly supported in our combined analysis. This lack of support is not caused by a lack of signal in our data, but rather by conflicting signals among the different partitions (see Intralinking Congruence discussion, above, and Fig. 3). Our sampling includes three species from the sudetica clade: C. pellucida, C. moupinensis, and C. sudetica. These species are primarily Asian; only C. sudetica extends west into Europe (Fig. 2C). Members of this clade are morphologically somewhat intermediate between C. montana (with which they share long-creeping rhizomes, widely-spaced leaves, and a tendency towards expanded basal pinnules on the lowermost pinnae) and C. bulbifera (with which they share an elongate-deltate leaf shape; see Fig. 1). In our sample, C. pellucida is sister to the rest of the sudetica clade (Fig. 4). An enigmatic species with a restricted range in central China (Fig. 2C), it differs from C. moupinensis in having more membranous leaves and a coarser leaf division (Wang 2008). Cystopteris moupinensis was recognized as a variety of C. sudetica by Blasdell (1963), and our analysis indicates that the two taxa are, indeed, very closely related. Our two accessions of C. sudetica are resolved as monophyletic, but these are very slightly diverged from—and form a polytomy with—our two C. moupinensis accessions (Fig. 4). The two species are allopatric (or very nearly so; Fig. 2C), and differ chiefly in the presence (C. sudetica) or absence (C. moupinensis) of glands on the indusia.

In our analysis, the bulbifera clade comprises all accessions of three species: C. bulbifera, C. tennesseensis, and C. utahensis. A limestone specialist of eastern North America with disjunct populations in the west, diploid C. bulbifera is hypothesized...
to be involved in the origin of tetraploid *C. tennesseensis* (through hybridization with *C. protrusa*; Shaver 1950; Haufler et al. 1990) and tetraploid *C. utahensis* (through hybridization with *C. reevesiana*; Haufler and Windham 1991). In our dataset, *C. bulbifera*, *C. tennesseensis*, and *C. utahensis* have nearly identical sequences (Fig. 4), confirming that *C. bulbifera* is the maternal parent of both tetraploids, and that they formed recently (as suggested by Haufler et al. (1990)). Our only sample of *C. laurentiana*, an allohexaploid also thought to contain a *C. bulbifera* genome (Wagner and Hagenah 1956), derived its plastid from a member of the *C. fragilis* complex (Fig. 4). Molecular confirmation that *C. bulbifera* was involved in the formation of *C. laurentiana* will therefore require data from the nuclear genome.

**Cystopteris fragilis Complex**—The *Cystopteris fragilis* complex has a special place in fern systematics, with Lovis (1978: 356) describing it as “perhaps the most formidable biosystematics problem in the ferns.” This species complex occurs on every continent except Antarctica (Fig. 2C, D), and includes ploidy levels ranging from diploid to octoploid (Blasdell 1963). Our plastid data, though preliminary, make some important new contributions. First, they permit a robust circumscription of the *C. fragilis* complex as the inclusive clade encompassing *C. fragilis* and *C. protrusa*, but not *C. bulbifera* or *C. sudetica* (Fig. 4). Second, they strongly support the position of the western North American, forest-dwelling diploid *C. protrusa* as sister to the rest of the complex (a position anticipated by Blasdell (1963) in his groundplan divergence scheme tree).

The bulk of the *C. fragilis* complex (i.e. excluding *C. protrusa*) is distinguished by a basal dichotomy. One side of the split is significantly supported, but barely so: it has 70% bootstrap support and 0.97 posterior probability (Fig. 4). This clade contains an interesting assemblage of taxa, including the alpine Eurasian hexaploid *C. alpina*, the endemic Australasian tetraploid *C. tasmanica*, both Hawaiian endemics (*C. sandwicensis* and *C. douglasii*), an eastern North American accession of *C. tenuis*, and three North American accessions of *C. fragilis* (including two unnamed hexaploids mentioned by Haufler and Windham (1991)). Most *Cystopteris* specimens with rugose spores—often called * dickieana* R. Sim. (Alston 1951; Bir and Trikha 1974; Nardi 1974; Wang 1983; Prada 1986; Parks et al. 2000) or *C. fragilis* subsp. * dickieana* (R. Sim) Hyl. (Fraser-Jenkins 2008)—also fall in this clade.

The sister clade is highly supported, and includes our only western North American accession of *C. tenuis*, all samples of *C. reevesiana* (a diploid), the North American putative allohexaploid *C. laurentiana*, various collections of *C. fragilis* from Europe and North America, *C. diaphana*, and single accessions identified as *C. membranifolia* and *C. millefolia* (Fig. 4). The latter three taxa form a weakly supported (61% bootstrap support and 0.99 posterior probability) clade that includes all Latin American samples of the *C. fragilis* complex, providing some support for Blasdell’s (1963) broad application of the name *C. diaphana* to plants from this region.

Our preliminary results for the *C. fragilis* clade confirm that its reputation for systematic complexity is well deserved. Except for *C. protrusa*, none of the species for which we included multiple accessions is monophyletic (Fig. 4). Each of the three known diploid members—*C. protrusa*, *C. reevesiana*, and a diploid cytotype of *C. diaphana* (Blasdell 1963)—lacks the morphological distinctiveness characteristic of species in the other clades of *Cystopteris*, making it difficult to diagnose polyploid taxa or to infer parentage based on morphological data alone. Though plastid DNA provide critical information, they tell only part of the story. Any progress on defining species limits in this group will depend on coordinated cytological and biparental molecular analyses, which are currently underway (Rothfels et al., unpubl.).

**Acknowledgments.** Great thanks are due to the many collectors who provided high quality materials that would otherwise be unavailable, and who thus made research of this scope possible. In particular, we would like to thank Marcelo Arana, David Barrington, Sam Brinker, Atsushi Ebihara, Alan Harris, Masahiro Kato, Ann Kelsey, Li-Yuang Kuo, Anders Larson, Fernando Matos, Jordan Metzgar, Michael Oldham, Daniel Palmer, Tom Rancer, Eric Schuettelpelz, Scott F. Smith, Erin Sigel, Mike Thorsen, K. Wood, and George Yatskievych. We dedicate this paper to the memory of Donald M. Britton (1923–2012), who introduced KMP to the Cystopteridaceae long before it was recognized as such. The curators and staff of the following herbaria provided critical assistance, in loaning their specimens for study and sampling: DUKE, MO, NY, P, TAFE, TNS, TRTE, and UC; particular thanks are due to Robbin Moran (NY), Alan R. Smith (UC) and George Yatskievych (MO) for their hospitality and assistance. Duke undergraduate Anne Johnson provided skilled assistance in the lab. The U.S. National Park Service gave permission to sample populations of *Gymnocarpium appalachianum* in Shenandoah National Park (permit SHEN-2010-SCI-0019) and critical Mexican material was secured under permit SCGA/DGGFS/712/13239/09 with the generous support of Universidad Nacional Autónoma de México (UNAM) and Mark Olson. This research was supported by funding from a Natural Sciences and Engineering Research Council (Canada) PGSD to CJR, an NSF DDIG award to KMP and CJR (DEB-1110767), a Duke Biology Department Grant-in-Aid to CJR, and a Duke University Center for Latin American and Caribbean Studies graduate student field research grant to CJR. The “HerbScan” scanner provided to the Duke herbarium under the Andrew W. Mellon Foundation Agreement #11100624 facilitated our production of the fern silhouettes shown in Fig. 1. This manuscript was greatly improved through comments from Editor-In-Chief Tom Rancer, Associate Editor Austin R. Mast, and two anonymous reviewers.

**Literature Cited**


Appendix 1. A list of accesses sampled in this study, presented in the following format: Species, Voucher (HERBARIUM ACRONYM), Fern


Cystopteris dickieana


Weisrock, D. W. 2012. Concordance analysis in mitogenomic phylo-

Zwickl, D. J. 2006. 


Molecular Phylogenetics and Evolution

Molecular Systematics

Cystopteris montana

Cystopteris tenuis

Cystopteris laevifolia

Cystopteris hypophylla

Cystopteris Dickieana

Cystopteris Dickieana

Cystopteris Dickieana

Cystopteris Dickieana

Cystopteris Dickieana

Cystopteris Dickieana