UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1182

Systematics of *Woodsia*

Ferns, bioinformatics and more

ANDERS LARSSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

ISSN 1651-6214
ISBN 978-91-554-9040-9
urn:nbn:se:uu:diva-232233

Dissertation presented at Uppsala University to be publicly examined in Lindahlsalen, Norbyvägen 18B, Uppsala, Friday, 31 October 2014 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Anne Krag Brysting (Centre for Ecological and Evolutionary Synthesis, Universitetet i Oslo).

Abstract

Larsson, A. 2014. Systematics of *Woodsia*. Ferns, bioinformatics and more. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1182. 36 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9040-9.

Ferns are one of the three main clades of vascular plats. They have few easily studied morphological characters, reflected in a historically unstable classification. The fern genus *Woodsia* is known to have a complex evolutionary history including numerous polyploid taxa and hybrids. It is a cosmopolitan group of small rock loving ferns mainly found in montane areas.

This thesis aims at analyzing the patterns of diploid and polyploid evolution in *Woodsia* and to resolve and classify the relationships of Woodsiaceae and the other families in the large fern clade Eupolypods II.

The Eupolypods II family relationships were inferred with DNA sequences from 81 specimens representing all major lineages. This resulted in the first well supported phylogeny of this clade and revealed Woodsiaceae to be non-monophyletic. The genera previously placed in this family were reclassified into five new or resurrected families. Swedish fern genera that have changed family classification are *Woodsia* (hällebräknar), now in the monogeneric family Woodsiaceae, *Athyrium* (majbräknar), now in Athyriaceeae and *Cystopteris* (stenbräknar) and *Gymnocarpium* (ekbräknar) now in Cystopteridaceae.

To analyze the evolution of *Woodsia*, phylogenies were produced from five plastid and two nuclear regions sequenced from 188 specimens. The results show that most taxa in *Woodsia* are polyploid. Polyploidization is the most common mode of speciation in the genus with an estimated polyploid speciation rate of 54%. The polyploids are mostly young and many of the polyploid taxa seem to have formed multiple times. The results also address several taxonomic and biogeographic questions.

In the process of the work we made methodological advancements and developed 20 new low copy nuclear marker regions as well as a software pipeline for finding primers in transcriptome datasets. The alignment editor software AliView was developed for handling the increasing size datasets in a user friendly way.

In conclusion this thesis provides new insights into the complexities of the evolution of a fern genus in which much of the diversity is accommodated in young species formed through polyploidization. It provides a framework of phylogenetic relationships at different levels that both answers long standing questions and generates new ones.

Keywords: ferns, Eupolypods II, Woodsia, phylogeny, biogeography, polyploidy, polyploid speciation, classification, alignment

Anders Larsson, Department of Organismal Biology, Systematic Biology, Norbyv. 18 D, Uppsala University, SE-75236 Uppsala, Sweden.

© Anders Larsson 2014

ISSN 1651-6214
ISBN 978-91-554-9040-9
urn:nbn:se:uu:diva-232233 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-232233)

Smilla återuppväckte mitt biologiintresse

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Rothfels, C. J.*, **Larsson, A.*,** Kuo, L.-Y*., Korall, P., Chiou, W.-L., & Pryer, K. M. 2012. Overcoming Deep Roots, Fast Rates, and Short Internodes to Resolve the Ancient Rapid Radiation of Eupolypod II Ferns. *Systematic Biology*, 61(3): 490–509.

II Rothfels, C. J., Sundue, M. A., Kuo, L.-Y., **Larsson, A**., Kato, M., Schuettpelz, E., & Pryer, K. M. 2012b. A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales). *Taxon*, 61(3): 515–533.

III Rothfels, C. J., **Larsson, A**., Li, F.-W., Sigel, E. M., Huiet, L., Burge, D. O., Ruhsam, M., Graham, S. W., Stevenson, D. W., Wong, G. K.-S., Korall, P., & Pryer, K. M. 2013. Transcriptome-Mining for Single-Copy Nuclear Markers in Ferns. *PLoS ONE*, 8(10).

IV **Larsson, A**. 2014. AliView: a fast and lightweight alignment viewer and editor for large data sets. *Bioinformatics*, btu531.

V **Larsson, A**., Windham, M., Korall, P. 2014. Phylogeny of *Woodsia* (Woodsiaceae): recent speciation through polyploidization is common in old diploid stock. Manuscript.

* Equally contributing authors.
AL was responsible for all parts of paper I, IV, V. AL was primarily responsible for the Woodsiaceae section of paper II. AL contributed to performing the experiments, analyzing the data, writing the manuscript, designing the analysis tool (lasseblaste) and aligning transcriptomes in paper III.

Reprints were made with permission from the respective publishers.

Contents

Abbreviations

DNA	Deoxyribonucleic acid
L.	Linnaeus
Ma	Mega annum (Million years ago)
MAFFT	Multiple sequence Alignment Fast Fourier Transform
ML	Maximum Likelihood
MLBS	Maximum Likelihood Bootstrap Support
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
PP	Posterior Probability
RAxML	Rapid Accelerated Maximum Likelihood
RNA	Ribonucleic acid
WGD	Whole Genome Duplication

Introduction

Ferns

Systematics/Classification

Ferns (including *Equisetum*) are one of the three main clades of vascular plants, the other two being seed plants and lycophytes. Ferns were for a long time grouped together with lycophytes, the only other vascular plant that has spores, as "Ferns and Fern Allies". This grouping of ferns and lycophytes together as one lineage was proven wrong by molecular analyses (Nickrent & al., 2000; Pryer & al., 2001), showing instead that ferns, lycophytes and seed plants are three distinct lineages. Ferns are relatively abundant in the fossil record (Taylor & Taylor, 2008) and with fossil and molecular dating we find that this lineage of plants diverged from the other lineages in the Devonian period about 400 Ma (Schneider & al., 2004; Fiz-Palacios & al., 2011). Today there are about 9000 recognized fern species, which can be compared to the more than 260 000 species of seed plants and ca 1200 species of lycophytes (Smith & al., 2006; Wood & al., 2009).

Without flowers and fruits, ferns have few easily studied morphological characters, something that probably influenced the historically unstable classification of ferns as reviewed in Smith (1995) and illustrated by the statement of Hennipman (1996) that "modern higher classifications of ferns are a jungle for the user".

During the last twenty years, molecular investigations have increased the knowledge and support for our understanding the evolutionary relationships among the ferns. A milestone in the studies of fern phylogeny is the 400 species molecular analyses of Schuettpelz & Pryer (2007) where many of the family level relationships were first resolved (Smith & al., 2006). Although many of the relationships were elucidated, it was not until the analysis included in this thesis that the family relationships among the large Eupolypods II clade, containing about 1/3 of all fern species, including *Woodsia*, were finally resolved.

Biology

A notable difference between ferns and flowering plants are the reproductive systems. Spores are a central part of the reproductive system of ferns, compared to seeds and flowers in the angiosperms. The typical fern reproductive

cycle starts with the production of sporangia on the underside of the leaves (Fig. 1). The developing sporangia are usually covered by a protective sheet of tissue (indusium). The characteristics of the indusium and the location of the sporangia on the leaf are very important morphological characters for genus or family level identification.

Diploid "mother cells" in the sporangia then undergo meiosis and transform into haploid spores. Mature spores are actively catapulted into the air by the annulus of the sporangia. Most fern spores are in the range of 30-70 µm and capable of long distance dispersal, as can be illustrated by the composition of the flora of the up to 25 Ma year old islands of Hawaii. In Hawaii the fern (and lycophyte) proportion of the flora is higher than the fern (and lycophyte) proportion of the flora in continental America, suggesting that ferns have better dispersal and establishment capabilities than the seed plants (Moran, 2008). Instead of producing a new fern, the haploid spore germinates into a small (ca 5 x 5 mm), green gametophyte. The gametophyte contains the reproductive organs antheridia and archegonia, which produce sperms and eggs respectively. Most gametophytes are bisexual but have barriers that reduce selfing and inbreeding (the most prominent is the temporal difference of maturing sperms and eggs) (Sheffield, 2008). Also most ferns show heterozygous genetic patterns that indicate that outcrossing is prevalent (Ranker & Geiger, 2008). If sexual outcrossing is to take place, spores from different individuals need to germinate closely together and sperms from one gametophyte need to swim in a film of water to the egg cells inside the archegonia of the other gametophyte. The zygote will then grow into a new fern sporophyte producing new spores in the sporangia on the back of the leaf, and one reproductive cycle has passed.

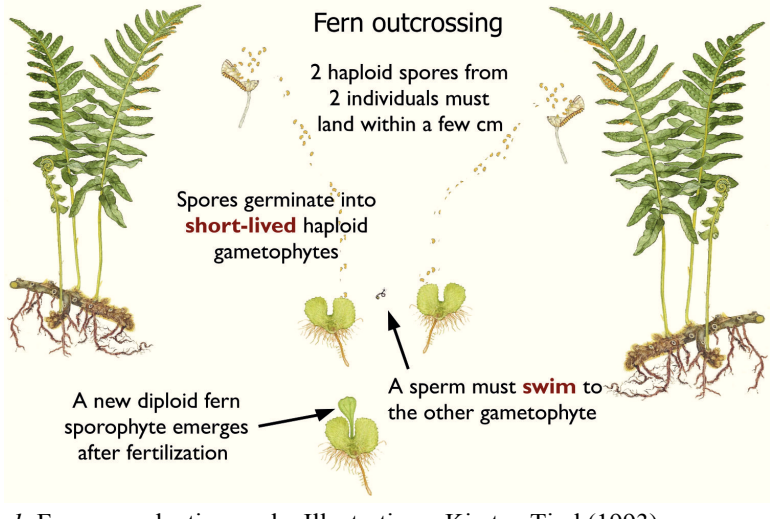

Figure 1. Fern reproductive cycle. Illustrations: Kirsten Tind (1993).

Although the reproductive system of seed plants is closely related to that of ferns and not a novel invention, the gametophyte "generation" of seed plants has been reduced and is an internal part of the pollen and the unfertilized seed, whereas the gametophyte generation of mosses is the most prominent one. Recognizing that the reproductive system with alternation of generations is the same for all land plants was initially done by the German amateur botanist William Hofmeister in 1851 (Kaplan & Cooke, 1996).

The differences in reproductive system due to the presence of a free living gametophyte instead of seeds most likely influences the biology of ferns in many ways relative to the seed plants. This calls for continued parallel research of biological phenomena in ferns and will make if less feasible to generalise some of the results from the more well studied flowering plants. Instead, the comparison of results between the two groups might generate new exciting insights into the biology of plants in general.

Main characters for identification

Most ferns have a subterranean stem and are primarily identified by the shape of the leaves, the properties and abundance of trichomes, location of sporangia, properties of the indusium and characteristics of the spores. In *Woodsia* and many other fern genera the size of the spores are correlated with ploidy level. Polyploid taxa have larger spores than closely related diploid taxa, making this a very good indication of ploidy level (Barrington & al., 1986; Beck & al., 2011; Dyer & al., 2012).

The fern genus Woodsia

Woodsia is a medium-sized genus (35-40 taxa) of small, tufted ferns growing on or in the vicinity of cliffs and rocks (Brown, 1964). They are relatively rare and mainly found in montane areas in the Northern Hemisphere, but also occur in South America and Southern Africa. The species richness is greatest in the Rocky Mountains in North America and the Himalayas in Asia. There are three circumboreal species, *W. glabella* R. Br., *W. ilvensis* (L.) R. Br. and *W. alpina* (Bolton) Gray (Fig. 2), but most species have a

Figure 2. Woodsia ilvensis, Woodsia glabella and Woodsia alpina. Photo: Anders L.

much more restricted distribution (Brown, 1964; Windham, 1993). Cytological studies and allozyme work have shown that *Woodsia* have a complex evolutionary history with numerous polyploid taxa (Brown, 1964; Ma, 1985, 1987; Windham, 1993; Mickel & Windham, 2004)

Polyploidy

Polyploidy is the condition of having more than two sets of chromosomes. It is the result of an abnormal meiosis or mitosis resulting in the duplication of the number of chromosomes in the new cell. Within plants, polyploidy is common. In a recent survey by Wood & al. (2009) it was found that as many as 15% of flowering plant species and 34% of leptosporangiate ferns are polyploid. Polyploidy is not only found in plants, but it is much less common in animals (Otto & Whitton, 2000).

The unmatched sets of chromosomes makes the reproductive isolation between the parent species and the polyploid offspring substantial, and the polyploids could be considered distinct species/biological entities (Mallet, 2007). Polyploids are commonly categorized as either auto- or allopolyploids. The definition of auto- and allopolyploids are usually following a cytological or taxonomic approach. The nature of the data in presented here makes a taxonomic approach the most appropriate, where an autopolyploid is defined as the result of a polyploidization within a species, and an allopolyploid is formed through hybridization between two species (Soltis & Rieseberg, 1986; Soltis & al., 2004).

Polyploidy has for a long time been considered an important factor for plant evolution (Soltis & al., 2004, 2009). This is not only because of the absolute number of species it generates (Wood & al., 2009), but also because of the evolutionary opportunities of the redundant duplicated genes. Most of the duplicated genes will over time be silenced and reduced but some will acquire new functionality (McGrath & Lynch, 2012).

As we get better knowledge of many more plant genomes it has been possible to detect ancient Whole Genome Duplications (polyploidizations), and it has been shown that most angiosperms are paleopolyploids (Soltis & al., 2009; Van de Peer, Fawcett, & al., 2009). A paleopolyploid lineage contains historical polyploidization events where most of the duplicated chromosomes have been reduced through various diploidization events, whereas neopolyploids still have extant closely related diploid species (Guerra, 2008).

It has been put forward that the historical polyploidizations have enabled bursts of adaptive radiations in the history of the angiosperms (Soltis & al., 2009; Van de Peer, Fawcett, & al., 2009) and vertebrates (Van de Peer, Maere, & al., 2009). Although we in some instances can correlate a burst of

speciation to the timing of a WGD, it is still difficult to prove that the polyploidization itself was the cause of the radiation (Soltis & al., 2009).

The high chromosome numbers in ferns have made researchers speculate about the existence of ancient polyploids in this lineage, but there has been no concrete evidence (Duncan & Smith, 1978), and we are still awaiting a genome wide analysis in the search for ancient WGD in ferns.

Research aims

The aim of **paper I** is to resolve the family-level relationships within the the large fern clade Eupolypods II. It was indicated in earlier studies that Woodsiaceae, one of the five families recognized by Smith & al. (2006), was a non-monophyletic assembly of genera. Furthermore, the interrelationships of the other families were partially unresolved. In **paper II** we aimed to make a new family classification based on the results of the analyses in paper I. The aim of **paper III** was to design several new fern primers that would amplify a variety of new low copy nuclear regions that can be used in phylogenetic analyses of ferns. A need for an improved alignment editor emerged in the primer design project and the aim of the newly created alignment editor software AliView, presented in **paper IV**, was to create a new fast alignment visualization and editing tool that could swiftly handle much larger datasets than previous software and at the same time have an intuitive user friendly interface. The principal aim of **paper V** was to resolve the patterns of diploid and polyploid evolution in *Woodsia* with the help of low copy nuclear markers and also to answer several taxonomic and biogeographical questions within this genus. A secondary aim was to describe speciation through polyploidization not only for *Woodsia* but also for the entire Eupolypods II clade.

Summary of papers I & II

The focus of this thesis, *Woodsia*, belongs to the Eupolypods II, a clade of ferns which has been very hard to resolve. As an important framework for the rest of this thesis project we started by resolving the previously unresolved family relationships of this large clade of ferns. In paper I we performed the phylogenetic analyses and in paper II we revised the classification of the families.

Although the relationships and circumscription of this clade of ferns, which includes some of the most common fern families (Aspleniaceae, Thelypteridaceae, Woodsiaceae, Blechnaceae), had not been fully resolved, there were analyses that indicated Woodsiaceae to be “unnatural”, containing a non-monophyletic assembly of genera (Schuettpelz & Pryer, 2007).

This group was the last major branch of the fern "tree of life" that was unresolved and it consists of about 2500 species, i.e. nearly a third out of the approximately 9000 species of ferns (Smith & al., 2006). Several common Swedish ferns are represented in this group, for example, *Gymnocarpium dryopteris* (ekbräken), *Cystopteris fragilis* (stenbräken), *Asplenium trichomanes* (svartbräken), *Thelypteris palustris* (kärrbräken), *Phegopteris connectilis* (hultbräken), *Woodsia ilvensis* (hällebräken), *Matteuccia struthiopteris* (strutbräken), *Athyrium filix-femina* (majbräken) and *Blechnum spicant* (kambräken).

Paper I

We identified three important factors that could have influenced why the relationships had resisted being resolved by either morphological or molecular data. They were (i) long branches among many of the ingroups in combination with very short shared internodes (the "ancient rapid radiation" model (Whitfield & Lockhart, 2007)), (ii) heterogeneous rates of substitution between different lineages, (iii) a long branch leading to the outgroup that apparently reduced the level of support in the ingroup.

To resolve the phylogeny and tackle the identified obstacles we assembled a dataset containing five chloroplast loci from a sample of 81 accessions representing all major lineages that have been previously detected by morphology or molecular analyses (Smith & al., 2006; Schuettpelz & Pryer, 2007). In our analyses we addressed the above mentioned issues by (i) using two modified implementations of Bayesian analyses (Phycas (Lewis & al., 2010) and "Yang prior MrBayes" (Yang, 2008)) that are developed to ac-

count for a "Bayesian star tree paradox" which can give false high support values on short internodes (Lewis & al., 2005); (ii) rate heterogenity effects in our dataset were assessed by comparing the results with a reduced consensus tree (Wilkinson, 1996; Burleigh & al. 2009), and also through a reduced data set where highly divergent Aspleniaceae was removed; (iii) performing the analyses with different outgroup compositions.

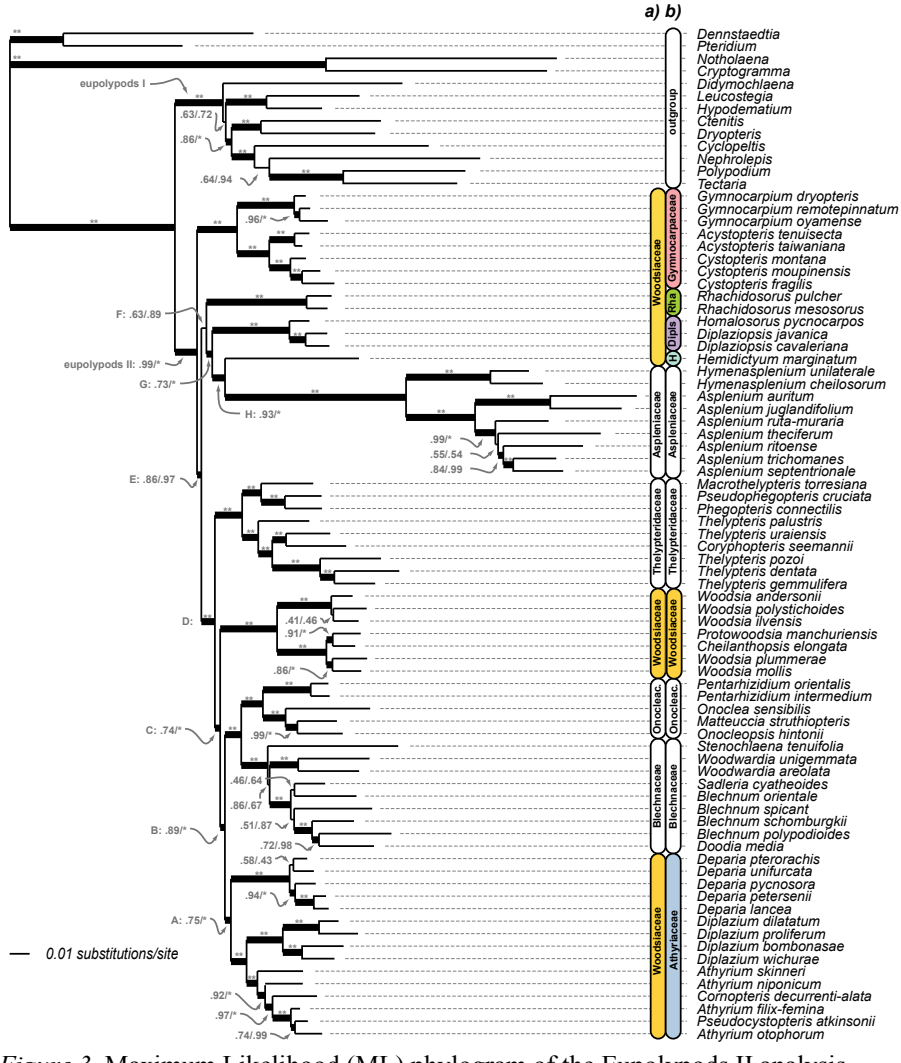

Figure 3. Maximum Likelihood (ML) phylogram of the Eupolypods II analysis. Support values indicated are ML-Bootstrap support followed by Bayesian PPs. Bold branches have 70% MLBS and 0.95 PP. Support values of 100% or 1.0 are indicated with an asterisk (*). a) Previous family designations (Smith et al. 2006), with the paraphyletic Woodsiaceae highlighted. b) The new classification, with the changed designations highlighted; Rha = Rhachidosoraceae; Dipls = Diplaziopsidaceae; H = Hemidictyaceae; Onocleac = Onocleaceae. Redrawn from Rothfels & al. (2010).

We inferred a phylogeny where the backbone for this group of ferns was for the first time strongly supported (Fig. 3). The results of the tests for various confounding factors showed that the phylogeny was robust and neither the "Bayesian star tree paradox" nor the heterogeneous substitution rates affected the support values very much. The factor that gave the strongest variation in node support values was the outgroup composition.

We found that Woodsiaceae with the previous circumscription was non-monophyletic, where for example *Cystopteris* and *Gymnocarpium*, previously included in Woodsiaceae, had a most common recent ancestor with the rest of the ferns in Woodsiaceae as long ago as 100 Ma.

This was also the first analysis to show that the sometimes segregated genera *Protowoodsia* and *Cheilanthopsis* are nested within *Woodsia*.

It was indicated by previous studies that *Woodsia* was probably non-monophyletic, but the placement of the genera *Diplaziopsis*, *Homalosorus* and *Rhachidosorus* as most closely related to Aspleniaceae was unanticipated. Previously they were almost unanimously considered to be part of the athyrioid ferns, and often even included in the genera *Athyrium* or *Diplazium*.

Paper II

After having created a robust and well supported phylogeny for this group of ferns for the first time, we could with confidence revise the family level classification.

In the classification from 2006 by Smith & al., this group of ferns was divided into five families (Fig. 3), and as illustrated by our phylogenetic analyses in paper I, Woodsiaceae was non-monophyletic.

When creating new family circumscriptions we created monophyletic entities while trying to keep nomenclatural stability and also recognizing evolutionary information such as deep splits in the phylogeny.

The circumscription of Aspleniaceae illustrates these sometimes conflicting aims. One option was to include the genera *Hemidictyum*, *Diplaziopsis*, *Homalosorus* and maybe also *Rhachidosorus*, previously placed in Woodsiaceae, into an expanded Aspleniaceae. Alternatively we could recognize a selection of new families with very few species in each. We did the latter and recognized new families to keep the well defined Aspleniaceae intact and also to highlight the long time separation of the previously misplaced genera. The single species *Hemidictyum marginatum* now belongs to the monotypic family Hemidictyaceae Christenh. & H. Schneid. Four to seven species of *Rhachidosorus* now belong to the monogeneric family Rhachidosoraceae X.C. Zhang. The two genera *Diplaziopsis* and *Homalosorus* containing three to five species belong to Diplaziopsidaceae X.C. Zhang & Christenh.

Another change in the circumscription which was made to deal with the paraphyly of Woodsiaceae as circumscribed by Smith et al. (2006), was the

move of the athyroids (e.g. *Athyrium* & *Diplazium*) to the resurrected family Athyriaceae Alston. *Gymnocarpium* and *Cystopteris* were included in the family Cystopteridaceae (Payer) Shmakov.

Woodsiaceae became monogeneric including only *Woodsia* and the new classification does not recognize the sometimes segregate genera *Cheilanthopsis* and *Protowoodsia*, since that would make *Woodsia* non-monophyletic.

Summary of paper III

Molecular fern phylogenies have until recently mainly been based on chloroplast sequences. Although extensively used for phylogenies, the chloroplast is maternally inherited as one single unit and does not necessarily reflect the phylogeny of the majority of the genome in instances when gene and species trees might be incongruent, due to for example incomplete lineage sorting (Rosenberg & Nordborg, 2002). This also makes chloroplast sequences on their own not suitable for resolving reticulate evolutionary events such as hybridization and polyploidization which have been shown to be important in ferns (Lovis, 1978). Instead, low copy nuclear markers are useful in resolving these patterns (Popp & Oxelman, 2001; Brysting & al., 2011), but very few nuclear markers have been available for ferns. The lack of nuclear markers can partly be attributed to the lack of genomic information for ferns. There has been no nuclear genome or even a mitochondrion sequenced for a fern, making it difficult to design nuclear primers.

A new opportunity for designing nuclear primers for ferns emerged when approximately 1000 green plant transcriptomes were sequenced by the One Thousand Plants Project (1KP; onekp.com). Among these transcriptomes were 62 fern accessions (many collected and extracted by myself and the other authors of paper III).

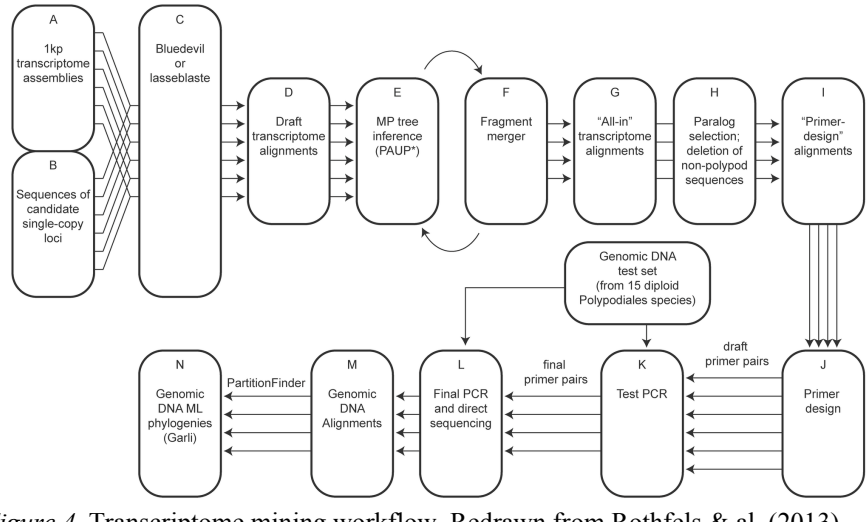

Figure 4. Transcriptome mining workflow. Redrawn from Rothfels & al. (2013).

The primer design workflow (Fig. 4) included the development of two automated pipelines for identifying candidate loci among these transcriptomes, Blue Devil (Li, 2013) and lasseblaste (Larsson, 2013). Blue Devil blasts a series of query sequences (we used nuclear regions the authors already had an interest in as seed sequences) against the pool of transcriptome assemblies and provides an alignment of the resulting hits. Lasseblaste also takes a series of query sequences as input (we used all putative single-copy markers listed by the 1KP project) and blasts each of these sequences against all the transcriptomes. It then takes the resulting hits and blasts them back to the full transcriptomes. From this final pool of hits, lasseblaste produces an alignment (with MAFFT) and provides an accompanying quality score. The scoring system rewards alignments that have broad representation across the included transcriptomes, indicating good taxon coverage, and penalizes alignments that have many hits per transcriptome, suggesting multiple paralogs and/or short read lengths. In the end we selected five of the top ten best-scoring of these alignments to pursue for primer design. The alignments were further refined by merging discontinuous or slightly overlapping sequences from the same specimen that appeared closely related in a preliminary maximum parsimony phylogenetic tree created with PAUP* (Swafford, 2003). Annotated *Arabidopsis thaliana* genomic DNA sequences from the TAIR database (Lamesch & al., 2012) were added to illustrate probable intron/exon boundaries. Primers were designed in promising conserved regions with help of the primer design tool in AliView (Paper IV). To increase the specificity of the primers we focused our effort on designing primers for the order Polypodiales which contain a majority of all the fern species.

Our efforts resulted in 20 new nuclear markers across ten protein coding genes, spanning both introns and exons. These new markers triple the available nuclear markers for ferns. To demonstrate the success of the primers we amplified and sequenced the regions and came up with a well supported phylogeny consistent with previous finds. Additionally the phylogeny were evaluated with various partitioning schemes that revealed strong effects on support values but lesser effects on topology. In our results a codon position partitioned model is favored. All the curated transcriptome alignments are published and will allow other researchers to develop their own primer sets meeting specific needs.

Summary of paper IV

The need for an improved alignment editor and viewer emerged while designing degenerate primers for a diverse set of fern species from the One Thousand Plants Project (1KP; onekp.com) transcriptome data (Paper III). What was lacking in the previously available programs was the combination of abilities to (i) have an overview of large nucleotide alignments, (ii) visually highlight various conserved regions, (iii) have a simple and intuitive way to align, rearrange, delete and merge sequences, and (iv) find and list degenerate primers in selected semi-conserved regions.

AliView (Fig. 5) was developed to meet the above requirements. The first of the key features implemented in AliView is the capability to draw millions of residues on the screen at once without slowing down. To increase the rendering speed well beyond previously available alignment editors, an optimized rendering function that copies the pixel representation of the residues to the screen buffer was implemented instead of only using the internal Java methods. This increased the drawing speed by 5-10 times and AliView renders alignments almost like a video-game. The speed in rendering even makes it possible to work with large alignments on older hardware.

Figure 5. Screenshot of AliView.

In the cases of very large data files (larger than the computer memory size), AliView will initially create an index of the sequences in the file. Only a part of the file will be cached in memory when the sequences are viewed. This way alignment files of unlimited size can be viewed. The index is created following the standards of Samtools FASTA index (Li & al., 2009), but AliView can also create indexes for other alignment file formats and will read unlimited size alignment files in Fasta, Phylip, Nexus, Clustal and MSF-format. AliView has unique functionality aimed at supporting the design of universal degenerate primers. In a selected region AliView will calculate all possible primers and present them to the user in a list sorted by the number of degenerate positions, self-binding values and melting temperature (TM) (Kämpke & al., 2001).
Another very important factor when developing AliView was the focus on usability and intuitive handling. The results of this effort can be illustrated by some "random" comments received from users:

"From first glance, this looks like it might be the best aligner viz tool out there" - Justin Blumenstiel

"I really like aliview! One big plus: primer design." - Dieter Best

"Har precis börjat använda AliView, funkar kanon! Har letat länge efter nåt som funkar att handjaga sangersekvenser med efter att Se-Al packade ihop och det här är ju Se-Al på steroider!" - Magnus Popp

"I am thrilled and delighted with AliView. Now I have an excellent editor to replace Se-Al, which doesn't run on new Macs. Zoomability, and the ability to edit while zoomed out, make it fantastically useful for large alignments." - Will Fisher

"I just really wanted thank you for developing Aliview - it is an excellent piece of software. It's hard to believe that out of all the editors out there, nobody seems to find it obvious that the features you have included are crucial." - Amr Aswad

"I love your AliView soooooo much! Thank you for developing such a wonderful program!" –Fay-Wei

In September 2014 there were more than 10 downloads of the program per day on average, and in the first three months that the program was available on-line more than 1000 unique downloads were generated.

Summary of paper V

Through previous cytological studies and allozyme work the fern genus *Woodsia* was known to have a complex evolutionary history and to comprise numerous polyploid taxa and hybrids (Brown, 1964; Ma, 1985, 1987; Windham, 1993; Mickel & Windham, 2004). Our main aim of paper V was to resolve these patterns of diploid and polyploid evolution in the genus, and also to answer taxonomic and biogeographic questions within *Woodsia*. We complemented our study of polyploidy in *Woodsia* with a broad analysis of polyploid frequencies and polyploid speciation frequencies in the entire Eupolypods II clade, comprising about 2500 species.

In detail we wanted to answer the questions: Which are the parent species of the polyploids? Do the polyploid taxa have single or multiple origins? How frequently have *Woodsia* species established over long distances? Are American and Asian taxa monophyletic groups? How did this mainly northern hemisphere genus reach South America and Africa and when did it happen? The questions we wanted to answer from the expanded polyploid analyses were: How extensive is polyploidy in *Woodsia* and related ferns? How frequent is speciation through polyploidization in *Woodsia* and its closest relatives? Are the polyploids young, being continuously replaced by new ones formed by the diploid stock? Are there old polyploid lineages?

To infer the phylogeny and polyploid evolution we sequenced five plastid (*atpA*, *atpB*, *matK*, *rbcL* and *trnG-R*) and two nuclear (*pgiC* and *RPA2*) regions from 188 specimens representing 36 taxa and six hybrids.

Our results (Fig. 6-7) reveal that the genus comprises two major well-supported clades, one with the circumboreal and most of the Asian species, and the other with American species and the rest of the Asian ones. In the dated analysis we estimate the split between these clades to be 45 Ma, and we found that *Woodsia* × *abbae* is a hybrid between members of these two clades. Few other records of hybridization events have been recorded involving species separated for such a long time (Feng & al., 2005; Pilotti & al., 2009).

We find that sister species in *Woodsia* have almost exclusively neighboring or overlapping distributions, indicating that although *Woodsia* spores are dispersed by wind, indicating that although *Woodsia* spores are dispersed by wind, long distance dispersal with establishment is rare. Despite this general pattern, there are some interesting distributions and disjunctions deviating from this.

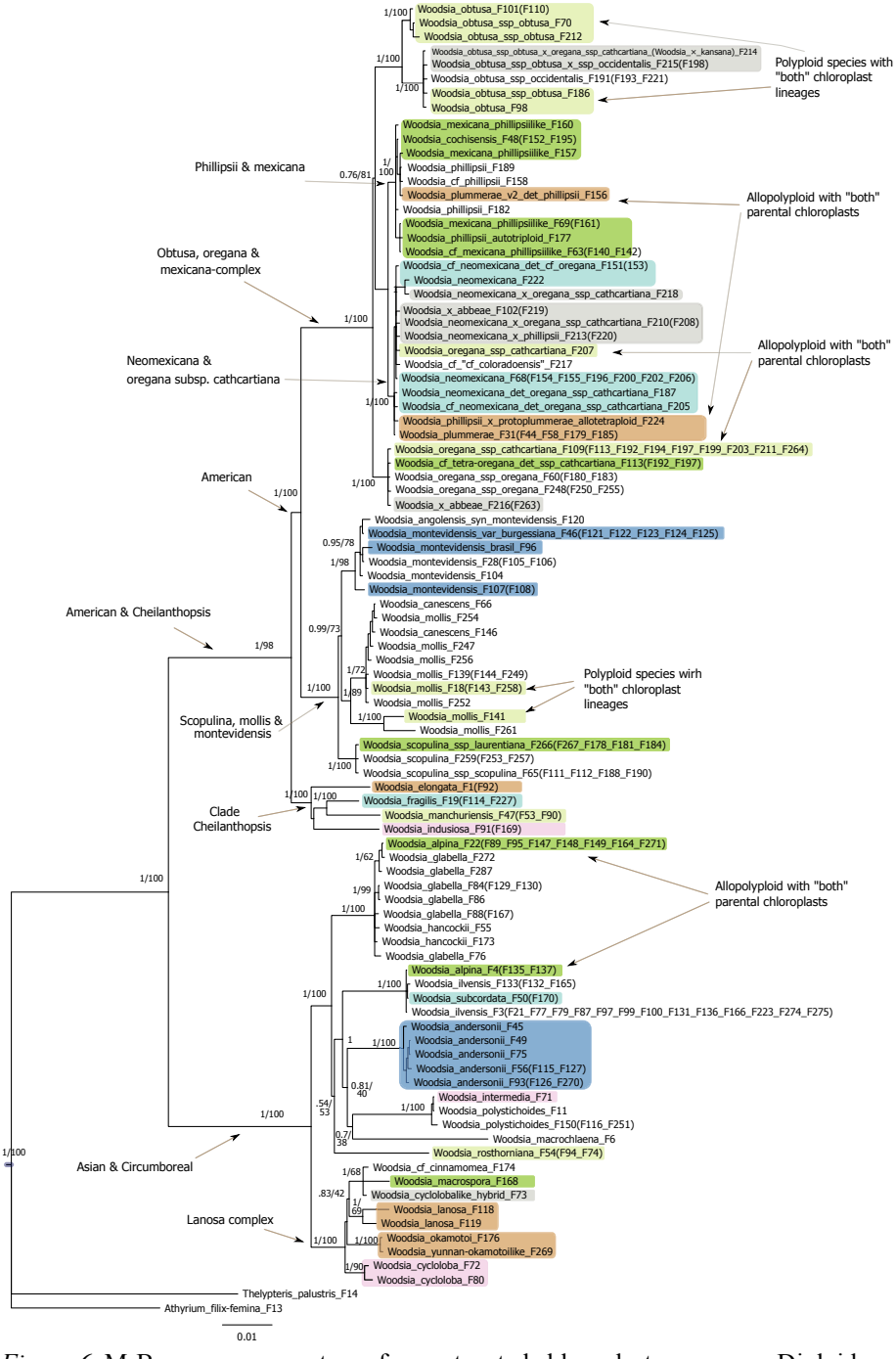

Figure 6. MrBayes consensus tree of concatenated chloroplast sequences. Diploid taxa are not colored. Colors indicate polyploid taxa. Grey indicate hybrid taxa. Numbers at nodes are Bayesian posterior probabilities (PP) followed by RAxML Maximum Likelihood bootstrap support. Branches are collapsed at 0.5 PP. Scale bar represents average number of substitutions per site.

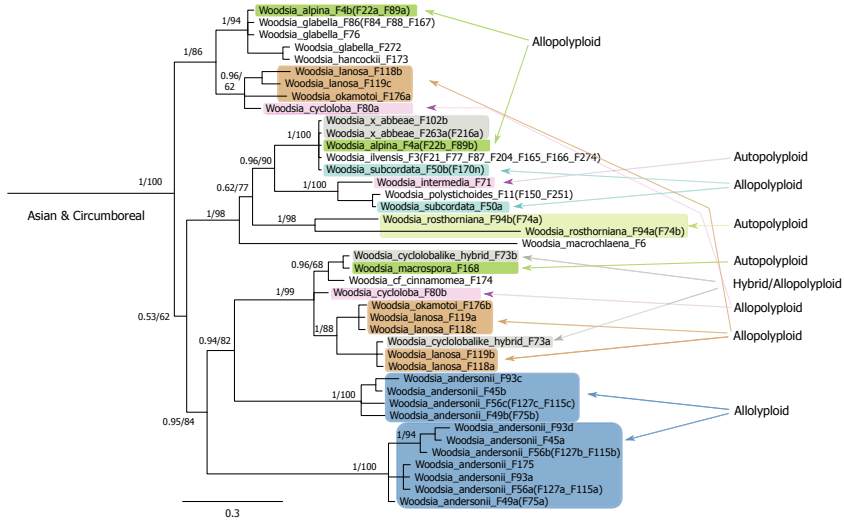

Figure 7. Asian & Circumboreal subsection of the MrBayes consensus phylogram of nuclear pgiC sequences. Names ending with a,b,c or d indicate specimens with multiple sequences, x indicates specimens with multiple sequences where only one has been retrieved. Diploid taxa are not colored. Colors indicate alleles belonging to the same polyploid taxon. Grey indicates hybrid taxa. Numbers at nodes are Bayesian posterior probabilities (PP) followed by RAxML Maximum Likelihood bootstrap support. Branches are collapsed at 0.5 PP. Scale bar represents average number of substitutions per site.

The circumboreal species do not form a monophyletic group and are nested among various Asian species, whereas the "American clade" is monophyletic and nested among other Asian species.

Within the American clade *W. montevidensis* has its main distribution in South America, but has also made a relatively recent leap to Southern Africa and Madagascar.

We find that most taxa in *Woodsia* are polyploid (22 out of 36) and that speciation through polyploidization is the most common mode of speciation in the genus. The polyploids are mostly young compared to the diploids. Several of the polyploid taxa seem to have been formed multiple times or have been involved in backcrossing and introgression, i.e. *W. alpina*, *W. obtusa*, *W. mollis*, *W. oregana* subsp. *cathcartiana* and *W. plummerae*. This can be illustrated with *W. alpina* specimens that have two copies of the nuclear pgiC sequences (Fig. 7), one that groups together with *W. glabella* and one that groups with *W. ilvensis*. In the chloroplast phylogeny (Fig. 6) on the other hand some specimens have the *W. ilvensis* haplotype while other specimens have the *W. glabella* haplotype.

The results of the Eupolypods II polyploid analyses agree with the *Woodsia* study in showing a high proportion of polyploids, and 46% of the speciation events are connected to a shift in ploidy level. Old polyploid lineages are rare.

Conclusions

In this thesis I have, together with colleagues, been able to clarify the phylogenetic relationships on the family level for the large Eupolypods II clade of ferns, as well as on the infrageneric scale within *Woodsia*. The use of low copy nuclear regions in the *Woodsia* phylogeny has revealed the complexity of polyploid and diploid evolution within this fern genus.

I have been able to contribute much needed methodological tools by creating the alignment software AliView. We also developed new low copy nuclear primers for use in fern phylogenetics, as well as developing a pipeline for designing primers from NGS transcriptome data.

In paper I & II we present for the first time with strong support the family level relationships of the Eupolypods II clade of ferns, encompassing almost 1/3 of the fern species. We inferred the position of the genus *Woodsia* and the non-monophyletic circumscription of Woodsiaceae. The genera previously placed in this family are now distributed in five new or resurrected families. There were several unanticipated relationships, such as the position of *Rhachidosorus*, *Hemidictyum* and *Diplaziopzis*.

As illustrated in the detailed phylogeny of *Woodsia*, presented in paper V, there is a need for low copy nuclear markers to disentangle the reticulate evolution within a fern genus. In paper III we published 20 new low copy nuclear marker regions to use in future fern phylogenies. We also published the curated alignments of these genes for other researchers to use, as well as a software pipeline for finding primers in other regions.

The alignment viewer and editor AliView developed in paper IV provides capabilities to handle the increasing size datasets in a user friendly way. The program has been very popular with more than 1000 downloads and is well received by users.

New insights into the complexities of the evolution of a fern genus were provided by our comprehensive study of *Woodsia* in paper V. The detailed phylogeny answers several taxonomic and biographical questions. We also find that most taxa in *Woodsia* are polyploid and that polyploidization is the most common mode of speciation in the genus with an estimated polyploid speciation rate of 54%. The polyploids are mostly young. Many of the polyploid taxa seem to have formed multiple times. We find that *Woodsia* × *abbae* is a remarkable triploid hybrid between two clades separated by as many as ca. 45 Ma.

Svensk sammanfattning

Systematik

Idag är vi säkra i vår uppfattning om att allt liv på jorden har ett gemensamt ursprung. Ju mer vi har lärt oss om cellernas funktion samt vetskapen att alla kända organismer har snarlikt organiserad arvsmassa (arvsmassa = DNA/RNA = ritningen för hur olika aminosyror skall sättas ihop till proteiner vilka i sin tur bildar organismen), desto mer osannolikt framstår det att samma system skulle utvecklas två gånger parallellt. Sammantaget leder oss detta till slutsatsen att alla organismer har ett gemensamt ursprung.

Utifrån denna urorganism, som antagligen uppstod för cirka 3,5 miljarder år sedan, har alla jordens organismer utvecklats. Arvsmassan kopieras från en generation till den nästa (exempelvis från förälder till barn). Ibland blir inte kopian perfekt utan i kopieringen uppstår små fel (mutationer) och dessa fel förs vidare till nästkommande generation. Organismerna blir åtskilda från varandra genom att exempelvis geografiskt avstånd gör att främst närbelägna organismer utbyter DNA (parar sig). Varefter tiden går och fler och fler mutationer ansamlas blir dessa populationer av organismer mer och mer olika och så småningom blir populationerna så olika att vi börjar kalla dem olika arter. Dessa nya arter genomgår samma process igen och ytterligare arter uppstår ur dessa arter.

Systematisk biologi är forskningsområdet där man kartlägger hur alla organismer har utvecklats ur gemensamma anfädrar och skapar ett ”livets träd” över släktskapet. Utöver den historiska aspekten där man kartlägger alla organismernas släktskap så definierar man också inom Systematisk biologi var gränsen går mellan olika arter och namnger nya arter. När man idag skapar artavgränsningar vill man att de följer släktskapshistorien och inte bara är baserade på yttre likhet.

För att spåra släktskap så har man länge förlitat sig på utseende (morfologi); två organismer som ser lika ut borde vara nära släkt. Dock har det visat sig att liknande utseende ibland uppstår flera gånger parallellt, och att tidigare systematiker tolkat vissa karaktärer fel. Man har även använt kemiska analyser för att se vilka arter som innehåller liknande substanser, och således borde vara närmast relaterade till varandra. Idag så använder man nästan helt uteslutande DNA-sekvenser för att jämföra organismernas släktskap. Man extraherar samma gener hos många olika organismer, exempelvis gener som används för tillverkningen av klorofyll hos växter. Därefter jämför man mu-

tationerna i generna och antar att de växter som har mest lika gener har en gemensam anfader närmast tillbaka i tiden, och därefter nästa o.s.v.

Ormbunkar

I den här avhandlingen har jag undersökt en grupp ormbunkars uppkomst och släktskap. Studierna omfattar släktet *Woodsia* (hällebräknar) och den större grupp av ormbunkar som *Woodsia* är närmast besläktad med (figur 2).

Ormbunkar är en grupp växter som skiljde sig från övriga gröna växter för ca 400 miljoner år sedan. Frön och blommor är exempel på anpassningar som uppstått senare i växthistorien och inget som finns hos ormbunkarnas anfader. Ormbunkar sprider sig med sporer istället för frön. När en spor gror så bildas inte en ny egentlig ormbunke, utan det gror fram en så kallad gametofyt, d.v.s. en liten 5 x 5 mm grön platta med könsorganen arkegon och anteridier. Anteridierna producerar spermier som vid fuktig väderlek kan simma till äggen i ett arkegon på en annan gametofyt i närheten. Ur de befruktade äggen växer sedan en ny ormbunke fram (figur 1).

I och med att förökning och spridning är väldigt annorlunda jämfört med de välstuderade blomväxterna så finns många unika upptäckter att göra hos ormbunkar.

Artiklar

I den **första och andra artikeln** har vi undersökt hur en stor grupp av ormbunkar (Eupolypods II) är besläktade. I denna stora samling ingår nästan 1/3 av alla världens ormbunksarter, och flertalet vanliga svenska arter såsom exempelvis majbräken (*Athyrium filix-femina*), hultbräken (*Phegopteris connectilis*), svartbräken (*Asplenium trichomanes*), ekbräken (*Gymnocarpium dryopteris*), stenbräken (*Cystopteris fragilis*) och hällebräken (*Woodsia ilvensis*).

De flesta ormbunkar ser nästan likadana ut och det finns tämligen få morfologiska detaljer (utseendemässiga) som ger indikationer på hur släkten och familjer relaterar till varandra (inom systematik är en familj en vidare gruppering än släkt, i en familj ingår ett eller flera släkten, i en släkt ingår en eller flera arter). Trots vissa försök så har man tidigare inte lyckats klassificera de närmast besläktade ormbunkarna i denna stora grupp till enhetliga familjer. Det fanns indikationer på att inte minst familjen Woodsiaceae var en onaturlig gruppering (innefattande arter som inte var närmast besläktade).

I den **första artikeln** genomförde vi analysen av släktskapet med hjälp av DNA-sekvensdata från fem kloroplastregioner från 81 arter inklusive alla samtliga tidigare svårplacerade släkten.

Vi utvärderade flera olika analysmetoder för att vara säkra på att våra resultat var tillförlitliga. Våra resultat beskriver tidigare okända relationer

inom Eupolypods II. Analysen visade exempelvis att stenbräken (*Cystopteris fragilis*) och ekbräken (*Gymnocarpium dryopteris*) ingår i systergruppen till övriga ormbunkar i Eupolypods II. Analysen visade att familjen Woodsiaceae var onaturlig och då vår undersökning hade starkt statistiskt stöd kunde vi känna oss säkra när vi i den **andra artikeln** skapade en ny klassifikation. Vi valde att utöka familjeantalet från 5 till 10 då vi ansåg att det bättre beskriver den evolutionära historien vi fann i den första artikelns analys (figur 3). Framför allt var det den tidigare familjen Woodsiaceae som vi förändrade. Vi flyttade de flesta arterna till familjerna Athyriaceae, Cystopteridaceae, Diplaziopsidaceae, Hemidictyaceae och Rhachidosoraceae. Slutligen blev endast *Woodsia* kvar i familjen Woodsiaceae.

De svenska arter som direkt påverkades av den nya klassifikationen var majbräken (*Athyrium filix-femina*) och fjällbräken (*A. distentifolium*) som flyttades från Woodsiaceae till Athyriaceeae samt *Cystopteris* (bl.a. stenbräken) och *Gymnocarpium* (bl.a. ekbräken) vilka flyttades till familjen Cystopteridaceae.

Tyvärr hade konkurrerande forskare några månader innan vår publikation redan skapat nya familjer utifrån de tidigare statistiskt osäkra analyser som fanns innan vår analys, och vi fick därför inte namnge de nya familjerna.

I den **tredje artikeln** letar vi efter nya lämpliga gener att använda för släktskapsanalyser hos ormbunkar. Ormbunkar är betydligt mindre undersökta än blomväxterna och det finns inte lika mycket kännedom om ormbunkarnas olika gener. De vanligaste gener man jämför är de som finns i kloroplasten. Kloroplasterna har dock den egenheten att de likt mitokondrierna överförs i sin helhet från mödernet till avkomman och fader-kloroplasten är inte inblandad. Det har tidigare påvisats att ormbunkar ofta hybridiserar och om man analyserar kloroplasterna hos en hybrid så finner man bara den ena av de två föräldraarterna. För att upptäcka båda föräldraarterna behöver man sekvensera gener från cellkärnan.

Genom att delta i insamlingen av material till One Thousand Plants Project (1KP; onekp.com) fick vi tillgång till ett unikt dataset bestående av 1000 växt-transkriptom (RNA-sekvenser från alla aktiva gener hos en organism). Vi utvecklade två datorprogram (lasseblaste och Blue Devil) för att leta efter lämpliga gener i denna gigantiska datamängd. Lämpliga gener är de som finns hos de allra flesta av ormbunksarterna och samtidigt är lagom muterade (tillräckligt mycket för att enkelt särskilja olika arter).

Vi skapade PCR- och sekvenseringsprimers (passbitar som behövs för at utvinna önskade DNA-sekvenser ur organismen) för de mest lovande generna vilka hade rankats av programmet lasseblaste. Dessa testades därefter på 15 olika ormbunkar för att utvärdera deras funktion. Slutligen lyckades vi publicera 20 nya gen-regioner lämpliga för fylogenetiska ormbunksanalyser, att jämföra med de 7 st som fanns sedan tidigare.

Det saknades bra dataprogram för att överblicka de stora mängder data som genereras med den senaste sekvenseringstekniken (NGS). I samband med arbetet att hitta nya lämpliga gener i den tredje artikeln blev denna brist påtaglig. I den **fjärde artikeln** presenteras programmet AliView (Alignment Viewer and Editor) som jag skapat för att råda bot på denna brist.

AliView (figur 5) hanterar mycket stora dataset bättre än något tidigare program. Användaren får en unik överblick av hela datasetet genom att zooma ut eller in utan begränsningar. En av de viktigaste aspekterna med programmet är att sekvenserna ritas ut väldigt fort på skärmen (likt ett dataspel), och att programmet inte laggar trots att miljontals nukleotider eller aminosyror visas på skärmen samtidigt. Den höga uppritningshastigheten uppnås genom en optimerad grafikmetod som kopierar pixelrepresentationer av nukleotiderna eller aminosyrorna direkt till skärmen istället för att endast använda Java:s interna grafikmetoder. Utöver förmågan att hantera stora dataset har även stor vikt lagts vid att programmet skall vara användarvänligt. AliView är programmerat i Java och fungerar därigenom oberoende av operativsystem.

Man kan nog säga att alla tidigare ansträngningar delvis var förberedelser för den **femte artikeln**.

Inom växtvärlden är det ganska vanligt att det uppstår nya arter genom kromosomtalsfördubbling (polyploidisering), inte minst i samband med hybridisering. I den **femte artikeln** undersöker vi hur omfattande hybridisering och kromosomfördubbling är inom ett ormbunksläkte. För blomväxter har det under det senaste årtiondet blivit vanligt att inte bara skapa släktskapsträd utifrån kloroplastgener utan även gener från cellkärnan. Kärngener finns i två varianter som ärvs från båda föräldrarna (till skillnad från kloroplasterna som endast ärvs från mödernet). De två varianterna innebär en möjlighet att finna eventuella hybridiseringar inom artbildningen. För ormbunkar finns sedan tidigare endast ett fåtal studier som inkluderar markörer från kärnan då det är mer komplicerat att sekvensera dessa regioner.

För att beskriva hällebräknarnas släktskap samt omfattningen av hybridisering och kromomfördubbling (polyploidisering) vid ormbunkars artbildning genomförde vi den hittills största studien med nukleära markörer av ormbunkars släktskap. För 188 olika kollekter fördelat över 36 arter sekvenserade vi två olika nukleära regioner (*pgiC* och *RPA2*), samt även fem kloroplastregioner.

Vi fann att arterna i *Woodsia* fördelar sig i två klader (klad = en gren på livets träd) vilka skildes åt för ca 45 millioner år sedan (figur 6). Extra anmärkningsvärt är att vi fann att två arter från var och en av dessa grupper hybridiserar. Det finns väldigt få exempel på andra organismer som kan hybridisera efter att ha utvecklats var för sig i 45 millioner år.

Arten *Woodsia montevidensis* finns både i Sydamerika och Afrika och har spridit sig från Sydamerika till Afrika relativt nyligen. Förfäderna har tidigare antagligen spridit sig från Asien via Nordamerika till Sydamerika.

Analysresultaten visar att det har skett och fortfarande pågår en omfattande hybridisering och kromosomfördubbling inom ormbunkssläktet *Woodsia*. Så många som 22 av 36 arter eller underarter av *Woodsia* är polyploider (kromosomfördubblade). Detta innebär att polyploidisering är den vanligaste artbildningsprocessen hos *Woodsia*. De flesta polyploida arter är relativt nybildade och det finns få äldre polyploida arter. Det verkar som att det kontinuerligt bildas nya polyploida arter från de äldre diploida arterna.

Åtminstonde fem av arterna uppvisar mönster i analysen som indikerar att de har uppstått flera gånger vid olika polyploidiseringstillfällen. Bland dessa finns den i Sverige förekommande fjällhällebräken (*W. alpina*) som är en allopolyploid hybrid *W. ilvensis* × *W. glabella* (figur 2). Att arten uppstått minst två gånger framgår av analysen av kloroplastsekvenser där vi kan se att vissa individer av *W. alpina* har kloroplaster nedärvda från *W. glabella* medan andra individer har kloroplaster från *W. ilvensis* (figur 6-7).

Vi utförde även en utökad studie där vi räknade antalet polyploider och analyserade frekvensen av nybildade arter genom polyploidisering för hela gruppen Eupolypods II. Resultaten är i linje med resultaten för *Woodsia* och vi fann att så stor andel som 54% av arterna och underarterna är polyploida och att 46% av artbildningen inom denna klad är kopplade till polyploidisering. Detta resultat påvisar en högre andel av polyploi än vad som uppskattats i tidigare undersökningar.

Acknowledgements

Tack till alla som hjälpt mig att genomföra det här projektet! Först ett stort tack till mina handledare Petra Korall och Mats Thulin för allt ni lärt mig, för att ni trott på mig, för all hjälp och allt samarbete. Tack till Nahid för att du lärt mig allt jag kan om labbarbete (och det är en hel del nu), och för att du är en bra kompis. Tack Afsaneh för all hjälp på labbet och för att du är en solstråle. Tack Agneta för alla kramar och din omtänksamhet (och även för alla biblioteksärenden du ordnat). Tack till Leif, Magnus, Martin, Micke, Sandie, Stefan, Thomas, Åsa och Hugo som läste igenom mina manus och kappa för att enas om det var gott nog. Tack Sandie för alla uppmuntrande ord och allt ditt jobb för avdelningen. Tack till mina nuvarande rumskompisar John, Karin och Paco som varit ett stort stöd och inspiration när jag varit less på att jobba. Och några tidigare rumskompisar som varit till stor glädje och inspirerande sällskap är givetvis Sunniva, Anneleen, Magnus, Maria L, Lars B och Hugo. Tack Allison för att du gjort lite extra med all korrekturläsning. Tack Magnus för ditt engagemang i manuskommentarer och mycket mer. Tack Martin för alla manus- och kappakommentarer och mycket mer. Tack Stina för att vi haft det så kul tillsammans när vi undervisat (och alla andra saker som du hjälpt mig med också). Tack Henrik för att jag fått lära mig en hel del om svamp, insekter och trädgårdsväxter. Tack Sarina (och Julius) för trevliga utflykter som gett mig mycket energi när jag varit lite trött på avhandlingsarbetet. Tack Ding för stödet i slutspurten. Tack Sanea för sockret som fått mig att orka jobba. Tack Anneleen för att du inspirerat mig att utveckla undervisningen. Tack Ping och Gemma för trevligt sällskap och att jag fick hjälpa er med servern. Tack Baset för att jag fick visa dig några svenska orkidéer. Tack Chengjie för att jag fått lära mig lite mer om horizontal transfer. Tack Sanja för att du drog tyngst i seminarie-arrangemangen och även för stöttande ord. Tack Inga för alla trevliga samtal. Tack Anja för allt du lärt mig om blommor och undervisning och mycket mer. Tack Åsa för skiten du ordnat med till kurslabbarna och mycket mer. Tack Anders R för att du var en mycket trevlig kursare. Tack Johanna för dina positiva råd! Tack till David för lite korrektur och tokiga råd. Tack alla tidigare kollegor som gjort avdelningen till en rolig plats att jobba på - Astrid, Andreas, Bozo, Cajsa, Catarina R, Catarina E, Christina W, Karolina, Katinka, Maria R, Omar och Rikke. Tack Katarina A och Henrik L för att ni lärt mig så mycket när vi undervisat tillsammans. Tack Niklas W för inspirerande diskussioner under luncherna. Tack Mariette för allt du lärt mig om Linné och för att du

såg till att jag inte missade att fullfölja den tredje uppgiften. Tack till alla mina medförfattare och inte minst till Carl för ett mycket trevligt och givande samarbete. Mina alltid positiva laotiska rumskompisar Chanda och Vichit förgyllde vardagen när de var här i Uppsala och jobbade. Några lärare, en del fortfarande kollegor och kompisar som jag vill tacka för att ni inspirerat mig lite extra är Per Erixon, Dick Andersson, Mattias Forshage, Leif Tibell och Mats Thulin. Tack Bengt för hjälpen med mitt magisterarbete, för att jag fick jobba som assistent i Sileneae-gruppen och för att din dialekt gör mig glad. Tack Elsbeth, Marianne, Lilian, Jan, Sven-Olof, Eva, Katariina och alla andra på IBG för att ni varit så positiva och hjälpt mig att få till undervisningen så bra. Elisabeth du får en egen rad för att du varit en så inspirerande kollega i undervisningen, hjälpt mig lite extra och för att du är en så älskvärd person. Tack Bosse, Margareta, Annika, Jesper och alla andra på trädgården för att ni är så trevliga och dessutom har tagit hand om mina ormbunksodlingar. Tack Monika, Mats, Jan, Stefan och alla andra på herbariet för att ni hjälpt mig med alla lån och mitt material. Tack Karin, Tove, Emma och Ulla för att ni ordnat med alla administrativa saker. Tack till Stefan, Micke, Håkan och alla andra datanördar som hjälp med min server, lånat ut prylar och program samt bistått i massor andra datorärenden. Tack Elisabeth för att du alltid är trevlig och hjälpt mig med mitt passerkort så många gånger. Tack till Johan, Stefan, Erik och alla andra på intendenturen för alla prylar jag fått låna och saker jag fått hjälp med.

Och så förstås en kram till min mamma, mina bröder och alla andra i min familj som ser till att jag har det bra. Tack till min andra familj (Sven, Olof Oskarsson, Camilla, Sofia, Edvin, Ludvig, Ester och Elsa) som har matat, stöttat och sporrat mig.

Tack för finansiella bidrag som hjälpt mig att genomföra arbetet, utdelade av: Linnéska stipendiestiftelsen, Regnells stiftelse, Helge Ax:son Johnson stiftelse, Tullbergs stipendiestiftelse, Knigges stipendiestiftelse, Sernanders stipendiestiftelse, KG Westmans stipendiestiftelse, K.V. Ossian Dahlgrens botaniska tidskriftsfond.

References

Barrington, D. S., Paris, C. A., & Ranker, T. A.. 1986. Systematic Inferences from Spore and Stomate Size in the Ferns. *American Fern Journal* 76:149-159. http://dx.doi.org/10.2307/1547723

Beck, J. B., Windham, M. D., & Pryer, K. M.. 2011. Do asexual polyploid lineages lead short evolutionary lives? A case study from the fern genus Astrolepis. *Evolution* 65:3217.

Brown, D. F. M. 1964. *A monographic study of the fern genus* Woodsia (Vol. 16). J. Cramer.

Brysting, A. K., Mathiesen, C., & Marcussen, T.. 2011. Challenges in polyploid phylogenetic reconstruction: A case story from the arctic-alpine *Cerastium alpinum* complex. *Taxon* 60:333-347.

Burleigh, J. G., Hilu, K. W., & Soltis, D. E.. 2009. Inferring phylogenies with incomplete data sets: a 5-gene, 567-taxon analysis of angiosperms. *BMC Evolutionary Biology* 9:61. http://dx.doi.org/10.1186/1471-2148-9-61

Duncan, T., & Smith, A. R.. 1978. Primary Basic Chromosome Numbers in Ferns: Facts or Fantasies? *Systematic Botany* 3:105-114. http://dx.doi.org/10.2307/2418535

Dyer, R. J., Savolainen, V., & Schneider, H.. 2012. Apomixis and reticulate evolution in the *Asplenium* monanthes fern complex. *Annals of Botany* 110:1515-1529. http://dx.doi.org/10.1093/aob/mcs202

Feng, Y., Oh, S.-H., Manos, P. S., & Zomlefer, W. B. 2005. Phylogeny and Historical Biogeography of the Genus *Platanus* as Inferred From Nuclear and Chloroplast DNA. *Systematic Botany* 30:786-799. http://dx.doi.org/10.1600/036364405775097851

Fiz-Palacios, O., Schneider, H., Heinrichs, J., & Savolainen, V. 2011. Diversification of land plants: insights from a family-level phylogenetic analysis. *BMC Evolutionary Biology* 11:341. http://dx.doi.org/10.1186/1471-2148-11-341

Guerra, M.. 2008. Chromosome numbers in plant cytotaxonomy: concepts and implications. *Cytogenetic and Genome Research* 120:339-350. http://dx.doi.org/10.1159/000121083

Kaplan, D. R., & Cooke, T. J. 1996. The Genius of Wilhelm Hofmeister: The Origin of Causal-Analytical Research in Plant Development. *American Journal of Botany* 83:1647-1660. http://dx.doi.org/10.2307/2445841

Kämpke, T., Kieninger, M., & Mecklenburg, M.. 2001. Efficient primer design algorithms. *Bioinformatics* 17:214-225. http://dx.doi.org/10.1093/bioinformatics/17.3.214

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., … Huala, E.. 2012. The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* 40:D1202-1210. http://dx.doi.org/10.1093/nar/gkr1090

Larsson, A.. 2013. lasseblaste. http://ormbunkar.se/phylogeny/lasseblaste/

Lewis, P.O., Holder, M. T., & Holsinger, K. E.. 2005. Polytomies and Bayesian Phylogenetic Inference. *Systematic Biology* 54:241-253. http://dx.doi.org/10.1080/10635150590924208

Lewis P.O., Holder M.T., Swofford D.L. 2010. Phycas. http://hydrodictyon.eeb.uconn.edu/projects/phycas/index.php

Li, F.-W.. 2013. Blue Devil - Blast Utility Seq Extr Devised By Li. http://pryerlab.biology.duke.edu/software

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R.. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078-2079. http://dx.doi.org/10.1093/bioinformatics/btp352

Lovis, J. D. 1978. Evolutionary Patterns and Processes in Ferns. In R.D. Preston and H.W. Woolhouse (Ed.), *Advances in Botanical Research* (Vol. Volume 4, pp. 229-415). Academic Press.

Ma, Y.-L.. 1985. Cytology and Taxonomy in Woodsiaceae. *Fern Gazette* 13:17-24.

Ma, Y.-L.. 1987. Studies on the Chinese Woodsiaceae (3)—Phylogeny and Speciation. Acta *Phytotaxonomica Sinica* 25:299-306.

Mallet, J. 2007. Hybrid speciation. *Nature* 446:279-283. http://dx.doi.org/10.1038/nature05706

McGrath, C. L., & Lynch, M.. 2012. Evolutionary Significance of Whole-Genome Duplication. In P. S. Soltis & D. E. Soltis (Eds.), *Polyploidy and Genome Evolution* (pp. 1-20). Springer Berlin Heidelberg.

Mickel, J., & Windham, M. D.. 2004. *Woodsia*. In The pteridophytes of Mexico (pp. 686-692). Bronx, N.Y.: New York Botanical Garden Press.

Moran, R. C. 2008. Diversity, biogeography, and floristics. In T. A. Ranker & C. H. Haufler (Eds.), *Biology and evolution of ferns and lycophytes*. Cambridge University Press Cambridge.

Nickrent, D. L., Parkinson, C. L., Palmer, J. D., & Duff, R. J. 2000. Multigene Phylogeny of Land Plants with Special Reference to Bryophytes and the Earliest Land Plants. *Molecular Biology and Evolution* 17:1885-1895.

Otto, S. P., & Whitton, J. 2000. Polyploid incidence and evolution. *Annual review of genetics* 34:401-437. http://dx.doi.org/10.1146/annurev.genet.34.1.401

Pilotti, M., Brunetti, A., Tizzani, L., & Marani, O. 2009. *Platanus × acerifolia* genotypes surviving to inoculation with *Ceratocystis platani* (the agent of canker stain): first screening and molecular characterization. *Euphytica* 169:1-17. http://dx.doi.org/10.1007/s10681-009-9884-9

Popp, M., & Oxelman, B.. 2001. Inferring the History of the Polyploid Silene aegaea (Caryophyllaceae) Using Plastid and Homoeologous Nuclear DNA Sequences. Molecular Phylogenetics and Evolution 20:474-481.

Pryer, K. M., Schneider, H., Smith, A. R., Cranfill, R., Wolf, P. G., Hunt, J. S., & Sipes, S. D. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409:618-622. http://dx.doi.org/10.1038/35054555

Ranker, T. A., & Geiger, J. M. O. 2008. Population genetics. In T. A. Ranker & C. H. Haufler (Eds.), *Biology and evolution of ferns and lycophytes*. Cambridge University Press Cambridge.

Rosenberg, N. A., & Nordborg, M.. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Reviews Genetics 3:380-390.

Rothfels, C. J., Sundue, M. A., Kuo, L.-Y., Larsson, A., Kato, M., Schuettpelz, E., & Pryer, K. M. 2012. A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales). *Taxon* 61:515-533.

Schneider, H., Schuettpelz, E., Pryer, K. M., Cranfill, R., Magallón, S., & Lupia, R. 2004. Ferns diversified in the shadow of angiosperms. *Nature* 428:553-557. http://dx.doi.org/10.1038/nature02361

Schuettpelz, E., & Pryer, K. M. 2007. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon* 56.

Sheffield, E. 2008. Alternation of generations. In T. A. Ranker & C. H. Haufler (Eds.), *Biology and evolution of ferns and lycophytes*. Cambridge University Press Cambridge.i

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. http://dx.doi.org/10.2307/25065646

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., … Soltis, P. S.. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* 96:336-348. http://dx.doi.org/10.3732/ajb.0800079

Soltis, D. E., Soltis, P. S., Schemske, D. W., Hancock, J. F., Thompson, J. N., Husband, B. C., & Judd, W. S.. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56:13-30.

Soltis, D. E., Soltis, P. S., & Tate, J. A.. 2004. Advances in the study of polyploidy since Plant speciation. *New Phytologist* 161:173-191. http://dx.doi.org/10.1046/j.1469-8137.2003.00948.x

Swofford, D. L.. 2003. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.

Taylor, T. N., & Taylor, E. L. 2008. Paleobotany, Second Edition: *The Biology and Evolution of Fossil Plants (2 edition.)*. Academic Press.

Van de Peer, Y., Fawcett, J. A., Proost, S., Sterck, L., & Vandepoele, K.. 2009. The flowering world: a tale of duplications. *Trends in Plant Science* 14:680-688. http://dx.doi.org/10.1016/j.tplants.2009.09.001

Van de Peer, Y., Maere, S., & Meyer, A.. 2009. The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics* 10:725-732. http://dx.doi.org/10.1038/nrg2600

Whitfield, J. B., & Lockhart, P. J. 2007. Deciphering ancient rapid radiations. *Trends in Ecology & Evolution* 22:258-265. http://dx.doi.org/10.1016/j.tree.2007.01.012

Wilkinson, M.. 1996. Majority-rule reduced consensus trees and their use in bootstrapping. *Molecular Biology and Evolution* 13:437-444.

Windham, M. D. 1993. *Woodsia*. In *Flora of North America* Editorial Committee (eds.) (Vol. 2, pp. 270-280). New York: Oxford University Press.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences* :pnas.0811575106. http://dx.doi.org/10.1073/pnas.0811575106

Yang, Z. 2008. Empirical evaluation of a prior for Bayesian phylogenetic inference. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363:4031-4039. http://dx.doi.org/10.1098/rstb.2008.0164

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1182

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-232233

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

Paper I

Syst. Biol. 61(3):490–509, 2012
© The Author(s) 2012. Published by Oxford University Press, on behalf of the Society of Systematic Biologists. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
DOI:10.1093/sysbio/sys001
Advance Access publication on January 4, 2012

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 KUO[3], PETRA KORALL[2], WEN-LIANG CHIOU[4], AND KATHLEEN M. PRYER[1]

[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 Taiwan 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;*
**Correspondence to be sent to: Department of Biology, Duke University, Box 90338, Durham, NC 27708, USA; E-mail: carl.rothfels@duke.edu.*
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 well-documented 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 character-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

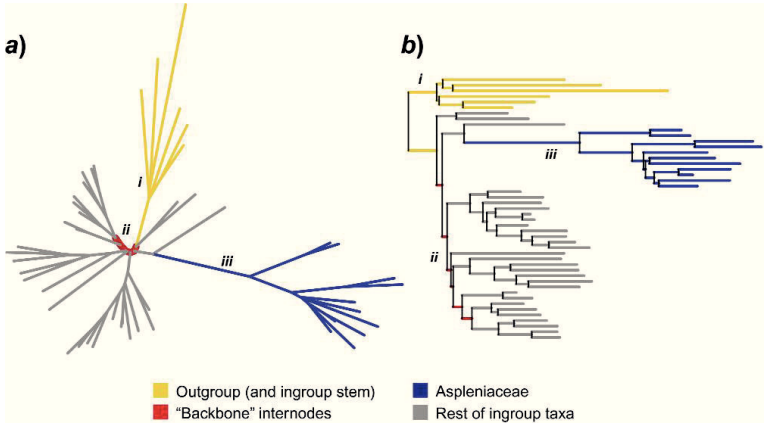

FIGURE 1. Challenges inherent in resolving the eupolypod II phylogeny. Eupolypods II phylogram modified from Schuettpelz and Pryer (2007), in (a) unrooted and (b) rooted form. (i) Outgroup taxa are on long branches. (ii) Backbone internodes are very short, suggesting an "ancient rapid radiation." (iii) 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 species-rich genera: *Thelypteris s. lat.* (~950 species); *Asplenium* (~700 species); *Diplazium* (~350 species); *Athyrium* (~220 species); and *Blechnum s. lat.* (~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

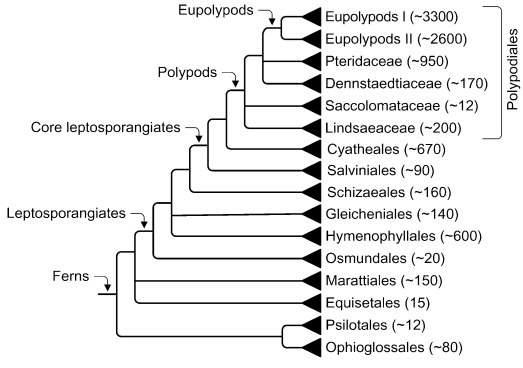

FIGURE 2. Broad phylogeny of ferns. Approximate number of species per clade is given in parentheses. Modified from Smith et al. (2006).

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 Eupolypods 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 g dried plant material, silica beads, 750 μL Carlson buffer, and 20 μL mercaptoethanol added to a 2-mL tube and ground for 45 s using a Mini-Beadbeater (BioSpec Products), followed by incubation at 65° 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: *atpA*, *atpB*, *matK*, *rbcL*, and the *trnG-trnR* intergenic spacer (henceforth "*trnG-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° C, followed by 35 cycles of 30 s denaturation at 95° C, 1 min annealing at 54° C, and 2 min elongation at 72° C, followed by a final elongation of 10 min at 72° C. The difficult amplification reaction, using Phusion High Fidelity DNA Polymerase (Finnzymes), was 1 min initial denaturation at 98° C, followed by 35 cycles of 10 s denaturation at 98° C, 30 s annealing at 58° C, and 1 min elongation at 72° C, followed by a final elongation of 8 min at 72° 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 μL of exonuclease I and 1 μL of Shrimp Alkaline Phosphatase (USB, Cleveland, OH); reaction tubes were incubated at 37° C for 15 min and then heated to 80° 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
atpA	ESATPA535F	F	S	ACAGCAGTAGCTACAGATAC	Schuettpelz et al. (2006)
atpA	ESATPA557R	R	S	ATTGTATCTGTAGCTACTGC	Schuettpelz et al. (2006)
atpA	ESATPA856F	F	S	CGAGAAGCATATCCGGGAGATG	Schuettpelz et al. (2006)
atpA	ESATPA877R	R	S	CATCTCCCGGATATGCTTCTCG	Schuettpelz et al. (2006)
atpA	ESATPF412F	F	A, S	GARCARGTTCGACAGCAAGT	Schuettpelz et al. (2006)
atpA	ESTRNR46F	R	A, S	GTATAGGTTCRARTCCTATTGGACG	Schuettpelz et al. (2006)
atpB	ATPB910R	R	S	TTCCTGYARAGANCCCATTTCTGT	Pryer et al. (2004)
atpB	ESATPB172F	F	A, S	AATGTTACTTGTGAAGTWCAACAAT	Schuettpelz and Pryer (2007)
atpB	ESATPB701F	F	S	TATGGTCAGATGAATGAACC	Schuettpelz and Pryer (2007)
atpB	ESATPE45R	R	A, S	ATTCCAAACWATTCGATTWGGAG	Schuettpelz and Pryer (2007)
matK	ASPmatKrLVV [a]	R	A, S	TTCGTGTCCRTAAAACAACCAA	This study
matK	Athymat rHTY [b]	R	A, S	CACACRAAGTTTYGTAYGTGTGAA	This study
matK	BLEmatKrDVP [c]	R	A, S	AATAGATGTRRAAATGGCACATC	This study
matK	CYSmatKrCGK [d]	R	A, S	AACTGAGTRACCTTTCCACACG	This study
matK	DematKrHTY [e]	R	A, S	ACGAAGTTTTGTACGTGTGAA	This study
matK	DimatKrTYK [f]	R	A, S	CCACACRAAGTTTTGTACGTGT	This study
matK	DIPZmatKrDSI [g]	R	A, S	GTCCATAAAACTACCAATATCGAATC	This study
matK	EuIImatKrHLL [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 [i]	R	A, S	GTRGAAATGGCACATCCCTAAT	This study
matK	PtmatKrIHY [j]	R	A, S	TTTCTMYATCTTSCRTARTGAAT	Kuo et al. (2011)
matK	THEmatKrVRL [k]	R	A, S	TCGACGAAACAAGCGAAC	This study
matK	WOOmatKrVRL [l]	R	A, S	TCKACGAAACAGGCGAAC	This study
rbcL	ESRBCL1361R	R	A, S	TCAGGACTCCACTTACTAGCTTCACG	Schuettpelz and Pryer (2007)
rbcL	ESRBCL1F	F	A, S	ATGTCACCACAAACGGAGACTAAAGC	Schuettpelz and Pryer (2007)
rbcL	ESRBCL663R	R	S	TACRAATARGAAACGRTCTCTCCAACG	Schuettpelz and Pryer (2007)
rbcL	PKRBCL556F	F	S	GGTAGRGCYGTCTAYGAATGYC	This study
trnG-R	trnG1F	F	A, S	GCGGGTATAGTTTAGTGGTAA	Nagalingum et al. (2007)
trnG-R	trnG353F	F	S	TTGCTTMTAYGACTCGGTG	Korall et al. (2007)
trnG-R	trnG63R	R	S	GCGGGAATCGAACCCGCATCA	Nagalingum et al. (2007)
trnG-R	trnGR43F1	F	S	TGATGCGGGTTCGATTCCCG	Nagalingum et al. (2007)
trnG-R	trnR22R	R	A, S	CTATCCATTAGACGATGGACG	Nagalingum et al. (2007)

Notes: F = forward; R = reverse; A = used in amplifications; S = used in sequencing reactions. While most primers were applied across the phylogeny, superscripts following primer names indicate lineage specificity: [a]Aspleniaceae; [b]athyriids; [c]Blechnaceae; [d]*Cystopteris*/*Gymnocarpium*; [e]depariids; [f]diplaziids; [g]*Diplaziopsis/Homalosorus*; [h]*Rhachidosorus*; [i]Onocleaceae; [j]Pteridaceae; [k]Thelypteridaceae; [l]*Woodsia* and allies.

Alignment and Tree Search

Sequences were manually aligned in *Mesquite 2.72* (Maddison and Maddison 2009). Indels (limited to *matK*, *trnG-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 *trnG-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+I+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 five-locus 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+I+G model as implemented in *PAUP* 4.0b10* 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

Data set	Taxa	Sites	Variable sites	Missing data (%)	Mean MLBS (%)	Bipartitions with ⩾70% MLBS (%)	Bipartitions with ⩾0.95 PP (%)
atpA	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
rbcL	78	1308	417	1.0	76	68	68
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* 4.0b10* 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"), and with only a single random-addition-sequence 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. Computation-intensive analyses were run on either the Duke Shared Cluster Resource (https://wiki.duke.edu/display/SCSC/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% (*matK*). The concatenated data have a mean MLBS of 92% and strongly support (i.e., have ⩾70% MLBS and ⩾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).

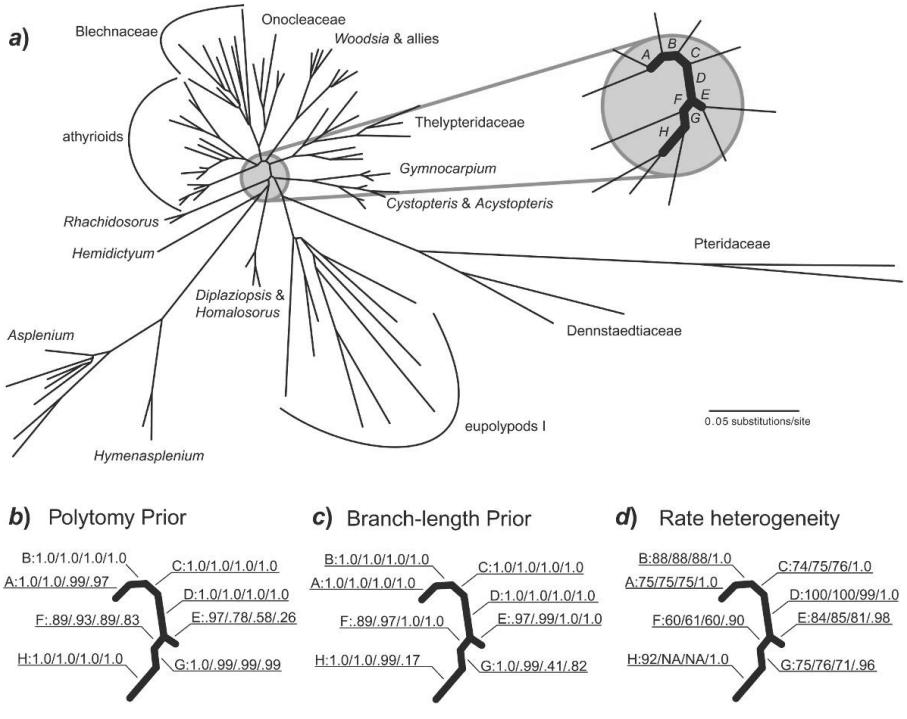

FIGURE 3. Phylogeny evaluation: rate heterogeneity and the Bayesian star-tree paradox artifact. a) Unrooted ML phylogram of the concatenated data, with the backbone internodes highlighted and labeled. b) Accounting for the impact of rate heterogeneity on backbone support values. The four values listed 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 mu1/mu0 = 0.01, *MrBayes 3.1.2* with branch-length prior mu1/mu0 = 0.001, and *MrBayes 3.1.2* with branch-length prior mu1/mu0 = 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 *atpA* and *atpB* place *Woodsia* and its allies within that clade (92% in *atpA*; 71% in *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-addition-sequence 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

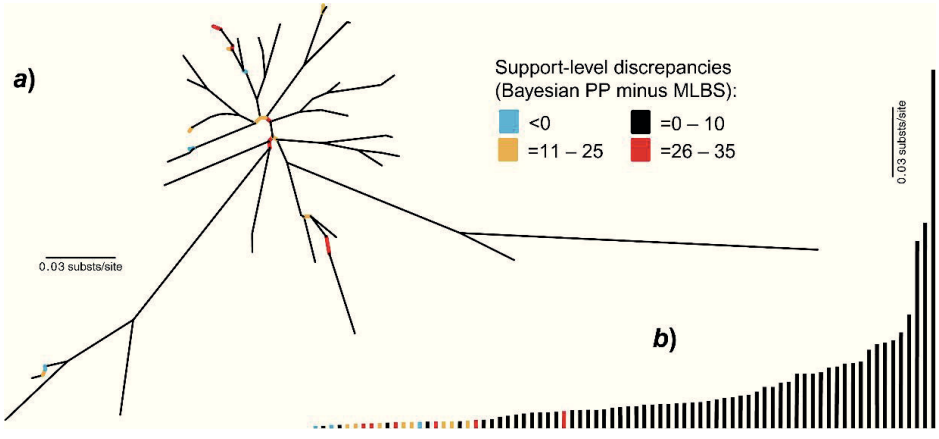

FIGURE 4. Discrepancies between ML and Bayesian support values. a) ML phylogram of the concatenated data, internal branches only (all tip branches have been deleted). Branches are colored according to the magnitude of the difference between their PP (from *MrBayes 3.1.1*) and their percent MLBS (from 5000 pseudoreplicates in *PAUP**). 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+I+G model, using a branch-length 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 *AWTY*-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 (PP = 0.93, 0.89, 0.83; Fig. 3b), whereas as the branch-length prior strength increased (mu0/mu1 = 0.01, 0.001, 0.0001), the PP on internode F also increased (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 branch-length 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 taxon-sample 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.

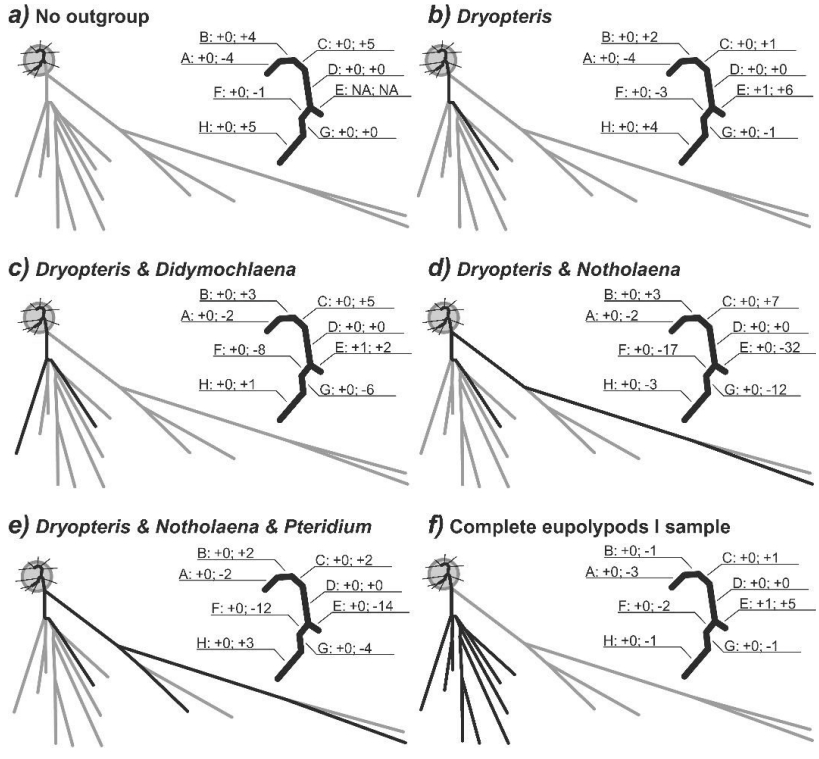

FIGURE 5. Effects of outgroup composition on ingroup backbone support values. Values on each internode indicate the percentage point difference in MLBS between the focal analysis and the analysis with the full taxon sample. The first values are those from the reduced consensus approach; second values 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. 5b,c,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/Gymnocarpium*; *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

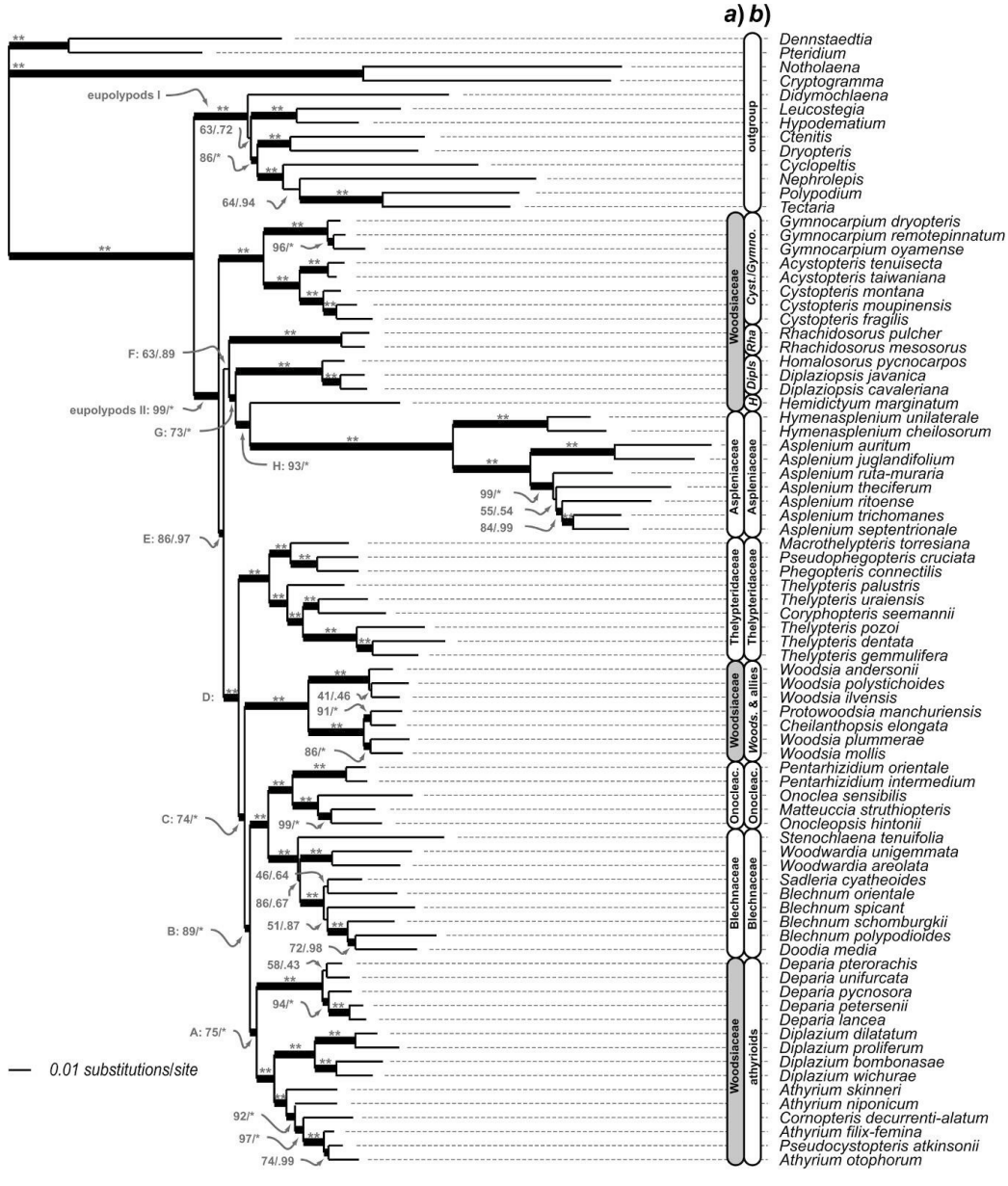

FIGURE 6. ML phylogram of the concatenated data. Support values indicated are MLBS followed by Bayesian PPs. Bold branches have ⩾70% MLBS and ⩾0.95 PP. Support values of 100% or 1.0 are indicated with an asterisk (*). a) Most recent family designations (Smith et al. 2006), with the paraphyletic Woodsiaceae 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 star-tree 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 mu0/mu1 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 branch-length 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 clades—Aspleniaceae, 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 (Mini-PEET 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 long-branch 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. 21:163–193.

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 *rbcL* 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 *rbcL* 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://mesquiteproject.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 *rbcL* 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., Morgan-Richards 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 *rbc*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 *rbcL* and *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 18S 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 *trn*L-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: *rbcL* 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 *rbc*L 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	*rbcL*	*atpB*	*atpA*	*trnG-R*	*matK*
Outgroup Dennstaedtiaceae								
Dennstaedtia dissecta (Sw.) T. Moore	2465	Schuettpelz 9 (DUKE)	Costa Rica: San Jose.	EF463166[j]	EF463373[j]	EF463652[j]	—	—
D. scabra(Wall. ex Hook.) T. Moore	—	Li 608 (TAIF)	Taiwan	—	—	—	—	JF303916[c]
Pteridium aquilinum (L.) Kuhn	—	Kuo 825 (TAIF)	Taiwan	—	—	—	JF832237	—
P. aquilinum (L.) Kuhn	—	Der 67 (UTC)	USA: California	—	—	—	—	EU223824[a]
P. esculentum (G. Forst.) Cockayne	125	Smith s.n. (UC)	Cult. (origin: Australia)	U05940[m]	—	EF452115[k]	—	—
Dryopteridaceae								
Ctenitis eatonii (Baker) Ching	—	Kuo 441 (TAIF)	Taiwan	—	—	—	—	JF303947[c]
C. sloanei (Poepp. ex Spreng.) C.V. Morton	3607	Schuettpelz 552 et al. (DUKE)	Cult. (Botanischer Garten München-Nymphenburg)	EF463172[j]	EF463383[j]	JF832106	JF832202	—
Didymochlaena truncatula (Sw.) J. Sm.	—	RBGE 19933685	Unknown	DQ508769[d]	—	—	—	—
D. truncatula (Sw.) J. Sm	2435	Schuettpelz 267 (DUKE)	Ecuador: Zamora-Chinchipe	—	EF452030[j]	JF832112	JF832210	—
D. truncatula (Sw.) J. Sm	—	No voucher taken	Cult. (Dr Cecilia Koo Bot. Cons. Center K017011)	—	—	—	—	JF303942[c]
Dryopteris filix-mas (L.) Schott	7057	Larsson 8 (UPS)	Sweden: Uppsala	JF832067	JF832164	JF832119	JF832217	—
D. sparsa (D. Don) Kuntze	—	Liu 481 (TAIF)	Taiwan	—	—	—	—	JF303945[c]
Hypodematium crenatum (Forssk.) Kuhn	6981	Yatskievych 09-072 et al. (MO)	Nepal: Kaski	JF832072	JF832169	JF832128	JF832225	—
H. crenatum (Forssk.) Kuhn	—	No voucher taken.	Cult. (Dr Cecilia Koo Bot. Cons. Center 013155; origin: Taiwan)	—	—	—	—	JF303944[c]
Leucostegia immersa Wall. ex C. Presl	—	Kuo 170 (TAIF)	Taiwan	—	—	—	—	JF303943[c]
L. pallida (Mett.) Copel.	3652	Schuettpelz 605 et al. (B)	Cult. (Botanischer Garten Berlin-Dahlem)	—	EF463425[j]	EF463709[j]	—	—
Lomariopsidaceae								
Cyclopeltis crenata (Fée) C. Chr.	—	No voucher taken	Cult. (Dr Cecilia Koo Bot. Cons. Center 0169403; origin: Thailand)	—	—	—	—	JF303954[c]
C. semicordata (Sw.) J. Sm.	3501	Barrington 2129 (VT)	Costa Rica: Punta Arenas	EF463234[j]	EF463480[j]	JF832107	JF832203	—
Nephrolepis cordifolia (L.) C. Presl	—	Li 335 (TAIF)	Taiwan	—	—	—	—	JF303953[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[k]	EF452027[k]	EU268740[h]	EU268687[h]	JF832265
Notholaena grayi Davenport	3187	Schuettpelz et al. 480 (DUKE)	USA: Arizona	EU268794[h]	JF832173	EU268749[h]	EU268697[h]	JF832280
Tectariaceae								
Tectaria zeylanica (Hout.) Sledge	3569	Schuettpelz 514 et al. (GOET)	Cult. (Alter Botanischer Garten Göttingen	EF463275[j]	EF463531[j]	JF832143	JF832242	—
T. zeylanica (Hout.) Sledge	—	No voucher taken.	Cult. (Dr Cecilia Koo Bot. Cons. Center 017295; origin: Taiwan)	—	—	—	—	JF303951[c]
Ingroup Aspleniaceae								
Asplenium auritum Sw.	3482	Schneider s.n. (GOET)	Cult. (Goett. Bot. Gard.)	EF463146[j]	EF463327[j]	EF463591[j]	—	—

Continued

APPENDIX 1. Continued

	DB#	Voucher	Provenance	*rbcL*	*atpB*	*atpA*	*trnG-R*	*matK*
A. juglandifolium Lam.	3465	Schneider s.n. (GOET)	Cult. (Goett. Bot. Gard.)	EF463151[j]	EF463333[j]	JF832092	JF832190	JF832252
A. ritoense Hayata	—	Nogami and K.O. 27 (KYO?)	Japan	AB014692[e]	—	—	—	—
A. ritoense Hayata	3479	Ranker 2063 (COLO)	Taiwan	—	EF463343[j]	EF463607[j]	—	—
A. ruta-muraria L.	—	K018516	Unknown	—	—	—	—	JF832253
A. ruta-muraria L.	2947	Christenhusz and Katzer 3869 (DUKE)	UK: Scotland	—	EF463344[j]	EF463608[j]	—	—
A. ruta-muraria L.	—	Unknown	Austria	AF525273[g]	—	—	—	—
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[j]	EF463349[j]	JF832095	—	JF832256
Hymenasplenium cheilosorum Kunze ex Mett.	3529	Schäfer 55 (GOET)	China: Yunnan	JF832071	EF463350[j]	JF832126	JF832223	—
H. unilaterale (Kunth) Mett.	3470	Ranker 2072 (COLO)	Taiwan	EF452140[k]	EF452020[k]	JF832127	JF832224	—
H. unilaterale (Kunth) Mett.	—	Li 619 (TAIF)	Taiwan	—	—	—	—	JF303924[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[c]
A. niponicum (Mett.) Hance	2852	Unknown	Japan	JF832057	JF832155	JF832097	JF832194	—
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[j]	EF463563[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[j]	EF463567[j]	JF832109	JF832207	—
D. lancea (Thunb.) R. Sano	—	Kuo 112 (TAIF)	Taiwan	—	—	—	—	JF303940[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	—	—	—	—	JF832269
D. pterorachis (H. Christ) M. Kato	—	Unknown	Unknown	AB021716[i]	—	—	—	—
D. pterorachis (H. Christ) M. Kato	—	Unknown	Unknown	D43907[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[j]	EF463569[j]	JF832111	JF832209	JF832271
Diplazium bombonasae Rosenst.	3764	Moran 7493 (NY)	Ecuador: Pastaza	EF463308[j]	EF463570[j]	JF832115	JF832213	JF832273
D. dilatatum Blume	3638	Schuettpelz 588 et al. (GOET)	Cult. (Botanischer Garten München-Nymphenburg)	EF463311[j]	EF463573[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[c]
D. proliferum (Lam.) Thouars	3639	Schuettpelz 590 and Schneider (GOET)	Cult. (Botanischer Garten München-Nymphenburg)	EF463315[j]	EF463577[j]	EF463918[j]	—	—
D. wichurae (Mett.) Diels	2874	Unknown	Cult. (UTBG 88-244)	—	EF463579[j]	JF832117	JF832215	—
D. wichurae (Mett.) Diels	—	Kuo 986 (TAIF)	Taiwan	—	—	—	—	JF832275
Pseudocystopteris atkinsonii (Bedd.) Ching	4837	Schuettpelz 1094A et al. (DUKE)	Taiwan: Nantou	JF832083	—	—	JF832235	—
P. atkinsonii (Bedd.) Ching	—	Kuo 477 (TAIF)	Taiwan: Nantou	—	—	—	—	JF832285

Continued

APPENDIX 1. Continued

	DB#	Voucher	Provenance	*rbcL*	*atpB*	*atpA*	*trnG-R*	*matK*
Blechnaceae								
Blechnum orientale L.	—	Kuo 827 (TAIF)	Taiwan: Taipei		—	JF304007[c]	—	JF303938[c]
B. orientale L.	—	Unknown	Unknown	AB040567[f]	—	—	—	—
B. polypodioides Raddi	2554	Schuettpelz 294 (DUKE)	Cult. (Duke U. Greenh.; origin: Costa Rica)	EF463159[j]	EF463352[j]	JF832100	JF832197	JF832260
B. schomburgkii (Klotzsch) C. Chr.	2410	Schuettpelz 242 (DUKE)	Ecuador: Zamora-Chinchipe	EF463160[j]	EF463353[j]	JF832101	JF832198	JF832261
B. spicant (L.) J. Sm.	6943	Windham 3395 (DUKE)	Canada: British Columbia	JF832059	JF832157	JF832102	JF832199	—
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[j]	JF832118	JF832216	JF832276
Sadleria cyatheoides Kaulf.	3432	Schuettpelz 507 (DUKE)	Cult. (Duke U. Greenhouse 87-166)	EF463161[j]	EF463356[j]	JF832141	JF832240	JF832288
Stenochlaena tenuifolia (Desv.) T. Moore	3429	Schuettpelz 504 (DUKE)	Cult. (Duke U. Greenhouse)	EF463163[j]	EF463358[j]	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[l]	—	—	—	—
W. unigemmata (Makino) Nakai	—	Kuo 493 (TAIF)	Taiwan	—	—	—	—	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[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	—	—	—	—	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[j]	EF463580[j]	JF832122	JF832220	—
G. remotepinnatum (Hayata) Ching	—	Kuo 149 (TAIF)	Taiwan	—	—	—	—	JF303926[c]
Dipls/Homal.								
Diplaziopsis cavaleriana (Christ) C. Chr.	—	Unknown	Unknown	D43909[i]	—	—	—	—
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	—	—	—	—	JF303928[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[c]
Hemidictyum								
Hemidictyum marginatum (L.) C. Presl	3054	Christenhusz 2476	French Guiana: Montagnes Tortue	EF463318[j]	EF463581[j]	JF832124	JF832221	JF303927[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[c]

Continued

APPENDIX 1. Continued

	DB#	Voucher	Provenance	*rbcL*	*atpB*	*atpA*	*trnG-R*	*matK*
Onoclea sensibilis L.	2998	Schuettpelz 353 (DUKE)	Cult. (Juniper Level Bot. Gard.; origin: USA)	—	EF463488[j]	JF832132	JF832229	—
O. sensibilis L.	—	Kuo 115 (TAIF)	USA: Massachusetts	—	—	—	—	JF303935[c]
O. sensibilis L.	—	Unknown	Unknown	JF832076[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[c]	JF832181	JF303998[c]	JF832239	JF303962[c]
Thelypteridaceae								
Coryphopteris seemannii Holttum	3774	Game 95/147 (UC)	Fiji: VitiLevu	EF463300[j]	EF463556[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[c]
Phegopteris connectilis (Michx.) Watt	7060	Larsson 17 (UPS)	Sweden: Uppsala	JF832080	JF832177	JF832136	JF832233	—
P. connectilis (Michx.) Watt	—	Kuo 151 (TAIF)	Taiwan	—	—	—	—	JF303932[c]
Pseudophegopteris cruciata (Willd.) Holttum	3559	Janssen 2724 (P)	France: Ile de la Reunion	EF463279[j]	EF463535[j]	JF832139	JF832236	JF832286
Thelypteris dentata (Forssk.) E.P. St. John	3654	Schuettpelz 607 et al. (B)	Cult. (Botanischer Garten Berlin-Dahlem)	EF463284[j]	EF463540[j]	JF832144	JF832243	JF832290
T. gemmulifera (Hieron.) A.R. Sm.	3747	Huiet s.n. et al. (UC)	Cult. (UCBG 83.0943; origin: Venezuala)	EF463285[j]	EF463541[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[c]	—	JF303933[c]
T. pozoi (Lag.) C.V. Morton	—	Unknown	Unknown	AB059579[n]	—	—	—	—
T. uraiensis (Rosenst.) Ching	—	Kuo 139 (TAIF)	Taiwan: Taipei	JF303972[c]	—	JF304002[c]	—	JF303934[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)	—	—	—	—	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[c]

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), [c]Kuo et al. (2011), [d]J.-M. Lu and D.-Z Li. (unpublished data), [e]N. Murikami, S. Nogami, M. Watanabe, and K. Iwatsuki. (unpublished data), [f]Nakahira (2000), [g]Pinter et al. (2002), [h]Rothfels et al. (2008), [i]Sano et al. (2000a), [j]Schuettpelz and Pryer (2007), [k]Schuettpelz et al. (2007), [l]Smith and Cranfill (2002), [m]Wolf et al. (1994), [n]Y. Yatebe and N. Murikami (unpublished data). Data missing or not applicable are indicated with a "—".

Paper II

A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales)

Carl J. Rothfels,[1] Michael A. Sundue,[2] Li-Yaung Kuo,[3] Anders Larsson,[4] Masahiro Kato,[5] Eric Schuettpelz[6] & Kathleen M. Pryer[1]

1 *Department of Biology, Duke University, Box 90338, Durham, North Carolina 27708, U.S.A.*
2 *The Pringle Herbarium, Department of Plant Biology, University of Vermont, 27 Colchester Ave., Burlington, Vermont 05405, U.S.A.*
3 *Institute of Ecology and Evolutionary Biology, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617, Taiwan*
4 *Systematic Biology, Evolutionary Biology Centre, Uppsala University, Norbyv. 18D, 752 36, Uppsala, Sweden*
5 *Department of Botany, National Museum of Nature and Science, Tsukuba 305-0005, Japan*
6 *Department of Biology and Marine Biology, University of North Carolina Wilmington, 601 South College Road, Wilmington, North Carolina 28403, U.S.A.*

Carl J. Rothfels and Michael A. Sundue contributed equally to this work.
Author for correspondence: *Carl J. Rothfels, carl.rothfels@duke.edu*

Abstract We present a family-level classification for the eupolypod II clade of leptosporangiate ferns, one of the two major lineages within the Eupolypods, and one of the few parts of the fern tree of life where family-level relationships were not well understood at the time of publication of the 2006 fern classification by Smith & al. Comprising over 2500 species, the composition and particularly the relationships among the major clades of this group have historically been contentious and defied phylogenetic resolution until very recently. Our classification reflects the most current available data, largely derived from published molecular phylogenetic studies. In comparison with the five-family (Aspleniaceae, Blechnaceae, Onocleaceae, Thelypteridaceae, Woodsiaceae) treatment of Smith & al., we recognize 10 families within the eupolypod II clade. Of these, Aspleniaceae, Thelypteridaceae, Blechnaceae, and Onocleaceae have the same composition as treated by Smith & al. Woodsiaceae, which Smith & al. acknowledged as possibly non-monophyletic in their treatment, is circumscribed here to include only *Woodsia* and its segregates; the other "woodsioid" taxa are divided among Athyriaceae, Cystopteridaceae, Diplaziopsidaceae, Rhachidosoraceae, and Hemidictyaceae. We provide circumscriptions for each family, which summarize their morphological, geographical, and ecological characters, as well as a dichotomous key to the eupolypod II families. Three of these families—Diplaziopsidaceae, Hemidictyaceae, and Rhachidosoraceae—were described in the past year based on molecular phylogenetic analyses; we provide here their first morphological treatment.

Keywords Athyriaceae; *Diplaziopsis*; ferns; *Rhachidosorus*; taxonomy; Woodsiaceae

■ INTRODUCTION

Despite intensive studies spanning the late 1930s to 1980s (Christensen, 1938; Ching, 1940, 1978a, b; Copeland, 1947; Holttum, 1947; Alston, 1956; Nayar, 1970; Pichi Sermolli, 1973; Sledge, 1973; Mickel, 1974; Tryon & Tryon, 1982; see Smith, 1995), evolutionary relationships within ferns remained obscure, and suprageneric treatments varied wildly. Holttum lamented in 1971 that "most family names of ferns have had such different meanings, as used by different authors, that such names are only intelligible if we associate them with the names of particular authors". He suggested "in the meantime it would best serve the ultimate stability of nomenclature if we regard all family names of ferns as informal and tentative (which in fact they have always been)" (Holttum, 1971a). Thirty-five years later, Hennipman (1996) voiced a similar sentiment, that "modern higher classifications of ferns are a jungle for the user". As recently as 1990, for example, the schizaeoid ferns (Schizaeales sensu Smith & al., 2006) and pteroid ferns (Polypodiales: Pteridaceae sensu Smith & al., 2006) were hypothesized to be each other's closest living allies (Tryon & al., 1990); current evidence, however, suggests these lineages shared a most recent common ancestor over 260 million years ago (Schuettpelz & Pryer, 2009, their table S3), and that pteroids are more closely related to other Polypodiales, the Cyatheales, and the Salviniales (in total, the vast majority of ferns) than they are to the schizaeoids. Suprageneric fern classifications had fallen into such disrepute that some recent Floras avoided them altogether, opting instead to present genera in alphabetical order (e.g., Smith, 1981; Palmer, 2002; Mickel & Smith, 2004; Zuquim & al., 2008).

For nearly two decades, renewed investigations using molecular (Hasebe & al., 1994, 1995; Manhart, 1994, 1995; Wolf & al., 1994, 1998, 1999; Wolf, 1995, 1997; Kranz & Huss, 1996; Pahnke & al., 1996; Vangerow & al., 1999; Sano & al., 2000a; Wang & al., 2003; Pryer & al., 2004; Schneider & al., 2004b; Wikström & Pryer, 2005; Korall & al., 2006a, b; Schuettpelz & al., 2006; Schuettpelz & Pryer, 2007), morphological (Schneider, 1996; Stevenson & Loconte, 1996), and combined molecular and morphological data (Pryer & al., 1995, 2001) have yielded increased

support for the relationships that shape the major branches of the fern tree of life. In 2006, these phylogenetic hypotheses were consolidated and presented in a revised classification for ferns (Smith & al., 2006).

Smith & al. (2006) recognized a monophyletic Polypodiales ("Polypods") within which the majority of species fall into two large "eupolypod" clades, sister to each other and christened Eupolypods I and Eupolypods II, respectively (Fig. 1) (Schneider & al., 2004b). Together, the eupolypod lineages include nearly 6000 species—more than half of extant fern diversity. The large eupolypod clades had been hinted at, rather presciently, by earlier workers, including Sledge (1973, his Aspidiaceae and Athyriaceae approximate the Eupolypods I and II, respectively) and Mickel (1974, who grouped members of what are now called Eupolypods together in a "derived" position on his tree, Polypodiaceae being the chief exception). The existence of the eupolypod clade was further suggested by early molecular (Hasebe & al., 1994, 1995), morphological (Stevenson & Loconte, 1996), and combined analyses (Pryer & al., 1995). Schneider & al. (2004b) were the first to adopt the names Eupolypods I and II for these two clades, and it was not until the Smith & al. (2006) compilation that their composition was broadly understood.

As currently circumscribed, Eupolypods II is a large clade, comprising over 2500 species, including those associated with the large genera *Asplenium* (~700 spp.), *Cyclosorus* (~650 spp.), *Diplazium* (~400 spp.), *Athyrium* (~180 spp.), and *Blechnum* (~150 spp.; estimates from Kramer & Viane, 1990; Kramer & al., 1990a, b; Smith, 1990). It encompasses great morphological and ecological variation (Fig. 2), including taxa as disparate as the diminutive dry-rock dwelling *Asplenium tenerrimum* Mett. ex Kuhn, large arborescent tropical *Blechnum auratum* (Fée) R.M. Tryon & Stolze, high-arctic plants of *Woodsia glabella* R. Br. ex Richardson, and the temperate floodplain understory (and frequently sautéed) *Matteuccia struthiopteris* (L.) Tod. Given its species richness, morphological disparity, and lack of historical recognition, it is not surprising that unequivocal morphological synapomorphies for Eupolypods II are lacking. However, some clear trends exist that are particularly useful for distinguishing Eupolypods II from Eupolypods I. Most eupolypod II taxa have two vascular bundles in the stipe (vs. many bundles in Eupolypods I), and many eupolypod II species have linear, indusiate sori (in the rare cases where members of Eupolypods I have linear sori, they are not indusiate; Fig. 3).

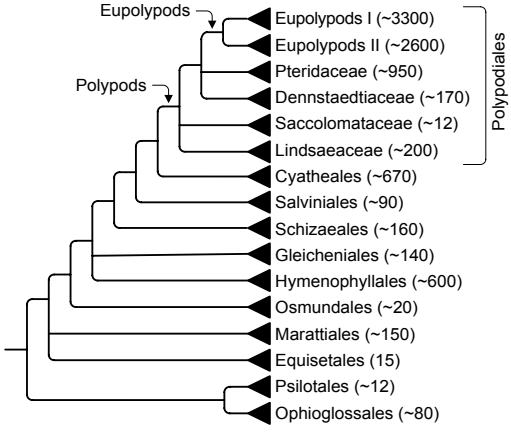

Fig. 1. Fern phylogeny. Numbers in parentheses indicate approximate species richness of each clade. Modified from Smith & al. (2006) and Rothfels & al. (2012).

In their treatment of Eupolypods II, Smith & al. (2006) recognized not only that the backbone relationships within the clade were unresolved, but that Woodsiaceae as it was then circumscribed was possibly not monophyletic; the data then available did not support a monophyletic Woodsiaceae, but they also did not support any alternative set of relationships (Hasebe & al., 1995; Sano & al., 2000a; Pryer & al., 2004; Schneider & al., 2004b). In recognizing a potentially non-monophyletic Woodsiaceae, Smith & al. (2006) issued the caveat that, while "it is premature to adopt the alternative of erecting (or resurrecting) numerous small families to house its constituent genera … further sampling will likely shed additional light on this subject, and the recognition of several additional families may be warranted" (Smith & al., 2006).

Further studies were rapidly forthcoming. In their 400-taxon, three-gene study, Schuettpelz & Pryer (2007) showed that three genera—*Cystopteris*, *Gymnocarpium*, *Hemidictyum*—tentatively placed in Woodsiaceae (Smith & al., 2006) were only distantly related to other members of Woodsiaceae sensu Smith & al. (2006). This general pattern—Woodsiaceae sensu Smith & al. (2006) not monophyletic and the backbone relationships within Eupolypods II only weakly supported—was also uncovered by the two-gene analyses of Wei & al. (2010), the three-gene analyses by Kuo & al. (2011), and the four-gene study of Li & al. (2011).

To directly address the composition of the major clades within Eupolypods II and the relationships among them, Rothfels & al. (2012) assembled an expanded molecular dataset (five plastid loci) for 67 eupolypod II species and 14 outgroup taxa. Their taxon sampling was designed to capture the deepest divergences across Eupolypods II and those within each major clade, as well as any potentially isolated lineages, as suggested by previous molecular (particularly Sano & al., 2000a; Tzeng, 2002; Schuettpelz & Pryer, 2007; Kuo & al., 2011) or morphological studies (chiefly Kato & Darnaedi, 1988; Wang & al., 2004). Although the results of Rothfels & al. (2012) were consistent with those of earlier studies, the more comprehensive taxon and data sampling provided higher levels of support for relationships and helped to resolve most of the taxonomic challenges in Eupolypods II. We base our classification on their inferred phylogeny (see Fig. 4), with the caveat that, like all phylogenetic studies of the Eupolypods II to date, their phylogeny is based solely on plastid data; no loci from the nucleus or mitochondrion were included. This classification is similar in outline to the linear sequence recently proposed by Christenhusz & al. (2011), but is further informed by the critical data of Kuo & al. (2011), Li & al. (2011), and Rothfels & al. (2012).

Fig. 2. Representative eupolypod II ferns. Photographers are credited after the species names. ASPLENIACEAE—**A,** *Asplenium nidus* L. s.l. [M. Sundue]; **B,** *Asplenium montanum* Willd. [S. Zylinski]. ATHYRIACEAE—**C,** *Athyrium asplenioides* (Michx.) A.A. Eaton [S. Zylinski]. BLECHNACEAE—**D,** *Woodwardia areolata* (L.) T. Moore [C. Rothfels]; **E,** *Blechnum schomburgkii* (Klotzsch) C. Chr. [M. Sundue]. CYSTOPTERIDACEAE—**F,** *Gymnocarpium remotepinnatum* (Hayata) Ching [L.-Y. Kuo]; **G,** *Cystopteris protrusa* (Weath.) Blasdell [C. Rothfels]; **H,** *Cystopteris fragilis* (L.) Bernh. [C. Rothfels]. DIPLAZIOPSIDACEAE—**I,** *Diplaziopsis javanica* (Blume) C. Chr. [L.-Y. Kuo]. HEMIDICTYACEAE—**J,** *Hemidictyum marginatum* (L.) C. Presl [M. Sundue]. ONOCLEACEAE—**K,** *Onocleopsis hintonii* F. Ballard [C. Rothfels]; **L,** *Matteuccia struthiopteris* (Hook.) Hayata [M. Sundue]. RHACHIDOSORACEAE—**M,** *Rhachidosorus mesosorus* (Makino) Ching [L.-Y. Kuo]. THELYPTERIDACEAE—**N,** *Thelypteris noveboracensis* (L.) Nieuwl. [C.W. Cook]. WOODSIACEAE—**O,** *Woodsia alpina* (Bolton) Gray [A. Larsson].

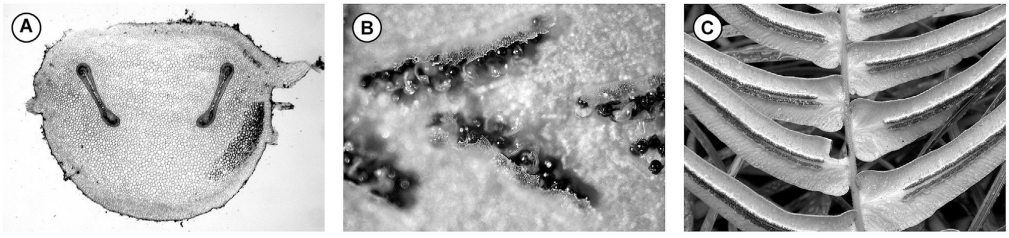

Fig. 3. Morphological characteristics of eupolypod II taxa. Photographers are credited within square brackets. **A,** Cross-section of *Diplaziopsis javanica* (Blume) C. Chr. (Diplaziopsidaceae) showing two vascular bundles at the base of the petiole [L.-Y. Kuo]. **B,** Close-up of abaxial leaf surface of *Asplenium platyneuron* (L.) Britton, Sterns & Poggenb. (Aspleniaceae), showing sporangia arranged in linear, indusiate sori. The sporangia are visible under the flap-like erose indusium, which opens away from the vein [C.J. Rothfels]. **C,** Abaxial leaf surface of *Blechnum occidentale* L. (Blechnaceae), again showing sporangia arranged in linear, indusiate sori. In this species the sori are contiguous along the main vein of each pinna, and the indusium opens towards the vein [R.C. Moran; modified with permission from www.plantsystematics.org].

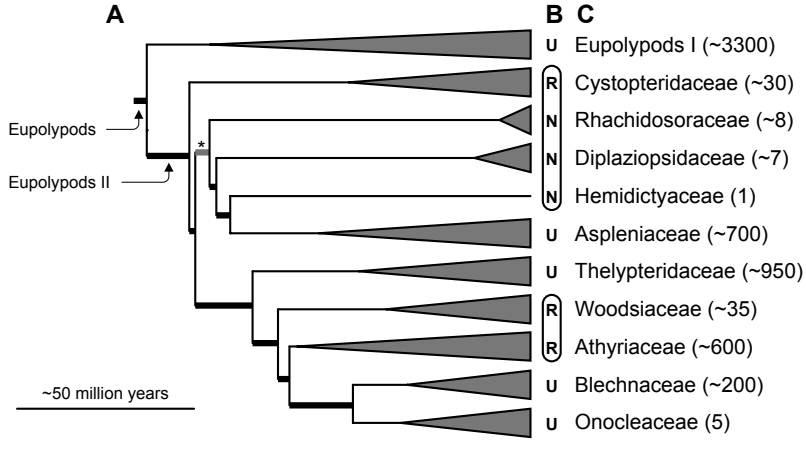

Fig. 4. Divergence and diversification in the Eupolypods II. **A,** Eupolypod phylogeny, with branch lengths approximately proportional to time (from the relaxed clock analyses of Rothfels & al., 2012, their Supplementary Fig. 1). The tip of the grey triangles along each branch marks the first sampled divergence within each family (Rothfels & al., 2012). All branches in this phylogeny are well supported (maximum likelihood bootstrap support ≥70% and Bayesian posterior probability ≥0.95) with the exception of the grey branch, marked with an asterisk (*), which had 63% maximum likelihood bootstrap support and 0.89 posterior probability. **B,** Family-level nomenclatural status: N, newly described since Smith & al. (2006); R, recircumscribed (family name existed, but was not adopted by Smith & al., 2006); U, unchanged from Smith & al. (2006). Letters that are encircled indicate those families that have been segregated from Woodsiaceae sensu Smith & al. (2006). **C,** Family names, and approximate species richness, for the classification adopted here.

■ CLASSIFICATION OF THE EUPOLYPODS II

The aim of our classification is to recognize families within the eupolypod II phylogeny that balance the somewhat conflicting criteria of maximizing evolutionary informativeness (we thus adhere to the principle of monophyly) and minimizing nomenclatural instability (we retain long-established circumscriptions as much as possible). This conflict is most difficult to reconcile for *Asplenium* and its allies. Both choices (to recognize an expanded Aspleniaceae that includes *Hemidictyum*, *Diplaziopsis*, and *Homalosorus*, or to create new families to accommodate the latter three genera) yield justifiable, monophyletic families. We take the latter approach—to recognize Aspleniaceae, Hemidictyaceae, and Diplaziopsidaceae—despite the addition of two small families, in order to preserve the long-standing use of Aspleniaceae in the more restricted sense, to highlight the deep divergence of each of the respective groups (Fig. 4), and because there are no clear morphological synapomorphies for the expanded family concept.

Many generic concepts in Eupolypods II are in flux and although not a focus of our classification, we attempt to account for all generic names in current general usage, and provide a

familial placement. For each family, we provide: a list of defining morphological characters (from the references cited in the family header, from our direct observations, and from the following general references: Wilson, 1959; Ogura, 1972; Tryon & Tryon, 1982; Gifford & Foster, 1989), nomenclatural data, and a list of included genera and the estimated number of species. In addition, we recommend possible English family names, and summarize information on ecology, geographic range, and phylogenetic relationships. Each family is accompanied by a concept map (see Franz & al., 2008), mapping our treatment onto previous classifications. For example, the following entry under Rhachidosoraceae: "=Athyriaceae: Rhachidosoroideae sensu Wang & al. (2004); <Woodsiaceae sensu Smith & al. (2006)" indicates that our treatment of that family is equivalent in composition to Wang & al.'s concept of subfamily Rhachidosoroideae of Athyriaceae, and is a subset of Smith & al.'s concept of Woodsiaceae. Family names are based on those in Hoogland & Reveal (2005), except for Diplaziopsidaceae and Rhachidosoraceae, which are from Christenhusz & al. (2011), and Hemidictyaceae, from Christenhusz & Schneider (2011).

■ CYSTOPTERIDACEAE (Payer) Shmakov, *Turczaninowia* 4: 60 (2001)

Cystopteridoids; Bladderferns, Brittleferns, Oakferns, and allies. Approximately 30 species in the genera *Acystopteris* Nakai (3 spp.), *Cystoathyrium* Ching (1 sp.), *Cystopteris* Bernh. (~20 spp.; incl. *Rhizomatopteris* A.P. Khokhr.), and *Gymnocarpium* Newman (~7 spp.; incl. *Currania* Copel.); (Tagawa, 1935; Blasdell, 1963; Vida, 1974; Sarvela, 1978; Haufler & Windham, 1991; Pryer & Haufler, 1993).

<Polypodiaceae: Asplenioideae + Polypodiaceae: Dryopteridoideae sensu Christensen (1938); <Dennstaedtiaceae: Dryopteridoideae + Athyrioideae sensu Holttum (1947); <Dryopteridaceae: Dryopteridoideae + Athyrioideae sensu Nayar (1970); <Athyriaceae sensu Pichi Sermolli (1977); <Athyriaceae sensu Ching (1978a); <Dryopteridaceae: Athyrioideae sensu Lovis (1978); <Dryopteridaceae: Physematieae sensu Tryon & Tryon (1982); <Dryopteridaceae: Athyrioideae: Physematieae sensu Kramer & al. (1990b); <Cystopteridaceae sensu Shmakov (2001); <Athyriaceae: Cystopterioideae sensu Wang & al. (2004); <Woodsiaceae sensu Smith & al. (2006); =Cystopteridaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial; roots blackish, wiry, inserted radially, non-proliferous; rhizomes epigeous or more often subterranean, short- to more often long-creeping, occasionally suberect (*Cystopteris*), commonly branched, bearing scales and sometimes golden hairs similar to the root-hairs (e.g., *C. protrusa* (Weath.) Blasdell); rhizome scales lanceolate, clathrate or non-clathrate, the margins glandular or not, without distinct pubescence, entire to ciliate, the teeth when present not formed by two adjacent cells; leaves green and not covered in mucilage during any stage of development, spirally arranged, monomorphic, bulbiferous in a few *Cystopteris*, closely spaced to distant, bearing scales and sometimes gland-tipped hairs, the scales sometimes reduced to filiform proscales (*Cystopteris*) or catenate hairs (*Acystopteris*); petioles stramineous throughout or proximally darkened, the base narrow, or conspicuously thickened and then starch-filled and persistent (trophopods), without conspicuous aerophores, without a proximal articulation, sometimes with golden hairs similar to the root hairs (e.g., *C. moupinensis* Franch.); petiolar vascular bundles two, the bundles with hippocampiform-shaped xylem, distally uniting to form a single V-shaped bundle; laminae thin-herbaceous, 2–3-pinnate-pinnatifid (pinnate-pinnatifid in *Cystoathyrium*), broadest at the base or lanceolate, the apex non-conform, the leaf marginal cells differentiated into nodulose hyaline cells (*Acystopteris*, *Cystopteris*) or not (*Gymnocarpium*); pinna axes distinctly articulate in *Gymnocarpium*, otherwise non-articulate, sulcate adaxially, lacking a free central ridge; the rachis grooves continuous or not, the sulcus wall of the rachis continuing as a prominent ridge onto the sulcus wall of the costa or not; veins free, terminating at the leaf margin, the vein endings not differentiated; sori dorsal along veins, not terminal, round or slightly elongate (*Gymnocarpium*), indusiate (*Acystopteris*, *Cystoathyrium*, *Cystopteris*) or exindusiate (*Gymnocarpium*); soral receptacle distinctly raised and hardened (*Acystopteris*, *Cystopteris*) or flat (*Gymnocarpium*); indusia basal (*Acystopteris*, *Cystoathyrium*, *Cystopteris*); sporangia with stalks two or three cells wide in the middle; spores monolete, non-chlorophyllous, tan (*Acystopteris*) or brown, the perispore echinate, tuberculate, or with broad folds, the folds sometimes perforate; chromosome base number $x = 40$ (*Gymnocarpium*; Kato & al., 1992; Pryer & Haufler, 1993) or 42 *(Acystopteris*, *Cystopteris*; Blasdell, 1963; Vida, 1974; Mitui, 1975). Reports of $x = 41$ (e.g., Christenhusz & al., 2011) are not substantiated.

Although the genera are distinctive, Cystopteridaceae as a whole are not easily characterized. Among families with petioles that contain two vascular bundles, they can be distinguished by an absent or hood-like indusium, usually long-creeping and subterranean rhizome, and veins that terminate at the leaf margin.

The indusia of *Acystopteris*, *Cystoathyrium*, and *Cystopteris* are unique in being attached at the base of the sporangia and curving, hood-like, around them. The sorus itself is situated upon a raised and hardened receptacle; we know of no other taxa within Eupolypods II with a similar receptacle. *Woodsia* also has a basally attached indusium, however, it can be distinguished by having a flat receptacle, the indusium encircling the sorus, usually dissected into multiple lobes, and veins that do not reach the leaf margin.

Gymnocarpium can be diagnosed by its articulate pinnae (that do not disarticulate) with a swollen protuberance at the base of each pinna. Among Eupolypods II, articulate pinnae also occur in *Stenochlaena* and *Woodwardia virginica* (L.) Sm. Those articulations differ, however, by lacking the basal protuberance present in *Gymnocarpium*. *Gymnocarpium* can also be distinguished from other Eupolypods II by having slightly elongate sori that lack an indusium. These sori are about twice as long as wide, and appear round until the sporangia are removed to reveal an elongate patch of sporangial stalks spreading along the vein.

Biology and phylogeny. — Cystopteridaceae are unusual in their primarily temperate distribution and tendency to occupy montane habitats. Both *Cystopteris* and *Gymnocarpium*

are common ferns of the north temperate zone, with *Cystopteris* also ranging south in montane habitats through the Andes and Himalayas, and to Australia, New Zealand, Hawaii, and southern Africa. Within the family, *Acystopteris* is the only genus found commonly in tropical areas; it is most rich in East Asia (Blasdell, 1963; Sarvela, 1978; Pryer, 1993).

The relationships of genera within Cystopteridaceae have been the subject of unusually strong disagreement; their affinities have been extremely difficult to infer from morphology, even more so than is typical for most eupolypod II taxa. Individually, both *Cystopteris* s.l. (i.e., including *Acystopteris*; Tagawa, 1935; Blasdell, 1963) and *Gymnocarpium* have been thought to be allied with Dryopteridaceae (in Eupolypods I) or Athyriaceae; in either position they were inevitably highlighted as being anomalous (see Sledge, 1973). Ching (1940) was an early exception, however, in placing both *Cystopteris* and *Gymnocarpium* together, but among the athyrioids.

Early molecular data supported *Cystopteris* and *Gymnocarpium* as sister genera (one accession each; *Acystopteris* was not sampled), and demonstrated their lack of close affinity to either Dryopteridaceae or *Athyrium*, but were unable to resolve their position within a broad assemblage of eupolypod II taxa (Wolf & al., 1994; Hasebe & al., 1995). In their landmark study, Sano & al. (2000a) included four representatives from this clade: one *Acystopteris*, one *Cystopteris*, and two *Gymnocarpium* species. Their within-clade relationships were consistent with earlier studies (*Acystopteris* + *Cystopteris* sister to *Gymnocarpium*), but they did not find support for the clade's placement within Eupolypods II. Conversely, Schuettpelz & Pryer (2007) in their broad study across ferns, included fewer taxa from this clade (a single *Gymnocarpium* and a single *Cystopteris*) but more character data; they were the first to find support for a sister-group relationship between this clade and the rest of Eupolypods II. Similarly, Kuo & al. (2011), using three plastid loci and single accessions of *Acystopteris* and *Gymnocarpium*, also recovered the sister-group relationship of this clade to the rest of Eupolypods II. The results of Rothfels & al. (2012) corroborate and strengthen that finding (Fig. 4).

Cystopteris and *Acystopteris* are strongly supported as sister, and are in turn sister to *Gymnocarpium*. The type species of all three genera have been included in molecular phylogenetic studies, as have those of the segregates *Rhizomatopteris* and *Currania*: *Rhizomatopteris* is sister to the remaining species of *Cystopteris* s.str.; *Currania* is embedded within *Gymnocarpium* (Sano & al., 2000a; Rothfels & al., 2012). Cystopteridaceae was first circumscribed by Shmakov (2001), who included *Pseudocystopteris* (which belongs in Athyriaceae) and omitted *Acystopteris* and *Cystoathyrium* (which were not in his geographic range).

The position of the monotypic genus *Cystoathyrium* is uncertain. In describing the genus, Ching (1966) emphasized its morphological intermediacy between *Cystopteris* and *Athyrium*, and little progress has since been made towards resolving its affinities. Wang & al. (2004, 2008) treated it as allied to *Cystopteris*, whereas Kramer & al. (1990b) place it in *Athyrium*, a position also advocated by Pichi Sermolli (1977). *Cystoathyrium* has yet to be included in any phylogenetic study, and we know it only from photographs and the illustration provided in the protologue; it is possibly extinct (X.-C. Zhang, pers. comm.). Other genera historically thought to be allied with *Cystopteris* (most notably *Pseudocystopteris*; Ching, 1964a) have been shown to be nested within Athyriaceae (Sano & al., 2000a; Liu, 2008; Rothfels & al., 2012), however, we tentatively include *Cystoathyrium* here in Cystopteridaceae based on four characters: round sori, hood-like indusium, strongly echinate spores, and veins that terminate at the leaf margin. Although homoplastic within Eupolypods II, these character states occur most frequently in Cystopteridaceae. More research is needed; *Cystoathyrium* may be an isolated lineage within the Eupolypods II.

■ RHACHIDOSORACEAE X.C. Zhang, *Phytotaxa* 19: 16 (2011)

Lacquer Ferns. Four to seven species of the genus *Rhachidosorus* Ching; (Ching, 1964a; Kato, 1975a; Li & al., 2011).

<Polypodiaceae: Asplenioideae sensu Christensen (1938); <Dennstaedtiaceae: Athyrioideae sensu Holttum (1947); <Dryopteridaceae: Athyrioideae sensu Nayar (1970); <Athyriaceae sensu Tagawa & Iwatsuki (1972); =“*Diplazium mesosorum* group” sensu Kato (1977); <Athyriaceae sensu Pichi Sermolli (1977); <Athyriaceae sensu Ching (1978a); <Dryopteridaceae: Athyrioideae sensu Lovis (1978); <Dryopteridaceae: Physematieae sensu Tryon & Tryon (1982); <Dryopteridaceae: Athyrioideae: Physematieae sensu Kramer & al. (1990b); =Athyriaceae: Rhachidosoroideae sensu Wang & al. (2004); <Woodsiaceae sensu Smith & al. (2006); =Rhachidosoraceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial; roots inserted radially, non-proliferous; rhizomes creeping or short-creeping, not commonly branched, bearing scales; rhizome scales lanceolate, clathrate, the margins entire, without distinct pubescence; leaves green and not covered in mucilage during any stage of development, spirally arranged, monomorphic, not articulate to the rhizome, closely spaced, sparsely scaly; petioles reddish to stramineous throughout, narrow at the base, not forming trophopods, without conspicuous aerophores, without a petiolar articulation; petiolar vascular bundles two, each with hippocampiform xylem, the bundles distally uniting to form a single U-shaped bundle; laminae herbaceous, 2–3-pinnate-pinnatifid, broadest at the base, the apex non-conform, the leaf marginal cells differentiated into nodulose hyaline cells; pinna axes not articulate, sulcate adaxially, lacking a free central ridge; the rachis grooves U-shaped, continuous, the sulcus wall of the rachis continuing as a prominent ridge onto the sulcus wall of the costa, and then departing on the costule of the first basiscopic segment; veins free, terminating before the leaf margin, the vein endings not differentiated; sori dorsal along veins, not terminal, elongate, indusiate; soral receptacle flat; indusia lateral, non-glandular; sporangia with stalks two or three cells wide in the middle; spores monolete, non-chlorophyllous, brown, the perispore echinate, tuberculate, or with broad folds, the folds sometimes perforate; chromosome base number $x = 41$ (Kato, 1975a; Kato & al., 1992; Takamiya & al., 2000). The count of $x = 40$ (Kurita, 1960) is unsubstantiated.

Rhachidosoraceae can be distinguished by the combination of subterranean creeping rhizomes, leaves without abundant anthocyanins or mucilage at any stage, petioles with two vascular bundles, elongate sori restricted to one side of the vein, with indusia, and laminae provided with narrow filiform scales, and lacking hairs. This suite of characters, however, renders it difficult to distinguish from either Aspleniaceae or Athyriaceae. With Aspleniaceae it shares clathrate scales and elongate sori that are largely confined to one side of the vein. It does not, however, have the pinna-costa architecture characteristic of Aspleniaceae—a non-sulcate petiole where wings are formed by a decurrent lamina margin. (See the key provided below for additional technical characters distinguishing Rhachidosoraceae from Aspleniaceae.) More difficult is distinguishing Rhachidosoraceae from Athyriaceae; both families have similar pinna-costa architecture. This architecture is characterized by a sulcate rachis that is not alate, and that has a prominent flange on the basiscopic side of the pinna costa formed by the sulcus wall as it continues from the rachis onto the pinna costa itself. In addition, *Rhachidosorus* has minute corniculae and scales adaxially at the junction of the pinna and rachis, which are similar to those of *Athyrium* and *Cornopteris*. The most useful characters for distinguishing between these two families are the clathrate scales and linear sori confined to one side of the vein in Rhachidosoraceae; most Athyriaceae have sori on two sides of a single vein, either back-to-back, or in a hooked arrangement.

Biology and phylogeny. — Endemic to Asia, *Rhachidosorus* is a genus of approximately eight species of understory terrestrial ferns, which are often found in limestone habitats, and are very similar in gross morphology to species of *Athyrium*. Based on morphology, *Rhachidosorus* was previously included in either *Athyrium* (Makino, 1899) or *Diplazium* (Kato, 1975a), or considered a closely allied segregate (Ching, 1964a; reviewed in Sano & al., 2000a). Early molecular phylogenies (Sano & al., 2000a; Tzeng, 2002; Wang & al., 2003), however, unexpectedly suggested that *Rhachidosorus* was not closely related to either *Athyrium* or *Diplazium*. These results were later corroborated by the three-gene study of Kuo & al. (2011) who resolved *Rhachidosorus* as sister to the large clade of Thelypteridaceae + *Woodsia* + Athyriaceae + Blechnaceae + Onocleaceae, but with only weak support, and by Li & al. (2011) who placed the genus as sister to the clade recognized here as Diplaziopsidaceae. These studies each included a single *Rhachidosorus* accession (*R. mesosorus* (Makino) Ching in Sano & al., 2000a; *R. consimilis* Ching in Wang & al., 2003; *R. pulcher* (Tagawa) Ching in Tzeng, 2002 and Kuo & al., 2011), except for Li & al. (2011) who included both *R. consimilis* and *R. blotianus* Ching. The five-locus dataset of Rothfels & al. (2012) also included two species (*R. mesosorus*, *R. pulcher*), but as with the other studies was unable to strongly support the phylogenetic position of the genus; it was weakly placed as sister to Diplaziopsidaceae + Hemidictyaceae + Aspleniaceae.

Molecular data from four species (including *R. mesosorus*, the type of the genus) and six independent studies consistently support the surprising finding that *Rhachidosorus* is not phylogenetically close to *Athyrium* or *Diplazium*, but instead comprises an isolated lineage within Eupolypods II. The phylogeny of Rothfels & al. (2012) suggests that *Rhachidosorus* diverged from its nearest relatives approximately 90 million years ago, long before, for example, the ancestors of *Blechnum* diverged from those of *Athyrium* (Fig. 4) (Rothfels & al., 2012, their Supplementary Fig. 1).

■ DIPLAZIOPSIDACEAE X.C. Zhang & Christenh., *Phytotaxa* 19: 15 (2011)

Glade Ferns. Approximately four to six species of the genera *Diplaziopsis* C. Chr. (2–4 spp.), *Homalosorus* Pic. Serm. (1 sp.), plus *Diplazium flavoviride* Alston; (Ching, 1964b; Kato, 1975b; Kato & Darnaedi, 1988; Wei & al., 2010; Li & al., 2011).

<Dennstaedtiaceae: Athyrioideae sensu Holttum (1947); <Dryopteridaceae: Athyrioideae sensu Nayar (1970); <Athyriaceae sensu Tagawa & Iwatsuki (1972); <"*Diplazium javanicum* group" sensu Kato (1977); <Athyriaceae sensu Pichi Sermolli (1977); <Athyriaceae sensu Ching (1978a); <Dryopteridaceae: Athyrioideae sensu Lovis (1978); <Dryopteridaceae: Physematieae sensu Tryon & Tryon (1982); <Dryopteridaceae: Athyrioideae: Physematieae sensu Kramer & al. (1990b); <Athyriaceae: Diplazioideae sensu Wang & al. (2004); <Woodsiaceae sensu Smith & al. (2006); <Diplaziopsidaceae + Athyriaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial or epipetric; roots fleshy, pale, inserted radially, non-proliferous; rhizomes erect to suberect (*Diplaziopsis*, *Diplazium flavoviride*) or short-creeping (*Homalosorus*), commonly unbranched, bearing scales, and sometimes golden hairs similar to the root hairs (*Homalosorus*); rhizome scales lanceolate, non-clathrate, the margins entire, non-glandular, without distinct pubescence; leaves green and not covered in mucilage during any stage of development, spirally arranged, monomorphic, non-bulbiferous, closely spaced, glabrous (*Diplaziopsis*) or with filiform proscales (*Homalosorus*); petioles stramineous throughout or proximally darkened, thin, without a proximal thickening, conspicuous aerophores, or proximal articulation, sometimes with golden hairs similar to the root hairs (*Homalosorus*); petiolar vascular bundles two, each with hippocampiform xylem, the bundles distally uniting to form a single V-shaped bundle; laminae soft-herbaceous, 1-pinnate, the apex conform (*Diplaziopsis*) or non-conform (*Homalosorus*), the leaf marginal cells differentiated into nodulose hyaline cells; pinna axes not articulate, sulcate adaxially, lacking a free central ridge; the rachis grooves V-shaped, not continuous, the sulcus wall of the rachis continuing as a prominent ridge onto the sulcus wall of the costa; veins free (*D. flavoviride*, *Homalosorus*) or anastomosing toward the pinna margins (*Diplaziopsis*), the areoles without free included veinlets, usually terminating before the leaf margin, however some veins reaching the leaf margin in *D. flavoviride* and *Homalosorus*, the vein endings differentiated, slightly raised and expanded; sori singular along one side of the vein, rarely paired back to back along the same vein, elongate, indusiate, not terminal; soral receptacle flat; indusia lateral, vaulted or essentially flat, glabrous or glandular (*Diplaziopsis*), opening along the lateral margin or sometimes rupturing irregularly (*Diplaziopsis*); sporangia with stalks two or three cells wide in the middle; spores monolete, non-chlorophyllous, brown, the

perispore folded with thin crests, the crests erose; chromosome base numbers $x = 40$ (*Diplazium flavoviride*, *Homalosorus*; Löve & al., 1977; Kato & Darnaedi, 1988) or 41 (*Diplaziopsis*; Mitui, 1975; Takamiya & Ohta, 2001).

Diplaziopsidaceae can be recognized by the combination of petioles with two vascular bundles, 1-pinnate laminae, elongate sori that are usually along one side of the vein only, vein endings that are thickened and raised adaxially, and by the sulcus wall of the rachis forming a ridge that connects to the pinna sulcus wall of the pinna costa. The thickened and raised vein endings are a particularly useful diagnostic character among Eupolypods II, because they are otherwise only found in *Hemidictyum* (Hemidictyaceae). *Hemidictyum* also has a 1-pinnate lamina, a conform terminal segment, and veins that anastomose towards the pinna margins, similar to *Diplaziopsis*; however, it differs in several other respects, most conspicuously by having a sub-marginal collecting vein, and pinna margins that have a broad, pale membranaceous edge.

As Price (1990) noted, leaves of Diplaziopsidaceae are conspicuously soft, green, and fleshy. The pale fleshy roots appear to be unique among Eupolypods II, and anatomical study may provide synapomorphies for the family. However, the claim by Price (1990) that the plants entirely lack sclerenchyma is overstated; sclerenchyma occurs in the cortex and xylem, as evidenced by staining with toluidine blue (Sundue & Rothfels, unpub. data). Diplaziopsidaceae are most likely to be confused with Athyriaceae, particularly *Diplazium*, which is morphologically similar. Diplaziopsidaceae differs from most *Diplazium* species, however, by the usually singular linear sori, non-continuous groove of the adaxial pinnae-costa junctions, the near absence of indument on the lamina, the narrow petiole bases that do not form trophopods, thickened vein endings, and vaulted indusium, when it is present.

Biology and phylogeny. — Diplaziopsidaceae are medium-sized ferns of mesic understory habitats. They show an interesting pattern of disjunction, with the monotypic *Homalosorus* being a common member of rich temperate forests of eastern North America, while the *Diplaziopsis* species and *Diplazium flavoviride* are found in Asia, extending east to the Pacific islands (Kato & Darnaedi, 1988).

The history of typification of *Diplaziopsis* is convoluted. Christensen (1906: XXXII) published it as a replacement name for *Allantodia* Wall., 1830 (a later homonym of *Allantodia* R. Br., 1810). As a replacement name, therefore, *Diplaziopsis* takes the type of *Allantodia* Wall., which is *A. brunoniana* Wall. However, Christensen did not publish a combination for *A. brunoniana* under *Diplaziopsis* (he considered *A. brunoniana* to be a synonym of *Asplenium javanicum* Blume); the combination *Diplaziopsis brunoniana* (Wall.) W.M. Chu was made only recently (Chu & Zhou, 1994). Christensen (1906) listed *Asplenium javanicum* (= *Diplaziopsis javanica* (Blume) C. Chr.) as the type of *Diplaziopsis*, but this is prohibited under Art. 7.3 of the *Vienna Code*. However, if *Diplaziopsis brunoniana* is regarded as a heterotypic synonym of *D. javanica*, as Christensen (1906: CCXXVII) indicated, then *D. javanica* has priority and must be used as the name of the species, and the type of the genus. Alternatively, if *D. brunoniana* and *D. javanica* are recognized as distinct (as by Chu & He, 1999), then the type of the genus remains *D. brunoniana*.

Prior to the availability of molecular data, members of this clade were consistently thought to belong with the athyrioids, and both *Diplaziopsis* and *Homalosorus* were typically treated as members of *Diplazium* (Ching, 1964b; Kato, 1975b, 1977, 1993; Kato & Darnaedi, 1988; Wang & al., 2004). The first indication that this placement might be inaccurate came from the study of Sano & al. (2000a), in which the monotypic *Homalosorus* was strongly supported as sister to *Diplaziopsis cavaleriana* (Christ.) C. Chr., with these two taxa forming an isolated lineage distant from *Diplazium*. The next molecular phylogenetic study to include members of this clade was by Wei & al. (2010), and their results placed the two genera together in an unresolved position within the Eupolypods. Kuo & al. (2011), with more character data and denser taxon sampling, again resolved *Diplaziopsis* as sister to *Homalosorus*. Their results showed that this combined lineage—*Diplaziopsis* + *Homalosorus*—diverged from the rest of the Eupolypods II at an unsupported position deep along the eupolypod II backbone. The results of Li & al. (2011) were similar (*Diplaziopsis* sister to *Homalosorus*, and that clade distant from *Diplazium*), with the exception that their study placed the *Diplaziopsis* + *Homalosorus* clade as sister to *Rhachidosorus*.

In the analyses of Rothfels & al. (2012), *Diplaziopsis cavaleriana* and *D. javanica* are strongly supported as sister, and together they are sister to *Homalosorus*. These data allow either for the recognition of a monotypic *Homalosorus*, or its treatment within *Diplaziopsis*, as *D. pycnocarpa* (Spreng.) M.G. Price (Price, 1990). Diplaziopsidaceae diverged from Hemidictyaceae + Aspleniaceae early in the diversification of Eupolypods II—these two lineages shared a most recent common ancestor some 90 million years ago (Fig. 4) (Rothfels & al., 2012, their Supplementary Fig. 1)—further supporting the recognition of Diplaziopsidaceae rather than merging it into an expanded Aspleniaceae. *Diplazium flavoviride* has not been included in any phylogenetic analyses, but is included here based on the arguments of Kato & Darnaedi (1988). *Hemidictyum*, however, does not fall in Diplaziopsidaceae; its inclusion in that family by Christenhusz & al. (2011) rendered their concept of Diplaziopsidaceae paraphyletic, an error they subsequently corrected (Christenhusz & Schneider, 2011).

■ HEMIDICTYACEAE Christenh. & H. Schneid., *Phytotaxa* 28: 51 (2011)

Hemidictyum. One species of the genus *Hemidictyum* C. Presl.; (Kato, 1975b).

<Dennstaedtiaceae: Athyrioideae sensu Holttum (1947); <Dryopteridaceae: Athyrioideae sensu Nayar (1970); <"*Diplazium javanicum* group" sensu Kato (1977); <Athyriaceae sensu Pichi Sermolli (1977); <Thelypteridaceae sensu Lovis (1978); <Dryopteridaceae: Physematieae sensu Tryon & Tryon (1982); <Dryopteridaceae: Athyrioideae: Physematieae sensu Kramer & al. (1990b); <Woodsiaceae sensu Smith & al. (2006); <Diplaziopsidaceae sensu Christenhusz & al. (2011); =Hemidictyaceae sensu Christenhusz & Schneider (2011).

Characters. — Plants terrestrial; roots inserted radially, proliferous; rhizomes erect or suberect, commonly unbranched, bearing scales; rhizome scales lanceolate, weakly-clathrate, the margins entire, non-glandular, without distinct pubescence; leaves green and not covered in mucilage during any stage of development, spirally arranged, monomorphic, non-bulbiferous, closely spaced, glabrous; petioles stramineous throughout or proximally darkened, thin, not forming trophopods, lacking conspicuous aerophores, without a petiolar articulation; petiolar vascular bundles two, each with hippocampiform xylem, the bundles distally uniting to form a single U-shaped bundle; laminae herbaceous, 1-pinnate, the apex conform, the lateral pinnae sub-opposite, the pinna bases cordate, the leaf margin differentiated into a broad membranaceous edge; pinna axes not articulate, sulcate adaxially, lacking a free central ridge; the rachis grooves not continuous, the sulcus wall of the rachis not continuing as a ridge along the costa; veins anastomosing toward the pinna margins, the areoles without free included veinlets, terminating before the leaf margin and forming a sub-marginal collecting vein, the vein endings differentiated, slightly raised and expanded; sori usually singular along one side of the vein, occasionally back-to-back along both sides of the vein, elongate, indusiate, not terminal; soral receptacle flat; indusia lateral, essentially flat, glabrous; sporangia with stalks two or three cells wide in the middle; spores monolete, non-chlorophyllous, brown, the perispore with broad folds and tubercles, the folds sometimes perforate; chromosome base number $x = 31$ (Walker, 1973a; F.S. Wagner, 1980).

Although its conform apical pinnae, pattern of anastomosing veins, and thickened and raised vein endings are shared with *Diplaziopsis*, Hemidictyaceae can be distinguished from all other Eupolypods II by the combination of its sub-marginal collecting vein and pinna margin that is differentiated into a broad membranaceous border. Hemidictyaceae are sister to Aspleniaceae, but the two families together share no known synapomorphies. One character that should be investigated further is whether roots are proliferous, yielding new plants asexually. Walker (1985: 217) reported such roots in *H. marginatum* (L.) C. Presl.; they also occur in some species of *Asplenium* (Mickel & Smith, 2004), although sporadically enough that a synapomorphy for the two families is unlikely.

Biology and phylogeny. — *Hemidictyum* is a monotypic genus of the New World tropics—from southern Mexico to southeastern Brazil—where it grows at low to mid elevations in wet forests. The genus has always been an awkward fit in fern classifications, with opinions alternating for an alliance with thelypteroid ferns (based on spore morphology, e.g., Lovis, 1978), *Diplaziopsis* (based on its sagenoid venation, e.g., Kato, 1975b), or with Dryopteridaceae (in Eupolypods I; e.g., Tryon & Tryon, 1982).

Kato's (1975b) study was the first to emphasize commonalities between *Hemidictyum* and *Diplaziopsis*, and he argued that they might be isolated from much of *Diplazium* (Kato, 1975b). Molecular data (Schuettpelz & Pryer, 2007; Kuo & al., 2011; Rothfels & al., 2012) corroborated this morphology-based hypothesis, in part; *Hemidictyum* (like *Diplaziopsis*) is not closely related to *Diplazium* s.str.—*Hemidictyum* and *Diplaziopsis*+*Homalosorus* are more closely related to each other than to any eupolypod II lineage outside of Aspleniaceae.

Given its sister relationship with *Asplenium*+*Hymenasplenium*, *Hemidictyum* could be subsumed within an expanded concept of Aspleniaceae while retaining the monophyly of the latter family (Fig. 4). However, we favor recognizing Hemidictyaceae, even though it is monotypic, because the most recent common ancestor of *Hemidictyum* and Aspleniaceae dates to the Late Cretaceous (approximately 85 million years ago, Fig. 4) (Rothfels & al., 2012, their Supplementary Fig. 1), *Hemidictyum* would be morphologically anomalous within Aspleniaceae, and Aspleniaceae has a long history of taxonomic treatment excluding *Hemidictyum*.

■ ASPLENIACEAE Newman, *Hist. Brit. Ferns*: 6 (1840)

Spleenworts. Approximately 700 species of one to ten genera, dominated by the large genus *Asplenium* L. (incl. *Anti gramma* C. Presl., *Asplenidictyum* J. Sm. in Hook., *Biropteris* Kümmerle, *Camptosorus* Link, *Ceterach* Willd., *Diellia* Brack. in Wilkes, *Diplora* Baker, *Holodictyum* Maxon, *Loxoscaphe* T. Moore, *Neottopteris* J. Sm., *Phyllitis* Hill, *Pleurosorus* Fée, *Schaffneria* Fée, *Scolopendrium* Adans., and *Sinephropteris* Mickel) and with a small genus *Hymenasplenium* Hayata (incl. *Boniniella* Hayata); (Murakami, 1995; Murakami & al., 1999; Gastony & Johnson, 2001; Schneider & al., 2004a).

<Polypodiaceae: Asplenioideae sensu Christensen (1938); =Dennstaedtiaceae: Asplenioideae sensu Holttum (1947); =Aspleniaceae sensu Nayar (1970); =Aspleniaceae sensu Tagawa & Iwatsuki (1972); =Aspleniaceae sensu Pichi Sermolli (1977); =Aspleniaceae sensu Lovis (1978); =Aspleniaceae sensu Tryon & Tryon (1982); =Aspleniaceae sensu Kramer & Viane (1990); =Aspleniaceae sensu Smith & al. (2006); =Aspleniaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial, epipetric, or epiphytic, sometimes rheophytic; roots blackish, wiry, inserted radially or ventrally (*Hymenasplenium*), proliferous or non-proliferous; rhizomes usually odorless, rarely with the odor of wintergreen (e.g., *A. longissimum* Blume), short- to long-creeping, or suberect, branched or more commonly unbranched, sometimes massive and forming a detritus-collecting basket (e.g., *A. nidus* L.), bearing scales; rhizome scales lanceolate, clathrate, usually with blackish cell walls and hyaline lumens, sometimes brown or golden-brown, the margins glandular or not, entire to dentate or ciliate, without distinct pubescence; leaves green and not covered in mucilage during any stage of development, usually monomorphic, sometimes hemidimorphic, spirally arranged or distichous and dorsal (*Hymenasplenium*), occasionally bulbiferous, the bulbils frequently at the leaf apex, in a distal pinna axil, or at the base of the lamina, leaves usually closely spaced, sparsely to densely scaly, occasionally pubescent, rarely glandular (e.g., *A. platyneuron* (L.) Britton, Sterns & Poggenb.), also frequently with minute filiform proscales; petioles dull and greenish, gray, or

brownish, or lustrous and castaneous, atropurpureous, or ebenous, the bases expanded in *Hymenasplenium*, otherwise not usually expanded, persistent or not, articulate to the rhizome in *Hymenasplenium*, otherwise not; petiolar vascular bundles two, each with C-shaped xylem, the bundles distally uniting to form a single X-shaped bundle; laminae soft-herbaceous to coriaceous, simple to 4-pinnate, the apex usually pinnatifid or non-conform, occasionally conform (e.g., *Asplenium davisii* Stolze), the leaf marginal cells usually not differentiated; pinna axes not articulate, among species with divided leaves the axes usually alate, with wings derived from a decurrent and thickened leaf margin, or the wings thin, fragile, and apparently derived from the rachis; the rachis axes usually not sulcate adaxially, without a free central ridge; veins free or anastomosing, the areoles without free included veinlets, reaching the leaf margin or terminating before it, some species with a sub-marginal collecting vein, the vein endings forming hydathodes, or not differentiated; sori elongate, along one side of the vein, rarely paired back to back, and then usually not along the same vein, and if so then usually where groups of veins converge, indusiate; soral receptacle flat; indusia lateral, essentially flat, glabrous or sometimes glandular, opening along the lateral margin; sporangia with stalks one cell wide in the middle; spores monolete, non-chlorophyllous, brown, the perispore with sharp ridges or broad folds, sometimes echinulate, fenestrate, or reticulate; chromosome base numbers $x = 36$ (most species, e.g., Bir & Shukla, 1967; Walker, 1973a; Smith & Mickel, 1977), 38 (*Hymenasplenium*; Murakami, 1995), and 39 (*Hymenasplenium*; Kato & al., 1990; Murakami, 1995).

Aspleniaceae exhibit a broad spectrum of morphological diversity, yet identification of the family is usually not difficult. Diagnostic for Aspleniaceae are the linear sori with lateral indusia restricted to one side of the vein. The so-called "back-to-back" or "diplazioid" sori occur in some Aspleniaceae, however, they tend to be restricted to small portions of the lamina. As Holttum (1954) pointed out, patterns of major vein groups suggest that these instances are likely the result of lamina fusion or reduction. Some species of *Hymenasplenium* have been confused with *Diplazium* (Smith, 1976). However, numerous technical apomorphies of Aspleniaceae serve to distinguish these two genera. See the first lead in the key to families provided below for a list of characters serving to separate Aspleniaceae from other eupolypod II families.

Biology and phylogeny. — Aspleniaceae are a distinctive element within Eupolypods II; the family has usually been regarded as readily definable, in its current circumscription, even before Eupolypods II or Polypodiales (sensu Smith & al., 2006; Pryer & al., 2008) were recognized as cohesive entities (e.g., Nayar, 1970). Aspleniaceae are somewhat unusual considering their species-richness, in that they show strong patterns of diversification in both temperate and tropical areas (rather than being predominantly tropical), and have approximately equal numbers of epiphytic and terrestrial species (Schneider & al., 2004a). These two major habit types—epiphytic versus terrestrial—are both scattered across the Aspleniaceae phylogeny, although there is some evidence that the most recent common ancestor of the Aspleniaceae crown clade was epiphytic (Schneider & al., 2004a). Our circumscription is identical to that of Smith & al. (2006), who include further information on this family.

■ THELYPTERIDACEAE Ching ex Pic. Serm., *Webbia* 24: 709 (1970)

Thelypteroids; Marsh Ferns, Beech Ferns, and allies. Approximately 950 species, divided among *Cyclosorus* Link (incl. *Amblovenatum* J.P. Roux, *Ampelopteris* Kunze, *Amphineuron* Holttum nom. illeg., *Chingia* Holttum, *Christella* H. Lév., *Christella* sect. *Pelazoneuron* Holttum, *Cyclogramma* Tagawa, *Cyclosorus* s.str., *Glaphyropteridopsis* Ching, *Goniopteris* C. Presl., *Meniscium* Schreb., *Menisorus* Alston, *Mesophlebion* Holttum, *Mesopteris* Ching, *Plesioneuron* (Holttum) Holttum, *Pneumatopteris* Nakai, *Pronephrium* C. Presl., *Pseudocyclosorus* Ching, *Sphaerostephanos* J. Sm., *Stegnogramma* Blume, *Steiropteris* C. Chr., *Trigonospora* Holttum), *Macrothelypteris* (H. Ito) Ching, *Phegopteris* (C. Presl.) Fée, *Pseudophegopteris* Ching, and *Thelypteris* Schmidel (incl. *Amauropelta* Kunze, *Coryphopteris* Holttum, *Metathelypteris* (H. Ito) Ching, *Oreopteris* Holub, *Parathelypteris* (H. Ito) Ching, *Thelypteris* s.str., and *Wagneriopteris* Á. Löve & D. Löve); (Holttum, 1947, 1971c, 1981; Smith, 1990; Smith & Cranfill, 2002).

<Polypodiaceae: Dryopteridoideae sensu Christensen (1938); =Thelypteridaceae sensu Holttum (1947); =Thelypteridaceae sensu Nayar (1970); =Thelypteridaceae sensu Tagawa & Iwatsuki (1972); =Thelypteridaceae sensu Pichi Sermolli (1977); =Thelypteridaceae sensu Ching (1978a); <Thelypteridaceae sensu Lovis (1978); =Thelypteridaceae sensu Tryon & Tryon (1982); =Thelypteridaceae sensu Smith (1990); =Thelypteridaceae sensu Smith & al. (2006); =Thelypteridaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial, sometimes epipetric or rheophytic, rarely scandent (*Thelypteris* subg. *Amauropelta* sect. *Lepidoneuron* A.R. Sm.); roots blackish, wiry, inserted radially, non-proliferous; rhizomes not usually branched, short- to long-creeping, suberect, or erect, rarely sub-arborescent, bearing scales; rhizome scales lanceolate, non-clathrate, grayish to tan or brown, entire or dentate, the margins and often the surfaces usually bearing distinct pubescence similar to that of the leaves; leaves usually greenish in all stages, occasionally reddish when young (e.g., some *Cyclosorus* species treated in *Mesophlebion* and *Pronephrium*, possibly others), sometimes covered in mucilage when young, usually monomorphic, sometimes sub-dimorphic, spirally arranged, closely to distantly spaced, occasionally bulbiferous, the bulbils usually distal or apical on the leaf, scaly or not, almost always pubescent, the hairs whitish or hyaline, acicular, or sometimes forked, stellate, stalked-stellate, or hamate, also often provided with sessile or stalked glands; petioles greenish to stramineous, sometimes darker, the bases not articulate to the rhizome, not expanded at the base, and generally not persistent on the rhizome; petiole with two vascular bundles (rarely more), the bundles with hippocampiform-shaped xylem, distally uniting to form a single U-shaped bundle; laminae thin-herbaceous to coriaceous,

simple and entire to 3-pinnate-pinnatifid, in divided leaves, the base with or without a series of reduced pinnae, the apex conform or non-conform, the leaf marginal cells not clearly differentiated; pinna axes not articulate, the pinna base often with a conspicuous aerophore, these usually appearing as a low elongate or orbicular protuberance, or erect and vermiform, up to ca. 1 cm long; the rachis axes sulcate adaxially or not, when present the sulcae not continuous onto the next order, lacking a free central ridge; veins reaching the leaf margin or terminating before it, free, connivent at or below the sinus in lobed pinnae, or anastomosing in various patterns, the areoles without free included veinlets, the vein endings expanded or not differentiated; sori circular or elongate, on top of veins, not terminal, indusiate or exindusiate; soral receptacle flat; indusia lateral, reniform, sometimes pubescent and or glandular; sporangia with stalks more than one cell wide in the middle, often bearing hairs or glands (paraphysate); spores usually monolete, sometimes trilete (*Cyclosorus* treated as *Trigonospora*), non-chlorophyllous, the perispore brown, often with sharp crests, or reticulate or echinulate; chromosome base numbers $x = 27$ (e.g., *Parathelypteris*; Weng & Qiu, 1988), 29 (e.g., *Amauropelta*; Walker, 1985), 30 (e.g., *Phegopteris*; Mitui, 1975), 31 (e.g., *Lastrea*, *Macrothelypteris*, *Pseudophegopteris*, *Wagneriopteris*; Mitui, 1975; Loyal, 1991; Tindale & Roy, 2002), 34 (e.g., *Oreopteris*; Holttum, 1981; Manton, 1950), 35 (e.g., *Metathelypteris*, *Pseudocyclosorus*, *Thelypteris;* Mitui, 1975; Loyal, 1991; Walker, 1985), or 36 (e.g., *Abacopteris*, *Ampelopteris*, *Amphineuron*, *Christella*, *Cyclogramma*, *Cyclosorus*, *Dictyocline*, *Goniopteris*, *Lastrea*, *Leptogramma*, *Meniscium*, *Pronephrium*, *Stenogramma*; Mitui, 1975; Walker, 1985; Loyal, 1991; Tindale & Roy, 2002).

Thelypteridaceae are a large and diverse family, however, they can usually be recognized by the presence of distinctive acicular hairs. These hairs are whitish or hyaline, and usually 1-celled. In addition to being on the leaves, these hairs also regularly occur upon the margins and faces of the rhizome scales. While dentate or ciliate scales are common, as far as we know, no other family in Eupolypods II has rhizome scales that bear hairs similar to those found upon the leaves. Hamate, forked, and stellate hairs also occur in Thelypteridaceae, which being uncommon in Eupolypods, are also useful diagnostic characters. Thelypteridaceae also frequently have conspicuous aerophores at the bases of their pinnae. These often differ in color and texture from the surrounding tissue, and are frequently raised. In some cases, elongate vermiform aerophores are present; these frequently occur in species in which the crosiers and young leaves are surrounded by thick mucilage.

Biology and phylogeny. — This large family is morphologically cohesive and has been long recognized as such, in its current circumscription (but cf. Hennipman, 1996). Within Thelypteridaceae, however, generic classifications vary widely, and only two molecular phylogenetic studies have included a substantial representation of the family (approximately 30 species each; Smith & Cranfill, 2002; Schuettpelz & Pryer, 2007). Our circumscription is identical to that of Smith & al. (2006), who discuss this family in further detail.

■ WOODSIACEAE Herter, *Revista Sudamer. Bot.* 9: 14 (1949)

Woodsias; Cliff Ferns. Approximately 35 species of the genus *Woodsia* R. Br. (incl. *Cheilanthopsis* Hieron., *Hymenocystis* C.A. Mey., and *Protowoodsia* Ching); (Brown, 1964; Kurita, 1965; Shmakov, 2003).

<Polypodiaceae: Woodsieae: Woodsiinae sensu Diels in Engler & Prantl (1897); <Polypodiaceae: Gymnogrammeoideae+Polypodiaceae: Woodsioideae sensu Christensen (1938); <Sinopteridaceae+Woodsiaceae sensu Ching (1940); <Dennstaedtiaceae: Dryopteridoideae sensu Holttum (1947); =Woodsiaceae sensu Herter (1949); <Dryopteridaceae: Dryopteridoideae sensu Nayar (1970); <Athyriaceae sensu Tagawa & Iwatsuki (1972); =Woodsiaceae sensu Pichi Sermolli (1977); =Woodsiaceae sensu Ching (1978a); <Dryopteridaceae: Athyrioideae sensu Lovis (1978); <Dryopteridaceae: Physematieae sensu Tryon & Tryon (1982); <Dryopteridaceae: Athyrioideae: Physematieae sensu Kramer & al. (1990b); =Woodsiaceae sensu Wu & Ching (1991); <Woodsiaceae sensu Smith & al. (2006); =Woodsiaceae sensu Christenhusz & al. (2011).

Characters. — Plants epipetric, or occasionally terrestrial; roots blackish, wiry, inserted radially, non-proliferous; rhizomes short-creeping, horizontal to suberect, commonly unbranched, bearing scales; rhizome scales lanceolate, non-clathrate, the margins glandular or eglandular, without distinct pubescence, entire to denticulate or ciliate, the teeth when present formed by two adjacent cells, or not; leaves green and not covered in mucilage during any stage of development, usually spirally arranged, monomorphic, closely spaced, bearing scales and hairs, the hairs catenate or terete, sometimes gland-tipped (e.g., *W. mollis* (Kaulf.) J. Sm.), sometimes the scales forming a reduction series that terminates in broad-based, catenate, hair-like scales (e.g., *W. mollis*); petioles stramineous, castaneous or dark purple throughout, or proximally darkened, the base thin, not forming trophopods, persistent, usually forming a thick mantle of old petiole bases, without conspicuous aerophores, in some species with a petiolar articulation, the articulation usually proximal (e.g., *W. ilvensis* (L.) R. Br.) or just below the lamina; petiolar vascular bundles two, the bundles with hippocampiform-shaped xylem, distally uniting to form a single U-shaped bundle; laminae herbaceous, 1-pinnate to 2-pinnate-pinnatifid, usually broadest in the middle, the base with a series of reduced pinnae or not, the apex non-conform, the leaf marginal cells differentiated into nodulose hyaline cells or not; pinna axes not articulate, sessile or slightly petiolate; the rachis axes sulcate adaxially, lacking a free central ridge, the grooves not continuous; veins free, terminating before the leaf margin, the vein endings usually expanded and forming hydathodes; sori dorsal along veins, sub-terminal, or terminal (e.g., *W. elongata* Hook.), round, indusiate; soral receptacle distinctly flat; indusia basal, composed of a series of scale-like or filamentous segments or sometimes sac-like globose, glandular, pubescent, or not; sporangia with stalks two or three cells wide in the middle; spores monolete, non-chlorophyllous, tan or brown, the perispore echinate, tuberculate, or with broad folds or narrow crests, these sometimes forming a reticulum; chromosome base numbers $x = 33$ (*W. manchuriensis* Hook.; Kurita, 1965), 38, 39, or 41 (Manton, 1950; Brown, 1964).

When fertile, Woodsiaceae are easily diagnosed by the unique basal indusium composed of multiple scale-like or filamentous segments (occasionally as a single globose structure enclosing the sorus), which is unique among ferns. Many Cyatheaceae (in Cyatheales; see Fig. 1) have basal scale-like indusia, but these are generally more robust, are often spherical or cup-shaped, and do not consist of multiple segments. Some other taxa in Cyatheaceae have sori protected by scaly indument (e.g., *Sphaeropteris* subsect. *Fourniera* (J. Bommer ex Fourn.) P.G. Windisch); in these cases the scales resemble those of the lamina whereas in *Woodsia* they do not. Sterile leaves of Woodsiaceae, however, are not as easily characterized. When present, the petiolar articulation is a powerful diagnostic character, because it is unique in Eupolypods II, and rare outside of this clade. However, its utility is hindered by its absence from most species. Nonetheless, all species tend to accumulate large mats of persistent petiole bases, which are characteristic. In addition, the combination of 1-pinnate to 2-pinnate-pinnatifid leaves with sessile or short-petioled pinnae, laminae usually bearing scales and hairs (that are not acicular), and veins that terminate before the leaf margin in hydathodes serve to diagnose Woodsiaceae. Sterile plants of *Cystopteris* can appear surprisingly similar to those of *Woodsia*, however, they can be distinguished by having veins that reach the leaf margin.

Biology and phylogeny. — As circumscribed here, species of Woodsiaceae typically occur in rocky, montane areas, predominantly in the Northern Hemisphere. Areas of particular species-richness include the mountains of Eurasia, and arid areas of Mexico and southwestern U.S.A.; one polymorphic species (*W. montevidensis* (Spreng.) Hieron.) extends south through the Andes and also occurs in South Africa. They have remarkable ecological and morphological resemblance to members of *Cystopteris* (in Cystopteridaceae), to which they are only distantly related (Figs. 2, 4).

This family—comprising *Woodsia* and its segregates—is an isolated lineage, not closely related to the other taxa frequently included in broad concepts of Woodsiaceae (e.g., Smith & al., 2006). *Protowoodsia* and *Cheilanthopsis* are nested within *Woodsia* s.l. (Rothfels & al., 2012), as is *Hymenocystis*, the other segregate genus recognized by Shmakov (2003; A. Larsson, unpub.). The molecular phylogeny of *Woodsia* is marked by a remarkably deep split between a clade of Old World or holarctic species, and a clade of predominantly New World species. This deep dichotomy essentially mirrors the results of Brown's (1964) groundplan divergence scheme (W.H. Wagner, 1980) based upon morphological characters.

■ ATHYRIACEAE Alston, *Taxon* 5: 25 (1956)

Athyrioids; Ladyferns, and allies. Approximately 600 species, in *Anisocampium* C. Presl. (4 spp.; incl. *Kuniwatsukia* Pic. Serm.), *Athyrium* Roth (~220 spp.; incl. *Pseudocystopteris* Ching), *Cornopteris* Nakai (9 spp.; incl. *Neoathyrium* Ching & Z.R. Wang), *Deparia* Hook. & Grev. (~70 spp.; incl. *Athyriopsis* Ching, *Lunathyrium* Koidz., *Dictyodroma* Ching, *Dryoathyrium* Ching, and *Triblemma* R. Br. ex C. Sprengel), and *Diplazium* Sw. (~300–400 spp.; incl. *Allantodia* R. Br., *Anisogonium* C. Presl., *Callipteris* Bory, *Monomelangium* Hayata; excl. *Diplazium flavoviride* Alston); (Kato, 1979, 1984; Tryon & Tryon, 1982; Pacheco & Moran, 1999; Sano & al., 2000b; Adjie & al., 2008; Liu, 2008; Liu & al., 2011).

<Polypodiaceae: Asplenioideae sensu Christensen (1938); <Dennstaedtiaceae: Athyrioideae sensu Holttum (1947); <Athyriaceae sensu Alston (1956); <Dryopteridaceae: Athyrioideae sensu Nayar (1970); <Athyriaceae sensu Tagawa & Iwatsuki (1972); <Athyriaceae sensu Ching (1978a); <Dryopteridaceae: Athyrioideae sensu Lovis (1978); <Dryopteridaceae: Physematieae sensu Tryon & Tryon (1982); <Dryopteridaceae: Athyrioideae: Physematieae sensu Kramer & al. (1990b); <Athyriaceae: Athyrioideae + Deparioideae + Diplazioideae sensu Wang & al. (2004); <Woodsiaceae sensu Smith & al. (2006); <Athyriaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial or epipetric, sometimes rheophytic; roots blackish, wiry, inserted radially, non-proliferous; rhizomes short- to long-creeping, or suberect to erect, branched or more commonly unbranched, bearing scales, and sometimes golden hairs similar to the root hairs (e.g., *Athyrium skinneri* (Baker) Diels); rhizome scales lanceolate, not or only weakly clathrate, the margins usually non-glandular, sometimes glandular (some *Deparia*), without distinct pubescence, entire or dentate, when present the teeth commonly formed by two adjacent cells (*Diplazium*); leaves sometimes internally mucilaginous (some *Deparia* and *Diplazium* spp., particularly those treated as *Callipteris*), not externally covered in mucilage during any stage of development, green in *Diplazium* and *Deparia*, the petiole and rachis frequently with a pink hue in *Athyrium* presumably due to anthocyanins, monomorphic, spirally arranged or sometimes distichous and dorsal (e.g., *Athyrium skinneri*), occasionally bulbiferous, closely to distantly spaced, sparsely to moderately scaly and occasionally pubescent (*Athyrium*, *Diplazium*) or with a reduction series beginning with scales at the base of the leaf that gradually reduce to catenate hairs distally (*Deparia*, some *Diplazium*); petioles castaneous, stramineous, or proximally darkened, often with a proximal thickening forming trophopods that may be starch-filled (*Athyrium*, *Cornopteris*, *Diplazium*), often with conspicuously elaborated aerophores (elsewhere termed pneumatophores; *Cornopteris*, some *Athyrium*, some *Deparia*), the bases usually persistent, rarely articulate to the rhizome (e.g., *Anisocampium* and *Athyrium skinneri*), and sometimes with golden hairs similar to the root hairs (e.g., *A. skinneri*); petiolar vascular bundles two (rarely more; Kato, 1972), each with hippocampiform xylem, the bundles distally uniting to form a single U- or V-shaped bundle; laminae soft-herbaceous to coriaceous, simple to 3-pinnate-pinnatifid, the apex usually pinnatifid or non-conform, sometimes conform in *Diplazium*, the leaf marginal cells differentiated into nodulose hyaline cells or not; pinna axes not articulate, sometimes muricate (*Diplazium*), sulcate adaxially, lacking a free central ridge; the rachis grooves V-shaped (*Anisocampium*, *Athyrium*, *Cornopteris*) or U-shaped (*Deparia*, *Diplazium*), continuous (*Anisocampium*, *Athyrium*, *Cornopteris*, most *Diplazium*) or not continuous (*Deparia*, some *Diplazium*), the sulcus wall of the rachis usually continuing as a prominent ridge onto the sulcus

wall of the costa (but not in *Deparia*); veins free or sometimes anastomosing (*Deparia* species treated as *Dictyodroma*, and some *Diplazium*), the areoles without free included veinlets, usually terminating before the leaf margin, the vein endings slightly raised and expanded, or forming hydathodes, or not differentiated; sori usually elongate, sometimes round, not terminal, on top of the vein, or along one side, singular or paired back-to-back along the same vein, or hooked in most *Athyrium* (i.e., paired back-to-back and crossing over the vein at one end in J- or U-shapes), or rarely sori marginal (e.g., *Deparia prolifera* (Kaulf.) Hook. & Grev.), at the tips of vein endings, usually indusiate; soral receptacle flat; indusia lateral, vaulted or essentially flat, glabrous or glandular, opening along the lateral margin; sporangia with stalks two or three cells wide in the middle; spores monolete, non-chlorophyllous, brown, the perispore nearly plain to coarsely tuberculate, echinate, or folded, the folds short and low, forming a rugate surface, or broad and wing-like; chromosome base numbers $x = 40$ (*Athyrium*, *Deparia*, some *Cornopteris*; Manton & Sledge, 1954; Kato, 1979; Sano & al., 2000a, b; Liu & al., 2011), or 41 (*Diplazium*, some *Cornopteris*; Manton & Sledge, 1954; Kato, 1979; Dawson & al., 2000). Reports of $x = 41$ for *Deparia* require confirmation (reviewed in Sano & al., 2000a), and reports for individual *Cornopteris* species are occasionally inconsistent (alternating between $x = 40$ and $x = 41$), indicating that further cytological study is needed.

Several genera of Athyriaceae can be diagnosed by unique or rare character states. However, character state reversals and homoplasy render these characters imperfect diagnostics for the family. *Deparia* typically has broad scales present at the base of the leaf that transition along a homologous series to septate hairs. Similar hairs occur in *Diplazium* (species treated in *Callipteris* by Pacheco & Moran, 1999), *Acystopteris*, and some species of *Woodsia* (e.g., *W. mollis*)—other species of *Woodsia* have septate hairs, but these are not reduced from broad scales. Many eupolypod II ferns have reduced filiform scales, but in most cases these never approach the septate hairs found in *Deparia* (see Fig. 3 in Sano & al., 2000b). *Deparia* also differs from most Eupolypods II in having sulcate rachises that are not continuous with the sulcae of the pinna costae. Many *Athyrium* and *Cornopteris* species have red-tinged leaves. This color is visible in live plants as well as on herbarium specimens. Blechnaceae also have reddish leaves and this has been attributed to the presence of anthocyanins (Crowden & Jarman, 1974). That family, and the other eupolypod II families with reddish leaves, differ in that the red coloration is present only in developing leaves and is not visible by maturity.

Another useful character of limited distribution is the corniculae/scales that are present adaxially at the junction of the pinna costa and the rachis in many *Athyrium* and *Cornopteris*. In addition, many *Athyrium* and some *Diplazium* species have small epidermal spinules along the adaxial pinna costae. Outside of the Athyriaceae, adaxial corniculae occur only in Rhachidosoraceae and *Onocleopsis* (Onocleaceae). In the Eupolypods I, similar structures also occur in *Didymochlaena*, and outside of the Eupolypods similar structures occur in *Pteris* (Pteridaceae). Athyriaceae also frequently have well-developed trophopods, which consist of a thickened petiole base that is often starch-filled, persistent upon the rhizome, and in some cases highly sclerified. The trophopods of some *Athyrium* and *Deparia* are additionally adorned with toothed or wing-like protuberances, referred to as pneumatophores by Iwatsuki (1970) and Kato (1984).

Biology and phylogeny. — Athyriaceae are mostly medium-sized understory terrestrial ferns, comprising three major clades that correspond to the subfamilies Athyrioideae, Diplazioideae, and Deparioideae of Wang & al. (2004; with the exception of *Diplaziopsis* and *Homalosorus*, which Wang & al. (2004) include in Diplazioideae, and which we place in Diplaziopsidaceae). This alliance of "athyriid", "diplaziid", and "depariid" ferns (Rothfels & al., 2012) has a long history; at some point they have all been treated in a broad concept of *Athyrium* (e.g., Copeland, 1947). The sister-group relationship of the athyriids and diplaziids, and they together as sister to the depariids, was anticipated first by Hiraoka (1978), suggested by the single-locus data of Sano & al. (2000a), and strongly supported by multi-locus molecular data (Schuettpelz & Pryer, 2007; Rothfels & al., 2012). Character evolution in Athyriaceae is complex, and the generic-level relationships within the two large clades (athyriids, diplaziids) need further investigation (e.g., Liu & al., 2011).

■ BLECHNACEAE Newman, *Hist. Brit. Ferns*, ed. 2: 8 (1844)

Blechnoids; Deer Ferns, Chain Ferns, and allies. Approximately 200 species in *Blechnum* L. s.l. (~150 spp.; incl. many potential segregates), *Brainea* J. Sm. (1 sp.), *Diploblechnum* Hayata (2 spp.), *Doodia* R. Br. (~15 spp.), *Pteridoblechnum* Hennipman (2 spp.), *Sadleria* Kaulf. (6 spp.), *Salpichlaena* J. Sm. (3 spp.), *Steenisioblechnum* Hennipman (1 sp.), *Stenochlaena* J. Sm. (8 spp.), and *Woodwardia* Sm. (14 spp.; incl. *Anchistea* C. Presl., *Lorinseria* C. Presl.); (Holttum, 1971b; Moran, 1990; Cranfill & Kato, 2003).

<Polypodiaceae: Pteridoideae + Polypodiaceae: Blechnoideae sensu Christensen (1938); = Dennstaedtiaceae: Blechnoideae sensu Holttum (1947); = Blechnaceae sensu Nayar (1970); <Blechnaceae + Pteridaceae sensu Tagawa & Iwatsuki (1972); = Blechnaceae sensu Pichi Sermolli (1977); = Blechnaceae + Stenochlaenaceae sensu Ching (1978a); = Blechnaceae sensu Lovis (1978); = Blechnaceae sensu Tryon & Tryon (1982); = Blechnaceae sensu Kramer & al. (1990a); = Blechnaceae sensu Smith & al. (2006); = Blechnaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial or climbing (by means of rhizomes in *Stenochlaena* and *Blechnum* sect. *Lomaria* (Willd.) Diels, or by leaves in *Salpichlaena*), rarely epiphytic or rheophytic; roots blackish, wiry, inserted radially, non-proliferous; rhizomes short- to long-creeping, suberect, or erect, sometimes massive and arborescent (*Blechnum* sect. *Lomariocycas* (J. Sm.) C.V. Morton, and *Sadleria*), branched, or more commonly unbranched, sometimes stoloniferous (*Blechnum* sect. *Eublechnum* Hook. & Baker), bearing scales; rhizome scales lanceolate to linear-lanceolate, non-clathrate, light-brown to blackish, the margins glandular or not, entire or dentate,

without distinct pubescence; leaves reddish when young, green at maturity, sometimes covered in mucilage when young (some *Blechnum*), monomorphic or dimorphic, spirally arranged, occasionally bulbiferous, the bulbils frequently in a distal pinna axil, leaves usually closely spaced, sparsely to densely scaly, sometimes pubescent, sometimes with glandular nectaries (e.g., *Stenochlaena palustris* (Burm. f.) Bedd., *Blechnum orientale* L.); petioles greenish to dark brown or atropurpureous, the bases not expanded, not articulate to the rhizome, usually not persistent; petiolar vasculature with two large bundles on the adaxial side of the petiole and an arc of smaller bundles on the abaxial side of the petiole, rarely petioles with only two bundles (e.g., *Woodwardia areolata* (L.) T. Moore), the larger bundles with hippocampiform-shaped xylem, distally uniting to form a single U-shaped bundle; laminae soft-herbaceous to more often coriaceous, pinnatifid to 2-pinnate-pinnatifid, the base with or without a series of reduced pinnae, the apex conform or not, the leaf marginal cells differentiated and scarious or membranaceous, or non-differentiated; pinna axes articulate (*Stenochlaena*), or usually non-articulate, pinna bases sometimes with conspicuous aerophores in *Blechnum*, the aerophores appearing as a low protuberance or elongate and vermiform (e.g., *Blechnum violaceum* (Fée) C. Chr.); rachis axes sulcate adaxially, rarely not (e.g., some *Woodwardia*), the sulcae not continuous onto the next order, lacking a free central ridge; veins anastomosing, or more commonly with costular areoles and otherwise free, the areoles without free included veinlets, reaching the leaf margin or terminating before it, the vein endings forming hydathodes, or not differentiated; sori elongate, along one side of the costular commissural vein, indusiate, or sori acrostichoid and exindusiate (*Stenochlaena*); soral receptacle flat; indusia lateral, essentially flat, glabrous or sometimes glandular, opening along the lateral margin with the opening facing the costa; sporangia with stalks more than one cell wide in the middle; spores monolete, occasionally chlorophyllous (e.g., *Blechnum nudum* (Labill.) Leurss.; Sundue & al., 2011), usually non-chlorophyllous, usually pale brown or tan, the perispore with sharp ridges, broad folds, echinulate, tuberculate, foliose, or nearly plain; chromosome base numbers $x = 27$ (*Pteridoblechnum*; Tindale & Roy, 2002), 29 (*Blechnum*; Walker, 1973a), 31 (*Blechnum*; Walker, 1973a), 32 (*Doodia*; Tindale & Roy, 2002), 33 (*Blechnum*, *Sadleria*; Walker, 1973a; Smith & Mickel, 1977; F.S. Wagner, 1995), 34 (*Blechnum*, *Woodwardia*; Manton & Sledge, 1954; Tryon & Tryon, 1982), 35 (*Brainea*, *Woodwardia*; Britton, 1964; W.H. Wagner, 1955; Aziz Bidin, 1995), 36 (*Blechnum*; Walker, 1973a), 37 (*Stenochlaena*; Manickam & Irudayaraj, 1988), or 40 (*Salpichlaena*; Walker, 1973b).

Blechnaceae are unique among ferns in having elongate sori along a sub-costular commissural vein that is parallel to the pinna costa, with an indusiate sorus that opens to face the costa. Other ferns with elongate sori lack a commissural vein, and have indusia that face the costa at a low angle (not parallel) or are exindusiate. Blechnaceae also differ from nearly all other Eupolypods II by having petioles with a vascular anatomy that resembles those of Eupolypods I. That is, in addition to the two large bundles on the adaxial side of the petiole, there is an arc of smaller bundles along the abaxial side of the petiole. However, *Woodwardia areolata* is aberrant among Blechnaceae in having only two.

Some genera of Onocleaceae, such as *Matteuccia* and *Pentarhizidium*, have a strong superficial resemblance to Blechnaceae. Those genera can be differentiated by fertile leaves with a modified leaf margin that opens to face the costa and the indusium itself, which is inconspicuous and faces away from the costa. *Plagiogyria* (Cyatheales; see Fig. 1) also appears similar; the shape of the lamina, dimorphic leaves, and young leaves covered in mucilage are all reminiscent of *Blechnum* sect. *Parablechnum* (C. Presl.) T. Moore. Other characters of *Plagiogyria*, however, support its position among the Cyatheales, namely the sporangial capsule with an oblique annulus, trilete spores, and a perispore with well-formed rodlets.

Biology and phylogeny. — Blechnaceae are cosmopolitan, with a wide range of growth habits, including tall arborescent species, near-annual roadside weeds, and tropical lianas. Like their sister group, Onocleaceae, Blechnaceae have a high frequency of fertile/sterile leaf dimorphism. Many of the smaller Blechnaceae genera nest within *Blechnum* s.l. (Nakahira, 2000; Cranfill, 2001; Schuettpelz & Pryer, 2007). While generic circumscription remains incomplete, the family is well defined and historically stable (with the exception of *Stenochlaena*, which was of uncertain affinity prior to the availablility of molecular data, a challenging situation further complicated by Christensen having included three widely divergent taxa in his original description of the genus; Christensen, 1906; Holttum, 1971b). Our circumscription is identical to that of Smith & al. (2006), who provide further information on this family.

■ ONOCLEACEAE Pic. Serm., *Webbia* 24: 708 (1970)

Onocleoids; Sensitive Fern, Ostrich Fern, and allies. Five species in *Matteuccia* Tod. (1 sp.), *Onoclea* L. (1 sp. with two varieties), *Onocleopsis* F. Ballard (1 sp.), and *Pentarhizidium* Hayata (2 spp.); (Gastony & Ungerer, 1997; Rothfels & al., 2012).

=Polypodiaceae: Onocleoideae sensu Christensen (1938); ="unplaced": Onocleoideae sensu Holttum (1947); =Dryopteridaceae: Onocleoideae sensu Nayar (1970); =Onocleaceae sensu Pichi Sermolli (1977); =Onocleaceae sensu Ching (1978a); =Dryopteridaceae: Onocleoideae sensu Lovis (1978); =Dryopteridaceae: Onocleeae sensu Tryon & Tryon (1982); =Dryopteridaceae: Athyrioideae: Onocleeae sensu Kramer & al. (1990b); =Dryopteridaceae: Onocleeae sensu Gastony & Ungerer (1997); =Onocleaceae sensu Smith & al. (2006); =Onocleaceae sensu Christenhusz & al. (2011).

Characters. — Plants terrestrial, often in wet or seasonally wet habitats; roots blackish, wiry, inserted radially, non-proliferous; rhizomes short-creeping, unbranched, and erect (up to 1 m tall in *Onocleopsis*), or long-creeping, and branched (*Onoclea*), rhizomes sometimes stoloniferous (*Matteuccia*), bearing scales; rhizome scales lanceolate, non-clathrate, brown, the margins eglandular, entire or dentate, without distinct pubescence; leaves greenish and not covered in mucilage during

any stage of development, dimorphic, spirally arranged, leaves usually closely spaced (sometimes distantly spaced in *Onoclea*), sparsely to densely scaly, sometimes pubescent; petioles greenish to stramineous, the bases not articulate to the rhizome, expanded and often starch-filled (forming trophopods), persistent on the rhizome, sometimes for decades, forming a massive protective sheath in *Matteuccia* and *Pentarhizidium*; petioles with two vascular bundles, the bundles with hippocampiform-shaped xylem, distally uniting to form a single U-shaped bundle; laminae herbaceous, pinnatifid to 1-pinnate-pinnatifid, the base with or without a series of reduced pinnae, the apex pinnatifid or non-conform, the leaf marginal cells scarious or not differentiated (*Matteuccia*); pinna axes not articulate; the rachis axes sulcate adaxially, the sulcae not continuous onto the next order, lacking a free central ridge; veins mostly reaching the leaf margin, or terminating before it in *Pentarhizidium*, free or anastomosing (*Onoclea*, *Onocleopsis*), the areoles without free included veinlets, the vein endings expanded in *Pentarhizidium*, otherwise not differentiated; sori orbicular, terminal on the vein, indusiate (except *P. intermedium* (C. Chr.) Hayata); soral receptacle raised, conical; indusia lateral, triangular, ephemeral; sporangia with stalks more than one cell wide in the middle; spores monolete, chlorophyllous, the perispore brown, perispore with broad folds and echinulae; chromosome base numbers $x = 37$ (*Onoclea*; Haufler & Soltis, 1986), 39 (*Matteucia*; Kurita, 1960), or 40 (*Onocleopsis*, *Pentarhizidium*; Tsai & Shieh, 1985; Gastony & Ungerer, 1997).

Onocleaceae can be diagnosed by having dimorphic leaves, petioles with two vascular bundles, and thickened petiole bases, chlorophyllous spores, and sori with conical receptacles. Blechnaceae appear similar, but differ by having petioles with more than two vascular bundles (except *Woodwardia areolata*, which has two) and that are not expanded at the base, leaves that are reddish when young, and indusia that open to face the costa.

Biology and phylogeny. — Onocleaceae are a small family, yet one of the most familiar to residents of the north-temperate zone. The family is noteworthy for the strong fertile/sterile leaf dimorphism of its members, their typically large size, chlorophyllous spores, and unusual distributions: *Matteuccia* is circumboreal; *Onoclea* is disjunct between eastern North America and eastern Asia; *Onocleopsis* is endemic to southern Mexico and Guatemala; and *Pentarhizidium* is limited to eastern Asia. Our circumscription is identical to that of Smith & al. (2006), who provide further information on this family.

■ KEY TO EUPOLYPOD II FAMILIES

1. Sori elongate, usually on one side of the vein, rarely paired back-to-back on a single vein, never curving over to the other side of the vein and forming a U- or J-shape**;** petioles with two vascular bundles, these united distally to form an X-shape as seen in cross-section, vascular bundles with xylem in the shape of a "C" as seen in cross-section; rhizome scales clathrate, rarely with darkened indurate lumens; sporangial stalks one cell wide in the middle … …… **Aspleniaceae**

1. Sori elongate or round, on top of the vein, on one side, paired back-to-back, or on one side of the vein and curving over to other side and forming a U- or J-shape; petioles with more than two vascular bundles, or if two, then these distally united to form a U- or V-shape as seen in cross-section, largest vascular bundles with hippocampiform-shaped xylem; rhizome scales non-clathrate (except Rhachidosoraceae, some Cystopteridaceae); sporangial stalks more than one cell wide in the middle, usually three cells wide…… 2
2. Petiole with more than two vascular bundles (two in some *Woodwardia*); sori elongate, parallel to the costa, on a subcostular commissural vein connecting lateral veins, indusiate, with the opening facing the costa, or sori acrostichoid (*Stenochlaena*); leaves reddish when young, not reddish at maturity …… **Blechnaceae**
2. Petiole with two vascular bundles (rarely more in Athyriaceae and Thelypteridaceae); sori on, along, or at the apex of a lateral vein, round or elongate, never acrostichoid, if elongate then usually at an angle to the costa, when parallel to the costa, the indusium opening to face the segment margin (away from the costa) or exindusiate; leaves green in all stages, or if reddish when young, then reddish at maturity as well (except some Thelypteridaceae and *Onoclea sensibilis* L., which are reddish only when young) …… 3
3. Fertile leaves strongly dimorphic with sori protected by contracted and inrolled segment margins; spores chlorophyllous…… **Onocleaceae**
3. Fertile leaves holomorphic or partially dimorphic (some Thelypteridaceae), the segments weakly contracted and not inrolled; spores not chlorophyllous…… 4
4. Leaves pubescent (rarely lacking hairs), the hairs acicular, forked, stellate, or hamate; rhizome scales often bearing similar hairs along the margin and surfaces; indusia, when present, reniform and attached laterally; pinna base usually with a prominent aerophore, the aerophore raised, orbicular, elongate, or vermiform; leaves sometimes mucilaginous when young …… **Thelypteridaceae**
4. Leaves glabrous or pubescent, the hairs not acicular, simple, never forked or hamate; rhizome scales ciliate or denticulate, but not bearing hairs similar to the leaves; indusia, when present, attached basally or laterally, if laterally then elongate or a minute scale, not reniform; pinna base without a prominent aerophore; leaves never mucilaginous…… 5
5. Indusium attached basally, encircling the sorus, globose, or composed of multiple scale-like or filamentous segments ……**Woodsiaceae**
5. Indusium attached laterally, or if attached basally then not encircling the indusium, and composed of a single scale-like segment, or exindusiate …… 6
6. Sori round, indusiate, and the receptacle slightly raised and hardened, or sori slightly elongate (not more than 2× longer than wide), exindusiate, the soral receptacle flat; veins reaching segment margin; indusium, when present, basal, a minute hood-like scale, arching over the sorus, frequently deciduous …… **Cystopteridaceae**
6. Sori usually elongate, several times longer than wide, sometimes round (some Athyriaceae), indusiate, the soral

receptacle flat; veins usually ending before segment margin; indusium lateral, vaulted or essentially flat, opening along the lateral margin, usually persistent 7

7. Rhizome scales clathrate; vein endings undifferentiated, neither expanded, raised nor forming hydathodes..... ..... **Rhachidosoraceae**
7. Rhizome scales non-clathrate; vein endings differentiated, either thickened, raised, or forming hydathodes..... 8
8. Veins forming a sub-marginal collecting vein; leaf margin with a broad membranaceous border; pinna bases subcordate, the basiscopic lobes overlapping the rachis..... ..... **Hemidictyaceae**
8. Veins free or anastomosing, but not forming a sub-marginal collecting vein; leaf margin scarious or undifferentiated, but not with a broad membranaceous border; pinna bases truncate, cuneate, or excavate, but not subcordate, and not overlapping the rachis..... 9
9. Sori usually along one side of the vein, rarely paired back-to-back; roots pale, fleshy; sori vaulted, the indusium often splitting apically prior to opening laterally; veins raised and cartilaginous on the adaxial side of the lamina..... ..... **Diplaziopsidaceae**
9. Sori usually along both sides of the vein, either paired back-to-back, or crossing over the vein and U- or J-shaped (on top of the vein in *Cornopteris* and some *Athyrium*); roots blackish, wiry; sori usually flat, sometimes vaulted, indusium never splitting apically prior to opening laterally; veins often expanded, but not raised or cartilaginous on the adaxial side of the lamina..... **Athyriaceae**

■ ACKNOWLEDGMENTS

We thank He Hai (College of Life Sciences, Chongqing Normal University) for providing photos of *Cystoathyrium*, and Art Gilman, Amanda Grusz, Layne Huiet, Anne Johnson, Fay-Wei Li, Yea-Chen Liu, Erin Sigel, Alan Smith, Michael Windham, editor Libing Zhang, and two anonymous reviewers for helpful comments on earlier drafts. This work was supported by funding from a National Science and Engineering Research Council (Canada) PGSD to CJR, a Duke University Graduate School Semester Fellowship to CJR, the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning ("Formas"; to AL as part of grant 2006-429 to Petra Korall), an NSF DDIG award to KMP and CJR (DEB-1110767), NSF DEB-1119695 to MAS, and an NSF CAREER award to KMP (DEB-0347840). The Alice & Rolla Tryon Pteridophyte Library at the University of Vermont provided important references.

■ LITERATURE CITED

Adjie, B., Takamiya, M., Ohto, M., Ohsawa, T.A. & Watano, Y. 2008. Molecular phylogeny of the lady fern genus *Athyrium* in Japan based on chloroplast *rbcL* and *trnL-trnF* sequences. *Acta Phytotax. Geobot.* 59: 79–95.

Alston, A. 1956. The subdivision of the Polypodiaceae. *Taxon* 5: 23–25.

Aziz Bidin, A. 1995. The ecology and cytology of *Brainea insignis* (Blechnaceae: Pteridophyta). *Fern Gaz.* 15: 21–24.

Bir, S.S. & Shukla, P. 1967. Cytology of some north Indian ferns. *Cytologia* 32: 24–30.

Blasdell, R.F. 1963. A monographic study of the fern genus *Cystopteris*. *Mem. Torrey Bot. Club* 21: 1–102.

Britton, D.M. 1964. Chromosome numbers of ferns in Ontario. *Canad. J. Bot.* 64: 1149–1154.

Brown, D.F.M. 1964. A monographic study of the fern genus *Woodsia Beih. Nova Hedwigia* 16: 1–156.

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.

Ching, R.-C. 1966. Three new fern genera. *Acta Phytotax. Sin.* 11: 11–30.

Ching, R.-C. 1978a. The Chinese fern families and genera: Systematic arrangement and historical origin. *Acta Phytotax. Sin.* 16: 1–19.

Ching, R.-C. 1978b. The Chinese fern families and genera: Systematic arrangement and historical origin (Cont.). *Acta Phytotax. Sin.* 16: 16–37.

Christenhusz, M.J.M. & Schneider, H. 2011. Corrections to Phytotaxa 19: Linear sequence of lycophytes and ferns. *Phytotaxa* 28: 50–52.

Christenhusz, M.J.M., Zhang, X.-C. & Schneider, H. 2011. A linear sequence of extant families and genera of lycophytes and ferns. *Phytotaxa* 19: 7–54.

Christensen, C. 1906. *Index filicum.* Copenhagen: Hagerup.

Christensen, C. 1938. Filicinae. Pp. 522–550 in: Verdoorn, F. (ed.), *Manual of pteridology.* The Hague: Nijhoff.

Chu, W.-M. & He, Z.-R. 1999. *Diplaziopsis* C. Chr. Pp. 499–509 in: Chu, W.-M. (ed.), *Flora Reipublicae Popularis Sinicae*, vol. 3(2). Beijing: Science Press.

Chu, W.-M. & Zhou, H.-G. 1994. Some pteridophytes new to Hainan and China. *Acta Bot. Yunnan.* 16: 123–130.

Copeland, E.B. 1947. *Genera filicum.* Waltham, Massachusetts: Chronica Botanica.

Cranfill, R. 2001. *Phylogenetic studies in the Polypodiales (Pteridophyta) with an emphasis on the family Blechnaceae.* Dissertation, University of California, Berkeley, California, U.S.A.

Cranfill, R. & Kato, M. 2003. Phylogenetics, biogeography, and classification of the woodwardioid ferns (Blechnaceae). Pp. 25–47 in: Chandra, S. & Srivastava, M. (eds.), *Pteridology in the new millennium*. Dordrecht: Kluwer Academic Publishers.

Crowden, R.K. & Jarman, S.J. 1974. 3-deoxyanthocyanins from the fern *Blechnum procerum*. *Phytochemistry* 13: 1947–1948.

Dawson, M.I., Brownsey, P. & Lovis, J.D. 2000. Index of chromosome numbers of indigenous New Zealand pteridophytes. *New Zealand J. Bot.* 38: 25–46.

Engler, A. & Prantl, K. 1897. *Die natürlichen Pflanzenfamilien nebst ihren Gattungen und wichtigeren Arten insbesondere den Nutzpflanzen*, I(4). Leipzig: Engelmann.

Franz, N.M., Peet, R.K. & Weakley, A.S. 2008. On the use of taxonomic concepts in support of biodiversity research and taxonomy. Pp. 63–85 in: Wheeler, Q. (ed.), *The New Taxonomy*. Boca Raton: CRC Press.

Gastony, G.J. & Johnson, W. 2001. Phylogenetic placements of *Loxoscaphe thecifera* (Aspleniaceae) and *Actiniopteris radiata* (Pteridaceae) based on analysis of *rbcL* nucleotide sequences. *Amer. Fern J.* 91: 197–213.

Gastony, G.J. & Ungerer, M.C. 1997. Molecular systematics and a revised taxonomy of the onocleoid ferns (Dryopteridaceae: Onocleeae). *Amer. J. Bot.* 84: 840–849.

Gifford, E.M. & Foster, A.S. 1989. *Morphology and evolution of vascular plants.* New York: Freeman.

Hasebe, M., Omori, T., Nakazawa, M., Sano, T., Kato, M. & Iwatsuki, K. 1994. *rbcL* gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. *Proc. Natl. Acad. Sci. U.S.A.* 91: 5730–5734.

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 *rbcL* nucleotide sequences. *Amer. Fern J.* 85: 134–181.

Haufler, C.H. & Soltis, D.E. 1986. Genetic evidence suggests that homosporous ferns with high chromosome numbers are diploid. *Proc. Natl. Acad. Sci. U.S.A.* 83: 4389–4393.

Haufler, C.H. & Windham, M.D. 1991. New species of North American *Cystopteris* and *Polypodium*, with comments on their reticulate relationships. *Amer. Fern J.* 81: 7–23.

Hennipman, E. 1996. Scientific consensus classification of Pteridophyta. Pp. 191–202 in: Camus, J.M., Gibby, M. & Johns, R. (eds.), *Pteridology in perspective*. Kew: Royal Botanical Gardens, Kew.

Herter, W.G. 1949. Flora del Uruguay. *Revista Sudamer. Bot.* 9: 14–15.

Hiraoka, A. 1978. Flavonoid patterns in Athyriaceae and Dryopteridaceae. *Biochem. Syst. Ecol.* 6: 171–175.

Holttum, R.E. 1947. A revised classification of leptosporangiate ferns. *J. Linn. Soc., Bot.* 53: 123–155.

Holttum, R.E. 1954. *A revised Flora of Malaya*. Singapore: Government Printing Office.

Holttum, R.E. 1971a. The family names of ferns. *Taxon* 20: 527–532.

Holttum, R.E. 1971b. The genus *Stenochlaena* J. Smith with description of a new species. *Amer. Fern J.* 61: 119–123.

Holttum, R.E. 1971c. Studies in the family Thelypteridaceae III: A new system of genera in the Old World. *Blumea* 19: 17–52.

Holttum, R.E. 1981. The genus *Oreopteris* (Thelypteridaceae). *Kew Bull.* 36: 223–226.

Hoogland, R.D. & Reveal, J.L. 2005. Index nominum familiarum plantarum vascularium. *Bot. Rev.* 71: 1–291.

Iwatsuki, K. 1970. Taxonomic studies of Pteriodophyta IX. *Acta Phytotax. Geobot.* 24: 182–188.

Kato, M. 1972. The vascular structure and its taxonomic significance in the Athyriaceae. *Acta Phytotax. Geobot.* 25: 79–91.

Kato, M. 1975a. On the systematic position of *Athyrium mesosorum*. *Acta Phytotax. Geobot.* 27: 56–60.

Kato, M. 1975b. Reticulate venation of *Diplazium heterophlebium*. *Acta Phytotax. Geobot.* 26: 160–163.

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, Sect. 3, Bot.* 13: 375–430.

Kato, M. 1993. *Diplazium*. Pp. 252–253 in: Flora of North America Editorial Committee (ed.), *Flora of North America north of Mexico.* New York: Oxford University Press.

Kato, M. & Darnaedi, D. 1988. Taxonomic and phytogeographic relationships of *Diplazium flavoviride, D. pycnocarpon*, and *Diplaziopsis*. *Amer. Fern J.* 78: 77–85.

Kato, M., Nakato, N., Akiyama, S. & Iwatsuki, K. 1990. The systematic position of *Asplenium cardiophyllum* (Aspleniaceae). *Bot. Mag. (Tokyo)* 103: 461–468.

Kato, M., Nakato, N., Cheng, X. & Iwatsuki, K. 1992. Cytotaxonomic study of ferns of Yunnan, southwestern China. *Bot. Mag. (Tokyo)* 105: 105–124.

Korall, P., Conant, D.S., Schneider, H., Ueda, K., Nishida, H. & Pryer, K.M. 2006a. On the phylogenetic position of *Cystodium*: It's not a tree fern—It's a polypod! *Amer. Fern J.* 96: 45–53.

Korall, P., Pryer, K.M., Metzgar, J.S., Schneider, H. & Conant, D.S. 2006b. Tree ferns: Monophyletic groups and their relationships as revealed by four protein-coding plastid loci. *Molec. Phylogenet. Evol.* 39: 830–845.

Kramer, K.U. & Viane, R. 1990. Aspleniaceae. Pp. 52–57 in: Kramer, K.U. & Green, P.S. (eds.), *The families and genera of vascular plants*, vol. 1 Berlin: Springer.

Kramer, K.U., Chambers, T.C. & Hennipman, E. 1990a. Blechnaceae. Pp. 60–68 in: Kramer, K.U. & Green, P.S. (eds.), *The families and genera of vascular plants*, vol. 1 Berlin: Springer.

Kramer, K.U., Holttum, R.E., Moran, R.C. & Smith, A.R. 1990b. Dryopteridaceae. Pp. 101–144 in: Kramer, K.U. & Green, P.S. (eds.), *The families and genera of vascular plants*, vol. 1 Berlin: Springer.

Kranz, H. & Huss, V. 1996. Molecular evolution of pteridophytes and their relationship to seed plants: Evidence from complete 18S rRNA gene sequences. *Pl. Syst. Evol.* 202: 1–11.

Kuo, L.-Y., Li, F.-W., Chiou, W.-L. & Wang, C.-N. 2011. First insights into fern *matK* phylogeny. *Molec. Phylogenet. Evol.* 59: 556–566.

Kurita, S. 1960. Chromosome numbers of some Japanese ferns. *J. Jap. Bot.* 35: 269–272.

Kurita, S. 1965. Chromosome numbers and systematic position of the genus *Woodsia*. *J. Jap. Bot.* 40: 358–362.

Li, C., Lu, S., Sun, X. & Yang, Q. 2011. Phylogenetic positions of the enigmatic asiatic fern genera *Diplaziopsis* and *Rhachidosorus* from analyses of four plastid genes. *Amer. Fern J.* 101: 142–155.

Liu, Y.-C. 2008. *A taxonomic study of* Athyrium *in Taiwan*. Dissertation, National Sun Yat-sen University, Kaohsiung, Taiwan.

Liu, Y.-C., Chiou, W.-L. & Kato, M. 2011. Molecular phylogeny and taxonomy of the fern genus *Anisocampium* (Athyriaceae). *Taxon* 60: 824–830.

Löve, A., Löve, D. & Pichi Sermolli, R.E.G. 1977. *Cytotaxonomical atlas of the Pteridophyta*. Vaduz: Strauss & Cramer.

Lovis, J. 1978. Evolutionary patterns and processes in ferns. *Advances Bot. Res.* 4: 229–415.

Loyal, D.S. 1991. Cytomorphological studies in the eastern Himalayan Thelypteridaceae. Pp. 171–248 in: Bhardwaja, T.N. & Gena C.B. (eds.), *Perspectives in pteridology: Present and future*. New Delhi: Today & Tomorrow's Printers and Publishers.

Makino, T. 1899. Plantae Japonenses novae vel minus cognitae. *Bot. Mag. (Tokyo)* 13: 79–82.

Manhart, J. 1994. Phylogenetic analysis of green plant *rbcL* sequences. *Molec. Phylogenet. Evol.* 3: 114–127.

Manhart, J. 1995. Chloroplast 16S rDNA sequences and phylogenetic relationships of fern allies and ferns. *Amer. Fern J.* 85: 182–192.

Manickam, V.S. & Irudayaraj, V. 1988. *Cytology of ferns of the Western Ghats (South India)*. New Delhi: Today & Tomorrow's Printers and Publishers.

Manton, I. 1950. *Problems of cytology and evolution in the Pteridophyta*. Cambridge: Cambridge University Press.

Manton, I. & Sledge, W. 1954. Observations on the cytology and taxonomy of the pteridophyte flora of Ceylon. *Philos. Trans., Ser. B* 238: 127–184.

Mickel, J.T. 1974. Phyletic lines in the modern ferns. *Ann. Missouri Bot. Gard.* 61: 474–482.

Mickel, J.T. & Smith, A.R. 2004. *The pteridophytes of Mexico*. New York: New York Botanical Garden Press.

Mitui, K. 1975. Chromosome number of Japanese pteridophytes. *Bull. Nippon Dent. Coll., Gen. Educ.* 4: 221–271.

Moran, R.C. 1990. Three new species of ferns from Mesoamerica. *Ann. Missouri Bot. Gard.* 77: 591–593.

Murakami, N. 1995. Systematics and evolutionary biology of the fern genus *Hymenasplenium* (Aspleniaceae). *J. Pl. Res.* 108: 257–268.

Murakami, N., Nogami, S., Watanabe, M. & Iwatsuki, K. 1999. Phylogeny of Aspleniaceae inferred from *rbcL* nucleotide sequences. *Amer. Fern J.* 89: 232–243.

Nakahira, Y. 2000. *A molecular phylogenetic analysis of the family Blechnaceae, using the chloroplast gene* rbcL. Dissertation, University of Tokyo, Tokyo, Japan.

Nayar, B. 1970. A phylogenetic classification of the homosporous ferns. *Taxon* 19: 229–236.

Newman, E. 1840. *A history of British ferns*. London: J. van Voorst.

Newman, E. 1844. *A history of British ferns and allied plants*. London: J. van Voorst.

Ogura, Y. 1972. *Comparative anatomy of vegetative organs of the pteridophytes*. Berlin: Gebruder Borntraeger.

Pacheco, L. & Moran, R.C. 1999. Monograph of the Neotropical species of *Callipteris* with anastomosing veins (Woodsiaceae). *Brittonia* 51: 343–388.

Pahnke, J., Goremykin, V., Bobrova, V., Troitsky, A., Antonov, A. & Martin, W. 1996. Utility of rDNA internal transcribed spacer sequences from the inverted repeat of chloroplast DNA in pteridophyte molecular phylogenetics. Pp. 217–230 in: Camus, J.M., Gibby, M. & Johns, R.J. (eds.), *Pteridology in perspective*. Kew: Royal Botanical Gardens, Kew.

Palmer, D.D. 2002. *Hawaii's ferns and fern allies*. Honolulu: University of Hawai'i Press.

Pichi Sermolli, R.E.G. 1970. Fragmenta pteridologiae II. *Webbia* 24: 699–722.

Pichi Sermolli, R.E.G. 1973. An historical review of the higher classification of the Filicopsida. Pp. 11–40 in: Jermy, A.C., Crabbe, J.A. & Thomas, B.A. (eds.), *The phylogeny and classification of the ferns*. London: Academic Press.

Pichi Sermolli, R.E.G. 1977. Tentamen Pteridophytorum genera in taxonomicum ordinem redigendi. *Webbia* 31: 313–512.

Price, M.G. 1990. Philippine fern notes. *Contr. Univ. Michigan Herb.* 17: 267–278.

Pryer, K.M. 1993. *Gymnocarpium*. Pp. 258–262 in: Flora of North America Editorial Committee (ed.), *Flora of North America north of Mexico*. New York: Oxford University Press.

Pryer, K.M. & Haufler, C.H. 1993. Isozymic and chromosomal evidence for the allotetraploid origin of *Gymnocarpium dryopteris* (Dryopteridaceae). *Syst. Bot.* 18: 150–172.

Pryer, K.M., Schneider, H., Smith, A.R., Cranfill, R., Wolf, P.G., Hunt, J.S. & Sipes, S.D. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409: 618–622.

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. *Amer. J. Bot.* 91: 1582–1598.

Pryer, K.M., Smith, A.R. & Rothfels, C.J. 2008. Polypodiidae Cronquist, Takht. & Zimmerm. 1966, leptosporangiate ferns. In: The Tree of Life Web Project, version 23 Dec. 2008. http://tolweb.org/Polypodiidae/21666/2008.12.23.

Pryer, K.M., Smith, A.R. & Skog, J.E. 1995. Phylogenetic relationships of extant ferns based on evidence from morphology and *rbcL* sequences. *Amer. Fern J.* 85: 205–282.

Rothfels, C.J., Larsson, A., Kuo, L.-Y., Korall, P., Chiou, W.-L. & Pryer, K.M. 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Syst. Biol.* 16: 490–509.

Sano, R., Takamiya, M., Ito, M., Kurita, S. & Hasebe, M. 2000a. Phylogeny of the lady fern group, tribe Physematieae (Dryopteridaceae), based on chloroplast *rbcL* gene sequences. *Molec. 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. Pl. Res.* 113: 157–163.

Sarvela, J. 1978. A synopsis of the fern genus *Gymnocarpium*. *Ann. Bot. Fenn.* 15: 101–106.

Schneider, H. 1996. *Vergleichende Wurzelanatomie der Farne*. Dissertation, University of Zurich, Zurich, Switzerland.

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 *rbcL* and *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., Magallon, S. & Lupia, R. 2004b. Ferns diversified in the shadow of angiosperms. *Nature* 428: 553–557.

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., Korall, P. & Pryer, K.M. 2006. Plastid *atpA* data provide improved support for deep relationships among ferns. *Taxon* 55: 897–906.

Shmakov, A. 2001. Synopsis of the ferns of Russia. *Turczaninowia* 4: 36–72.

Shmakov, A. 2003. Review of the family Woodsiaceae (Diels) Herter of Eurasia. Pp. 49–64 in: Chandra, S. & Srivastava, M. (eds.), *Pteridology in the new millennium*. Dordrecht: Kluwer Academic Publishers.

Sledge, W. 1973. Generic and family boundaries in the Aspidiaceae and Athyriaceae. *Bot. J. Linn. Soc.* 67: 203–210.

Smith, A.R. 1976. *Diplazium delitescens* and the Neotropical species of *Asplenium* sect. *Hymenasplenium*. *Amer. Fern J.* 66: 116–120.

Smith, A.R. 1981. *Flora of Chiapas*, pt. 2, *Pteridophytes*. San Francisco: California Academy of Sciences.

Smith, A.R. 1990. Thelypteridaceae. Pp. 263–272 in: Kramer, K.U. & Green, P.S. (eds.), *The families and genera of vascular plants*, vol. 1. Berlin: Springer.

Smith, A.R. 1995. Non-molecular phylogenetic hypotheses for ferns. *Amer. Fern J.* 85: 104–122.

Smith, A.R. & Cranfill, R.B. 2002. Intrafamilial relationships of the thelypteroid ferns (Thelypteridaceae). *Amer. Fern J.* 92: 131–149.

Smith, A.R. & Mickel, J. 1977. Chromosome counts for Mexican ferns. *Brittonia* 29: 391–398.

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.

Stevenson, D. & Loconte, H. 1996. Ordinal and familial relationships of pteridophyte genera. Pp. 435–467 in: Camus, J.M., Gibby, M. & Johns, R. J. (eds.), *Pteridology in perspective*. Kew: Royal Botanic Gardens, Kew.

Sundue, M.S, Vasco, A. & Moran, R.C. 2011. Cryptochlorophyllous spores in ferns: Nongreen spores that contain chlorophyll. *Int. J. Pl. Sci.* 172: 1110–1119.

Tagawa, M. 1935. A review of the genus *Cystopteris* of Japan. *Acta Phytotax. Geobot.* 4: 51–57.

Tagawa, M. & Iwatsuki, K. 1972. Families and genera of the pteridophytes known from Thailand. *Mem. Fac. Sci. Kyoto Univ., Ser. Biol.* 5: 67–88.

Takamiya, M. & Ohta, N. 2001. Cytological and reproductive studies of Japanese *Diplazium* (Woodsiaceae; Pteridophyta). III. The cytological complexity of species groups with simply pinnate to bipinnatifid leaves. *J. Pl. Res.* 114: 443–457.

Takamiya, M., Ohta, N., Fujimaru-Takaoka, C. & Uki, K. 2000. Cytological and reproductive studies of Japanese *Diplazium* (Woodsiaceae; Pteridophyta). II. Polyploidy and hybridity in the species group with summer-green bi- to tripinnate leaves. *J. Pl. Res.* 113: 203–215.

Tindale, M.D. & Roy, S.K. 2002. A cytotaxonomic survey of the Pteridophyta of Australia. *Austral. Syst. Bot.* 15: 839–937.

Tryon, R.M. & Tryon, A.F. 1982. *Ferns and allied plants with special reference to tropical America*. New York: Springer.

Tryon, R.M., Tryon, A.F. & Kramer, K.U. 1990. Pteridaceae. Pp. 230–256 in: Kramer, K.U. & Green, P.S. (eds.), *The families and genera of vascular plants*, vol. 1 Berlin: Springer.

Tsai, J.L. & Shieh, W.C. 1985. A cytotaxonomic survey of the fern family Aspidiaceae (sensu Copeland) in Taiwan. *J. Sci. Engin.* 40: 121–144.

Tzeng, Y.-H. 2002. *Phylogenetic relationships of athyrioid ferns inferred from chloroplast DNA sequences*. Master's Thesis, National Sun Yat-Sen University, Kaohsiung, Taiwan.

Vangerow, S., Teerkorn, T. & Knoop, V. 1999. Phylogenetic information in the mitochondrial *nad5* gene of pteridophytes: RNA editing and intron sequences. *Pl. Biol.* 1: 235–243.

Vida, G. 1974. Genome analysis of the European *Cystopteris fragilis* complex. 1: Tetraploid taxa. *Acta Bot. Acad. Sci. Hung.* 20: 181–192.

Wagner, F.S. 1980. New basic chromosome numbers for genera of neotropical ferns. *Amer. J. Bot.* 67: 733–738.

Wagner, F.S. 1995. The chromosomes of *Sadleria* (Blechnaceae). *Contr. Univ. Michigan Herb.* 20: 239–240.

Wagner, W.H., Jr. 1955. Cytotaxonomic observations on North American ferns. *Rhodora* 57: 219–240.

Wagner, W.H., Jr. 1980. Origin and philosophy of the groundplan-divergence method of cladistics. *Syst. Bot.* 5: 173–193.

Walker, T.G. 1973a. Additional cytotaxonomic notes on the pteridophytes of Jamaica. *Trans. Roy. Soc. Edinburgh* 69: 109–135.

Walker, T.G. 1973b. Evidence from cytology in the classification of ferns. *Bot. J. Linn. Soc.* 67: 91–110.

Walker, T.G. 1985. Cytotaxonomic studies of the ferns of Trinidad 2. The cytological and taxonomic implications. *Bull. Brit. Mus. (Nat. Hist.), Bot.* 13: 149–249.

Wang, M.-L., Chen, Z.D., Zhang, X.-C., Lu, S.-G. & Zhao, G.-F. 2003. Phylogeny of the Athyriaceae: Evidence from chloroplast *trnL-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]. In: Flora of China Editorial Committee (ed.). Flora of China. http://www.flora.ac.cn/volume2/Athyriaceae-MO-part1_original.htm.

Wei, R., Zhang, X.-C. & Qi, X.-P. 2010. Phylogeny of *Diplaziopsis* and *Homalosorus* based on two chloroplast DNA sequences: *rbcL* and *rps4+rps4-trnS* IGS. *Acta Bot. Yunnan.* 17: 46–54.

Weng, R.-F. & Qiu, S.-P. 1988. Chromosome counts of some ferns from Zhejiang. *Invest. Stud. Nat.* 8: 43–52.

Wikström, N. & Pryer, K.M. 2005. Incongruence between primary sequence data and the distribution of a mitochondrial *atp1* group II intron among ferns and horsetails. *Molec. Phylogenet. Evol.* 36: 484–493.

Wilson, K.A. 1959. Sporangia of the fern genera allied to *Polypodium* and *Vittaria*. *Contr. Gray Herb.* 185: 97–127.

Wolf, P.G. 1995. Phylogenetic analyses of *rbcL* and nuclear ribosomal RNA gene sequences in Dennstaedtiaceae. *Amer. Fern J.* 85: 306–327.

Wolf, P.G. 1997. Evaluation of *atpB* nucleotide sequences for phylogenetic studies of ferns and other pteridophytes. *Amer. J. Bot.* 84: 1429–1440.

Wolf, P.G., Pryer, K.M., Smith, A.R. & Hasebe, M. 1998. Phylogenetic studies of extant pteridophytes. Pp. 541–556 in: Soltis, D.E., Soltis, P.S. & Doyle, J.J. (eds.), *Molecular systematics of plants II: DNA sequencing*. Boston: Kluwer Academic Publishers.

Wolf, P.G., Sipes, S., White, M., Martines, M., Pryer, K.M., Smith, A.R. & Ueda, K. 1999. Phylogenetic relationships of the enigmatic fern families Hymenophyllopsidaceae and Lophosoriaceae: Evidence from *rbcL* nucleotide sequences. *Pl. Syst. Evol.* 219: 263–270.

Wolf, P.G., Soltis, P.S. & Soltis, D.E. 1994. Phylogenetic relationships of dennstaedtioid ferns: Evidence from *rbcL* sequences. *Molec. Phylogenet. Evol.* 3: 383–392.

Wu, S.-H. & Ching, R.-C. 1991. *Fern families and genera of China*. Beijing: Science Press.

Zuquim, G., Costa, F.R.C., Prado, J. & Tuomisto, H. 2008. *Guide to the ferns and lycophytes of REBIO Utamã, Central Amazonia*. St. Louis: Missouri Botanical Garden Press.

Paper III

OPEN ACCESS Freely available online

PLOS ONE

Transcriptome-Mining for Single-Copy Nuclear Markers in Ferns

Carl J. Rothfels[1,2*], Anders Larsson[3], Fay-Wei Li[1], Erin M. Sigel[1], Layne Huiet[1], Dylan O. Burge[4], Markus Ruhsam[5], Sean W. Graham[4], Dennis W. Stevenson[6], Gane Ka-Shu Wong[7,8], Petra Korall[3], Kathleen M. Pryer[1]

1 Department of Biology, Duke University, Durham, North Carolina, United States of America, **2** Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada, **3** Systematic Biology, Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden, **4** Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada, **5** Royal Botanic Garden Edinburgh, Edinburgh, Scotland, **6** New York Botanical Garden, Bronx, New York, United States of America, **7** Department of Medicine, University of Alberta, Edmonton, Alberta, Canada, **8** BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen, China

Abstract

Background: Molecular phylogenetic investigations have revolutionized our understanding of the evolutionary history of ferns—the second-most species-rich major group of vascular plants, and the sister clade to seed plants. The general absence of genomic resources available for this important group of plants, however, has resulted in the strong dependence of these studies on plastid data; nuclear or mitochondrial data have been rarely used. In this study, we utilize transcriptome data to design primers for nuclear markers for use in studies of fern evolutionary biology, and demonstrate the utility of these markers across the largest order of ferns, the Polypodiales.

Principal Findings: We present 20 novel single-copy nuclear regions, across 10 distinct protein-coding genes: *ApPEFP_C, cryptochrome 2, cryptochrome 4, DET1, gapCpSh*, *IBR3, pgiC, SQD1, TPLATE*, and *transducin*. These loci, individually and in combination, show strong resolving power across the Polypodiales phylogeny, and are readily amplified and sequenced from our genomic DNA test set (from 15 diploid Polypodiales species). For each region, we also present transcriptome alignments of the focal locus and related paralogs—curated broadly across ferns—that will allow researchers to develop their own primer sets for fern taxa outside of the Polypodiales. Analyses of sequence data generated from our genomic DNA test set reveal strong effects of partitioning schemes on support levels and, to a much lesser extent, on topology. A model partitioned by codon position is strongly favored, and analyses of the combined data yield a Polypodiales phylogeny that is well-supported and consistent with earlier studies of this group.

Conclusions: The 20 single-copy regions presented here more than triple the single-copy nuclear regions available for use in ferns. They provide a much-needed opportunity to assess plastid-derived hypotheses of relationships within the ferns, and increase our capacity to explore aspects of fern evolution previously unavailable to scientific investigation.

Citation: Rothfels CJ, Larsson A, Li F-W, Sigel EM, Huiet L, et al. (2013) Transcriptome-Mining for Single-Copy Nuclear Markers in Ferns. PLoS ONE 8(10): e76957. doi:10.1371/journal.pone.0076957

Editor: Keith A Crandall, George Washington University, United States of America

Received June 12, 2013; **Accepted** August 27, 2013; **Published** October 8, 2013

Copyright: © 2013 Rothfels et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by funds from the National Science Foundation (www.nsf.gov) DDIG DEB-1110767 to KMP and CJR (co-authors), DDIG DEB-1110775 to KMP and EMS (co-authors), and DEB-1145614 to KMP and LH; a Natural Science and Engineering Research Council (Canada) Postgraduate Scholarship – Doctoral of NSERC (Natural Sciences and Engineering Research Council) to CJR and Discovery grant to SWG; grants from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas) to PK (2006-429 and 2010-585); GKSW acknowledges the support of Alberta Enterprise and Advanced Education, Genome Alberta, Alberta Innovates Technology Futures iCORE, Musea Ventures, and BGI-Shenzhen. The funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare that they have the following interest: This study was partly funded by Musea Ventures. There are no patents, products in development or marketed products to declare. This does not alter their adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: crothfels@yahoo.ca

Introduction

Over the past twenty years, molecular phylogenetic approaches have radically altered our understanding of relationships in the fern tree of life. Arguably the most important finding (and among the most contentious) is that ferns, including the horsetails (*Equisetum*) and whisk ferns (Psilotaceae), form a clade sister to seed plants [1]. Ferns are therefore one of the great vascular plant radiations; only the angiosperm clade has more extant species. Broad studies of fern phylogeny, e.g., [1-18] have increasingly found stronger resolution and support across the majority of the backbone nodes, many of which were unanticipated based on morphological data. This rewriting of fern phylogeny has resulted in a novel emerging consensus of deep fern relationships [19-22]. Similarly, new molecular data have greatly facilitated inquiries into fern relationships at much finer scales, such as within genera, e.g., [23-46].

The vast majority of these studies, however, have been limited to data from the plastid genome; nuclear and mitochondrial data have not been widely used, cf. [4,9,47-49]. This dependence upon plastid data reflects a general absence of genomic resources available for ferns [50-53]—for example, no fern mitochondrial or nuclear genome has been sequenced yet—which has impeded the development of novel markers. To date only seven nuclear regions have been used in fern phylogeny investigations: ITS [13,54-57]; ribosomal 18S [4,58]; *LEAFY* [59-62]; *gapCpSh* (*gapCp "short"*) [36,42,43,49,57,63-67]; *gapCpLg* (*gapCp "long"*) [67]; *cam* [67]; and *pgiC* [57,59,66,68-75].

This strong reliance on the plastid genome makes fern phylogenetics vulnerable to misleading inferences, such as failures of this linkage group to track the organismal divergences (e.g., due to deep coalescence or reticulation). In addition, plastid data are poorly suited for species-level work in the many fern groups that have reticulate evolutionary histories [76-79]. Polyploidy and hybridization are common in ferns [80], and fully unraveling relationships in these groups will require the development of multiple unlinked markers. Regions from the nucleus are particularly attractive for this purpose because that genome has multiple linkage groups that are expected to have an elevated rate of evolution in ferns, e.g., [49].

The recent sequencing of approximately 1000 green plant transcriptomes by the One Thousand Plants Project (1KP; onekp.com) provides an unprecedented opportunity to facilitate the development of novel low-copy nuclear markers for use in ferns. There is no fern nuclear genome that has been sequenced to date, and only a handful of EST libraries, sequenced plastomes, or transcriptomes, e.g., [81-89]. Included in the 1KP sampling (as of January 2013, when we finished the sampling for this project) are 62 fern accessions, comprising 60 unique species. Our sampling from this time point is particularly rich in members of the leptosporangiate order Polypodiales, especially Pteridaceae (*sensu* [20,90]) and eupolypods II (*sensu* [8,16,19,91]), but also includes representatives of each of the major eusporangiate clades (Ophioglossales, Psilotales, Equisetales, Marattiales), as well as each of the leptosporangiate orders, except for the Osmundales [20]. Additional taxa, including representatives of the Osmundales, were sequenced in the 1KP project after we had finished our sampling. The full list—including algae, bryophytes, lycophytes, ferns, and seed plants—is available at http://www.onekp.com/samples/list.php.

Here, we utilize these transcriptome data to design primers for 20 nuclear markers across ten protein-coding genes. Our primary goals are: 1) to provide primers that will amplify single-copy nuclear markers across the majority of the Polypodiales; 2) to demonstrate the relative success of those primers in amplifying the desired region from genomic DNA, using a test set of 15 diploid Polypodiales species, and; 3) to characterize the resulting sequences and their efficacy in inferring relationships at various phylogenetic depths. In addition, we provide transcriptome alignments—curated broadly across ferns—for each of our target loci (including closely related paralogs in the case of gene families) to assist other investigators in designing primers for fern taxa of interest outside of the Polypodiales.

Results

Primer Development

We developed primer pairs for 20 regions across a total of ten distinct single-copy protein-coding genes: *ApPEFP_C, cryptochrome 2, cryptochrome 4, DET1, gapCpSh*, *IBR3, pgiC, SQD1, TPLATE*, and *transducin* (Tables 1, 2, 3; Figure 1). Each primer pair successfully amplifies the majority of taxa in our genomic DNA test set (comprising DNA from 15 diploid Polypodiales species; Table 3; Figure 2; Appendix S1). In general, we only attempted to sequence PCR products that had strong single bands (viewed with agarose gel electrophoresis). Many of the missing sequences are likely to be attainable by applying cloning protocols (e.g., see 64). Those sequences that we did attain by cloning are noted in the Methods section and in Table 3.

ApPEFP_C. We developed primers for three regions of *ApPEFP_C* (Figure 1A), and for one of those regions (Region 1) we designed non-overlapping internal primers that can amplify two smaller subset regions (Figure 1A). Region 1 is approximately 700–1000bp long in ferns (Table 3) and spans introns 2, 3, and 4, exons 3 and 4, and half of exon 5 (Figure 1A). It could be direct-sequenced for most taxa in our genomic DNA test set, although cloning was necessary for *Polypodium* (due to a hypothesized gene duplication in the Polypodiaceae; see Figure 2) and *Cystopteris protrusa*. Within Region 1, the additional reverse primer 4218Cr3 allows for the amplification of the subset Region 1a, and the forward primer 4218Cf6 yields Region 1b (Figure 1A). Both these smaller regions are approximately 200–300 bp long. Region 1a is the more variable of the two, whereas Region 1b has good length conservation among taxa and is easy to align across the complete breadth of the Polypodiales (Table 3). Region 2 overlaps with the 3' end of Region 1; it includes a portion of exon 4, introns 4 and 5, exon 5, and most of exon 6 (Figure 1A). Finally, Region 3 is intermediate in length and includes much of the large exon 8, intron 8, and most of exon 9. In the eupolypods, Region 3 ranges in length from approximately 400 to 500 bp, but is

Table 1. Summary of the genes for which we designed primers.

	Abbreviation	Protein Name	Length (CDS; in bp)		*Arabidopsis*		
			Arabid.	Ferns	TAIR Gn#	# of Introns	Chromo. #
1	*ApPEFP_C*	appr-1-p processing enzyme family protein	1689	~1650-1743	AT1G69340	13	1
2	*CRY2*	cryptochrome 2	2046	~2000	AT4G08920	3	4
3	*CRY4*	cryptochrome 4	1839	~2100	AT1G04400	3	1
4	*DET1*	Nuclear-localized regulator of plant development	1632	~1600-2700	AT4G10180	9	4
5	*gapCpSh*	Plastid-localized GAPDH, short copy	1266; 1260	1315	AT1G79530; AT1G16300	13	1
6	*IBR3*	IBA-Response 3 (acyl-CoA dehydrogenase)	2475	~2445-2490	AT3G06810	16	3
7	*pgiC*	glucose-6-phosphate isomerase / sugar isomerase family protein	1683	--[a]	AT5G42740	21	5
8	*SQD1*	Sulfoquinovosyldiacylglyerol 1	1431	~1515-1521	AT4G33030	1	4
9	*TPLATE*	a cytokinesis protein targeted to the cell plate	3531	--[a]	AT3G01780	6	3
10	*transducin*	transducin family protein / WD-40 repeat family protein	2868	--[a]	AT3G21540	11	3

For each gene, we list its length in ferns and in *Arabidopsis*, provide the TAIR accession number for the *Arabidopsis* sequence (as well as its number of introns and chromosomal position). The TreeBASE accession number for our "all-in" fern alignments is S14616. Comparisons with *Arabidopsis thaliana* are based on the most closely related homolog(s). [a] These loci were trimmed to a focal region prior to completion, so the full length of the coding DNA sequence (CDS) is unknown.

doi: 10.1371/journal.pone.0076957.t001

Table 2. Priming details for 20 novel nuclear markers.

		Primers (Forward, Reverse)		
Protein Region		Name	Sequence (5'–3')	PCR Program
ApPEFP_C	1	4218Cf4, 4218Cr12	GGACCTGGSCTYGCTGARGAGTG, GCAACRTGAGCAGCYGGTTCRCGRGG	6512035
ApPEFP_C	1a	4218Cf4, 4218Cr3	GGACCTGGSCTYGCTGARGAGTG, TCGTAAGCRTTYGTTACTTTDGCC	5506035
ApPEFP_C	1b	4218Cf6, 4218Cr6	AAAGTTATACATACTGTTGGTCC, GCAACATGAGCAGCTGGTTCACGAGG	5506035
ApPEFP_C	2	4218f25, 4218r7	AATGCTCTRAGTCAYTGYTAYMGATC, TTGTAAATCTCTGTRTCRGATGYYGT	5509035
ApPEFP_C	3	4218f26, 4218r13	CAAAGGCGCAARGAACARTGGGARAGRGTTGC, TCAAGACAYCGTAGCAGRAARTGBGCYCC	6512035
CRY2	1	CRY2F3289_Pt, CRY2R3838_Pt	AGGATGARYTGGAGAAAGGYAGCAATG, GTRTCCCAGAAATAYTTCATACCCC	5209035
CRY4	1	CRY2F3289_Pt, CRY2R3838_Pt	AGGATGARYTGGAGAAAGGYAGCAATG, GTRTCCCAGAAATAYTTCATACCCC	5209035
DET1	1	det1-335all, det1-906all	TATGAYGTGGARTGCCCAGAT, TCTCTGCAGAAHKGYCCAA	5506035
gapCpSh	1	gapCpShF1, gapCpShR2	TGCACMACHAACTGCCTTGCRCCBCTTGC, CCATTYARCTCTGGRAGCACCTTTCC	6512035
IBR3	1	4321F2, 4321R2	TCTGCMCATGCMATTGAAAGAGAG, CCCARKGTYGAAAGYTCCCAATC	6312035
IBR3	2	4321F5, 4321R6	ATGACYGAACCAGATGTKGCDTCVTCRGATGC, TGRTGGAGYCTKCCTGGGCCTA	6512035
pgiC	1	pgic_1156F, pgiC_1900R	GGYCTYYTRAGYGTYTGGAATGT, GGTGAAATYGAYTTYGGDGARC	5812035
SQD1	1	EMSQD1E1F6, EMSQD1E1R2	GCAAGGGTACHAAGGTHATGATCATAGG, CCTTTDCCRTARACTGTAAGAGGATG	5512035
SQD1	1a	EMSQD1E1F6, EMSQD1E1R4	GCAAGGGTACHAAGGTHATGATCATAGG, GCGTGARTCRTGCACTTTGCTRAGATG	5512035
SQD1	2	EMSQD1E2F4, EMSQD1E2R8	CGHGTRTTYAATCARTTYACAGAAC, GTCACTGTHACAGGTTTYACDCCAGC	5512035
TPLATE	1	6560_1630F, 6560_2329R	TGCYTAGTSGARAGYTGYTTTCA, AATGTAGCAACTAACAGGCTTCAAGA	5812035
TPLATE	2	6560_3136F, 6560_3686R	AAYCTYCARCATCTYCAGTCTC, GCAACKGCHGCDGTBGAAAG	5812035
transducin	1	6928_850F, 6928_1357R	TTRCGBGGRCAYARAGATCA, GGAWCSTTARTSGGYTGCCAA	5812035
transducin	2	6928_1955F, 6928_2816R	AAGGCDGGRAARCTNGAGAT, ATGGAYATYTCCWCYGATGC	5812035
transducin	3	6928_3406F, 6928_3802R	TCBATTCGRMGATGGGAGCG, CAAACYCARGARWCYSTGAC	5812035

The first two digits of the PCR program is the annealing temperature, followed by a three-digit elongation time (in seconds), followed by the number of cycles.

doi: 10.1371/journal.pone.0076957.t002

consistently larger in the Pteridaceae and Dennstaedtiaceae (reaching 802 bp in *Dennstaedtia*). It amplified and direct-sequenced well (Figure 2), but required cloning for *P. glycyrrhiza* and *C. protrusa* (see above); however, the *P. amorphum* Region 3 sequence was clean (not double-peaked when directly sequenced), and did not require cloning.

CRY2 and CRY4. We designed primers to target a 516 bp region in the third exon of *CRY2* (Figure 1B). However, these primers also cross-amplify the same region in *CRY4* (also 516 bp; Figure 1C), and, less frequently, in *CRY3* and *CRY5* (see Figure S2). The PCR products thus cannot be directly sequenced. Nevertheless, after cloning, we recovered *CRY2* and *CRY4* for 14 and 11 (respectively) of the 15 taxa in our genomic DNA test set. These two loci are sufficiently divergent that assigning sequences to the correct copy is straightforward. *CRY2* appears to have higher sequence variation, with nine nucleotide differences between the *Cystopteris* species pair, whereas there are only two nucleotide differences between the

Table 3. Sequence characteristics for the single-copy regions developed in this study.

		# of Differences		Length of Amplified Region													
Protein Region		**Cys. pair**	**Poly. pair**	**Cya.**	**Lin.**	**Sac.**	**Adi.**	**Che.**	**Cry.**	**Den.**	**Dry.**	**P.am.**	**P.gl.**	**Ath.**	**C.bu.**	**C.pr.**	**The.**	**Woo.**
ApPEFP_C	1	>38; >25[a]	50; 25[b]	?	?	>751	>690	>829	721	931	761	846; >871[c]	843; 837[c]	>849	>866	932; 934[c]	687	939
ApPEFP_C	1a	>13; >12[a]	10; 9[b]	145	?	258[d]	321[d]	281	300[d]	293	267	~257	~255	201[d]	245	161	>225	255
ApPEFP_C	1b	6; 6[a]	11; 3[b]	?	?	?	224	209[d]	223[d]	269	223	224; 223[d]	216; 223[d]	224	214[d]	223[d]	223	222
ApPEFP_C	2	5	?	382	334	360	360	340	357	406	359	?	?	360	359	358	359	357
ApPEFP_C	3	14	11; 25	?	377	?	>506	>776	770	802	454	385	378; 384[c]	490	482	461[c]	489	457
CRY2	1	9	14	516	516	516	516	516	516	?	516	516	516	516	516	516	516	516
CRY4	1	2	?	?	516	516	516	516	?	516	516	?	516	516	516	516	516	?
DET1	1	~5	6	?	?	?	~630	?	?	?	667	668	668	~670	665	664	669	669
gapCpSh	1	?	14	?	455	459	?	?	482	476	522[c]	531	525	?	?	466	517	592
IBR3	1	>16	29; 31	?	?	870	817	>700	827	836	819; 828[c]	843[c]	844	815	>819	840	>910	821
IBR3	2	~6	~21	~600	>766	611	~574	581	568	~1196	?	~590	595	586	~580	~582	579	588
pgiC	1	>32	>19	674	?	?	?	?	?	?	625	>615	678	>474	>664	>581	619	620
SQD1	1	10	>8[e]	?	700	668	700	700	700	700	700	?	700	700	700	700	700	685
SQD1	1a	8	8	530	530[f]	530[f]	530[f]	530[f]	530[f]	530[f]	530[f]	529	530[f]	530[f]	530[f]	530[f]	530[f]	530[f]
SQD1	2	1	3	264	263	264	264	256	264	263	264	264	264	264	264	233	264	?
TPLATE	1	12	14	719	>657	?	646	>561	>529	627	711	696	698	>638	710	>662	>692	687
TPLATE	2	5	10	>327	?	?	?	551	?	529	512	497	493	302	427	424	722	541
transducin	1	11	?	?	?	?	402	397	426	416	?	>381	?	436	437	435	420	421
transducin	2	11	>7	>521	?	?	?	>426	?	539	534	529	>445	518	525	517	514	504
transducin	3	6	7	?	?	?	227	308	231	?	268	251	251	242	244	244	243	242

Cys: *Cystopteris*, Poly: *Polypodium*, Cya: Cyatheales, Lin.: *Lindsaea*, Sac.: *Saccoloma*, Adi: *Adiantum*, Che: *Cheilanthes*, Cry: *Cryptogramma*, Den: *Dennstaedtia*, P.am.: *Polypodium amorphum*, P.gly.: *Polypodium glycyrrhiza*, Ath.: *Athyrium*, C.bu.: *Cystopteris bulbifera*, C.pr.: *Cystopteris protrusa*, The: *Thelypteris*, Woo: *Woodsia*. [a] The two values come from comparing the single incomplete *Cystopteris bulbifera* sequence against two sequences cloned from *C. protrusa*. [b] This locus has a duplication in *Polypodium*; these values are the number of bp changes between each of the ortholog pairs. [c] Required cloning. [d] These lengths are derived from the corresponding portion of the *APPEFP_C* Region 1 alignment (we did not attempt to amplify Region 1a or 1b for all taxa). [e] For *P. amorphum* we were not able to amplify Region 1 for this locus, only Region 1a. [f] These lengths are derived from the corresponding portion of the *SQD1* Region 1a alignment (Region 1a for all taxa).

doi: 10.1371/journal.pone.0076957.t003

same species pair in *CRY4* (*Cystopteris protrusa* and *C. bulbifera* constituted one of the two pairs of closely related species that we used as a metric for informativeness at shallow phylogenetic depths—see Methods and Table 3).

DET1. We focused on a single region of *DET1* (Figure 1D). The primer pair 4321F2-4321R2 amplifies a ~670 bp region that includes most of the second exon (in *Arabidopsis*; ferns contain an additional intron within this region). All sequences were obtained by direct-sequencing.

GapCpSh. We designed primers for a single region of *gapCpSh*. The forward primer is situated just before intron 8 and the reverse priming site is just after intron 10 (Figure 1E); this region ranges in length from ~450 to 590 bp in our genomic DNA test set. In general *GapCpSh* amplified and direct-sequenced well, although we were not able to obtain clean sequences for *Alsophila, Adiantum, Cheilanthes, Athyrium*, or *Cystopteris bulbifera* (cloning not attempted) and the *Dryopteris* sequence required cloning. This region physically overlaps with the *gapCp* region amplified with the primers of Schuettpelz et al. [64], but differs in that our primers are specific to the *gapCp Short* (*sensu* [64]) copy in Polypodiales, and amplify a region slightly shorter than that of Schuettpelz et al. [64].

IBR3. We designed primers for two regions of *IBR3*. Region 1 spans introns 2–5 and exons 3–5 (Figure 1F); it is approximately 900bp long in the Polypodiales species we examined (Table 3). Region 2, at the 3' end of the gene, is shorter, at around 600 bp in most species; however, it is much larger (1200 bp) in *Dennstaedtia*. It spans introns 12–14, exons 13 and 14, and the end of exon 12. Both regions amplified well, and gave clean direct sequences for the majority of taxa in our test set.

PgiC. We developed one novel primer pair for *pgiC*—a locus already known to work well in fern phylogenetics [57,66,71-75,92]. Our primers are situated in exons 14 and 16, amplifying introns 14, 15, and exon 15 (Figure 1G). The amplified region ranges in length between 600 and 700bp across our test set (Table 3), and the range of variation observed is appropriate for resolving infrageneric relationships (Table 3 and see citations above). It amplified and direct-sequenced well for 10 of the 15 test set taxa; *Lindsaea, Saccoloma, Dennstaedtia, Cheilanthes* and *Adiantum* failed amplify and/or direct-sequence cleanly.

SQD1. Primers were designed for two regions of *SQD1*: a 700 bp region within the first exon and a 264 bp region within the second exon (Figure 1H; Table 3). The Region 1 forward and reverse primers—EMSQDE1F6 and EMSQDE1R2—produced successful amplifications for 13 of the 15 taxa in our test set. An additional reverse primer, EMSQDE1R4, was designed to amplify a 530 bp subset (henceforth designated Region 1a; Figure 1H, Table 3), which resulted in successful

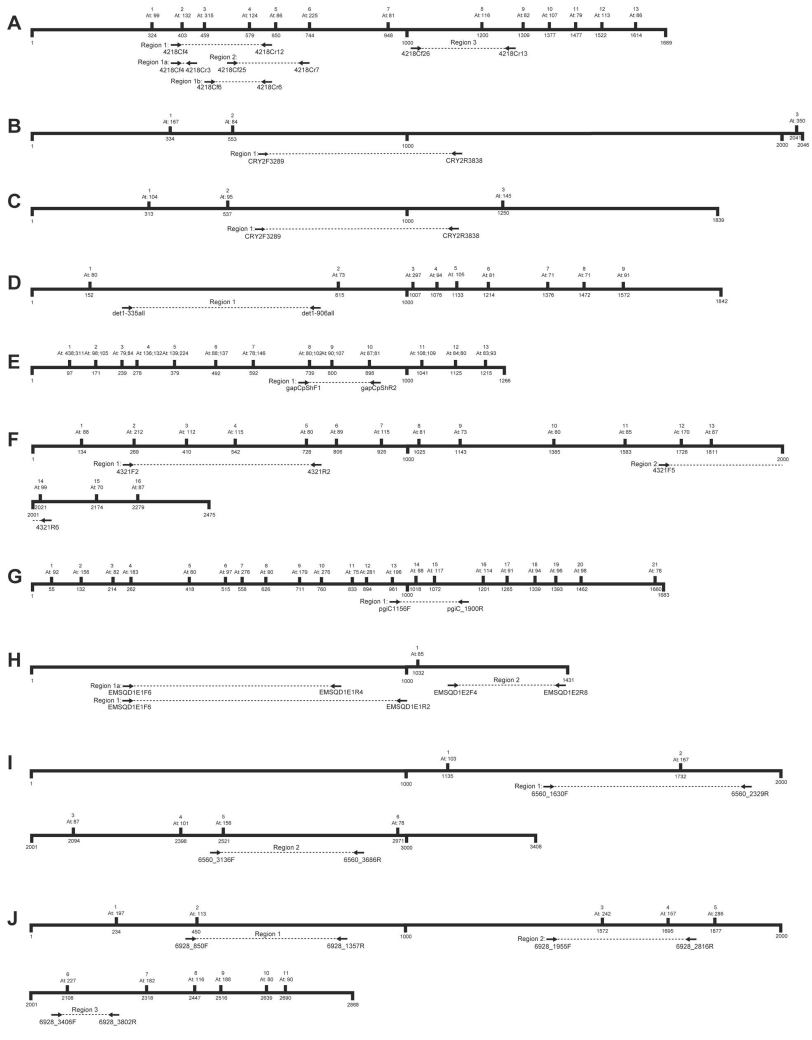

Figure 1

Figure 1. Schematic diagrams of the ten nuclear genes for which we developed fern-specific primers. (A) *ApPEFP_C*; (B) *CRY2*; (C) *CRY4*; (D) *DET1*; (E) *gapCpSh*; (F) *IBR3*; (G) *pgiC*; (H) *SQD1*; (I) *TPLATE*; (J) *transducin*. Each subset of the figure represents one protein-coding locus, using the most closely related *Arabidopsis thaliana* homolog as the template. The coding sequence is measured (in base pairs) along the bottom of the thickened horizontal line, with each locus wrapping onto a new line every 2000 base pairs, when necessary. Intron location, number, and length (in base pairs in *Arabidopsis*) are given above the line. Also shown below the line are the priming locations for each of the markers we developed. For *gapCpSh*, intron locations are based on *Arabidopsis gapCp1*: the first two exons of *Arabidopsis gapCp2* are each one codon shorter than in *gapCp1*.

doi: 10.1371/journal.pone.0076957.g001

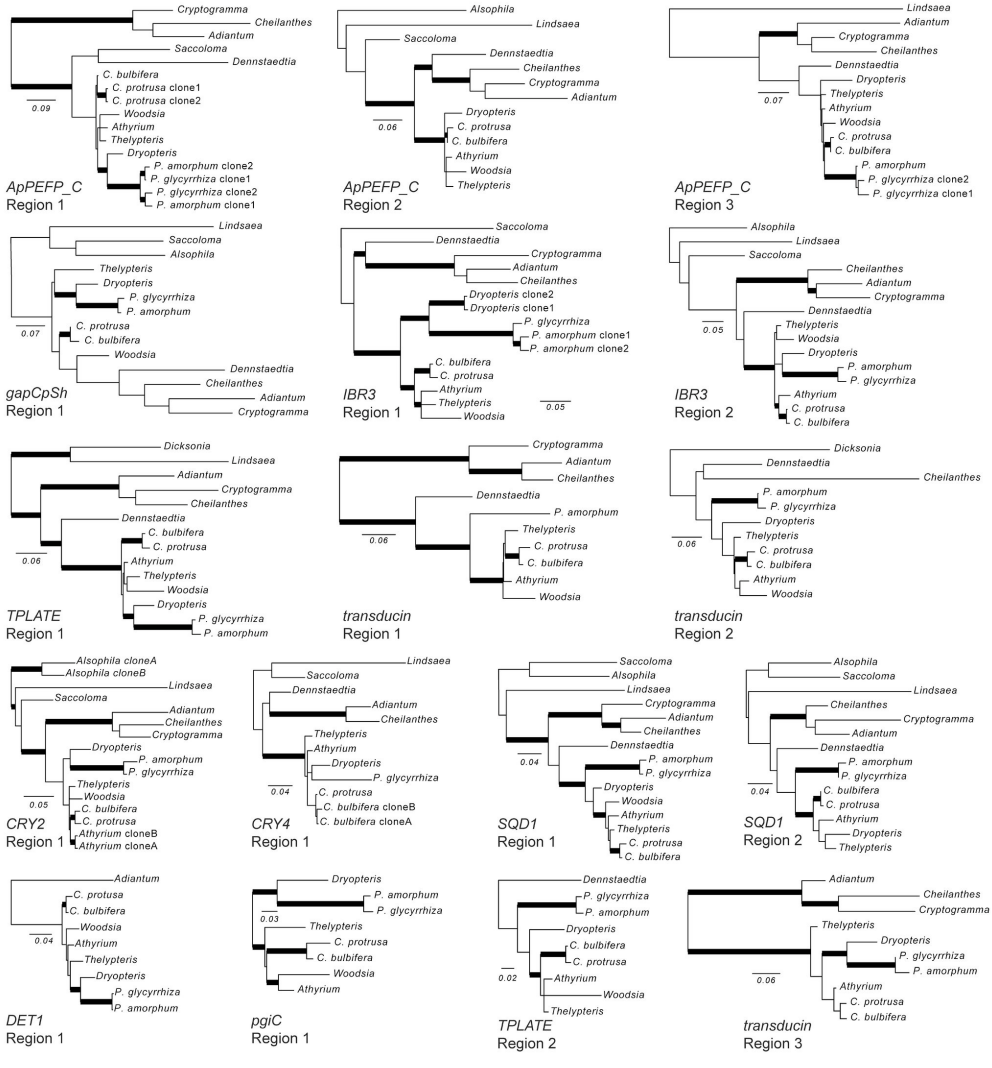

Figure 2

Figure 2. Maximum likelihood phylograms for each region, including only those taxa that were successfully sequenced from our 15-taxon genomic DNA test set. Bold branches indicate strong support (≥70% bootstrap support). Scale bars are in units of substitutions per site. In the taxon names, "*C.*" and "*P.*" refer to *Cystopteris* and *Polypodium*, respectively. These phylograms are unrooted, but oriented as if rooted by the Cyatheales (or our best guess, when the Cyatheales accession did not sequence successfully), when space permits.

doi: 10.1371/journal.pone.0076957.g002

Table 4. Model comparison, by the Bayesian Information Criterion (BIC).

					Support Values by Branch (ML bootstrap percent)											
Model	**lnL**	**BIC**	**Subsets**	**Param.**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
1: Pos. & Locus	-40574.4	83316.0	30	238	56	**100**	53	**100**	**100**	**100**	**100**	**99**	**100**	**100**	**87**	58
2a: Locus (each)	-42767.5	86819.0	19	141	49	**100**	**74**	**100**	**100**	**100**	**100**	**96**	**100**	**100**	**88**	61
2b: Locus (scheme)	-42748.0	86470.3	11	107	52	**100**	**72**	**100**	**100**	**100**	**100**	**97**	**100**	**100**	**87**	67
3: Pos.	-40875.4	82370.0	4	68	55	**100**	41	**100**	**100**	**100**	**100**	**98**	**100**	**100**	**82**	42
4: Unpartitioned	-43190.2	86717.3	1	37	66	**100**	61	**100**	**100**	**100**	**100**	**95**	**100**	**100**	**80**	41

Values in bold face indicate strong support (≥70%). Branch designations (A – L) refer to Figure 3. Model 1 is the best *PartitionFinder* scheme given each codon position, for each locus, as the data blocks. In model 2a each locus gets its own partition, across codon positions. Model 2b is the best *PartitionFinder* scheme given the loci as the data blocks. Model 3 is partitioned by codon position, across loci. Model 4 is not partitioned. For substitution model parameterization, see Appendix S2. Subsets = the final number of subsets ("partitions") for that model. Param. = number of free parameters.

doi: 10.1371/journal.pone.0076957.t004

amplification and direct-sequencing of the two remaining accessions (*Alsophila* and *Polypodium amorphum*). Primers designed for Region 2 resulted in the successful sequencing of all taxa in our test set, except *Woodsia*.

TPLATE. We designed primers for two regions of *TPLATE* (Figure 1I). Our primers for Region 1 were highly successful (only *Saccoloma* failed to amplify). It spans part of exon 2, all of intron 2, and part of exon 3, ranging in length among taxa in our test set from 650-720bp. Region 1 had moderate levels of variation (16 differences for the *Cystopteris* species pair, and 15 for *Polypodium*; Table 3). Region 2 is 400-550bp long (Table 3), and slightly less variable than Region 1. Its primers are situated in exon 5 and exon 6 and span intron 5 (Figure 1I). We managed to sequence this region for 11 of the 15 test set taxa (*Adiantum, Cryptogramma, Lindsaea*, and *Saccoloma* were unsuccessful). In our phyogenetic analyses of these data, we had to exclude *Cheilanthes* and *Dicksonia* because they were too divergent from the other taxa to align confidently.

Transducin. For *transducin* we designed primers for three regions. Region 1 extends from the 5' end of intron 2 through to the middle of exon 3 (Figure 1J), and is approximately 400-450 bp long (Table 3). Amplification and sequencing was successful for 10 of our 15 test set species (*Lindsaea, Saccoloma, Dicksonia, Dryopteris*, and *Polypodium glycyrrhiza* were unsuccessful). Region 2 is approximately 550bp long, spanning exons 3 to 5 (Figure 1J, Table 3). It was amplified and sequenced successfully for 11 of the test set species (*Adiantum, Cryptogramma, Lindsaea*, and *Saccoloma* failed). Region 3 is 250-300 bp long, amplifying intron 6 and portions of exons 6 and 7 (Figure 1J, Table 3). It was successfully amplified and sequenced for all test set taxa except *Lindsaea* and *Saccoloma*.

Model selection and the Polypodiales phylogeny

The combined alignment of our 19 newly developed regions (*SQD1* Region 1 and Region 1a were merged for these analyses) across our 15-taxon Polypodiales genomic DNA test set (the set of genomic DNAs that we used to test our new primer sets) is 9007 base pairs long; 42 percent of the sites are variable. Twenty-eight percent of the characters in this alignment are missing (i.e., gaps or question marks). We investigated five models for these data (where "model" refers to the product of the partitioning scheme and the substitution model applied to each subset of the data), which ranged from one subset and 37 free parameters to 30 subsets and 238 free parameters (see Methods; Table 4; Appendix S2). In their extremes, these models differed by over 2600 in their log likelihood scores, and by nearly 4500 Bayesian information criterion (BIC) points (Table 4). Model selection based on the BIC favored a relatively simple model for these data: four data subsets corresponding to the three codon positions and the noncoding sites, respectively, with the first two codon positions optimized under a GTR+G model and the two other subsets including an additional proportion invariant parameter (GTR+I +G; Table 4; Appendix S2). No parameters, other than relative branch lengths, were linked across partitions.

Model parameterization had strong effects on the fit to our data and on our subsequent inference. In general, the models without a codon-position component to their partitioning schemes (the unpartitioned model—model 4, and the two models partitioned by locus—models 2a and 2b) performed very poorly. The addition of codon-position-based partitions dramatically improved model fit (Table 4), such that the subsequent addition of locus-based partitions resulted in a decline in model fit. For example, the BIC favored the simple by-position partitioning scheme (model 3: four subsets, 68 free parameters) over the best by-position-and-locus scheme (model 1: 30 subsets, 238 free parameters; Table 4).

Model choice impacted the ML estimate of topology, but only slightly: model 2a resolved *Lindsaea* as sister to the rest of the Polypodiales, whereas all other models put *Saccoloma* in that position. However, model choice had a stronger effect on support values. The most extreme example of this effect was branch C (Figure 3), which ranged from 41 percent ML bootstrap support under model 3 (our best-fitting model) to 74 percent support under model 2a (our worst-fitting model; Table 4).

Despite the high proportion of missing data, ML analyses of this alignment under our best-fitting model yielded a well supported phylogeny, with only three branches lacking strong support: Branch A (the earliest divergence in the Polypodiales), Branch C (the position of *Dennstaedtia* with respect to Pteridaceae and the eupolypods), and Branch L (the

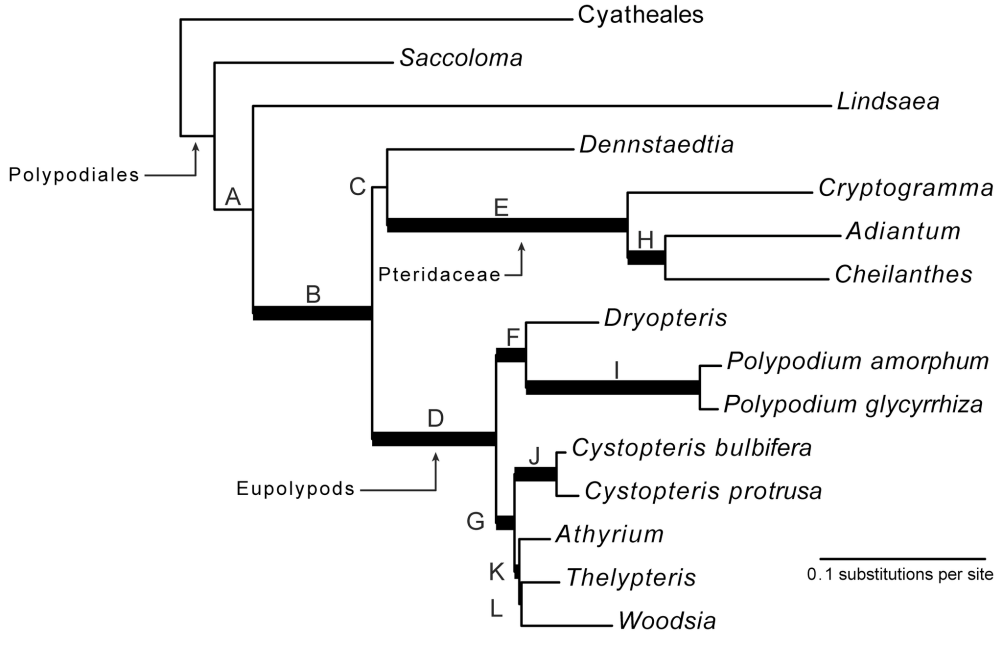

Figure 3

Figure 3. Combined data maximum likelihood phylogram of our 15-taxon genomic DNA test set. Analyses were performed under our best-fitting model (model 3, see Table 3). Bold branches indicate strong support (≥70% bootstrap support); internal branches are labeled A – L for ease of discussion.

doi: 10.1371/journal.pone.0076957.g003

relationships among the three non-Cystopteridaceae eupolypod II accessions; branch labels refer to Figure 3). The transcriptome alignments themselves also contain rich phylogenetic data. The analysis of these data is beyond the scope of this paper, and is the focus of a forthcoming manuscript.

Discussion

Single-copy locus identification, and alignment inference

Our approach to single-copy locus identification and development was highly effective, albeit labor-intensive. Specifically, we combined repeated rounds of sequence-merging and tree-building for each of our candidate loci. This approach allowed us to build compact matrices (long reads for each accession with relatively little missing data) despite the fragmentary nature of the source assemblies. Our hands-on method of alignment development and curation also allowed us to identify novel gene duplication events, to distinguish among paralogs, and to detect both contaminants and misidentifications. We are therefore confident that our final alignments are both of high data quality (data density, taxon representation, and alignment inference quality) and high accuracy (free from contaminants, inter-paralog chimaeras, etc.; see Figures S1-S9). This approach was only possible given the modest amount of data that we worked with (the "moderate data approach", see 6), and would not scale to large genomic datasets [93,94].

Nuclear genes with newly designed primer sets

ApPEFP_C is poorly characterized and does not appear to have a history as a phylogenetic marker; in *Arabidopsis* it is described simply as an "appr-1-p processing enzyme family protein" [95]. *ApPEFP*, formerly thought to be single-copy across much of the green plants (1KP data, Norman Wicket, pers. comm.), appears to have duplicated early within leptosporangiate ferns (Figure S1). We designated the pre-

duplication version *ApPEFP_A* and the two post-duplication copies *ApPEFP_B* and *ApPEFP_C*, respectively. The *ApPEFP_B*/*C* duplication may have taken place early within leptosporangiate diversification; *Dipteris* and *Lygodium* each have both copies. Further sampling (particularly of the Osmundales) will be necessary to refine the timing of this duplication.

There appears to be an additional duplication of *ApPEFP_A* in the Equisetales, and probably at least one other duplication in the Marattiales (Figure S1). The *ApPEFP_B* phylogeny is well resolved, with no additional apparent duplications in this paralog. *ApPEFP_C* was the best represented in our transcriptome sampling, and is the only copy that we pursued for primer generation. Within *ApPEFP_C* there is an apparent duplication in the Polypodiaceae. This duplication occurred in the ancestry of *Polypodium*, after the divergence of *Phlebodium* and *Pleopeltis* (Figure S1).

CRY2 and ***CRY4*** are members of the cryptochrome family of blue light photoreceptors, a gene family known in both prokaryotes and eukaryotes. In *Arabidopsis*, cryptochromes are responsible for circadian clock entrainment, flower induction, and de-etiolation [96,97]. Their function in ferns is not entirely clear, although some copies may be involved in inhibition of spore germination under blue light [98]. There are three cryptochrome copies in *Arabidopsis* [97], and five copies described in *Adiantum capillus-veneris* [98]. In our data, we recover these five copies (which we denote as *CRY1* through *CRY5*) from the majority of the polypod fern transcriptomes (Figure S2). The gene family appears to have evolved via an initial duplication on the fern stem lineage, producing the ancestral *CRY1*/*2* and *CRY3*/*4* paralogs. CRY5 originated around this time, too, perhaps from duplication of the CRY1/2 paralog. Two additional duplications followed, producing *CRY1* and *CRY2* on the stem branch of Cyatheales + Polypodiales, and *CRY3* and *CRY4* on the stem branch of Polypodiales, after the divergence of Cyatheales (Figure S2).

The first intron of fern *CRY2* is currently being developed as a phylogenetic marker in *Deparia* (Li-Yaung Kuo pers. comm.) and *Adiantum* (Wanyu Zhang pers. comm.). We designed our primer pair to target the third exon instead, and found that it recovers the corresponding region from both *CRY2* and *CRY4* for most of our test set.

The ***DET1*** protein is an important regulator in the ubiquitin-proteasome system as part of the CDD (COP10-DET1-DDB1) complex. It also has been found to be a transcriptional co-repressor recruited to target genes by specific transcription factors [99]. The gene appears to be single copy in polypod ferns (Figure S3) and is present in other eukaryotes, including humans [100]. Of all the nuclear regions for which we designed primers, DET1 is the most conserved (Table 3).

GapCpSh is a member of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene family and is one of the most frequently used nuclear loci in ferns, following the pioneering work of Ebihara et al. [63] and Schuettpelz et al. [64]. Land plants have four deeply divergent GAPDH genes—*gapA, gapB, gapC*, and *gapCp*—each of which is nuclear encoded. The first two are originally of mitochondrial origin, and the latter two were plastid encoded prior to their relocation to the nucleus [101-104]. Although we used only fern *gapCp* sequences as queries to build our all-in transcriptome alignments, our pool of transcriptome hits included representatives of each of the four main copies, as well as a fifth clade of uncertain identity (Figure S4). This myster copy appears to be a member of the GAPDH family (it is readily alignable to other members of the family and all our sequences in this clade have well-characterized members of the GAPDH family as their closest blast hits), but is deeply divergent from the known copies. It appears to be most closely related to the *gapC/gapCp* copies, but diverged from their ancestor prior to the *gapC*/*gapCp* duplication event (Figure S4a). Our transcriptome hits included a good representation of this mystery copy from across the Polypodiales, with an additional hit in *Anemia* (Schizaeales). Presumably it has been either lost from other ferns or was transcribed at insufficient levels to be captured in many of our source transcriptomes.

GapA and *gapB* are very poorly represented in our blast hits, as might be expected, given their phylogenetic distance from our query sequences. *GapC* sequences, however, are well represented, with a broad sample of sequences from the Ophioglossales and Polypodiales, and sparser representation from the Cyatheales (Culcita) and Salviniales (*Azolla* and *Pilularia*). Within the *gapC* portion of the phylogeny (Figure S4a), species are generally in their expected phylogenetic position, with two main exceptions. The first is the position of *Culcita* (a member of the Cyatheales) within the Pteridaceae (in the Polypodiales). This *Culcita* sequence, however, is very short (128 bp) and its position is likely an artifact due to limited data. The second irregularity is more difficult to explain: a clade of three Pteridaceae sequences is effectively sister to the rest of the leptosporangiate sequences and far from the Pteridaceae. The relationship among these three sequences corresponds with their expected species relationships, and each of the three species also has a "good" *gapC* sequence in the appropriate phylogenetic position. These three anomalous sequences may represent an otherwise uncaptured *gapC* duplication early in the leptosporangiate fern evolution.

The position of the Equisetales sequences is also ambiguous. Each of the two *Equisetum* accessions (*E. diffusum* and *E. hyemale*) has multiple *gapC/Cp* type sequences, but they fall together in a clade that is resolved in our maximum likelihood (ML) analyses as sister to the fern + seed plant *gapCp* clade, rather than in separate *gapC* and *gapCp* clades (Figure S4). Based on this result, we tentatively treat them as *gapCp* copies, with an *Equisetum*-specific *gapCp* duplication. Consistent with the results of Schuettpelz et al. [64], we recovered three main *gapCp* copy types in the ferns: a pre-duplication copy, and a duplication in the leptosporangiates forming *gapCp* "*short*" (*gapCpSh*) and *gapCp* "*long*" (*gapCpLg*). Schuettpelz et al. [64] hypothesized that the *gapCpSh*/*Lg* duplication event occurred near the base of the Polypodiales, or possibly more deeply (with subsequent losses, based on their sampling). Our transcriptome data suggest that the duplication very likely occurred at a point after the divergence of the Hymenophyllales, Gleicheniales, and Schizaeales, but prior to the divergence of the Salviniales, from the remaining leptosporangiates (Figure S4). Within the *gapCp*

clade there is one group of sequences that is difficult to reconcile with the organismal phylogeny: a clade of five Cyatheales sequences (three from *Thyrsopteris*, and one each from *Plagiogyria* and *Culcita*) that appear to have diverged before the *gapCpSh*/*Lg* duplication (Figure S4b). The three species represented also have "good" *gapCpSh* and *gapCpLg* sequences, so it is unclear what paralog this anomalous clade represents.

GapCpLg is represented in our transcriptome sample by a single Salviniales sequence (*Pilularia*), and by sequences from the majority of our sampled species of Cyatheales and Polypodiales. The phylogeny of these sequences is consistent with the currently accepted fern topology [8,20], and does not show any indication of subsequent duplication. *GapCpSh* is even better represented, with sequences from both *Pilularia* and *Azolla*, plus broad representation across Cyatheales and Polypodiales. As with the *Adiantum*-specific *gapCpSh* duplication found by Rothfels and Schuettpelz [49], the two *Astrolepis*-specific duplications found by Beck et al. [42], and the *gapCp* duplication documented in the evolution of *Arabidopsis* [105] our data suggest at least two more duplications of *gapCpSh*: one in a common ancestor of *Culcita* and *Plagiogyria*, and another in the Lindsaeaceae (Figure S4b).

IBR3 has not been previously used as phylogenetic marker. It is related to acyl-CoA dehydrogenases and, while its subcellular location has not been confirmed, it contains a peroxisomal targeting sequence and likely is localized to that organelle [95,106]. *IBR3* appears to be present as a single copy throughout the fern tree, and is thought to be single copy across land plants (1KP data; Norman Wicket, pers. comm.). One possible exception in our data is in the Psilotaceae, where there may be a duplication (Figure S5).

PgiC is one of the most extensively used nuclear markers in ferns (e.g., [57,59,66,71-75]). It also has a history in angiosperm phylogenetics, e.g., [107,108], was one of the most frequently used enzymes in allozyme studies, e.g., [109,110], and is single-copy in ferns [71] (Figure S6). The gene codes for phosphoglucose isomerase, an enzyme active in the glycolysis of glucose-6-phosphate isomerase [95]. In our phylogenetic analyses, we excluded the *Dicksonia* sequence because it was too divergent from the other taxa to align confidently.

The ***SQD1*** gene encodes a protein required for synthesis of sulfoquinovosyldiacylglycerol (SQDG), a well-characterized sulfolipid found in chloroplast membranes, and is widely distributed across land plants, green algae, and cyanobacteria [111]. It is hypothesized that *SQD1* permits proper functioning of photosystem II under phosphorous limited conditions [112]. Studies utilizing Southern hybridization demonstrated that *SQD1* is single-copy in *Arabidopsis thaliana* and the chlorophyte algae *Chlamydomonas reinhardtii* [113,114]. In silico analysis of fully annotated genomes indicated that *SQD1* is also present as a single copy in *Oryza sativa* and *Populus trichocarpa*, prompting the development of angiosperm-specific primers [115,116]. Additional genomic analyses confirmed single copies of *SQD1* in *Physcomitrella patens, Selaginella moellendorffii, Vitis vinifera, Zea mays*, and *Sorghum bicolor* [117]. Our study suggests that *SQD1* is a single copy gene for the majority of fern taxa (Figure S7). A notable exception is the presence of an apparent duplication in an ancestor of the Marattiaceae. Several other more-recent duplications have occurred in isolated genera or species such as *Lindsaea, Culcita macrocarpa* and *Nephrolepis exaltata*. Notably, our *Ophioglossum* (Ophioglossales) *SQD1* sequence is resolved as sister to *Lygodium* (Schizeaceae), a position incompatible with the current, accepted understanding of fern phylogeny [20]. It is possible that this is an alignment artifact. We do not suspect contamination, because an *Ophioglossum* + *Lygodium* clade is not recovered in phylogenies of any of other loci in this study.

TPLATE has been identified as a cytokinesis protein involved in the formation of the cell plate [95,118]. Van Damme et al. [119] have also shown that it is important for the formation of viable pollen. It is a member of the group of putatively single-copy markers identified by the 1KP project (Norman Wicket pers. comm.), and to our knowledge has not previously been used as a phylogenetic marker. It is single-copy in our transcriptome sample except for possible duplications in the Ophioglossaceae (Figure S8).

The function of the ***transducin*** protein in plants is not well described. It belongs to the G-protein complex, which is involved in signaling across the cell membrane [120]. In *Arabidopsis* this complex is thought to be involved in the export and import of mRNA and protein to the nucleus [95]. It is a member of the group of putatively single-copy markers identified by the 1KP project (Norman Wicket pers. comm.), and is single-copy in our sample (Figure S9). To our knowledge it has not previously been used as a phylogenetic marker.

Model selection and the Polypodiales phylogeny

The strong improvement in fit to our data that is provided by selecting a model with codon-position based partitions, and the correspondingly weak (or negative) contribution of locus-based partitions, is consistent with other studies [121-124]. This result both emphasizes the importance of including codon position information in model selection procedures, and suggests that our loci share organismal histories: the absence of strong by-locus effects on model fit suggests congruence among the gene trees. Also notable is the strong effect of model choice on ML bootstrap support levels (Table 4). Each of our five models was the best of its "class," in the sense that each represented the optimal parameterization for the chosen partitioning scheme (by the Akaike information criterion—AIC), and each was at least moderately parameterized (had a minimum of 37 free parameters; Table 4). One might thus naïvely expect that these models, on the same data, would perform similarly. Instead, they resulted in the inference of quite divergent levels of bootstrap support for some nodes (up to a 33 percentage point difference in support; Table 4). Interestingly, the poorer-fitting models tended to find higher levels of support for both branch C of Figure 3 (a branch that was unsupported or weakly supported in earlier phylogenetic investigations [7,8,16]) and for branch L (a branch that is inconsistent with the strongly supported—74 percent ML bootstrap support and 1.0 posterior probability—results of Rothfels et al. [6]). The importance of adequate partitioning schemes for accurate phylogenetic

inference has been long acknowledged [125,126], and our results mirror those of other recent empirical studies that found strong effects—both on inference of topology and support levels—of partitioning methods [122,123,127-129].

The resulting phylogenetic conclusions under our best-fitting model (model 3; see Table 4; Figure 3; Appendix S2) are comfortingly consistent with earlier results (e.g., [7,8,90]). The nine branches that are highly supported in our analyses have been inferred with high support in earlier studies, and the three branches that lack support in our data likewise have historically resisted resolution [7,8,16]. The sole exception to this pattern is the relationship among the three non-Cystopteridaceae members of the eupolypods II, which our data do not support (branch L in Figure 3), but which Rothfels et al. [6] were able to resolve with strong support (using much denser taxon sampling).

Conclusions

The 1KP fern transcriptomes provide a powerful means to generate new single-copy nuclear regions for use by evolutionary biologists. The 20 primer pairs presented here (amplifying regions across 10 protein-coding genes) more than triple the number of such regions available for ferns. Moreover, across most of our Polypodiales genomic DNA test set (Appendix S1), the majority of these primer pairs yield PCR products that can be directly sequenced. Our sample spans the phylogenetic breadth of the Polypodiales, which includes approximately two thirds of extant fern species. Our test set, however, was focused on diploid species; researchers working with polyploids, questions of hybridization, or heterozygous individuals will need to clone their PCR products.

These newly available markers vary in their degree of variation and phylogenetic informativeness at a range of evolutionary depths (see Table 3, Figures 2, 3). In combination they yield the first broad multi-gene nuclear phylogeny for ferns. This phylogeny features strong levels of support, is consistent with the results of earlier studies, and thus provides critical evidence for the general consistency of inferences from these two genomic compartments.

For researchers working on groups outside of the Polypodiales (or those with a narrower focus within the Polypodiales), our new primers may not be directly applicable, but serve instead as a proof of concept. For these researchers, our fern-wide "all-in" alignments (see Figures S1-S9; TreeBASE accession number S14616) will provide an opportunity to design primers for their study group of choice, regardless of the position of that group within the fern phylogeny.

Methods

Extracting transcriptome sequences of interest and creating "all-in" alignments

As of January 2013, the 1KP project (www.onekp.com) had sequenced 62 fern transcriptomes, spanning the deepest branches in the fern phylogeny. RNA extraction protocols used here varied [130] although we found that the Spectrum Total Plant RNA Kit (Sigma-Aldrich, St. Louis, Missouri, U.S.A.) was effective for use with ferns. The sequencing was performed on Illumina's GAIIx (earlier samples) or HiSeq (later samples) sequencing platforms at BGI-Shenzhen, and the 2x75 bp (GAIIx) or 2x90 bp (HiSeq) paired-end reads were assembled with *SOAPdenovo* (http://soap.genomics.org.cn/soapdenovo.html [131]) and *SOAPdenovo-trans* (http://soap.genomics.org.cn/SOAPdenovo-Trans.html); for further details on RNA extractions, transcriptome sequencing, and assembly, see Johnson et al. [130]. We took a top-down approach to finding single-copy loci in the transcriptome data. Potential single-copy loci were first selected based on personal interest or from a list of markers (generated by the 1KP project) that are putatively single-copy across a broad sample of land plants (Norman Wickett, pers. comm.). Subsequently, for each of these loci we used a combination blast [132] and tree-searching approach (Figure 4), which allowed us to confirm that the loci were single-copy (in the transcriptome data), and to focus on those with particularly good representation in the transcriptomes available to us.

We inferred fern-wide alignments for our candidate loci using one of two broad approaches (Figure 4). The first utilized the python script *Blue Devil* v0.6 [133], which detects the longest open reading frames (ORFs) in a series of query sequences, blasts those ORFs against a pool of transcriptome assemblies and provides a *MUSCLE*-based [134] alignment of the resulting hits. *Blue Devil* provides the options of using either *blastn* or *tblastx* [135], of varying the blast significance cut-off values, and of using *CAP3* [136] to re-assemble the blast hits prior to producing the alignment. *CAP3* was particularly useful in our pipeline because it allowed the *SOAPdenovo* and *SOAPdenovo-trans* assemblies of each transcriptome to be assembled together into one "master" assembly.

Our second main approach to producing transcriptome alignments was based on a nested series of blast searches using *lasseblaste* [137]. This script takes a series of query sequences as input (we used the entire pool of putatively single-copy markers listed by the 1KP project; Norman Wicket pers. comm.) and blasts each of these sequences against the pool of transcriptomes. It then takes the resulting hits and blasts them back to the full transcriptomes. From this final pool of hits, *lasseblaste* utilizes *MAFFT* [138] to produce a separate alignment of the hit sequences obtained for each query sequence and provides an accompanying quality score. The scoring system rewards alignments that have broad representation across the included transcriptomes, indicating good taxon coverage and penalizes alignments that have many hits per transcriptome, suggesting multiple paralogs and/or short read lengths. We selected five of the top 10 best-scoring of these alignments to pursue for primer design.

Regardless of whether we used *Blue Devil* or *lasseblaste* to infer the initial alignment, we subsequently refined that alignment manually, in an iterative manner. First, we inferred a preliminary phylogenic tree from that alignment using maximum parsimony (MP) in PAUP* v4.0a125 [139]. Groups of discontinuous (or slightly overlapping) sequences from a given accession that appeared closely related in the resulting tree and did not have any conflicts with each other were combined

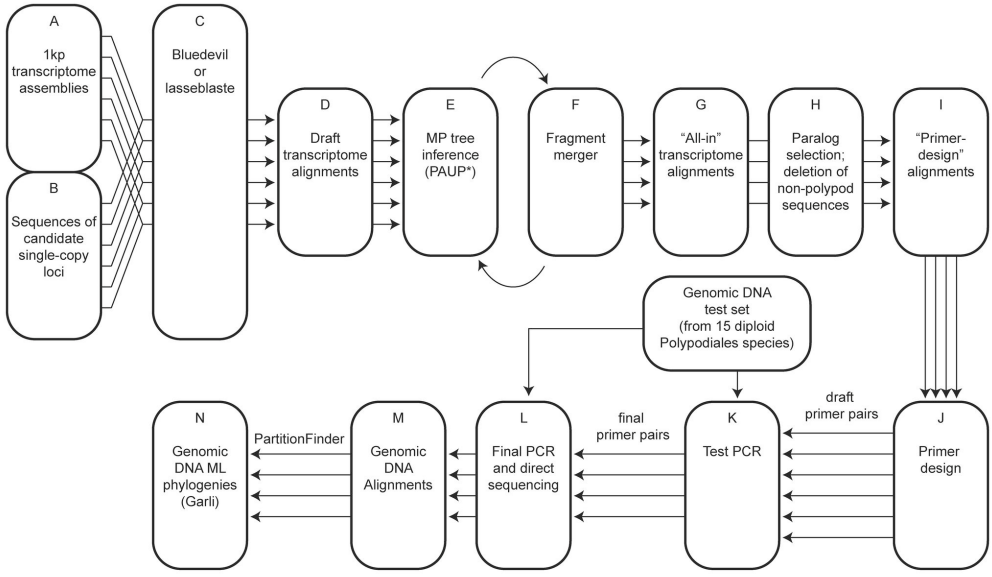

Figure 4

Figure 4. Flowchart of our transcriptome-mining pipeline.

doi: 10.1371/journal.pone.0076957.g004

into a single sequence in Mesquite v2.75 [140] (Figure 5). We then repeated the MP analyses on this new alignment. The resulting tree had fewer terminals, and was inferred from longer average sequences, and so provided greater power to place previously uncertain fragments. We continued this "infer-tree, group-sequences" approach until no further fragments met our criteria for merging. This process allowed us to produce an alignment with minimal missing data, and to effectively distinguish among paralogs. The final alignments generated in this way are referred to as our "all-in" alignments (see TreeBASE study number S14616).

Despite our targeting putatively single-copy genes, some of the transcriptome queries returned a variety of paralogs. In these cases, our sequence pools occasionally included two or more sequence fragments of different paralogs from a single individual taxon, where it was unclear which fragments belonged together. For example, an accession might have two fragments from the 5' end of the protein that conflict with each other, and two conflicting sequences from the 3' end, without any indication of which one of the 5' sequences corresponds to which of the 3' sequences. In this case, we created two sequences by merging the non-conflicting fragments arbitrarily (Figure 5). All sequence variation is thus preserved for primer generation purposes, but the resulting sequences may be chimeras, and their fine-scale phylogenetic relationships incorrect.

We inferred a final phylogenetic tree from each all-in alignment by ML, using *Garli* 2.0 [141], under the best-fitting model and partitioning scheme as determined by *PartitionFinder* v1.0.1 [124]. In each case we designated three data blocks (one for each codon position), and used *PartitionFinder* to evaluate all partitioning schemes, with the best selected according to the AIC. The subsequent tree searches (in *Garli*) were each run ten times, independently, from different random addition starting trees (see Figures S1-S9).

Polypod-only alignment and primer design

From each all-in alignment we identified the copy (if multiple paralogs were present) that included the best representation of polypod sequences, and extracted those sequences to produce a new, polypod-only alignment. To this alignment we added the related *Arabidopsis thaliana* genomic DNA and cDNA sequences based on blast searches of TAIR [95] using Mesquite's pair-wise alignment tool with a high gap-opening penalty (40). We were able to use the comparison of

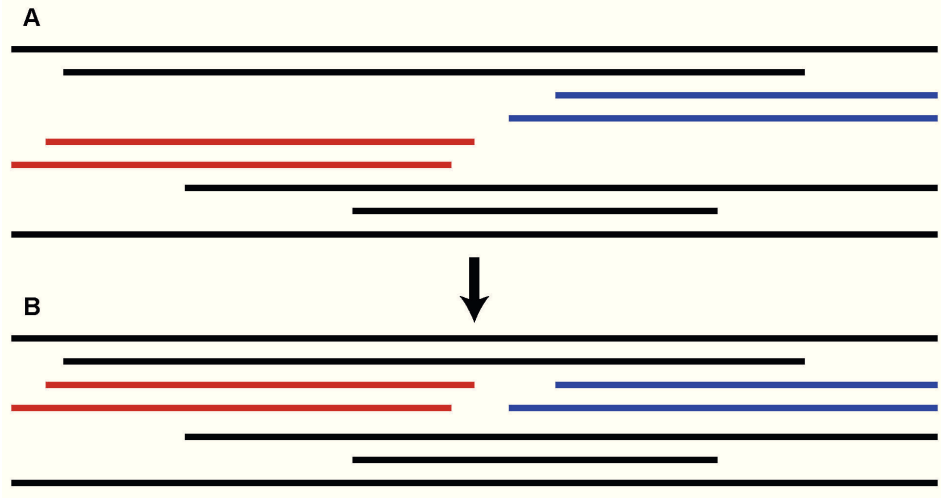

Figure 5

Figure 5. Example of our sequence-merging protocol. (A) In this schematic of a transcriptome alignment, aligned sequence fragments are indicated by the horizontal bars. Included are four fragments (colored) from our focal accession, which group together in the maximum parsimony tree. However, the two fragments from the 5' end of the protein (in red) have some base pair conflicts with each other, as do the fragments from the 3' end (in blue). Since the two sets of fragments do not overlap, and they group in the same area of the MP tree, it is not possible to determine which 5' fragment belongs with which 3' one. In this case we merged the sequences arbitrarily (B). The resulting alignment retains the full nucleotide data for primer-design purposes, but the relationships at the tips of the tree may be erroneous due to the two potentially chimaeric sequences.

doi: 10.1371/journal.pone.0076957.g005

Arabidopsis genomic and cDNA sequences to estimate the location of exon-intron boundaries in the fern transcriptome sequences. In cases where the exact beginning and end of the *Arabidopsis* introns were ambiguous, we refined the boundaries to match known exon-intron boundary sequence signatures as closely as possible (e.g., see 142).

The resulting alignments are our "primer-design" alignments —they contain all available information for our taxonomic target (the Polypodiales) for each region of choice. Using the primer-design alignment we searched for conserved sites for primer design. Each primer pair was checked for hairpins, melting point, self-dimers, and hetero-dimers with Integrated DNA Technologies' *OligoAnalyzer* v3.1 (http://www.idtdna.com/analyzer/applications/oligoanalyzer/).

Amplification of genomic DNA and sequence characterization

Primer pairs were assayed against the test set of genomic DNA from 15 fern taxa, spanning the major polypod divergences (Appendix S1). PCR conditions followed published protocols [143] with two adjustments: (1) We incorporated one additional microliter of each primer (to compensate for primer degeneracy) and (2) reduced the volume of water by two microliters (to keep reaction size constant). Total reaction size was 21 microliters. The initial PCRs were performed across a temperature gradient, with the final optimal thermocycling conditions listed in Table 2.

For each region that amplified consistently (produced strong single bands for the majority of the test genomic DNAs), we purified and direct sequenced the products following established protocols [6,8]. For high priority targets that gave poor sequencing results, we cloned the PCR products following established protocols [64], and sequenced them as listed in Table 2. For the cryptochrome loci (*CRY2* and *CRY4*), the PCR products were gel-extracted using the QIAquick Gel Extraction Kit (QIAGEN Inc., Gaithersburg, MD) prior to cloning. We aligned the resulting sequences by hand or *MAFFT* [138] and used *Garli* v2.0 [141] to infer the best ML phylogenetic tree under a GTR+I+G model. Support was assessed via 1000 bootstrap pseudoreplicates, with each bootstrap tree search performed twice, from different random addition starting trees (Figure 2).

Due to the breadth of our taxon sample, much of the intron data could not be unambiguously aligned and thus were excluded prior to tree-searching, which reduced our ability to

assess the utility of these markers at shallower phylogenetic depths. To overcome this weakness, we chose two pairs of closely related species (*Cystopteris bulbifera* and *C. protrusa* and *Polypodium amorphum* and *P. glycyrrhiza*) to provide metrics for the variability of each region. For each species pair we computed the total number of base differences between the sequences of the two species (with each indel counted as a single "difference" regardless of its length) for each region (Table 3). All newly generated genomic sequences are available in GenBank (Appendix S1).

Polypodiales combined data phylogeny

To demonstrate the utility of our markers across various phylogenetic depths (the earliest divergences in the Polypodiales occurred approximately 190 million years ago [144]) and to attempt to resolve polytomies in the backbone of the Polypodiales phylogeny [8,20] we combined the genomic DNA alignments for our loci and inferred their phylogeny by ML. Some of the locus alignments contained multiple sequences for individual accessions (representing paralogs, or allelic variation; see Figure 2). In these cases, the longest sequence was retained. In the event of a predicted duplication affecting multiple accessions, the copy that had the greatest average length was kept, rather than the longest sequence within each copy. The resulting alignments were combined into a single alignment using *abioscripts* (available at http://ormbunkar.se/phylogeny/abioscripts/). This script produces a concatenated alignment, inserting blank characters for accessions not represented in a particular locus, while maintaining exclusion set, codon position, and character set information.

We used *PartitionFinder* v1.0.1 [124] to find the best model for the analysis of these data. We performed three *PartitionFinder* runs to investigate a spectrum of possible models (Table 4). The first had four predefined data blocks (one for each codon position, and one for the noncoding sequences), the second had 19 data blocks (one for each locus), and the third had 72 data blocks (each codon position/noncoding sequence considered separately, for each locus). For each of these three runs, we set *PartitionFinder* to find the best partitioning scheme while considering all possible substitution models (with subset-specific substitution models selected by the AIC), testing all possible schemes in the first case, and using a greedy heuristic for the latter two runs. We selected the final model (optimal partitioning scheme with accompanying substitution models for each subset) by fit, as assessed by the BIC. To this set of three models, we added two others (see Table 4). The simplest is an unpartitioned GTR +I+G model, and the more complicated is partitioned by locus, with each locus given its own best-fit substitution model (manually derived from the subset output files from the by-locus *PartitionFinder* run).

We performed ML tree searches under these five models in *Garli* 2.0 [141]. For each model we did 10 best-tree searches, from different random-addition sequence starting trees, and assessed support via 1000 bootstrap pseudoreplicates, each from a single random-addition starting tree (Table 4, Figure 3). These bootstrap runs, and other computation-intensive analyses, were run on the Duke Shared Cluster Resource (https://wiki.duke.edu/display/SCSC/DSCR).

Supporting Information

Appendix S1. Voucher data and GenBank accession numbers for our Polypodiales genomic DNA test set. Numbers in parenthesis following the species names are Fern Lab Database accession numbers (fernlab.biology.duke.edu); letters in parentheses are acronyms for the herbaria where the vouchers are deposited, from Index Herbariorum [145]. Missing data are indicated by an n-dash ("-").
(DOCX)

Appendix S2. Full description of partitioning schemes and substitution models applied for the five models investigated (1, 2a, 2b, 3, and 4). In the "Subset Contents" field for model 2a, terminal digits refer to codon position: _1= First codon position; _2= Second codon position; _3= Third codon position; _N= Non-coding sequence.
(XLSX)

Figure S1. *ApPEFP* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Figure S2. *CRY* all-in maximum likelihood transcriptome phylogeny. a) preduplication *CRY3/4*, *CRY3*, and *CRY4*; b) *CRY5*, preduplication *CRY1/2*, and *CRY2*; c) *CRY1*, and a cartoon "map" of the entire cryptochrome fern phylogeny.
(PDF)

Figure S3. *DET1* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Figure S4. *GAP* all-in maximum likelihood transcriptome phylogeny. a) *gapA*, *gapB*, mystery *gap*, and *gapC*; b) *gapCp* (including *Cp Short* and *Cp Long*), and a cartoon map of the *GAP* family phylogeny.
(PDF)

Figure S5. *IBR3* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Figure S6. *pgiC* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Figure S7. *SDQ1* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Figure S8. *TPLATE* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Figure S9. *transducin* all-in maximum likelihood transcriptome phylogeny.
(PDF)

Acknowledgements

We thank the many people responsible for the availability of transcriptomes through the 1KP project, particularly Eric Carpenter for his many contributions, the staff at BGI-Shenzhen for the transcriptome sequencing, Norman Wicket for providing the draft list of single-copy loci, and M. Barker, A. Calcedo, L. DeGironimo, M. Deyholos, N. Stewart, T. Shen, and D. Soltis for providing important fern material. The Juniper Level Botanic Garden at Plant Delights Nursery (www.juniperlevelbotanicgarden.org) generously allowed sampling of their collections for RNA extractions, as did the North Carolina Botanical Garden at Chapel Hill, Royal Botanic Gardens at Sydney, and the Duke University Live Plant Collection (http://liveplantcollections.biology.duke.edu). Computations were enabled by facilities at WestGrid, through Compute/Calcul Canada, and at Texas Advanced Computing Center, through iPlant Collaborative. Finally, we thank David Swofford for assistance with analyses, Li-Yaung Kuo for providing unpublished sequence information on fern cryptochromes, and Editor Keith Crandall and two anonymous reviewers for their comments and edits.

Author Contributions

Conceived and designed the experiments: CJR. Performed the experiments: CJR AL FWL EMS LH. Analyzed the data: CJR AL FWL EMS LH. Contributed reagents/materials/analysis tools: CJR AL EMS DOB MR DS SWG GW PK KMP. Wrote the manuscript: CJR AL FWL EMS LH MR DOB SWG PK KMP. Designed analysis tools (lasseblaste and BlueDevil): AL FWL. Aligned transcriptomes: CJR AL EMS..

References

1. Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG et al. (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409: 618–622. doi:10.1038/35054555. PubMed: 11214320.
2. Hasebe M, Wolf PG, Pryer KM, Ueda K, Ito M et al. (1995) Fern phylogeny based on *rbcL* nucleotide sequences. Am Fern J 85: 134–181. doi:10.2307/1547807.
3. Korall P, Pryer KM, Metzgar JS, Schneider H, Conant DS (2006) Tree ferns: Monophyletic groups and their relationships as revealed by four protein-coding plastid loci. Mol Phylogenet Evol 39: 830–845. doi: 10.1016/j.ympev.2006.01.001. PubMed: 16481203.
4. Pryer KM, Schuettpelz E, Wolf PG, Schneider H, Smith AR et al. (2004) Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. Am J Bot 91: 1582–1598. doi: 10.3732/ajb.91.10.1582. PubMed: 21652310.
5. Pryer KM, Smith AR, Skog JE (1995) Phylogenetic relationships of extant ferns based on evidence from morphology and *rbcL* sequences. Am Fern J 85: 205–282. doi:10.2307/1547810.
6. Rothfels CJ, Larsson A, Kuo L-Y, Korall P, Chiou W-L et al. (2012) Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. Syst Biol 61: 490–509. doi: 10.1093/sysbio/sys001. PubMed: 22223449.
7. Schuettpelz E, Korall P, Pryer KM (2006) Plastid *atpA* data provide improved support for deep relationships among ferns. Taxon 55: 897–906. doi:10.2307/25065684.
8. Schuettpelz E, Pryer KM (2007) Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. Taxon 56: 1037–1050. doi:10.2307/25065903.
9. Wikström N, Pryer KM (2005) Incongruence between primary sequence data and the distribution of a mitochondrial *atp1* group II intron among ferns and horsetails. Mol Phylogenet Evol 36: 484–493. doi:10.1016/j.ympev.2005.04.008. PubMed: 15922630.
10. Hasebe M, Omori T, Nakazawa M, Sano T, Kato M et al. (1994) *rbcL* gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. Proc Natl Acad Sci USA 91: 5730–5734.
11. Sano R, Takamiya M, Ito M, Kurita S, Hasebe M (2000) Phylogeny of the lady fern group, tribe Physematieae (Dryopteridaceae), based on chloroplast *rbcL* gene sequences. Mol Phylogenet Evol 15: 403–413. doi:10.1006/mpev.1999.0708. PubMed: 10860649.
12. Wolf PG (1995) Phylogenetic analyses of *rbcL* and nuclear ribosomal RNA gene sequences in Dennstaedtiaceae. Am Fern J 85: 306–327. doi:10.2307/1547812.
13. Wolf PG (1996) Pteridophyte phylogenies based on analyses of DNA sequences: A multiple gene approach. In: JM CamusM GibbyRJ Johns. Pteridology in perspective. Kew. Royal Botanical Gardens. pp. 203–215.
14. Wolf PG (1997) Evaluation of *atpB* nucleotide sequences for phylogenetic studies of ferns and other pteridophytes. Am J Bot 84: 1429–1440. doi:10.2307/2446141. PubMed: 21708550.
15. Wolf PG, Soltis PS, Soltis DE (1994) Phylogenetic relationships of dennstaedtioid ferns: Evidence from *rbcL* sequences. Mol Phylogenet Evol 3: 383–392. doi:10.1006/mpev.1994.1044. PubMed: 7697195.
16. 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. doi:10.1016/j.ympev.2011.03.010. PubMed: 21402161.
17. Rai HS, Graham SW (2010) Utility of a large, multigene plastid data set in inferring higher-order relationships in ferns and relatives (monilophytes). Am J Bot 97: 1444–1456. doi:10.3732/ajb.0900305. PubMed: 21616899.
18. Lehtonen S, Wahlberg N, Christenhusz MJM (2012) Diversification of lindsaeoid ferns and phylogenetic uncertainty of early polypod relationships. Bot J Linn Soc 170: 489–503. doi:10.1111/j.1095-8339.2012.01312.x.
19. Rothfels CJ, Sundue MA, Kuo L-Y, Larsson A, Kato M et al. (2012) A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales). Taxon 61: 515–533.
20. Smith AR, Pryer KM, Schuettpelz E, Korall P, Schneider H et al. (2006) A classification for extant ferns. Taxon 55: 705–731. doi: 10.2307/25065646.
21. Christenhusz MJM, Zhang X-C, Schneider H (2011) A linear sequence of extant families and genera of lycophytes and ferns. Phytotaxa 19: 7–54.
22. Schuettpelz E, Pryer KM (2008) Fern phylogeny. In: TA RankerCH Haufler. Biology and evolution of ferns and lycophytes. New York: Cambridge University Press. pp. 395–416.
23. Haufler CH, Ranker TA (1995) *RbcL* sequences provide phylogenetic insights among sister species of the fern genus *Polypodium*. American Fern Journal 85: 361-374.
24. Nagalingum NS, Schneider H, Pryer KM (2007) Molecular phylogenetic relationships and morphological evolution in the heterosporous fern genus *Marsilea*. Syst Bot 32: 16–25. doi: 10.1600/036364407780360256.
25. Metzgar J, Skog JE, Zimmer EA, Pryer KM (2008) The paraphily of *Osmunda* is confirmed by phylogenetic analyses of seven plastid loci. Syst Bot 33: 31–36. doi:10.1600/036364408783887528.
26. Des Marais DL, Smith AR, Britton DM, Pryer KM (2003) Phylogenetic relationships and evolution of extant horsetails, *Equisetum*, based on chloroplast DNA sequence data (*rbcL* and *trnL-F*). Int J Plant Sci 164: 737–751. doi:10.1086/376817.
27. Dubuisson JY, Hennequin S, Douzery EJP, Cranfill RB, Smith AR et al. (2003) *rbcL* phylogeny of the fern genus; Trichomanes (Hymenophyllaceae), with special reference to neotropical taxa. Int J Plant Sci 164: 753–761. doi:10.1016/S0168-9452(03)00060-8.
28. Rothfels CJ, Windham MD, Grusz AL, Gastony GJ, Pryer KM (2008) Toward a monophyletic *Notholaena* (Pteridaceae): Resolving patterns of evolutionary convergence in xeric-adapted ferns. Taxon 57: 712–724.

29. Labiak PH, Rouhan G, Sundue MA (2010) Phylogeny and taxonomy of *Leucotrichum* (Polypodiaceae): A new genus of grammitid ferns from the Neotropics. Taxon 59: 911–921.
30. McKeown M, Sundue MA, Barrington DS (2012) Phylogenetic analyses place the Australian monotypic *Revwattsia* in *Dryopteris* (Dryopteridaceae). Phytokeys 14: 43–56. doi:10.3897/phytokeys.14.3446. PubMed: 23170072.
31. Sundue MA (2010) A monograph of *Ascogrammitis*, a new genus of grammitid ferns (Polypodiaceae). Brittonia 62: 357–399. doi:10.1007/s12228-009-9108-6.
32. Zhang L-B, Zhang L, Dong S-Y, Sessa EB, Gao X-F et al. (2012) Molecular circumscription and major evolutionary lineages of the fern genus *Dryopteris* (Dryopteridaceae). BMC Evol Biol 12: 180. doi:10.1186/1471-2148-12-180. PubMed: 22971160.
33. Adjie B, Takamiya M, Ohto M, Ohsawa TA, Watano Y (2008) Molecular phylogeny of the lady fern genus *Athyrium* in Japan based on chloroplast *rbcL* and *trnL-trnF* sequences. Acta Phytotaxonomica Geobotanica 59: 79–95.
34. Wang M-L, Chen ZD, Zhang X-C, Lu S-G, Zhao G-F (2003) Phylogeny of the Athyriaceae: Evidence from chloroplast *trnL-F* region sequences. Acta Phytotaxonomica Sin 41: 416–426.
35. Liu Y-C, Chiou W-L, Kato M (2011) Molecular phylogeny and taxonomy of the fern genus *Anisocampium* (Athyriaceae). Taxon 60: 824–830.
36. Li F-W, Pryer KM, Windham MD (2012) *Gaga*, a new fern genus segregated from *Cheilanthes* (Pteridaceae). Syst Bot 37: 845–860. doi:10.1600/036364412X656626.
37. Windham MD, Huiet L, Schuettpelz E, Grusz AL, Rothfels CJ et al. (2009) Using plastid and nuclear DNA sequences to redraw generic boundaries and demystify species complexes in cheilanthoid ferns. Am Fern J 99: 68–72.
38. Cranfill R, Kato M (2003) Phylogenetics, biogeography, and classification of the woodwardioid ferns (Blechnaceae). In: S ChandraM Srivastava. Pteridology in the New Millennium. Dordrecht: Kluwer Publishing House Academic Publishers. pp. 25–47.
39. He L-J, Zhang X-C (2012) Exploring generic delimitation within the fern family Thelypteridaceae. Mol Phylogenet Evol, 65: 1–8. PubMed: 22877644.
40. Smith AR, Cranfill RB (2002) Intrafamilial relationships of the thelypteroid ferns (Thelypteridaceae). Am Fern J 92: 131–149. doi:10.1640/0002-8444(2002)092[0131:IROTTF]2.0.CO;2.
41. Sigel EM, Windham MD, Huiet L, Yatskievych G, Pryer KM (2011) Species relationships and farina evolution in the cheilanthoid fern genus *Argyrochosma* (Pteridaceae). Syst Bot 36: 554–564. doi:10.1600/036364411X583547.
42. Beck JB, Windham MD, Yatskievych G, Pryer KM (2010) A diploids-first approach to species delimitation and interpreting polyploid evolution in the fern genus *Astrolepis* (Pteridaceae). Syst Bot 35: 223–234. doi:10.1600/036364410791638388.
43. Grusz AL, Windham MD, Pryer KM (2009) Deciphering the origins of apomictic polyploids in the *Cheilanthes yavapensis* complex (Pteridaceae). Am J Bot 96: 1636–1645. doi:10.3732/ajb.0900019. PubMed: 21622350.
44. Kreier H-P, Schneider H (2006) Phylogeny and biogeography of the staghorn fern genus *Platycerium* (Polypodiaceae, Polypodiidae). Am J Bot 93: 217–225. doi:10.3732/ajb.93.2.217. PubMed: 21646182.
45. Schneider H, Russell S, Cox C, Bakker F, Henderson S et al. (2004) Chloroplast phylogeny of asplenioid ferns based on *rbcL* and *trnL-F* spacer sequences (Polypodiidae, Aspleniaceae) and its implications for biogeography. Syst Bot 29: 260–274. doi:10.1600/036364404774195476.
46. Hennequin S, Hovenkamp P, Christenhusz MJM, Schneider H (2010) Phylogenetics and biogeography of *Nephrolepis*–A tale of old settlers and young tramps. Bot J Linn Soc 164: 113–127. doi:10.1111/j.1095-8339.2010.01076.x.
47. Qiu YL, Cho Y, Cox JC, Palmer JD (1998) The gain of three mitochondrial introns identifies liverworts as the earliest land plants. Nature 394: 671–674. doi:10.1038/29286. PubMed: 9716129.
48. Vangerow S, Teerkorn T, Knoop V (1999) Phylogenetic information in the mitochondrial *nad5* gene of pteridophytes: RNA editing and intron sequences. Plant Biol 1: 235–243. doi:10.1111/j.1438-8677.1999.tb00249.x.
49. Rothfels CJ, Schuettpelz E (2013) Accelerated rate of molecular evolution for vittarioid ferns is strong and not driven by selection. Syst Biol. In press
50. Pryer KM, Schneider H, Zimmer EA, Banks JA (2002) Deciding among green plants for whole genome studies. Trends Plant Sci 7: 550–554. doi:10.1016/S1360-1385(02)02375-0. PubMed: 12475497.
51. Barker MS, Wolf PG (2010) Unfurling fern biology in the genomics age. BioScience 60: 177–185.
52. Nakazato T, Barker MS, Rieseberg LH, Gastony GJ (2008) Evolution of the nuclear genome of ferns and lycophytes. In: TA RankerCH Haufler. Biology and Evolution of Ferns and Lycophytes. Cambridge: Cambridge University Press. pp. 175–198.
53. Zimmer EA, Wen J (2012) Using nuclear gene data for plant phylogenetics: Progress and prospects. Mol Phylogenet Evol 65: 774–785. doi:10.1016/j.ympev.2012.07.015. PubMed: 22842093.
54. Gastony GJ, Rollo DR (1998) Cheilanthoid ferns (Pteridaceae: Cheilanthoideae) in the southwestern United States and adjacent Mexico--A molecular phylogenetic reassessment of generic lines. Aliso 17: 131–144.
55. Reid JD, Plunkett GM, Peters GA (2006) Phylogenetic relationships in the heterosporous fern genus *Azolla* (Azollaceae) based on DNA sequence data from three noncoding regions. Int J Pl Sci 167: 529–538. doi:10.1086/501071.
56. Maggini F, Marrocco R, Gelati MT, Dominicis RI (1998) Lengths and nucleotide sequences of the internal spacers of nuclear ribosomal DNA in gymnosperms and pteridophytes. Plant Syst Evol 213: 199–205. doi:10.1007/BF00985200.
57. Schneider H, Navarro-Gomez A, Russell SJ, Ansell SW, Grundmann M et al. (2012) Exploring the utility of three nuclear regions to reconstruct reticulate evolution in the fern genus *Asplenium*. J Syst Evolution 51: 142–153.
58. Kranz H, Huss V (1996) Molecular evolution of pteridophytes and their relationship to seed plants: Evidence from complete 18S rRNA gene sequences. Plant Syst Evol 202: 1–11. doi:10.1007/BF00985814.
59. Wang L, Schneider H, Wu Z, He L, Zhang X et al. (2012) Indehiscent sporangia enable the accumulation of local fern diversity at the Qinghai-Tibetan Plateau. BMC Evol Biol 12: 158. doi:10.1186/1471-2148-12-158. PubMed: 22929005.
60. Shepherd LD, Perrie LR, Brownsey PJ (2008) Low-copy nuclear DNA sequences reveal a predominance of allopolyploids in a New Zealand *Asplenium* fern complex. Mol Phylogenet Evol 49: 240–248. doi:10.1016/j.ympev.2008.06.015. PubMed: 18640280.
61. Adjie B, Masuyama S, Ishikawa H, Watano Y (2007) Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. J Plant Res 120: 129–138. doi:10.1007/s10265-006-0032-5. PubMed: 16955374.
62. Chen C-W, Kuo L-Y, Wang C-N, Chiou W-L (2012) Development of a PCR primer set for intron 1 of the low-copy gene *LEAFY* in Davalliaceae. Am J Bot, 99: e223–e225. PubMed: 22623608.
63. Ebihara A, Ishikawa H, Matsumoto S, Lin S-J, Iwatsuki K et al. (2005) Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. Am J Bot 92: 1535–1547. doi:10.3732/ajb.92.9.1535. PubMed: 21646171.
64. Schuettpelz E, Grusz AL, Windham MD, Pryer KM (2008) The utility of nuclear *gapCp* in resolving polyploid fern origins. Syst Bot 33: 621–629. doi:10.1600/036364408786500127.
65. Nitta JH, Ebihara A, Ito M (2011) Reticulate evolution in the *Crepidomanes minutum* species complex (Hymenophyllaceae). Am J Bot, 98: 1–19. PubMed: 22012924.
66. Sessa EB, Zimmer EA, Givnish TJ (2012) Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). BMC Evol Biol 12: 104. doi:10.1186/1471-2148-12-104. PubMed: 22748145.
67. Zhang R, Liu T, Wu W, Li Y, Chao L et al. (2013) Molecular evidence for natural hybridization in the mangrove fern genus *Acrostichum*. BMC Plant Biol 13: 74. doi:10.1186/1471-2229-13-74. PubMed: 23634934.
68. Lee S-J, Park C-W (2013) Relationships and origins of the *Dryopteris varia* (L.) Kuntze species complex (Dryopteridaceae) in Korea inferred from nuclear and chloroplast DNA sequences. Biochem Syst Ecol 50: 371–382.
69. Chang Y, Li J, Lu S, Schneider H (2013) Species diversity and reticulate evolution in the *Asplenium normale* complex (Aspleniaceae) in China and adjacent areas. Taxon. In press
70. Dyer RJ, Savolainen V, Schneider H (2012) Apomixis and reticulate evolution in the *Asplenium monanthes* fern complex. Ann Bot 110: 1515–1529. doi:10.1093/aob/mcs202. PubMed: 22984165.
71. Ishikawa H, Watano Y, Kano K, Ito M, Kurita S (2002) Development of primer sets for PCR amplification of the *PgiC* gene in ferns. J Plant Res 115: 65–70. doi:10.1007/s102650200010. PubMed: 12884051.
72. James KE, Schneider H, Ansell SW, Evers M, Robba L et al. (2008) Diversity arrays technology (DArT) for pan-genomic evolutionary studies of non-model organisms. PLOS ONE 3: e1682. doi:10.1371/journal.pone.0001682. PubMed: 18301759.
73. Chang H-M, Chiou W-L, Wang J-C (2009) Molecular evidence for genetic heterogeneity and the hybrid origin of *Acrorumohra subreflexipinna* from Taiwan. Am Fern J 99: 61–77. doi:10.1640/0002-8444-99.2.61.

74. Chao Y-S, Dong S-Y, Chiang Y-C, Liu H-Y, Chiou W-L (2012) Extreme multiple reticulate origins of the *Pteris cadieri* complex (Pteridaceae). Int J Mol Sci 13: 4523–4544. doi:10.3390/ijms13044523. PubMed: 22605994.
75. Juslén A, Väre H, Wikström N (2011) Relationships and evolutionary origins of polyploid *Dryopteris* (Dryopteridaceae) from Europe inferred using nuclear *pgiC* and plastid *trnL-F* sequence data. Taxon 60: 1284–1294.
76. Vogel JC, Russell SJ, Rumsey FJ, Barrett JA, Gibby M (1998) Evidence for the maternal transmission of chloroplast DNA in the genus *Asplenium* (Aspleniaceae, Pteridophyta). Bot Acta 111: 247–249.
77. Stein DB, Barrington DS (1990) Recurring hybrid formation in a population of *Polystichum* x *potteri*: Evidence from chloroplast DNA comparisons. Ann Mo Bot Gard 77: 334–339. doi:10.2307/2399548.
78. Guillon JM, Raquin C (2000) Maternal inheritance of chloroplasts in the horsetail *Equisetum variegatum* (Schleich.). Curr Genet 37: 53–56. doi: 10.1007/s002940050008. PubMed: 10672445.
79. Gastony GJ, Yatskievych G (1992) Maternal inheritance of the chloroplast and mitochondrial genomes in cheilanthoid ferns. Am J Bot 79: 716–722. doi:10.2307/2444887.
80. Lovis J (1978) Evolutionary patterns and processes in ferns. Adv Bot Res 4: 229–415. doi:10.1016/S0065-2296(08)60371-7.
81. Der J, Barker M, Wickett NJ, dePamphilis CW, Wolf PG (2011) *De novo* characterization of the gametophyte transcriptome in bracken fern, *Pteridium aquilinum*. BMC Genomics 12.
82. Wolf PG, Karol KG (2012) Plastomes of bryophytes, lycophytes and ferns. Genomics Chloroplasts Mitochondria Advances Photosynth Respiration 35: 89–102. doi:10.1007/978-94-007-2920-9_4.
83. Karol KG, Arumuganathan K, Boore JL, Duffy AM, Everett KD et al. (2010) Complete plastome sequences of *Equisetum arvense* and *Isoetes flaccida*: Implications for phylogeny and plastid genome evolution of early land plant lineages. BMC Evol Biol 10: 321. doi: 10.1186/1471-2148-10-321. PubMed: 20969798.
84. Wolf PG, Der JP, Duffy AM, Davidson JB, Grusz AL et al. (2010) The evolution of chloroplast genes and genomes in ferns. Plant Mol Biol 76: 251–261. PubMed: 20976559.
85. Salmi ML, Bushart TJ, Stout SC, Roux SJ (2005) Profile and analysis of gene expression changes during early development in germinating spores of *Ceratopteris richardii*. Plant Physiol 138: 1734–1745. doi: 10.1104/pp.105.062851. PubMed: 15965014.
86. Yamauchi D, Sutoh K, Kanegae H, Horiguchi T, Matsuoka K et al. (2005) Analysis of expressed sequence tags in prothallia of Adiantum capillus-veneris. J Plant Res 118: 223–227. doi:10.1007/s10265-005-0209-3. PubMed: 15940394.
87. Grewe F, Guo W, Gubbels EA, Hansen AK, Mower JP (2013) Complete plastid genomes from *Ophioglossum californicum*, *Psilotum nudum*, and *Equisetum hyemale* reveal an ancestral land plant genome structure and resolve the position of Equisetales among monilophytes. BMC Evol Biol 13: 8-. PubMed: 23311954.
88. Roper JM, Kellon Hansen S, Wolf PG, Karol KG, Mandoli DF et al. (2007) The complete plastid genome sequence of *Angiopteris evecta* (G Forst.) Hoffm. (Marattiaceae). Am Fern J 97: 95–106. doi: 10.1640/0002-8444(2007)97[95:TCPGSO]2.0.CO;2.
89. Gao L, Wang B, Wang ZW, Zhou Y, Su YJ et al. (2013) Plastome sequences of *Lygodium japonicum* and *Marsilea crenata* reveal the genome organization transformation from basal ferns to core leptosporangiates. Genome Biol Evolution 5: 1403–1407. doi: 10.1093/gbe/evt099. PubMed: 23821521.
90. Schuettpelz E, Schneider H, Huiet L, Windham MD, Pryer KM (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. doi:10.1016/j.ympev.2007.04.011. PubMed: 17570688.
91. Schneider H, Schuettpelz E, Pryer KM, Cranfill R, Magallón S et al. (2004) Ferns diversified in the shadow of angiosperms. Nature 428: 553–557. doi:10.1038/nature02361. PubMed: 15058303.
92. Larsson A, Windham MD, Korall P (2013) Phylogeny of *Woodsia* (Woodsiaceae). In prep.
93. Wong KM, Suchard MA, Huelsenbeck JP (2008) Alignment uncertainty and genomic analysis. Science 319: 473–476. doi:10.1126/science.1151532. PubMed: 18218900.
94. Rannala B, Yang Z (2008) Phylogenetic inference using whole genomes. Annu Rev Genomics Hum Genet 9: 217–231. doi:10.1146/annurev.genom.9.081307.164407. PubMed: 18767964.
95. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C et al. (2011) The Arabidopsis Information Resource (TAIR): Improved gene annotation and new tools. Nucleic Acids Res 40: D1202–D1210. PubMed: 22140109.
96. Möglich A, Yang X, Ayers RA, Moffat K (2010) Structure and function of plant photoreceptors. Annu Rev Plant Biol 61: 21–47. doi:10.1146/annurev-arplant-042809-112259. PubMed: 20192744.
97. Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T et al. (2011) The cryptochromes: Blue light photoreceptors in plants and animals. Annu Rev Plant Biol 62: 335–364. doi:10.1146/annurev-arplant-042110-103759. PubMed: 21526969.
98. Imaizumi T, Kanegae T, Wada M (2000) Cryptochrome nucleocytoplasmic distribution and gene expression are regulated by light quality in the fern *Adiantum capillus-veneris*. Plant Cell 12: 81–96. doi:10.2307/3871031. PubMed: 10634909.
99. Lau OS, Deng XW (2012) The photomorphogenic repressors COP1 and DET1: 20 years later. Trends Plant Sci 17: 584–593. doi:10.1016/j.tplants.2012.05.004. PubMed: 22705257.
100. Wertz IE, O'Rourke KM, Zhang Z, Dornan D, Arnott D et al. (2004) Human De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase. Science 303: 1371–1374. doi:10.1126/science.1093549. PubMed: 14739464.
101. Petersen J, Teich R, Becker B, Cerff R, Brinkmann H (2006) The *GapA/B* gene duplication marks the origin of Streptophyta (charophytes and land plants). Mol Biol Evol 23: 1109–1118. doi:10.1093/molbev/msj123. PubMed: 16527864.
102. Martin W, Cerff R (1986) Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). Eur J Biochem 159: 323–331. doi:10.1111/j.1432-1033.1986.tb09871.x. PubMed: 3530755.
103. Brinkmann H, Martinez P, Quigley F, Martin W, Cerff R (1987) Endosymbiotic origin and codon bias of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. J Mol Evol 26: 320–328. doi:10.1007/BF02101150. PubMed: 3131533.
104. Petersen J, Brinkmann H, Cerff R (2003) Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. J Mol Evol 57: 16–26. doi:10.1007/s00239-002-2441-y. PubMed: 12962302.
105. Muñoz-Bertomeu J, Cascales-Miñana B, Mulet JM, Baroja-Fernández E, Pozueta-Romero J et al. (2009) Plastidial glyceraldehyde-3-phosphate dehydrogenase deficiency leads to altered root development and affects the sugar and amino acid balance in *Arabidopsis*. Plant Physiol 151: 541–558. doi:10.1104/pp.109.143701. PubMed: 19675149.
106. Zolman BK, Nyberg M, Bartel B (2007) *IBR3*, a novel peroxisomal acyl-CoA dehydrogenase-like protein required for indole-3-butyric acid response. Plant Mol Biol 64: 59–72. doi:10.1007/s11103-007-9134-2. PubMed: 17277896.
107. Ford VS, Lee J, Baldwin BG, Gottlieb LD (2006) Species divergence and relationships in *Stephanomeria* (Compositae): *PgiC* phylogeny compared to prior biosystematic studies. Am J Bot 93: 480–490. doi: 10.3732/ajb.93.3.480. PubMed: 21646207.
108. Kamiya K, Harada K, Tachida H, Ashton PS (2005) Phylogeny of *PgiC* gene in *Shorea* and its closely related genera (Dipterocarpaceae), the dominant trees in Southeast Asian tropical rain forests. Am J Bot 92: 775–788. doi:10.3732/ajb.92.5.775. PubMed: 21652457.
109. Charlesworth D, Yang Z (1998) Allozyme diversity in *Leavenworthia* populations with different inbreeding levels. Heredity 81(4): 453–461. doi:10.1046/j.1365-2540.1998.00415.x. PubMed: 9839439.
110. Witter MS (1990) Evolution in the Madiinae: Evidence from enzyme electrophoresis. Ann Mo Bot Gard 77: 110–117. doi:10.2307/2399630.
111. Sato N (2004) Roles of the acidic lipids sulfoquinovosyl diacylglycerol and phosphatidylglycerol in photosynthesis: Their specificity and evolution. J Plant Res 117: 495–505. doi:10.1007/s10265-004-0183-1. PubMed: 15538651.
112. Mizusawa N, Wada H (2012) The role of lipids in photosystem II. Biochim Biophys-Bioenerget 1817: 194–208. doi:10.1016/j.bbabio.2011.04.008. PubMed: 21569758.
113. Essigmann B, Güler S, Narang RA, Linke D, Benning C (1998) Phosphate availability affects the thylakoid lipid composition and the expression of *SQD1*, a gene required for sulfolipid biosynthesis in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 95: 1950–1955. doi: 10.1073/pnas.95.4.1950. PubMed: 9465123.
114. Sato N, Sugimoto K, Meguro A, Tsuzuki M (2003) Identification of a gene for UDP-sulfoquinovose synthase of a green alga, *Chlamydomonas reinhardtii*, and its phylogeny. DNA Res 10: 229–237. doi:10.1093/dnares/10.6.229. PubMed: 15029954.
115. Li M, Wunder J, Bissoli G, Scarponi E, Gazzani S et al. (2008) Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. Cladistics 24: 727–745. doi:10.1111/j.1096-0031.2008.00207.x.

116. Bruni I, De Mattia F, Galimberti A, Galasso G, Banfi E et al. (2010) Identification of poisonous plants by DNA bar coding approach. Int J Leg Med 124: 595–603. doi:10.1007/s00414-010-0447-3. PubMed: 20354712.
117. Moummou H, Kallberg Y, Tonfack LB, Persson B, van der Rest Bt (2012) The plant short-chain dehydrogenase(SDR) superfamily: Genome-wide inventory and diversification patterns. BMC Plant Biol 12
118. Van Damme D, Gadeyne A, Vanstraelen M, Inzé D, Van Montagu MCE et al. (2011) Adaptin-like protein *TPLATE* and clathrin recruitment during plant somatic cytokinesis occurs via two distinct pathways. Proc Natl Acad Sci U S A 108: 615–620. doi:10.1073/pnas.1017890108. PubMed: 21187379.
119. Van Damme D, Coutuer S, De Rycke R, Bouget FY, Inzé D et al. (2006) Somatic cytokinesis and pollen maturation in *Arabidopsis* depend on *TPLATE*, which has domains similar to coat proteins. Plant Cell 18: 3502–3518. doi:10.1105/tpc.106.040923. PubMed: 17189342.
120. Neves SR, Prahlad TR, Iyengar R (2002) G protein pathways. Science 296: 1636–1639. doi:10.1126/science.1071550. PubMed: 12040175.
121. Shapiro B, Rambaut A, Drummond AJ (2005) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol 23: 7–9. doi:10.1093/molbev/msj021. PubMed: 16177232.
122. Strugnell J, Norman M, Jackson J, Drummond AJ, Cooper A (2005) Molecular phylogeny of coleoid cephalopods (Mollusca: Cephalopoda) using a multigene approach; the effect of data partitioning on resolving phylogenies in a Bayesian framework. Mol Phylogenet Evol 37: 426–441. doi:10.1016/j.ympev.2005.03.020. PubMed: 15935706.
123. Brandley MC, Schmitz A, Reeder TW (2005) Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst Biol 54: 373–390. doi:10.1080/10635150590946808. PubMed: 16012105.
124. Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol 29: 1695–1701. doi:10.1093/molbev/mss020. PubMed: 22319168.
125. Bull JJ, Huelsenbeck JP, Cunningham CW, Swofford DL, Waddell PJ (1993) Partitioning and combining data in phylogenetic analysis. Syst Biol 42: 384–397. doi:10.2307/2992473.
126. Brown JM, Lemmon AR (2007) The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst Biol 56: 643–655. doi:10.1080/10635150701546249. PubMed: 17661232.
127. Ward PS, Brady SG, Fisher BL, Schultz TR (2010) Phylogeny and biogeography of dolichoderine ants: Effects of data partitioning and relict taxa on historical inference. Syst Biol 59: 342–362. doi:10.1093/sysbio/syq012. PubMed: 20525640.
128. Li C, Lu G, Ortí G (2008) Optimal data partitioning and a test case for ray-finned fishes (Actinopterygii) based on ten nuclear loci. Syst Biol 57: 519–539. doi:10.1080/10635150802206883. PubMed: 18622808.
129. McGuire JA, Witt CC, Altshuler DL, Remsen JV, [!(surname)!] (2007) Phylogenetic systematics and biogeography of hummingbirds: Bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. Syst Biol 56: 837–856. doi:10.1080/10635150701656360. PubMed: 17934998.
130. Johnson MTJ, Carpenter EJ, Tian Z, Bruskiewich R, Burris JN et al. (2012) Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. PLOS ONE 7: e50226. doi:10.1371/journal.pone.0050226. PubMed: 23185583.
131. Luo R, Liu B, Xie Y, Li Z, Huang W et al. (2012) SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. Gigascience 1: 18. doi:10.1186/2047-217X-1-18. PubMed: 23587118.
132. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410. doi:10.1016/S0022-2836(05)80360-2. PubMed: 2231712.
133. Li F-W (2013) Blue Devil Blast Utility Seq Extr Devised By Li. Available: http://pryerlab.biology.duke.edu/software.
134. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797. doi: 10.1093/nar/gkh340. PubMed: 15034147.
135. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J et al. (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10: 421. doi:10.1186/1471-2105-10-421. PubMed: 20003500.
136. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9: 868–877. doi:10.1101/gr.9.9.868. PubMed: 10508846.
137. Larsson A (2013) lasseblaste. Available: ormbunkar.se/phylogeny/lasseblaste/.
138. Katoh K, Misawa K, Kuma K-i, Miyata T (2002) MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059–3066. doi:10.1093/nar/gkf436. PubMed: 12136088.
139. Swofford DL (2002) PAUP*: Phylogenetic analysis using parsimony (* and other methods). 4.0 ed Sunderland, MA: Sinauer.
140. Maddison WP, Maddison DR (2009) Mesquite: A modular system for evolutionary analysis. Version 2.72 ed: http://mesquiteproject.org.
141. Zwickl DJ (2006) GARLI. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. [PhD]. Austin: the University of Texas.
142. Brown JWS, Smith P, Simpson CJ (1996) *Arabidopsis* consensus intron sequences. Plant Mol Biol 32: 531–535. doi:10.1007/BF00019105. PubMed: 8980502.
143. Rothfels CJ, Windham MD, Pryer KM (2013) A plastid phylogeny of the cosmopolitan fern family Cystopteridaceae (Polypodiopsida). Syst Bot 38: 295–306. doi:10.1600/036364413X666787.
144. Schuettpelz E, Pryer KM (2009) Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. Proc Natl Acad Sci U S A 106: 11200–11205. doi:10.1073/pnas.0811136106. PubMed: 19567832.
145. Thiers B ([continuously updated]) Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium.

Paper IV

BIOINFORMATICS APPLICATIONS NOTE 2014, pages 1–3
doi:10.1093/bioinformatics/btu531

Phylogenetics

Advance Access publication August 5, 2014

AliView: a fast and lightweight alignment viewer and editor for large datasets

Anders Larsson

Systematic Biology, Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Uppsala 75236, Sweden

Associate Editor: David Posada

ABSTRACT

Summary: AliView is an alignment viewer and editor designed to meet the requirements of next-generation sequencing era phylogenetic datasets. AliView handles alignments of unlimited size in the formats most commonly used, i.e. FASTA, Phylip, Nexus, Clustal and MSF. The intuitive graphical interface makes it easy to inspect, sort, delete, merge and realign sequences as part of the manual filtering process of large datasets. AliView also works as an easy-to-use alignment editor for small as well as large datasets.

Availability and implementation: AliView is released as open-source software under the GNU General Public License, version 3.0 (GPLv3), and is available at GitHub (www.github.com/AliView). The program is cross-platform and extensively tested on Linux, Mac OS X and Windows systems. Downloads and help are available at http://ormbunkar.se/aliview

Contact: anders.larsson@ebc.uu.se

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on May 28, 2014; revised on July 30, 2014; accepted on August 1, 2014

1 INTRODUCTION

As DNA and protein datasets are getting larger, the demand for a refined and fast alignment editor increases. The need for an improved alignment editor and viewer, therefore, emerged in the 1000 plants project (1KP, www.onekp.com) while designing degenerate primers for a diverse set of ferns from transcriptome data (Rothfels *et al.*, 2013).

What was lacking in the previous available programs was the combination of abilities to (i) get an overview of large nucleotide alignments, (ii) visually highlight various conserved regions, (iii) have a simple and intuitive way to align, rearrange, delete and merge sequences and (iv) find degenerate primers in selected semiconserved regions. Although some of these features are individually present in current alignment editors, the combination is not.

In addition to the core functionality meeting these specific needs, AliView (Fig. 1) is designed with a complete set of intuitive general functions meeting the most common demands for preparing a multiple sequence alignment.

Here, AliView is introduced as an alignment viewer and editor with a unique combination of features that allows the user to work with large datasets. The intuitive user interface provides easy visual overview and navigation and works with unlimited size alignments.

2 IMPLEMENTATION

AliView is cross-platform, built in Java and thoroughly tested on Linux, Mac OS X and Windows operating systems. It uses the Java Evolutionary Biology Library v2.0 (available at http://code.google.com/p/jebl2/) for parsing files in Nexus format.

3 FEATURES

3.1 Large alignments, speed and more

The key features of AliView include the ability to swiftly handle large alignments with low memory impact (see Table 1 for comparison with other popular free cross-platform alignment viewers). AliView loads large alignment files 2–14 times faster and demands less than half of the memory resources than comparable alignment editors (Table 1; Supplementary Table S1A–C). AliView will read unlimited size alignment files in FASTA, Phylip, Nexus, Clustal and MSF-format (Table 2). This works through an indexing process where the sequences in the file initially are indexed and only cached in memory when viewed. Aside from the built-in indexing of large files, the program also reads and saves Fasta index files (.fai) as implemented by Samtools (Li *et al.*, 2009). The program either reads the whole alignment into memory or leaves parts on file, depending on memory resources available on the specific computer. This way any alignment file can be opened regardless of the memory resources of the computer.

Another important feature of AliView is the speed in rendering large alignments. The speed, together with the mouse wheel zoom feature, makes it possible to get a quick overview and easily navigate in large alignments.

AliView can merge overlapping sequences into a consensus sequence. This feature is useful when working with multiple read NGS-generated sequences. Sometimes the overlap of different sequences or contigs falls outside of the tolerance of assembly programs, and a manually merged sequence is needed.

AliView has unique functionality aimed at supporting the design of universal degenerate primers. It is possible to select an alignment region and have AliView calculate all possible primers (Kämpke *et al.*, 2001). To make it easy to select which

© The Author 2014. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contactjournals.permissions@oup.com

Table 1. Time to open alignment file and memory usage. Comparison of AliView with popular free and cross-platform alignment editors

Alignment		Dimension (sequence × character)	Program				
Size	Format		AliView	JalView 2	SeaView	ClustalX	Mesquite
22.4 GB	FASTA	479 726 × 46 512	5–110 s (88 MB)[a,b]	Not supported	Not supported	Not supported	Not supported
22.4 GB	FASTA	479 726 × 46 512	0.6 s (88 MB)[a,c]	Not supported	Not supported	Not supported	Not supported
2.1 GB	FASTA	39 407 × 54 103	17 s (2.2 GB)	73 s (4.7 GB)	51 s (5.7 GB)	Memory error	>10 min
1.3 GB	FASTA	11 792 × 107 401	5.6 s (1.2 GB)	33 s (3.3 GB)	23 s (3.6 GB)	Memory error	>5 min
1.3 GB	PHYLIP	11 799 × 107 401	5.9 s (1.2 GB)	Not supported	17 s (2.7 GB)	Memory error	>5 min
1.3 GB	NEXUS	11 792 × 107 401	5.7 s (1.2 GB)	Not supported	18 s (3.5 GB)	Not supported	>5 min
317 MB	FASTA	361 874 × 4958	2.1 s (608 MB)	31 s (3.1 GB)	9.5 s (3.8 GB)	Memory error	>5 min
42.2 MB	FASTA	5441 × 7682	0.6 s (53 MB)	2.8 s (160 MB)	1.2 s (145 MB)	20 s (1GB)	>5 min

Note: Test results shown were performed on Linux Ubuntu 12.04, Intel Core i7 2700K 3.5 Ghz, 16 GB internal memory and Intel 520 SSD. Similar results were obtained on Mac OS X and Windows systems. For a more extensive comparison including the test methodology, see Supplementary Table S1A–C.
[a]The 22.4 GB FASTA file was not read completely into memory but instead accessed as an indexed file. In all other tests the files were read into memory.
[b]Times depending on how many sequences being indexed at once.
[c]With alignment file already indexed.

Table 2. Comparison of AliView features with popular free and cross-platform alignment editors

Feature / Program	AliView	JalView 2	SeaView	ClustalX	Mesquite
Open alignments of unlimited size (read from disk)	Yes	–	–	–	–
Maximum number of sequences visible at once[a]	Unlimited	495 or overview window	106	120	68
Maximum sequence length visible at once[a]	Unlimited	1830 or overview window	305	345	1650
Merge sequences	Yes	–	–	–	–
Find degenerate primers in selected areas	Yes	–	–	–	–
Define exon boundaries and codon positions for translating nucleotides	Yes	–	–	–	Yes
Highlight difference from consensus or 'trace sequence'	Yes	–	–	Yes	–
Highlight consensus residues	Yes	Yes	Only protein	Only protein	Yes

Note: A more thorough comparison is included as Supplementary Table S2.
[a]Maximum number of sequences and maximum sequence length visible were tested at 1920 × 1200 screen resolution.

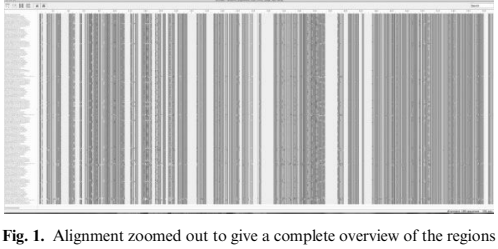

Fig. 1. Alignment zoomed out to give a complete overview of the regions

primer to use, they are presented as an ordered list sorted by the number of degenerate positions, self-binding values and melting temperature.

3.2 Other features

Apart from the key features, AliView also has several other alignment program functions. Alignment can be done by calling any external alignment program. AliView includes and has MUSCLE integrated as the default alignment program (Edgar, 2004), but the user can incorporate other programs if desired. Other features include, for example, manual editing capabilities to insert, delete, change, move or rename sequences in an alignment; undo/redo functionality; several visual cues to highlight consensus characters or characters deviating from the consensus; ClustalX conserved region color scheme (Larkin *et al*., 2007); search functionality that finds patterns across gaps and follows IUPAC codes; implementation of the Nexus specification of Codonpos, Charset and Excludes.

AliView is intended to be a simple easy-to-use alignment editor, and not a complete program for phylogenetic analyses. Instead, the 'external interface' function is aimed to ease the use of AliView as one program in a chain of software, making it possible to call other programs from within AliView with the current alignment or selected sequences as arguments. As a proof of concept, AliView comes with a preset code that adds a button for directing the alignment to FastTree (Price *et al*., 2010) that calculates a phylogenetic tree that is then automatically opened in FigTree (Rambaut, 2012).

For comparison of the key features of AliView with other free cross-platform editors such as Jalview 2 (Waterhouse *et al.*, 2009), SeaView (Gouy *et al.*, 2010), ClustalX (Larkin *et al.*, 2007) and Mesquite (Maddison, and Maddison, 2011) see Table 2. For a more comprehensive comparison of features see Supplementary Table S2.

3.3 User interface and usability

Because an alignment editor is an everyday tool for many researchers, AliView was designed with extensive focus on usability and intuitive handling, implemented by following the logical standards of commonly used software such as text-editors, word processors, browsers and, of course, other alignment viewers.

ACKNOWLEDGEMENTS

Thanks to my colleagues at the Systematic Biology department for immense beta-testing, bug-reporting and mostly good suggestions. Thanks to Allison Perrigo, John Petterson and Martin Ryberg for comments on the manuscript. I also would like to thank the three reviewers and the associate editor David Posada, who gave me very valuable feedback on the previous versions of the manuscript.

Funding: This work was supported by grants from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas) to Petra Korall (2006-429 and 2010-585).

Conflict of interest: none declared.

REFERENCES

Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.

Gouy,M. *et al.* (2010) SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.*, **27**, 221–224.

Kämpke,T. *et al.* (2001) Efficient primer design algorithms. *Bioinformatics*, **17**, 214–225.

Larkin,M.A. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Li,H. *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Maddison,W.P. and Maddison,D.R. (2011) Mesquite: a modular system for evolutionary analysis. Version 2.75, http://mesquiteproject.org (18 August 2014, date last accessed).

Price,M.N. *et al.* (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*, **5**, e9490.

Rambaut,A. (2012) Figtree 1.4.0. http://tree.bio.ed.ac.uk/software/figtree/ (18 August 2014, date last accessed).

Rothfels,C.J. *et al.* (2013) Transcriptome-mining for single-copy nuclear markers in ferns. *PLoS One*, **8**e76957.

Waterhouse,A.M. *et al.* (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.

Paper V

Phylogeny of *Woodsia* (Woodsiaceae): recent speciation through polyploidization is common in old diploid stock

Anders Larsson,[1] Michael D. Windham [2] & Petra Korall [1]

1 *Systematic Biology, Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden*

2 Department of Biology, Duke University, Durham, North Carolina 27708, U. S. A.
Author for correspondence: Anders Larsson, anders.larsson@ebc.uu.se

Abstract *Woodsia* is a genus of small tufted ferns growing in rocky habitats. It is widely distributed in temperate and montane regions of the world. Through previous cytological studies and allozyme work it is known to have a complex evolutionary history and includes numerous polyploid taxa and hybrids. We present here for the first time detailed phylogenetic analyses of a wide range of taxa within *Woodsia*, including samples from many chromosome-counted specimens. Five plastid (*atpA, atpB, matK, rbcL* and *trnG-R*) and two nuclear (*pgiC* and *RPA2*) regions have been used from an ingroup of 188 samples representing 36 taxa and six hybrids. In a complementary expanded analysis the polyploid speciation frequency is estimated within the 10 families most closely related to Woodsiaceae (the Eupolypods II). *Woodsia* is monophyletic if the often segregated genera *Protowoodsia, Cheilanthopsis* and *Hymenocystis* are included. The genus comprises two major well-supported clades, one including the circumboreal and most of the Asian species, and the other including all American and the remaining Asian species. The split between these clades is estimated to 45 Ma. *Woodsia* × *abbae* is a remarkable triploid hybrid between members of these two clades. Most taxa in *Woodsia* are polyploid and polyploidization is the most common mode of speciation in the genus with an estimated polyploid speciation rate of 54%. The polyploids are mostly young. Some of the polyploid taxa, such as *W. alpina* and *W. obtusa*, seem to have been formed multiple times. The Eupolypod II study agrees with the *Woodsia* study in showing a high proportion of polyploids with a polyploid speciation rate of 46%. Old polyploid lineages are rare. The circumboreal species do not form a monophyletic group and are nested among various Asian species, whereas the "American clade" is monophyletic and nested among Asian species. Within the American clade *W. montevidensis* has its main distribution in South America, but also has made a recent leap to Southern Africa and Madagascar.

Keywords biogeography; chromosome numbers; Eupolypods II; hybrids; nuclear and plastid DNA; polyploid speciation

INTRODUCTION

Woodsia R. Br. is a medium-sized genus of small, tufted and relatively rare ferns growing on or in the vicinity of cliffs and rocks (Brown, 1964). They are mainly found in montane areas in the Northern Hemisphere, but also in South America and Southern Africa. The species richness is greatest in the Rocky Mountains in North America and the Himalayas in Asia. There are three circumboreal species, *W. glabella* R. Br., *W. ilvensis* (L.) R. Br. and *W. alpina* (Bolton) Gray, but most species have a more restricted distribution (Brown, 1964; Windham, 1993). In a recent study of the family relationships within the eupolypods II clade by Rothfels & al. (2012), *Woodsia* was placed in the monogeneric family Woodsiaceae, and the often segregated genera *Cheilanthopsis* Hieron. and *Protowoodsia* Ching. were found to be nested within *Woodsia.*

Species delimitation within *Woodsia* has been a problem due to the occurrence of many plants with an intermediate morphology (Brown, 1964; Windham, 1993). Donald Brown in his monograph of the genus surveyed ca. 4500 specimens with focus on morphology, but also including cytological work (Brown, 1964). He recognized 22 species, two varieties and three hybrids, stating that he used a "broad" species concept, whereas other authors have together recognized up to about 40 species (Windham, 1993; Mickel & Windham, 2004; Zhang & al., 2013).

In the period spanning 1985-2004 four important studies put the substantial polyploidization within *Woodsia* in focus by using chromosome counts and allozyme electrophoresis. Ma (1985, 1987) made and reviewed chromosome counts for most of the Asian species and found many of them to be tetraploid or sometimes octoploid, and only six species diploid. He also proposed a hypothesis for the reticulate ancestry of the polyploids (Ma, 1987). Windham (1993) and Mickel & Windham (2004), in their extensive cytological, allozyme and morphological work on the North American and Mexican species, identified several new species. The eight American species recognized by Brown were expanded into 16 species and subspecies, more than half of them tetraploids. Most previously recognized species were found to have two cytotypes. The *W. mexicana* Fée complex was split into several new species based on allozyme and morphological characters. Also *Woodsia oregana* D.C. Eaton was proposed to be part of this complex via the polyploid *W. oregana* subsp. *cathcartiana* (B.L. Rob.) Windham. *W. obtusa* (Spreng.) Torr. and *W. scopulina* D.C. Eaton were found to have both diploid and tetraploid cytotypes. *W. canescens* (Kunze) Mett. was resurrected from the morphologically heterogeneous *W. mollis* (Kaulf.) J. Sm.

In light of these studies much of the taxonomic difficulties within *Woodsia* can be attributed to its reticulate history and possible multiple origins of polyploids that may give rise to a range of intermediate morphologies between parent taxa. In this study we present for the first time phylogenetic analyses of a wide range of taxa within *Woodsia*, including many of the specimens that were chromosome counted by M. D. Windham (Windham, 1993; Mickel & Windham, 2004). Five chloroplast markers and two nuclear markers are used. The chloroplast markers are uniparentally inherited from one of the parents when a polyploid is formed, and single copy nuclear markers are needed as a complement to detect hybridization and allopolyploidization (Vogel & al., 1998; Popp & Oxelman, 2004; Brysting & al., 2011). Despite this, molecular studies of ferns have not involved nuclear markers until the last few years (reviewed in Rothfels & al. 2013).

Our main aim is to resolve the patterns of diploid and polyploid evolution in the genus, but also to answer taxonomic and biogeographic questions within *Woodsia*, such as: How frequently have

Woodsia species established over long distances? Are the American *Woodsia* species a monophyletic group? Have *Woodsia* crossed the oceans multiple times? Is it fruitful and valid to recognize the segregate genera *Cheilanthopsis, Hymenocystis* and *Protowoodsia*? Do the polyploid taxa have single or multiple origins? Do the octoploid Asian species have evolutionary histories involving multiple diploid ancestors? How is this mainly northern hemisphere genus connected to the South American and African species and when did this disjunction happen?

We complement our study of polyploidy in *Woodsia* with a broad analysis of polyploid frequencies and polyploid speciation frequencies in the entire Eupolypods II sensu Rothfels & al. (2012), comprising about 2500 species in the families Cystopteridaceae, Rhachidosoraceae, Diplaziopsidaceae, Hemidictyaceae, Aspleniaceae, Thelypteridaceae, Woodsiaceae, Athyriaceae, Blechnaceae, and Onocleaceae. Cytological information from this clade was compiled and mapped on a phylogeny, where the results could be compared to the results of the *Woodsia* study. The questions we want to answer from this comparison are: How extensive is polyploidy in *Woodsia* and related ferns? How frequent is speciation through polyploidization in *Woodsia* and its closest relatives? Are the polyploids young, being continuously replaced by new ones formed by the diploid stock? Are there old polyploid lineages?

Definitions. — We use the term neopolyploid for a recently formed polyploid where most of the duplicated chromosomes are still present (plus/minus a few chromosomes), as opposed to a paleo-/ancient polyploid which is a lineage that contains historical polyploidization events where most of the duplicated chromosomes have been reduced through various diploidization events.

We use a taxonomic approach when defining an autopolyploid as the result of polyploidization within a species, whereas an allopolyploid is formed by two species (Soltis & Rieseberg, 1986; Soltis & al., 2004). The reproductive isolation between polyploids and their diploid progenitors make us count autopolyploids as distinct entities in our analyses, although we do not name them formally, as is suggested by Soltis & al. (2007),.

MATERIALS AND METHODS

Taxon sampling. — The ingroup includes a total of 188 specimens representing 36 species and subspecies and 6 homoploid hybrids (Appendix 1). The sampling was made to include as many taxa of *Woodsia* as possible. Furthermore, when available, we included several specimens per taxon in order to cover morphological and geographical variation. For the American taxa a denser sampling was made to elucidate the reticulate evolution of the American polyploid taxa (Windham, 1993; Mickel & Windham, 2004). We paid special attention to include specimens that had previously been cytologically investigated in Windham (1993) and Mickel & Windham (2004). *Athyrium filix-femina* and *Thelypteris palustris* were selected as outgroups based on their close relationship to *Woodsia* in previous studies (Rothfels et al. 2012).

Molecular data. — Two nuclear regions, *pgiC* and *RPA2*, together with five chloroplast regions, the protein coding *atpA*, *atpB*, *matK*, and *rbcL* and the intergenic spacer *trnG-trnR* (hereafter referred to as *trnG-R* also including the *trnG* intron), were selected for amplification and sequencing. The plastid regions were already shown to be informative by Rothfels & al. (2012). The nuclear *pgiC* region also has a history of successful usage in fern systematics (Juslén & al., 2011; Dyer & al., 2012; Rothfels & al., 2013) whereas the nuclear *RPA2* region has previously only been used for flowering plants (Popp & Oxelman, 2004; Rautenberg & al., 2012).

DNA isolation, amplification, and sequencing. — DNA was extracted from silica-dried tissue or herbarium material (up to ca. 85 years old), using a modified Carlson-Yoon protocol (Yoon & al., 1991). Less than 0.01 g of dried plant material, silica beads, 750 µl Carlson buffer, and 20 µl 2-mercaptoethanol were added to a 2 ml tube and ground for 45 s using a Mini-Beadbeater (BioSpec Products), followed by incubation at 65°C for 30–60 min. Chloroform/isoamyl-alcohol 24:1 was added and put on slow shake for 30–60 min, and the water phase was obtained after 10 min of centrifugation. The extracted DNA water phase was either left to precipitate overnight in 2/3 volumes of isopropyl alcohol, where the DNA-pellet was collected from the bottom of the tube after centrifugation and washed in 75% EtOH, or purified by Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare).

PCR reactions were performed with standard Taq polymerase and the following cycle: 3 min initial denaturation at 95°C, followed by 30–40 cycles of 30 s denaturation at 95°C, 1 min annealing at 54–60°C (primer dependent), and 1–2 min elongation at 72°C, and a final elongation of 10 min at 72°C. PCR products were purified either by using MultiScreen Plates in a vacuum manifold (Millipore) or ExoSAP-IT (Affymetrix) and sequenced by Macrogen Inc. (South Korea). Information on amplification and sequencing primers are found in (Table 1). For the *RPA2* region, specific *Woodsia* primers were designed using a nested PCR approach with unspecific universal primers as described in Popp & Oxelman (2004). Sequences for all 190 specimens were obtained for the most variable plastid region *matK*, and 47 specimens representing 33 species were sequenced for all regions (Appendix 1).

Table 1. Primers used in amplification and sequencing.

Locus	Primer	Direction	Usage	Sequence (5' – 3')	Reference
atpA	ESATPA535F	F	S	ACAGCAGTAGCTACAGATAC	Schuettpelz et al. 2006
atpA	ESATPA557R	R	S	ATTGTATCTGTAGCTACTGC	Schuettpelz et al. 2006
atpA	ESATPA856F	F	S	CGAGAAGCATATCCGGGAGATG	Schuettpelz et al. 2006
atpA	ESATPA877R	R	S	CATCTCCCGGATATGCTTCTCG	Schuettpelz et al. 2006
atpA	ESATPF412F	F	A, S	GARCARGTTCGACAGCAAGT	Schuettpelz et al. 2006
atpA	ESTRNR46F	R	A, S	GTATAGGTTCRARTCCTATTGGACG	Schuettpelz et al. 2006
atpB	ATPB910R	R	S	TTCCTGYARAGANCCCATTTCTGT	Pryer et al. 2004
atpB	ESATPB172F	F	A, S	AATGTTACTTGTGAAGTWCAACAAT	Schuettpelz & Pryer 2007
atpB	ESATPB701F	F	S	TATGGTCAGATGAATGAACC	Schuettpelz & Pryer 2007
atpB	ESATPE45R	R	A, S	ATTCCAAACWATTCGATTWGGAG	Schuettpelz & Pryer 2007
matK	EuIImatKfSIH	F	A, S	TCRAAAATBTCRCAGTCTATTCATTC	Rothfels et al. 2012
matK	FERmatKrAGK	R	A, S	CGTRTTGTACTYYTRTGTTTRCVAGC	Kuo et al. 2011
matK	matK490R	R	A, S	TCKACGAAACAGGCGAAC	this study
matK	matK3F	F	A, S	GAAGATAACTTKTRGATCKCTTCCT	this study
matK	matK435F	F	A, S	ATCTAGCCATGTTTTAGAG	this study
matK	matK9F	F	A, S	GATCCCTTCCTAAATTTGAT	this study
matK	matkK1273R	R	A, S	TATCTCAATCTACGCAATCCAT	this study
rbcL	ESRBCL1361R	R	A, S	TCAGGACTCCACTTACTAGCTTCACG	Schuettpelz & Pryer 2007
rbcL	ESRBCL1F	F	A, S	ATGTCACCACAAACGGAGACTAAAGC	Schuettpelz & Pryer 2007
rbcL	ESRBCL663R	R	S	TACRAATARGAAACGRTCTCTCCAACG	Schuettpelz & Pryer 2007
rbcL	PKRBCL556F	F	S	GGTAGRGCYGTCTAYGAATGYC	Rothfels et al. 2012
trnG-R	trnG1F	F	A, S	GCGGGTATAGTTTAGTGGTAA	Nagalingum et al. 2007
trnG-R	trnG353F	F	S	TTGCTTMTAYGACTCGGTG	Korall et al. 2007
trnG-R	trnG63R	R	S	GCGGGAATCGAACCCGCATCA	Nagalingum et al. 2007
trnG-R	trnGR43F1	F	S	TGATGCGGGTTCGATTCCCG	Nagalingum et al. 2007
trnG-R	trnR22R	R	A, S	CTATCCATTAGACGATGGACG	Nagalingum et al. 2007
pgiC	pgiC14F	F	A, S	GTGCTTCTGGGTCTTTTGAGTG	Ishikawa et al. 2002
pgiC	pgiC16R	R	A, S	GTTGTCCATTAGTTCCAGGTTCCCC	Ishikawa et al. 2002
pgiC	pgiC14F2	F	A, S	GAGTGTTTGGAATGTTTCATTCCT	Juslén et al. 2011
pgiC	pgiC16R2	R	A, S	CAAAGWCAATTTCACCAGCTTCA	Juslén et al. 2011
pgiC	pgiC14F4	F	A, S	AATGTCTCATTCCTTGGYCAC	this study
pgiC	pgiC16R4	R	A, S	CCCTTTCCATTGCTTTCCATAC	this study
rpa2	rpa2-22F	F	A, S	GATATAGACATGCCTTTTTCACAC	rhis study
rpa2	rpa2-40F	F	A, S	TCACACGTCACCGGCATG	rhis study
rpa2	rpa2-529R	R	A, S	AGTCCACTGTACATAACTTCGGT	rhis study
rpa2	rpa2-577R	R	A, S	TCTGGTAATACACAACTCCTATGT	rhis study

F = Forward, R = Reverse, A = PCR Amplification, S = Sequencing.

Cloning or allele specific primers. — When the initial read of the nuclear regions *pgiC* and *RPA2* was polymorphic, this was taken as evidence for multiple homoeologues or alleles. We separated the multiple sequences either by cloning or by sequencing with specifically designed allele and homoeologue specific primers. Cloning reactions were performed with TOPO TA Cloning kit (Invitrogen) following the standard protocol but performing 1/2 size reactions. 10–24 clones were selected for sequencing and these clone sequences were assembled and separate consensus sequences were created if at least three clones were in agreement. Single nucleotide substitutions that was unique to a single read was regarded as PCR or cloning errors. This sequence assembly was performed with Staden package version 2.0 (Staden & al., 2000). Sequences were manually screened

and removed if they showed signs of PCR recombination. When doing many samples at once, cloning was most efficient technique. When we during the progress of the work added a few supplementary specimens to the study, we applied instead a strategy of designing two allele/homoeologue specific sequencing primers (13–14bp long) at the first polymorphic site found in the mixed trace as described by Scheen & al. (2012).

To verify that all sequence copies had been found, we compared the initial mixed trace with a mixed summary trace of the sequences we extracted. This analysis was performed with the software SplitTrace that takes separate sequence traces and creates a visual mixed trace that can be compared to the initial mixed trace (Larsson, 2013). The mixed traces of specimens belonging to the same taxon were also initially compared in SplitTrace and if the mixed traces were identical, we regarded that sequence as a duplicate and only one was used in the analyses (the other specimen identifier was added to the name in the phylogenetic trees).

Alignment and phylogenetic analyses. — Sequence fragments were assembled and edited using the Staden package version 2.0 (Staden & al., 2000). Sequences were aligned with MAFFT 6.903b (Katoh & al., 2002) Some indel regions that after visual inspection seemed to have ambiguous alignment were manually excluded with the alignment editor AliView 1.0.1 (Larsson, 2014).

Before concatenating the five chloroplast regions into a single dataset, each of the five chloroplast regions (*atpA*, *atpB*, *matK*, *rbcL*, and *trnG-R*) was analyzed individually in a Maximum Likelihood framework using RAxML 7.4.7 (Stamatakis, 2006). For each region a bootstrap analysis was run with 1000 bootstrap replicates under a GTR+G model. The resulting topologies were manually checked for strongly supported conflicts (>0.70 bootstrap support values). No such conflicts were found and the separate regions were concatenated with abioscript (Larsson, 2010).

The chloroplast dataset contained 35 terminals for which sequence data were available for all five regions and an additional 57 terminals that had sequence data from at least the *matK* region. The matrix contains 92 terminals and an aligned length of 6564 bp without ambiguous regions and represents 190 specimens. If a specimen had a *matK* sequence identical to another specimen they were treated as duplicate haplotypes and the specimen with the most regions sequenced was retained, while the other specimen identifier was added to the name in the phylogenetic trees. For full dataset statistics see Table 2.

The two nuclear regions were analyzed separately. The aligned nuclear *pgiC* dataset is 709 bp long and contains 169 terminals representing 143 specimens (some specimens include multiple alleles). The nuclear *RPA2* dataset is the smallest. As it turned out to be less informative it was kept in the study as a supplementary dataset, as it is sometimes useful to verify or refute the results from the other two datasets. This dataset is 503 bp long and contains 92 sequences representing 52 specimens.

Our alignments are available in TreeBase [####]. All new sequences are deposited in GenBank (see Appendix 1).

The best fit DNA-substitution model for each dataset was evaluated based on the corrected Akaike information criterion (AICc) as implemented in MrAic (Nylander, 2009). The selected models for the chloroplast regions were GTR+G for *atpA*, *atpB*, *matK*, and *trnG-R*, and HKY+G for *rbcL*. The selected models for the *pgiC* and *RPA2* datasets were GTR+G and K2P, respectively.

Bayesian inference tree searches were performed with MrBayes 3.2.1 (Ronquist & Huelsenbeck,

2003). For all datasets we ran four runs of four chains each (one cold and three heated), for 15 million generations. In the chloroplast dataset each locus was assigned a separate partition with model substitution parameters unlinked, and branch lengths linked. The posterior was sampled every 1000 generations. Inspection of log-files in Tracer (Rambaut & Drummond, 2007) and AWTY (Wilgenbusch & al., 2004; Nylander & al., 2008) revealed that all the runs converged within the first 1 million generations. We excluded the first 1.5 million generations (10%) of each run as burn-in before summarizing the posterior. We also performed Maximum Likelihood bootstrap analyses using RAxML 7.4.7 (Stamatakis, 2006), with the following settings: 1000 bootstrap replicates, the GTR+G model with the regions in the chloroplast dataset assigned as separate partitions. The bootstrap tree files were summarized onto the MrBayes phylogram with SumTrees 3.3.1 (Sukumaran & Holder, 2010). Nodes with posterior probabilities (PP) of 0.95 or above or with a maximum likelihood bootstrap support of 70% or above were regarded as well supported. Visual tree inspection and tree design was done with FigTree 1.4.0 (Rambaut, 2012). Tree layout was done with the open source software Inkscape (The Inkscape Team, 2013). Computation-intensive analyses were run at the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX).

Region	Number of sequences	Sites	Variable sites	Missing data or Gap (%)
atpA	39	1752	322	4,2
atpB	40	1316	199	2,6
matK	92	1247	381	10,6
rbcL	39	1309	191	3,1
trnG-R	37	940	302	6,4
rpa2	92	503	188	18,0
pgiC	169	709	313	29,7

Molecular dating. — We estimated the node ages in the *Woodsia* chloroplast phylogeny (Fig. 3) with BEAST 1.8.0 (Drummond & al., 2012). The dataset was created by adding our chloroplast dataset to a broadly sampled dataset of the Eupolypods II clade that included 81 species (Rothfels & al., 2012). We used the same methodology and settings as in Rothfels & al. (2012). The phylogeny was calibrated using age estimates of six different nodes from Schuettpelz & Pryer (2009). We used indirect dating since there are no direct ages such as fossils available among Woodsiaceae or the closest relatives. The age parameters were given a normal distribution and a standard deviation of 10% of the estimated mean ages. The node ages used were: the root of the tree = 165.6 Mega annum (Ma), Dennstaedtiaceae (119.3 Ma), Eupolypods (116.7 Ma), Pteridaceae (110.8 Ma), Eupolypods II (103.1 Ma) and Eupolypods I (98.9 Ma).

The analyses were performed with a relaxed molecular clock model with the following settings: birth-death tree prior, lognormal uncorrelated relaxed clock, and GTR+I+G substitution model. Priors were set to default values except the previously described node ages. We ran two independent runs each for 100 million generations and sampled the posterior every 1000 generations. Visual inspection in Tracer (Rambaut & Drummond, 2007) revealed that the runs converged within 5 million generations. Therefore we removed the first 10 million generations (10%) as burn in before we combined the result of the two runs. For the combined sample, the sample size of each parameter was above 200.

We also expanded this dataset into a comprehensive chloroplast dataset of the whole Eupolypods II (Electr. Suppl. Fig. S2), by adding sequences from all related taxa from the dataset of Lehtonen (2011). This matrix comprised 523 taxa (including 21 outgroup taxa) and DNA sequence data from five chloroplast regions (*atpA*, *atpB*, *rbcL*, *matK*, *trnG-R*) with an aligned length of 6613 bp. The topology and node ages for the entire Eupolypods II clade were estimated using the same approach as above.

Estimating ploidy levels. **—** We estimated the ploidy level of all included specimens and species (Appendix 1; Table 3) through one or several of the approaches described below.

Chromosomal counts. The most direct approach is to count the chromosomes. For 30 of the American specimens used in this study chromosome counts are available from the same specimens that are analysed here. Additional chromosome counts have been collected from the literature (Kurita, 1965; Shimura & Matsumo, 1975; Löve, Löve, & Pichi Sermolli, 1977; Ma, 1985; Wagner, 1987; Windham, 1993).

Spore measurements. Spore measurements have proven useful for determining ploidy levels in *Woodsia* and other fern genera. In closely related taxa the relative size of the spores is correlated to ploidy level (Barrington & al., 1986; Windham, 1993; Beck & al., 2011; Dyer & al., 2012). For this study spore measurements were made on selected specimens, especially on specimens belonging to species lacking chromosome counts. Measurements were made only if mature and well-formed spores were present. For each specimen 10–30 spores were measured with a Leitz Dialux 20 microscope at a magnification of 100x (Electr. Suppl. Table S2). When spore measurements were lacking for a species we estimated ploidy level in that species by comparing the spore size ranges of diploids and polyploids of closely related species. Spore size ranges for different taxa were collected from Brown (1964) and Windham (1993).

Spore count. The number of spores per sporangium in 3–10 sporangia were counted for a selection of specimens (Electr. Suppl. Table S2). The presence of 32 spores per sporangia instead of the normal 64 spores has been shown to indicate an apomictic life cycle and polyploidy (Manton, 1950; Beck & al., 2011).

Nuclear sequences count. The number of nuclear sequences retrieved from a specimen was used as a parameter when estimating the ploidy level (Appendix 1). We found no indication that the nuclear region *pgiC* has more than one copy in diploid specimens of *Woodsia* (maximum of two alleles found per specimen). The presence of more than two sequence copies in one specimen was treated as evidence of duplication through polyploidization.

Gametophyte sequencing. In some cases when viable spores were available, we were able to sequence gametophytes grown from spores to distinguish between homoeologues and heterozygous alleles. Since gametophytes of sexual diploid fern species have haploid tissues, we assumed that a specimen had heterozygous alleles when we found two alleles in the diploid sporophytic tissue but only a single allele in the gametophytic tissue. We grew spores on 9 cm petri dishes on sterile standard Murashige and Skoog medium with 1% plant agar (Sigma-Aldrich). Before sowing, the spores were sterilized with 30% chlorine solution for 5 min and thereafter thoroughly rinsed. Petri dishes were sealed with Parafilm and placed in growth chambers at 20°C with the light cycle 16 h light/8 h dark.

Estimating frequency of polyploid speciation in *Woodsia* and Eupolypods II. — We compiled a database with species chromosome counts and included data mainly from the online

version of Index of Plant Chromosome Numbers (IPCN) and the Cytotaxonomical Atlas of the Pteridophyta (Löve, Löve, & Sermolli, 1977). The total number of references is 3455 involving 635 names. Minimum and maximum generic and family chromosome base numbers were collected from Rothfels & al. (2012) and Smith & al. (2006). We created a Java program that assigned ploidy level to all entries in the database using the following algorithm: If generic base number min/max was available, a taxon with a chromosome number within 2× base number (± 5%) was considered diploid. If no generic base number was available, family base number min/max was used instead (none of the included families have a span of min/max base numbers that would be able to hold both diploid and polyploid counts). The database entries were thereafter assigned ploidy levels of 3x, 4x, 5x, 6x, 7x, 8x or more. Specimens that fell outside the limits ± 5% stayed unassigned and were manually inspected, and these specimens could all be dismissed as erroneous records (Electr. Suppl. Table S1).

We annotated the tree file from the extended Eupolypods II molecular dating analysis with ploidy and chromosome base number information from the chromosome count database. If ploidy information was available in the database, the taxa in the tree were assigned with this. Synonyms were accounted for through a cross match with The Plant List Database (The Plant List, 2013).

For both the *Woodsia* and the Eupolypods II phylogeny we calculated all polyploid speciation events by traversing through the tree starting at each tip with a polyploid taxon and then finding the most recent node that contains a diploid progeny; if there is a diploid progeny we assume that the ancestor was diploid since polyhaploidy is not widely accepted (Mayrose & al., 2011). If a taxon had both diploid and polyploid chromosome counts, the tip node was counted as a polyploid speciation event (Electr. Suppl. Fig. S2; Table S3). Non-assigned taxa were excluded from the calculations. The final number of taxa with ploidy information in the tree was 251. In the *Woodsia* phylogeny we only calculated speciation events involving the taxa we recognize in Table 3, and also excluded homoploid hybrids.

Ancestral reconstruction of chromosome base numbers. — The extended *Woodsia* and Eupolypods II phylogeny with chromosome base number annotations was analyzed with the most parsimonious reconstruction of ancestral states in Mesquite 2.75 (Maddison & Maddison, 2011). Chromosome base numbers were considered unordered categorical characters and therefore a reduction or increase of single or multiple chromosomes along a branch between two nodes was treated equally. The ancestral states of all nodes were annotated onto the tree shown in Electr. Suppl. Fig. S2 with FigTree 1.40 (Rambaut, 2012).

Distribution maps. — The geographic ranges of the species as shown in Figs. 4–13 are compiled from published revisions, floras and checklists (Brown, 1964; Ohwi & al., 1965; Li, 1975; Hultén & al., 1986; Windham, 1993; Tutin, 1993; Marticorena & Rodríguez, 1995; Schmakov, 2001; Crouch & al., 2012; CFD Galapagos Species Checklists, 2013; FloraGREIF Team, 2013).

RESULTS

Chloroplast phylogeny. — For our chloroplast dataset of *Woodsia*, the maximum likelihood bootstrap analysis and Bayesian analysis (Fig. 1) resulted in similar topologies without any well supported conflicts between the analyses, and with most of the clades at or above species level well supported. The tree can be divided into two major, well-supported clades, one, the "American &

Cheilanthopsis" clade containing the American (and African) species together with the Asian "*Cheilanthopsis*" clade, the other, the "Asian & Circumboreal" clade containing the rest of the Asian species together with the circumboreal taxa (Fig. 1).

Within the "American & *Cheilanthopsis"* clade, the Asian "*Cheilanthopsis*" clade is sister to the "American" clade with strong support. The "American" clade can be further divided into two distinct and well supported clades, one of them comprising *W. montevidensis*, *W. mollis* and *W. scopulina*, the other including *W. obtusa*, *W. phillipsii, W. oregana* and their closest relatives.

Within the American species *W. obtusa* and *W. mollis* two distinct and well supported subclades are found.

In the "Asian & Circumboreal" clade there is a well-supported sister relationship between the *W. lanosa*-complex and the rest of the species. The circumboreal species (*W. glabella, W. ilvensis* and *W. alpina*) do not form a monophyletic group.

Three of the polyploid taxa (*Woodsia alpina*, *W. oregana* subsp. *cathcartiana* and *W. plummerae*) have specimens with different *matK* chloroplast haplotypes representing both of the proposed diploid parent species.

Nuclear *pgiC* phylogeny. — The analysis of the *pgiC* dataset resulted in a relatively well supported phylogeny (Fig. 2). Many of the clades and their relationships are congruent with the chloroplast tree, but there are also some incongruencies. Whereas the "*Cheilanthopsis*" clade is monophyletic in the chloroplast phylogeny, it is split into two clades in the nuclear phylogeny. One comprise *W. elongata* and *W. indusiosa* and is sister to the American clade, while the other, comprising *W. fragilis* and *W. manchuriensis*, is sister to a clade of *W. elongata*/*W. indusiosa* and the American clade.

There is a well supported conflict between the nuclear *pgiC* and the plastid phylogenies in the placement of *W. andersonii*. In the *pgiC* phylogeny *W. andersonii* is resolved into two subclades that are successive sister clades to the *W. lanosa*-complex, whereas in the plastid phylogeny it is monophyletic and nested among other species in the Asian & Circumpolar clade.

The number of *pgiC* sequence copies found in diploid species was always one or two and in polyploid species 1–4, making it sometimes impossible to distinguish allelic variation from homoeologues (Appendix 1). We do not suspect any duplications in the *pgiC* gene apart from polyploidization, as it seems to be single copy in all the diploids.

The American diploid taxon *W. oregana* subsp. *oregana* consists of two very divergent clades of *pgiC* sequences that probably are alleles, verified by sequenced gametophytes.

Nuclear *RPA2* phylogeny. — The phylogeny based on nuclear *RPA2* is the one that is least resolved (Electr. Suppl. Fig. S1). There are multiple sequence copies in the American clade of species. Of the diploid species, *W. oregana* subsp. *oregana* and *W. phillipsii* have one shorter and one longer sequence, only short sequences are found in *W. scopulina* subsp. *laurentiana, W. mollis* and *W. montevidensis*, and only long a sequence in *W. obtusa* subsp. *occidentalis*, whereas all Asian species have long sequences only.

Dated *Woodsia* phylogeny. — The topology of the dated *Woodsia* phylogeny was congruent with the topology of the analyses performed with MrBayes (Fig. 3). The median age of the deep split between the American & *Cheilanthopsis* clade and the Asian & Circumboreal clade is estimated to 45 Ma (28–63 Ma, 95% highest posterior density, HPD). The median age of the most recent common ancestor (MRCA) of the American & *Cheilanthopsis* clade is estimated to 20 Ma (14–27 Ma, 95%

HPD). The MRCA of the American clade is estimated to 17 Ma (12–23 Ma, 95% HPD). The inferred ages of *Woodsia* polyploids are listed in Table 3 as the divergence times of the polyploid taxa from the most closely related diploid lineages.

Polyploid frequency and speciation. — Out of the 36 taxa of *Woodsia* sampled (homoploid hybrids excluded), 22 (61%) are polyploid and 14 (39%) are diploid (Table 3). In the expanded Eupolypods II analysis, 346 taxa (54%) were estimated to be polyploid and 291 taxa (46%) were estimated to be diploid (Table 3; Electr. Suppl. Table S1).

In the *Woodsia* phylogeny, 20 speciation events out of a total of 39 (54%) involve a shift from a diploid to a polyploid cytotype (Electr. Suppl.: Fig S3; Table S3). In the expanded Eupolypods II analysis 133 speciation events out of 289 (46%) were connected with a shift in ploidy level (counting double cytotypes in terminal taxa as two taxa and as one additional speciation event), or alternatively 94 speciation events out of 250 (38%) were connected with a shift in ploidy level (counting double cytotypes in terminal taxa only once and not as a speciation event) (Electr. Suppl.: Fig S4.; Table S4). Estimated ages of all polyploidy speciation events are listed in Electr. Suppl. Table S4.

Table 3. *Woodsia* taxa included in this study. *Ploidy level* = referenced or estimated ploidy level. *Chromosome base number* = referenced chromosome number.
Multiple origin of polyploids = our estimation whether taxon has originated multiple times. Most recent diploid ancestor = ages are compiled from the dated Beast analysis (Fig. 3).

Taxon	Ploidy level	Chromosome base number	Chromosome reference	Auto or allopolyploid, our classification	Multiple origin of polyploid	Most recent diploid ancestor. Chloroplast min age (Ma) 95% HPD	Most recent diploid ancestor. Chloroplast max age (Ma) 95% HPD	Our classification of polyploidization age. Recent = min age 95% HPD < 2 Ma
Woodsia cinnamomea Christ	2x* or 4x*	?	-					
Woodsia glabella R. Br.	2x	39	Löve et al. (1977) and references therein.					
Woodsia hancockii Baker	2x*	39?	-					
Woodsia ilvensis (L.) R. Br.	2x	41	Ma, Yi-Lun (1985) and references therein., Löve et al. (1977) and references therein.					
Woodsia macrochlaena Mettenius ex Kuhn	2x	41	Kurita (1965)					
Woodsia obtusa subsp *occidentalis* Windham	2x	38	Windham, FNA (1993)					
Woodsia oregana subsp *oregana* D. C. Eaton	2x	38	Windham, FNA (1993)					
Woodsia phillipsii Windham	2x	38	Windham, FNA (1993)					
Woodsia polystichoides D. C. Eaton	2x	41	Ma, Yi-Lun (1985) and references therein.					
Woodsia scopulina subsp *scopulina* D. C. Eaton	2x	38	Windham, FNA (1993)					
Woodsia elongata Hook.	2x and 4x	41,82	Ma, Yi-Lun (1985) and references therein.	auto		0	0	recent
Woodsia manchuriensis Hook.	2x and 4x	33,66	Ma, Yi-Lun (1985) and references therein.	auto		0	0	recent
Woodsia mollis (Kaulf.) J. Sm.	2x and 3x	38,n=2n=114	MDW (unpublished)	auto	yes, 2 distinct chloroplast	0	0,8	recent
Woodsia montevidensis (Spreng.) Hieron.	2x* and 4x*	-	-	auto		0,2	1,3	recent
Woodsia cf *phillipsii* Windham	3x	114	MDW (unpublished)	auto or maybe hybrid		0	0	recent
Woodsia alpina (Bolton) Gray	4x	80	Löve et al. (1977) and references therein.	allo	yes, reciprocal chloroplast	0	0,33	recent
Woodsia cochisensis Windham	4x	76	Windham, FNA (1993)	allo – unknown second diploid parent		0,1	1,13	recent
Woodsia cf. mexicana (see W. cochisensis)								
Woodsia fragilis (Trevir.) T.Moore	4x*	-	-	allo – unknown parents		4(12.5)	14.3(43.4)	old
Woodsia indusiosa Christ	4x	148	Ma, Yi-Lun (1985)	allo		8,2	19,1	old
Woodsia intermedia Tagawa	4x	82	Shimura & Matsumo (1975)	auto		0	0,32	recent
Woodsia macrospora C.Chr. & Maxon	4x*	-	-	auto		0,4	4,3	recent
Woodsia neomexicana Windham	4x	76	Windham, FNA (1993)	auto		1,6	4,5	prob. recent
Woodsia obtusa subsp *obtusa* (Spreng.) Torr.	4x	76	Windham, FNA (1993)	auto	yes, 2 distict chloroplast	0,13	1,7	recent
Woodsia oregana subsp *cathcartiana* D. C. Eaton	4x	76	Windham, FNA (1993)	allo	yes, reciprocal chloroplast	0,01	0,88	recent
Woodsia plummerae Lemmon	4x	76	Windham, FNA (1993)	allo	yes, reciprocal chloroplast	1,6	4,5	prob. recent
Woodsia rosthorniana Diels	4x	82	Ma, Yi-Lun (1985)	auto		11,7	24	old
Woodsia scopulina subsp *laurentiana* Windham	4x	76	Windham, FNA (1993)	auto		0,01	1,7	recent
Woodsia subcordata Turcz.	4x	c.160 (prob.164)	Ma, Yi-Lun (1985)	allo	maybe 2 version of chloroplast	0	0,3	recent
Woodsia andersonii (Bedd.) Christ	8x	164	Ma, Yi-Lun (1985)	allo	no, but great variation in pgiC	7,6	19,3	old
Woodsia cycloloba Hand.-Mazz.	8x	164	Ma, Yi-Lun (1985)	allo		4,3	10,7	old
Woodsia lanosa Hook.	8x	164	Ma, Yi-Lun (1985)	allo		3,5	9,4	old
Woodsia okamotoi Tagawa	8x*	-	-	allo		3,5	9,4	old
Woodsia neomexicana x phillipsii	3x	114	MDW (unpublished)	hybrid		0	0	recent-hybrid
Woodsia obtusa subsp *obtusa x* subsp *occidentalis*	3x	2n=114	MDW (unpublished)	hybrid		0	0	recent-hybrid
Woodsia x abbeae	3x	117(41+76)	Wagner (1987)	hybrid	yes, 2 distinct chloroplast	0	0	recent-hybrid
Woodsia neomexicana x oregana subsp *cathcartiana*	4x*	-	-	hybrid		0	0	recent
Woodsia obtusa subsp *obtusa x oregana* subsp *cathcartiana (Woodsia × kansana)*	4x*	-	-	hybrid		0	0	recent
Woodsia cyclolobalike hybrid	8x*	-	-	hybrid or allo		0,1	2,6	recent

* Ploidy level estimated in this study, not from chromosome counts (as described in Materials and Methods).

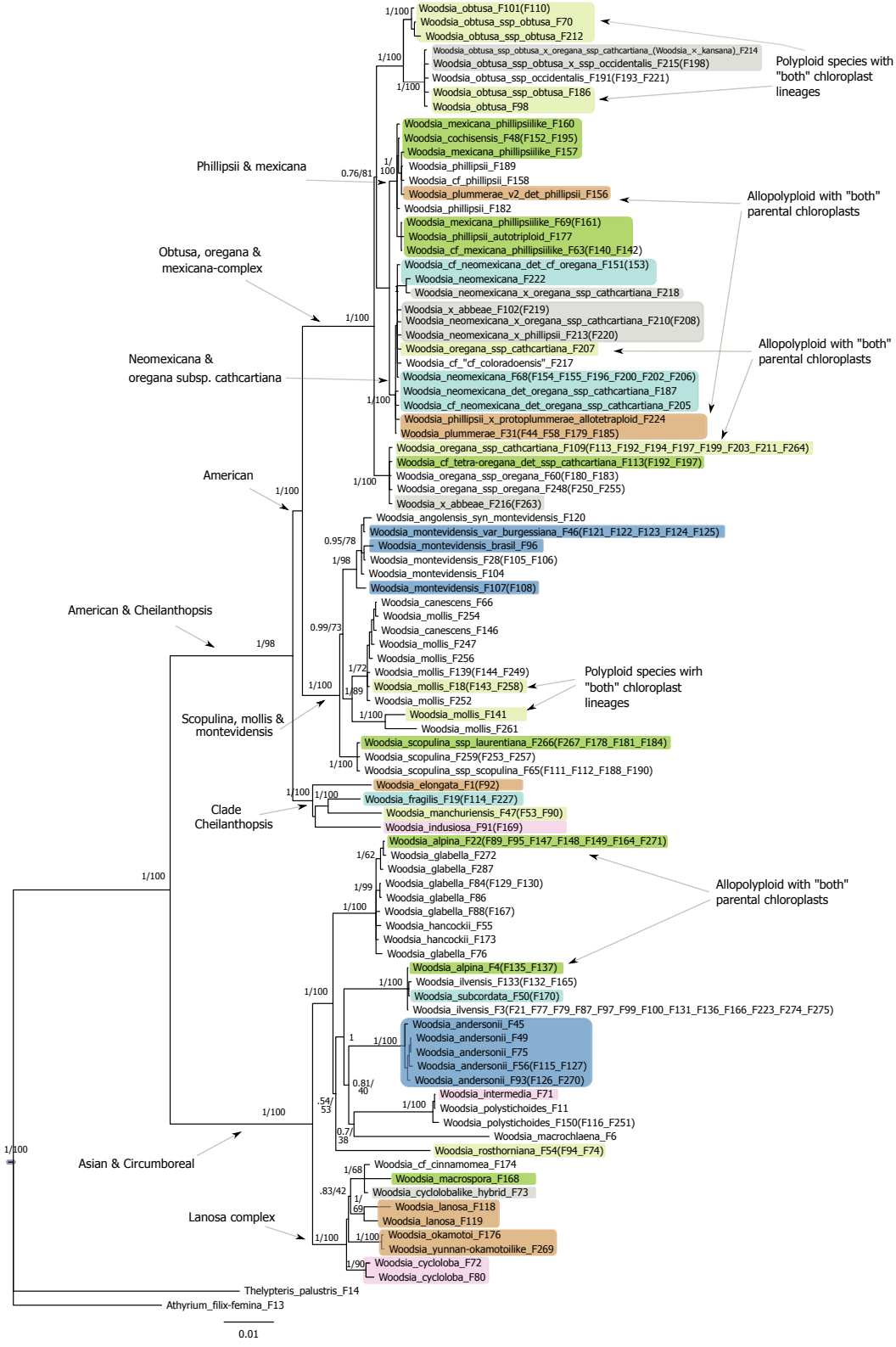

Fig. 1. MrBayes consensus tree of concatenated chloroplast sequences. Diploid taxa are not colored. Colors indicate polyploid taxa. Grey indicate hybrid taxa. Numbers at nodes are Bayesian posterior probabilities (PP) followed by RAxML maximum likelihood bootstrap support. Branches are collapsed at 0.5 PP. Scale bar represents average number of substitutions per site.

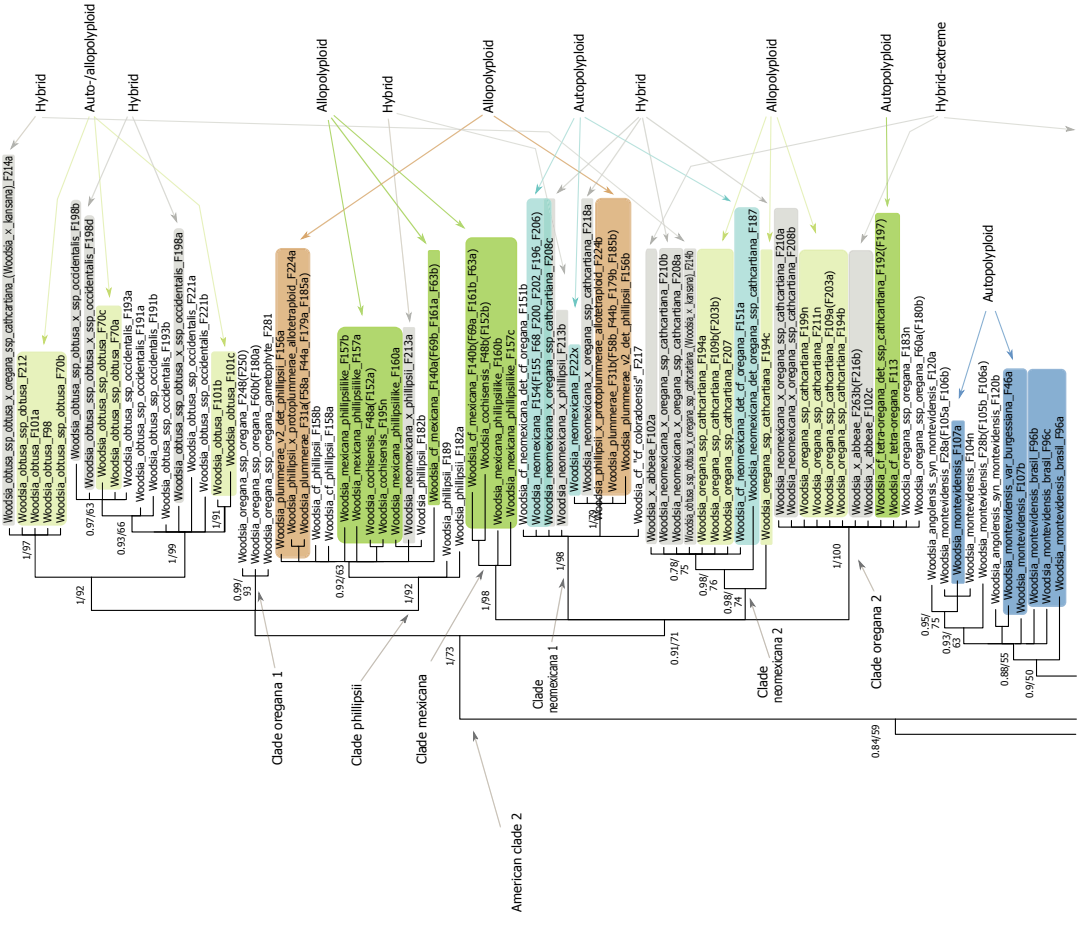

Fig. 2. MrBayes consensus tree of nuclear pgiC sequences. Names ending with a,b,c or d indicate specimens with multiple sequences, x indicates specimens with multiple sequences where only one has been retrieved. Diploid taxa are not colored. Colors indicate alleles belonging to the same polyploid taxon. Grey indicates hybrid taxa. Numbers at nodes are Bayesian posterior probabilities (PP) followed by RAxML maximum likelihood bootstrap support. Branches are collapsed at 0.5 PP. Scale bar represents average number of substitutions per site.


American clade 1
American & Cheilanthopsis
Cheilanthopsis
Asian & Circumboreal
Woodsia_canescens_F66a
Woodsia_canescens_F66b
Woodsia_canescens_F146a
Woodsia_mollis_F141a
Woodsia_mollis_F144
Woodsia_canescens_F146b
Woodsia_mollis_F18a
Woodsia_mollis_F18c
Woodsia_mollis_F258n
Woodsia_mollis_F252n
Woodsia_mollis_F256
Woodsia_mollis_F143a
Woodsia_mollis_F143c
Woodsia_mollis_F139b
Woodsia_mollis_F139a
Woodsia_mollis_gametophyte_F277
Woodsia_mollis_F247n
Woodsia_mollis_F141b
Woodsia_mollis_F141c
Woodsia_mollis_F18b
Woodsia_mollis_F261
Woodsia_mollis_F143b
Autopolyploid
Woodsia_scopulina_ssp_scopulina_F188a
Woodsia_scopulina_ssp_scopulina_F112
Woodsia_scopulina_ssp_scopulina_F190a
Woodsia_scopulina_ssp_laurentiana_F178a(F181a_F184a)
Woodsia_scopulina_ssp_scopulina_F65a(F159a)
Woodsia_scopulina_ssp_laurentiana_F178b(F181b_F184b)
Woodsia_scopulina_ssp_scopulina_F65b(F159b)
Woodsia_scopulina_ssp_scopulina_F190b
Woodsia_scopulina_ssp_scopulina_F188b
Autopolyploid
Woodsia_elongata_F92
Autopolyploid+diploid
Woodsia_indusiosa_F169a
Woodsia_indusiosa_F169b
Woodsia_indusiosa_F91b
Woodsia_indusiosa_F91a
Allopolyploid
Woodsia_fragilis_F19a
Woodsia_fragilis_F114a(F227a)
Woodsia_fragilis_F19b
Allopolyploid
Woodsia_fragilis_F19c
Woodsia_fragilis_F114b(F227b)
Woodsia_manchuriensis_F90b
Woodsia_manchuriensis_F90a
Autopolyploid+diploid
Woodsia_alpina_F4b(F22a_F89a)
Woodsia_glabella_F86(F84_F88_F167)
Woodsia_glabella_F76
Woodsia_glabella_F272
Woodsia_hancockii_F173
Allopolyploid
Woodsia_lanosa_F118b
Woodsia_lanosa_F119c
Woodsia_okamotoi_F176a
Woodsia_cycloloba_F80a
Woodsia_x_abbeae_F102b
Woodsia_x_abbeae_F263a(F216a)
Woodsia_alpina_F4a(F22b_F89b)
Woodsia_ilvensis_F3(F21_F77_F87_F204_F165_F166_F274)
Woodsia_subcordata_F50b(F170n)
Autopolyploid
Woodsia_intermedia_F71
Allopolyploid
Woodsia_polystichoides_F11(F150_F251)
Woodsia_subcordata_F50a
Woodsia_rosthorniana_F94b(F74a)
Autopolyploid
Woodsia_rosthorniana_F94a(F74b)
Woodsia_macrochlaena_F6
Woodsia_cyclolobalike_hybrid_F73b
Woodsia_macrospora_F168
Autopolyploid
Woodsia_cf_cinnamomea_F174
Hybrid/Allopolyploid
Woodsia_cycloloba_F80b
Woodsia_okamotoi_F176b
Woodsia_lanosa_F119a
Woodsia_lanosa_F118c
Allopolyploid
Allopolyploid
Woodsia_cyclolobalike_hybrid_F73a
Woodsia_lanosa_F119b
Woodsia_lanosa_F118a
Woodsia_andersonii_F93c
Woodsia_andersonii_F45b
Woodsia_andersonii_F56c(F127c_F115c)
Woodsia_andersonii_F49b(F75b)
Allolyploid
Woodsia_andersonii_F93d
Woodsia_andersonii_F45a
Woodsia_andersonii_F56b(F127b_F115b)
Woodsia_andersonii_F175
Woodsia_andersonii_F93a
Woodsia_andersonii_F56a(F127a_F115a)
Woodsia_andersonii_F49a(F75a)
Athyrium_filix-femina_F13a
Athyrium_filix-femina_F13b
0.3

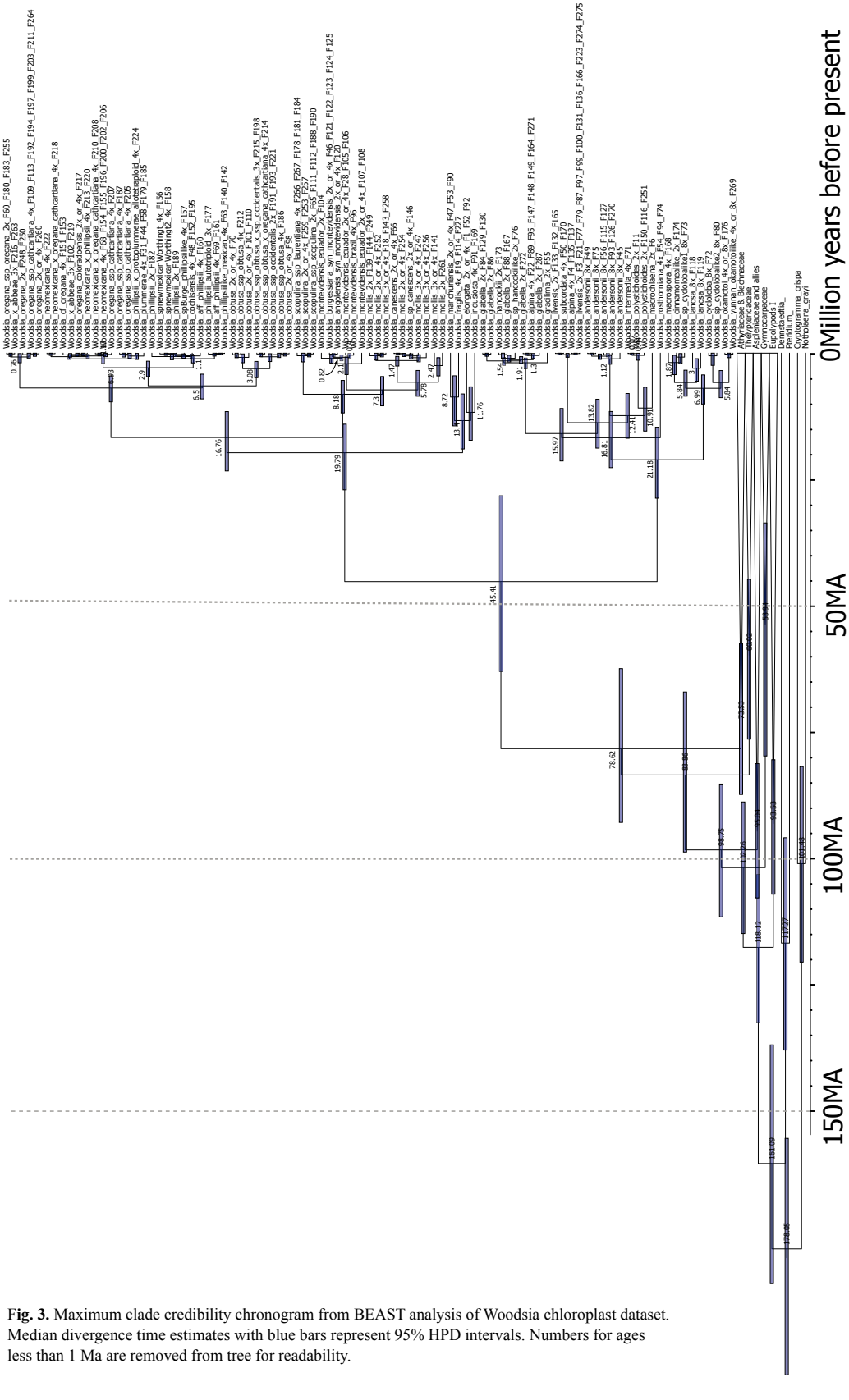

Fig. 3. Maximum clade credibility chronogram from BEAST analysis of Woodsia chloroplast dataset. Median divergence time estimates with blue bars represent 95% HPD intervals. Numbers for ages less than 1 Ma are removed from tree for readability.

SPECIES DISCUSSION

The phylogenetic analyses of the chloroplast and the nuclear datasets show that *Woodsia* includes an assemblage of diploid and polyploid taxa with an evolutionary history that is often very complex and sometimes very hard to interpret with certainty. Below we discuss the species and their evolution, and for each major clade (Figs. 1, 2) we first discuss the diploid species followed by the polyploids.

Asian & Circumboreal clade: Articulated stipe group

Within the Asian & Circumboreal clade a number of species are characterized by having an articulated stipe, a very distinctive morphological feature. Among the species sampled in this study this feature is shared by the diploids *W. ilvensis*, *W. glabella*, *W. hancockii*, *W. polystichoides* and *W. macrochlaena*, and it is also present in the polyploids *W. alpina*, *W. subcordata* and *W. intermedia.* Other polyploid taxa (i.e., *W. andersonii* and *W. rosthorniana*), which are also found in this clade, lack the articulated stipe. This could be a result of a reticulate origin involving both articulated and non-articulated parents (see below under *W. andersonii*). If we only consider the diploids, this group of species is well supported as monophyletic (Figs. 1, 2), suggesting that the articulated stipe is probably of a single origin. There are a few other rare and closely related Chinese species described as having an articulated stipe, but we have not been able to include them in the study (see below under Unsampled species).

***Woodsia ilvensis*. —**We have sequenced 16 specimens of *Woodsia ilvensis* from different parts of its circumboreal distribution (Fig. 4). In both the chloroplast (Fig. 1) and the nuclear (Fig. 2) phylogeny all accessions form distinct lineages with very little sequence variation, and are intermixed with the allopolyploids *W. alpina* and *W. subcordata*.

Two different chromosome numbers have been reported for *W. ilvensis*, the most common being n = 41 (Löve & al., 1977; Windham, 1993), and it is uncertain if the other number, n = 39 (Löve & al., 1977; Ma, 1985), is a second cytotype or based an erroneous counts. N = 39 is otherwise only found in *W. glabella,* a species involved in allopolyploidizations with *W. ilvensis.*

W. ilvensis is involved in more allopolyploidization and hybridization events than any other species in the phylogeny. *W. alpina* is the allopolyploid between *W. ilvensis* and *W. glabella*, *W. subcordata* is the allopolyploid between *W. ilvensis* and *W. polystichoides*, and *W. ilvensis* also hybridizes with *W. alpina* to form triploids with malformed spores (Manton, 1950; Windham, 1993).

We can also verify molecularly that *W.* × *abbeae* is a hybrid between diploid *W. ilvensis* and tetraploid *W. oregana* subsp. *cathcartiana*. This origin was proposed by Wagner (1987), who found it to be a sterile (it has malformed spores) triploid hybrid with 117 chromosomes (41 + 76). We find that the age of the deep split in the phylogeny between these two parent species is ca. 45 Ma (28–63 Ma, 95% HPD). This means that despite a total of at least 56 Ma (28 Ma × 2) of evolutionary changes these taxa are still able to hybridize. Very few other records of hybridization can match this, but one hybrid with similarly distantly related parent species is the London plane, *Platanus* × *hispanica* (*P. orientalis* × *P. occidentalis*), where the estimated divergence time of the parent lineages is ca. 45 Ma (Feng & al., 2005; Pilotti & al., 2009). Another example is the split between the parent lineages of the tetraploid fern hybrid *Cystopteris* × *Gymnocarpium*, estimated to ca. 58 Ma (Rothfels & al., 2014).

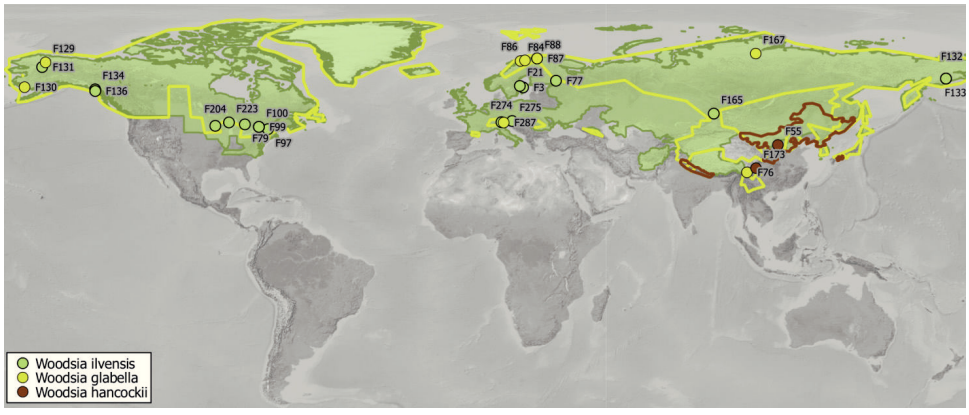

Fig. 4. Distributions of *W. ilvensis, W. glabella* and *W. hancockii* with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

***Woodsia glabella* and *W. hancockii*. —** In the analyses of the chloroplast (Fig. 1) as well as the nuclear (Fig. 2) datasets the accessions of *W. glabella* are found in a well-supported clade together with the accessions of *W. hancockii*, including the type specimen of *W. gracillima*. *W. gracillima* has already been placed in the synonymy of *W. hancockii* (Brown, 1964; Zhang & al., 2013). The seven accessions of *W. glabella* were sampled in Alaska, Central Siberia, Northern Scandinavia, Central Europe and Himalaya, representing its geographical distribution. As in the other circumboreal species the genetic variation is minimal.

The analysis of the nuclear *pgiC* dataset further shows that the accession of the East Asian *W. hancockii* is nested within *W. glabella*. These two species are very similar, but *W. hancockii* is said to have larger indusia and larger pinnae that are both more distantly spaced and more dissected (Brown, 1964; Zhang & al., 2013). Our results indicate that the recognition of *W. hancockii* can be questioned, but further sampling is needed to settle this issue.

Chromosome counts of *W. glabella* have found it to be diploid with the unique number 2n = 78 (Löve & al., 1977; Windham, 1993). The chromosomes of *W. hancockii* have never been counted.

Plants of *W. glabella* from the European Alps have sometimes, because of minor differences in pinnae shape and glands, been recognized as *W. glabella* subsp. *pulchella* (Tutin, 1993). Our results do not support this segregation.

In nuclear *pgiC* there is a sister relationship between *W. glabella*/*W. hancockii* and some of the *pgiC* alleles of the octoploid species in the *W. lanosa* complex (see below under *W. lanosa*).

W. glabella is one of the parent species of the allopolyploid *W. alpina*.

***Woodsia macrochlaena*. —** *W. macrochlaena* is represented by a single sample and an evaluation of its monophyly is therefore not possible. We have found a single sequence for both the nuclear *pgiC* and *RPA2*, and this is supported by the chromosome counts where the species has been identified as a diploid with n = 41 (Kurita, 1965; Mitui, 1966). Based on morphology *W. macrochlaena* has been thought to have its closest affinity with *W. polystichoides*. Both these species have a jointed stipe where the articulation is situated just below the first pinna pair, whereas *W. ilvensis* has the articulation further below (Brown, 1964). However, our study does not give a

conclusive answer with regards to the phylogenetic position of *W. macrochlaena*. The three diploids *W. ilvensis*, *W. polystichoides* and *W. macrochlaena* are found in a well-supported clade, but their interrelationships are only weakly supported and the phylogenies resulting from the chloroplast and nuclear datasets give contradictory topologies (Figs. 1, 2). *W. macrochlaena* is the only diploid species in the genus that is not involved in polyploidization or where a recent ancestor has not been part of a polyploidization event. Its geographic distribution spans North Eastern China, Japan, Korea and Russia.

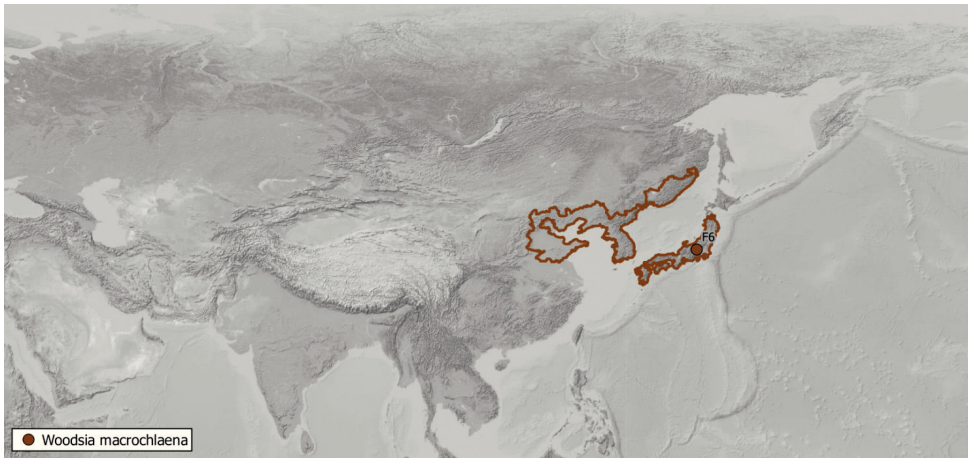

Fig. 5. Distribution of *W. macrochlaena* with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

***Woodsia polystichoides* and autopolyploid *W. intermedia*. —** *Woodsia polystichoides* is represented in this study by four accessions that are retrieved with high support as a monophyletic group in both the chloroplast and nuclear analyses (Figs. 1, 2). It has been identified as diploid with 2n = 82 (Ma, 1985), and in nuclear *pgiC* and *RPA2* we have found all specimens to be homozygous. *W. polystichoides* is widely distributed in East Asia (Fig. 6), and although our samples span Japan, China and Taiwan the molecular variation is very limited. *W. polystichoides* is one of the parent species of the allopolyploid *W. subcordata.*

Woodsia intermedia was described from China by Tagawa (1936), but overlooked by Brown (1964). Chromosome counts have found it to be tetraploid with 2n = 164 (Shimura & Matsumo, 1975) and our data firmly suggest it to be an autopolyploid of *W. polystichoides*. We have only found one nuclear allele in both *RPA2* and *pgiC*, and this pattern is congruent with a genome duplication of a *W. polystichoides* individual that was homozygous in these loci.

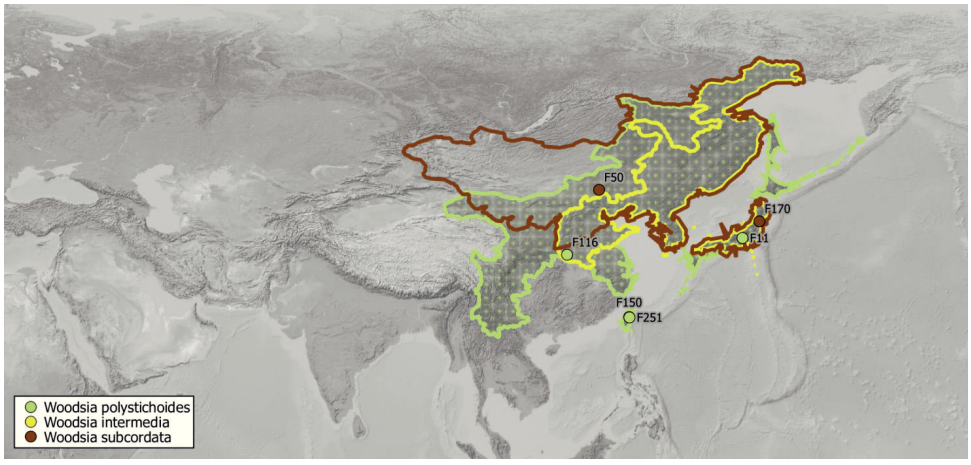

Fig. 6. Distributions of *W. polystichoides, W. intermedia and W. subcordata* with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

***Woodsia subcordata*. —** Our nuclear phylogenies corroborate the view that the East Asian *W. subcordata* is an allopolyploid species recently formed by *W. polystichoides* and *W. ilvensis*. A tetraploid chromosome count together with the intermediate morphology of *W. subcordata* has made earlier researchers suggest this origin (Brown, 1964; Ma, 1985). In nuclear *pgiC* and *RPA2* we have found alleles that are almost identical to both parent species suggesting that the formation of *W. subcordata* is a recent event.

***Woodsia alpina*. —** Our results confirm previous findings that the circumboreal *W. alpina* (Fig. 7) is an allotetraploid between *W. ilvensis* and *W. glabella* (Manton, 1950; Windham, 1993). Furthermore, the results indicate that *W. alpina* has originated at least twice by different allopolyploidization events. The chloroplast sequences from eight of the eleven *W. alpina* sequences are nested among the *W. glabella* accessions, and three are found in the lineage containing the *W. ilvensis* accessions (Fig. 1). This implies that *W. glabella* was the maternal donor for the first eight accessions and *W. ilvensis* for the last three and thus indicating at least two polyploidization events. An alternative interpretation for multiple chloroplast haplotypes could be introgressive hybridization between *W. alpina* and the parent species (Thórsson & al., 2001; Slotte & al., 2008; Oberle & al., 2012). However, all documented cases of the hybrid *W. alpina* × *W. ilvensis* have had malformed spores and such plants would not be able to back-cross (Windham, 1993). Either way, there have been multiple gene transfer events from diploid parental taxa to *W. alpina*.

All eight accessions of *W. alpina* (including two from Alaska) that are nested among the *W. glabella* specimens in the chloroplast phylogeny (Fig. 1) have the same chloroplast *matK* haplotype as one of the *W. glabella* specimens collected from the European Alps (F272). In *W. glabella* we have found another four different haplotypes but none of these is found in a *W. alpina* specimen. This indicates a limitation in the gene flow between diploid *W. glabella* and *W. alpina,* i.e., although the species have originated more than once, or have been the subject of introgression, it does not seem to be a very frequent process.

All *W. alpina* specimens have chloroplast haplotypes identical to specimens of *W. ilvensis* or *W. glabella* and nuclear *pgiC* and *RPA2* that are identical to specimens of both *W. ilvensis* and *W. glabella*. The limited sequence divergence between *W. alpina* and its parent species indicates that the allopolyploidization is recent.

There is disagreement with regards to the exact chromosome number in *W. alpina.* Counts of 2n = 156–164 have been reported (Löve & al., 1977). However, considering that it is the allopolyploid between *W. ilvensis* (n = 41) and *W. glabella* (n = 39), 2n = 160 seem most likely, although the number 2n = 156 would be in line with the sporadic counts of n = 39 in *W. ilvensis* (see above under *W. ilvensis*).

As in its two circumboreal parent species, the sequence variation among the *W. alpina* accessions is none or very minor.

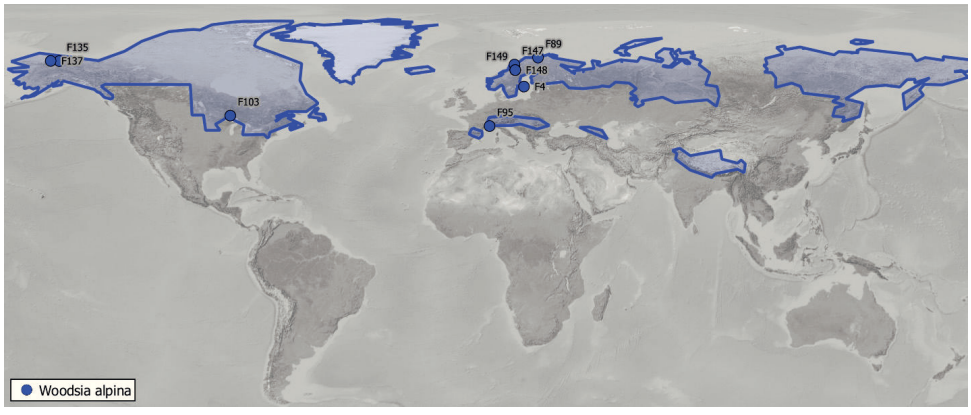

Fig. 7. Distribution of *W. alpina* with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

Asian & Circumboreal clade: Species with non-articulated stipe

We have only detected one presumably diploid species among the non-articulated *Woodsia* in the Asian & Circumboreal clade, provisionally identified as *W.* cf. *cinnamomea*. The remaining species in this non-monophyletic assembly of species are the tetraploid or octoploid *W. rosthorniana*, *W. andersonii*, *W. lanosa*, *W. cycloloba*, *W. okamotoi* and *W. macrospora*.

***Woodsia* cf. *cinnamomea* and autopolyploid *W. macrospora*. —** *W. cinnamomea* was described by Ching (1932), who noted that "it is a very rare species only collected once". Brown (1964) was unable to separate the type of *W. cinnamomea* from *W. macrospora,* but noted that some of the specimens previously determined as *W. macrospora* had small spores (average 49.5 µm), while the type of *W. macrospora* had large spores (average 65.6 µm). This made him suggest that *W. macrospora* might be an autopolyploid *W. cinnamomea* (Brown, 1964).

We have been able to obtain DNA sequence data from an isotype of *W. macrospora* and have found this specimen (F168) to have sequences very similar to a recently collected specimen from Sichuan (F174). The Sichuan specimen has small spores (49.4 µm) and fits very well with the description of *W. cinnamomea,* apart from the color of the stipe that is purplish black as opposed to

the straw-colored stipe of typical *W. macrospora* and *W. cinnamomea*. The spore size of specimen F168 is 59.6 µm. Specimens F168 and F174 have almost identical sequences. We suggest that *W. macrospora* is an autopolyploid *W. cinnamomea*.

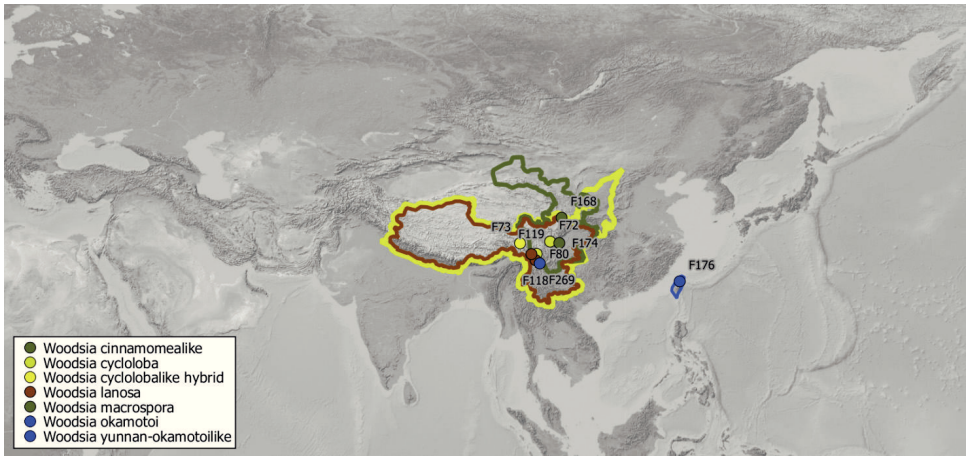

Fig. 8. Distributions of *W.* cf. *cinnamomea* and *W. lanosa* polyploid complex with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

***Woodsia lanosa* polyploid complex (*W. lanosa, W. cycloloba, W. okamotoi*). —** This group of polyploids is, together with *W.* cf. *cinnamomea* and *W. macrospora,* a well-supported clade with four subclades in the chloroplast phylogeny (Fig. 1). The nuclear *pgiC* phylogeny instead reveals a pattern of reticulate evolution where multiple sequence copies form two groups. One group is related to *W.* cf. *cinnamomea* and *W. andersonii* and the other to *W. glabella* (Fig. 2). *W. cycloloba* and *W. lanosa* have been found to be octoploids (Ma, 1985), which is in agreement with the multiple gene copies we have found in the nuclear phylogeny. The ancestor of the diploid *W. glabella* linage has probably been involved in an allopolyploidization event with another of these species ancestors.

All these species (as well as *W.* cf *cinnamomea* and *W. macrospora*) are morphologically similar and have also previously been thought to be closely related (Ching, 1932; Brown, 1964; Ma, 1987; Zhang & al., 2013). Most characters used to separate them are continuous rather than discrete (e.g., shape of pinnae, color of stipe, amount of hair and scales, width and length of indusial segments). These authors have also included *W. andersonii* in this complex. In our analyses the nuclear *pgiC* sequences support a sister relationship between the *W. lanosa*-complex and *W. andersonii,* but in the chloroplast analysis there is no support for such a relationship, whereas the analysis of nuclear *RPA2* is inconclusive.

The Taiwan endemic *W. okamotoi* seems to be very closely related to *W. lanosa* and *W. cycloloba*. Our *W. okamotoi* sample from Taiwan (F176) has a chloroplast haplotype that is identical to the sample F269 from Yunnan, and the nuclear *pgiC* is also almost identical. Morphologically this Yunnan specimen fits well with the description of *W. okamotoi* and along with the molecular similarity, this indicates that the dispersal to Taiwan has happened recently or is a continuous process

keeping the genetic variation between populations low. When *W. okamotoi* was described it was compared to *W. cinnamomea* and *W. cycloloba* but not *W. lanosa* (Tagawa, 1938). Although we have not seen the type material of *W. okamotoi*, the original description fits very well with *W. lanosa*. Molecularly *W. okatomoi* shares one copy of the nuclear pgiC sequence with *W. lanosa*, but lacks the other copy. The relationship between *W. okatomoi* and *W. lanosa* should be further investigated, *W. okamotoi* could, for example, represent a different cytotype than octoploid *W. lanosa.*

We have found a new and previously undetected species (accession F73) that apparently is a recent allopolyploid between *W. lanosa* and *W. cinnamomea/macrospora*. It has *pgiC* alleles that are identical to the two parent species and the morphology is also intermediate.

Woodsia andersonii. — Chromosome counts have identified *W. andersonii* as octoploid (Ma, 1985) and we find two to three divergent nuclear *pgiC* sequences in our specimens. In nuclear *pgiC* there is strong support for *W. andersonii* as sister to the *W. lanosa*-complex, although the molecular distance is substantial to any of the other species. In the chloroplast phylogeny on the other hand, the non-articulated *W. andersonii* is with strong support nested within the Articulated stipe group. *W. andersonii* therefore probably has a reticulate origin involving both articulated and non-articulated ancestors.

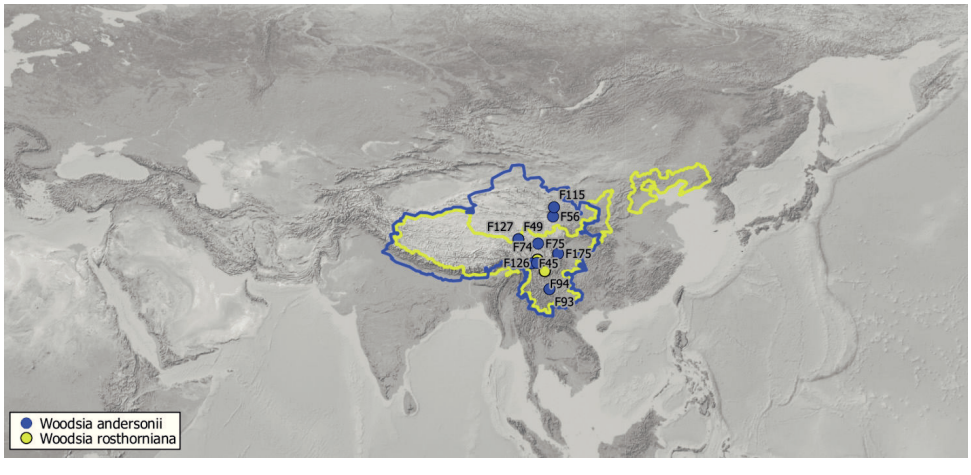

Fig. 9. Distributions of the *W. andersonii* and *W. rosthorniana* with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

Woodsia rosthorniana. — *W. rosthorniana* is a tetraploid (Ma, 1985) that in both the *pgiC* and *RPA2* phylogenies has two very divergent sequences sister to each other, indicating either an autopolyploid origin where the homoeologues now have diverged, or an allopolyploidization between two closely related diploid lineages/species that have gone extinct or have not been sampled.

In both the nuclear and chloroplast phylogenies (Figs. 1, 2) *W. rosthorniana* is situated on a long branch without a resolved sister relationship. This species has previously been suggested to belong to a complex of species together with *W. lanosa* - *W. andersonii* (Ching, 1932; Brown, 1964; Ma, 1985), a placement not supported by our data.

Cheilanthopsis* clade: *Woodsia manchuriensis, W. elongata, W. indusiosa, W. fragilis

In the chloroplast phylogeny, the four Asian species *W. manchuriensis, W. elongata, W. indusiosa* and *W. fragilis* form a well-supported clade, sister to the American clade (Fig. 1). The nuclear *pgiC* phylogeny (Fig. 2), however, puts them in two separate clades, one with *W. manchuriensis* and *W. fragilis* and another with *W. elongata* and *W. indusiosa,* that are successive sister clades to the American clade.

This conflict between the chloroplast and nuclear phylogenies could be explained by incomplete lineage sorting if, for example, the progenitor of the American and the *Cheilanthopsis* clade had a single chloroplast haplotype but two different *pgiC*-alleles (a similar pattern is present in extant *W. scopulina*). Then one of these *pgiC* lineages became fixed as *W. elongata* and the American clade, whereas the other became fixed in *W. manchuriensis* and *W. fragilis*. *W. indusiosa* later evolved as the allopolyploid of the *W. elongata*-lineage and the *W. manchuriensis*-lineage.

The four species in the *Cheilanthopsis* clade are all medium to large, with pinnate fronds and large globose indusia. They have a morphological similarity to some of the species in the American clade, such as *W. mollis*, *W. montevidensis* and *W. scopulina*. Due to differences in morphology and chromosome number from the other Asian species of *Woodsia*, they have often been classified in the segregate genera *Protowoodsia* Ching (for *W. manchuriensis*), *Cheilanthopsis* Hieron. (including *W. elongata* and *W. indusiosa*) and *Hymenocystis* C.A. Mey (for *W. fragilis*) (Ching, 1932, 1945; Brown, 1964; Shmakov, 2003; Zhang & al., 2013). However, according to our results, the segregation of any of these genera would also require segregation of the entire American clade of *Woodsia* species.

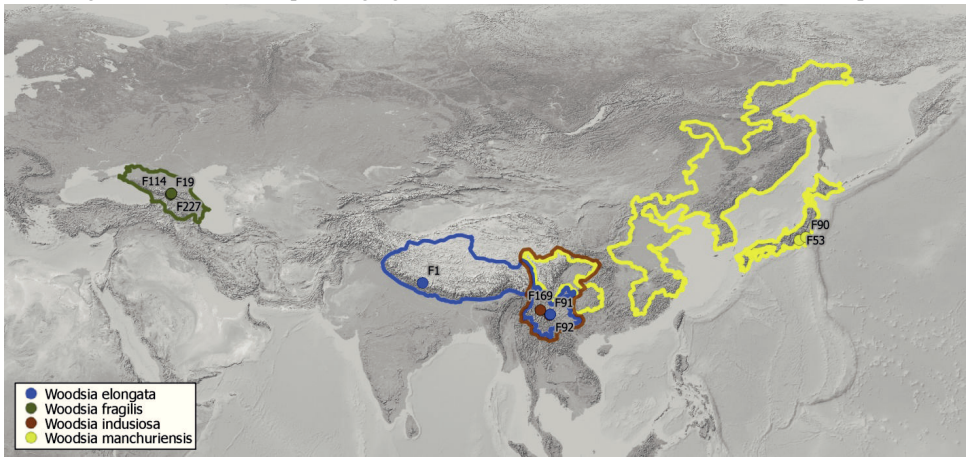

Fig. 10. Distribution of *Cheilanthopsis* clade (*W. elongata*, *W. fragilis*, *W. indusiosa* and *W. manchuriensis*) with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

Woodsia manchuriensis (= *Protowoodsia manchuriensis* (Hook.) Ching)**. —** Both diploid and tetraploid cytotypes are known in this species, either with n = 33 recorded in Japanese specimens (Kurita, 1965; Mitui, 1965) or n = 66 in one Chinese specimen (Ma, 1985). These numbers are

unique to this species within *Woodsia*. Two accessions are included in our phylogeny, both from Japan. In *pgiC* there are two relatively divergent alleles (or homoeologues), and we cannot say if our samples are diploid or polyploid.

An ancestor of *W. manchuriensis* has probably been involved in allopolyploid hybridization together with a species with n = 41 to create *W. indusiosa* (see below under *W. indusiosa*). Nuclear *pgiC* indicates that the *W. manchuriensis* lineage could also be involved in a polyploid hybridization event resulting in *W. fragilis* (see below under *W. fragilis*).

Woodsia elongata (= *Cheilanthopsis elongata* (Hook.) Copel.)**. —** *Woodsia. elongata* includes both diploid and tetraploid cytotypes with chromosome counts of n = 41 or n = 82 (Mehra & Singh, 1955; Ma, 1985). It is homozygous in nuclear *pgiC* and *RPA2* (only one specimen sequenced), and the specimen included in our analyses could be either diploid or autotetraploid. The recurved margin of the frond forming a false indusium is the main reason why this species has sometimes been placed in the segregate genus *Cheilanthopsis* (Zhang & al., 2013). The ancestor of *W. elongata* might be one of the parent species of the polyploid *W. indusiosa*.

Woodsia indusiosa (= *Cheilanthopsis indusiosa* (Christ) Ching)**. —** This species has most frequently been included in the often monotypic genus *Cheilanthopsis* (Ching, 1932; Brown, 1964; Zhang & al., 2013), and its false indusium is even more conspicuous than in *W. elongata* (Ching, 1932; Zhang & al., 2013).

W. indusiosa is probably the result of an allopolyploidization between the *W. manchuriensis* and *W. elongata* lineages. The chromosome number of *W. indusiosa* is n = 74 (Ma, 1985) and considering the base numbers of other *Woodsia* species, the most probable explanation of this polyploid count is 41 + 33. In our nuclear analyses *W. indusiosa* is pairing up with *W. elongata* that has n = 41 or n = 82 (Mehra & Singh, 1955; Ma, 1985). The only known *Woodsia* species with a base number of n = 33 is *W. manchuriensis* (Kurita, 1965; Mitui, 1965; Ma, 1985). However, neither in *RPA2* nor in *pgiC* do we find a second homoeologue that is closely related to those of contemporary *W. manchuriensis*.

Woodsia fragilis (= *Hymenocystis fragilis* (Trev.) Askerov)**. —** This species is isolated in the Caucasus Mountains in contrast to the otherwise East-Asian distribution of the *Cheilanthopsis*-clade. Although there is no chromosome count available for this species, our nuclear sequences (two alleles found in *RPA2* and three alleles in *pgiC*) together with large spores of 55 µm (Electr. suppl. Table S4; Brown, 1964) makes us hypothesize that *W. fragilis* is an allopolyploid of the progenitor of *W. manchuriensis* and another unsampled or extinct diploid linage. Both analyses of the nuclear and chloroplast datasets group together *W. fragilis* and *W. manchuriensis*

American clade 1: *Woodsia scopulina, W. mollis, W. montevidensis*

W. scopulina, W. mollis and *W. montevidensis* form a well-supported clade both in the chloroplast and the nuclear phylogenies. In the chloroplast phylogeny (Fig. 1) *W. scopulina* is sister to the other two species, whereas in the *pgiC* phylogeny (Fig. 2) the *W. scopulina* specimens are found in an unresolved polytomy together with the two other species. All three species have both diploid and polyploid cytotypes. Morphologically they are similar but their non-overlapping distribution areas help in identification.

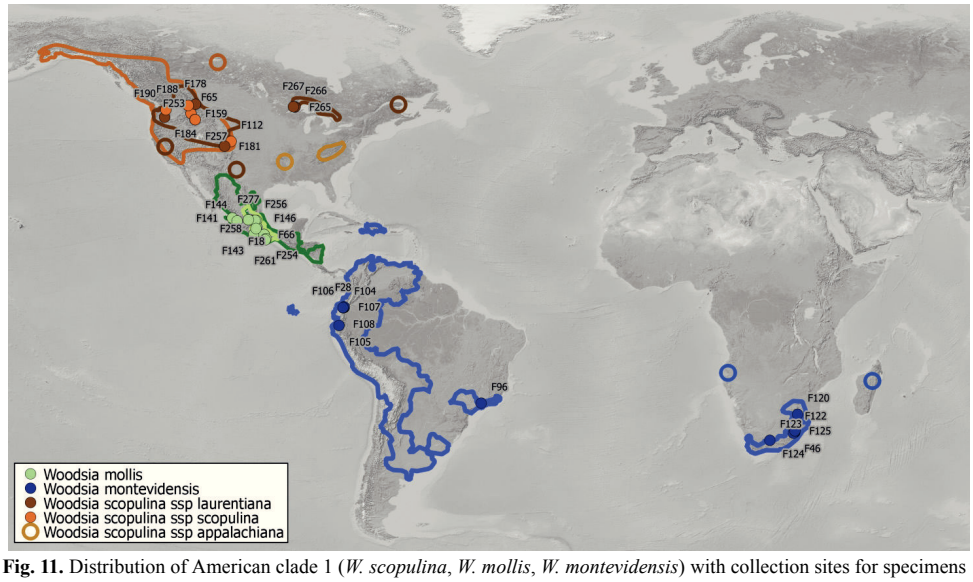

Fig. 11. Distribution of American clade 1 (*W. scopulina*, *W. mollis*, *W. montevidensis*) with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

***Woodsia scopulina*. —** The North American *W. scopulina* has been subdivided into three subspecies that reflect cytotypes and geographic ranges (Windham, 1993). *W. scopulina* subsp. *scopulina* is a diploid with a western distribution, *W. scopulina* subsp. *appalachiana* is a diploid with an eastern distribution, whereas *W. scopulina* subsp. *laurentiana* is a tetraploid distributed in the Great Lakes region and also in the west, where it is sympatric with *W. scopulina* subsp. *scopulina.* (Windham, 1993).

In the chloroplast analysis, all 13 sequenced specimens have the same haplotype (we have included specimens from the western distribution and the Great Lakes region, but no subsp. *appalachiana*). Even though the phylogenetic analyses of the pgiC dataset fail to resolve the *W. scopulina* specimens as a monophyletic group, some of the diploid accessions of subsp. *scopulina* and tetraploid accessions of subsp. *laurentiana* share the same allele. This indicates that *W. scopulina* subsp. *laurentiana* is an autopolyploid of subsp. *scopulina,* or an allopolyploid between subsp. *scopulina* and subsp. *appalachiana* where subsp. *appalachiana* is closely related to subsp. *scopulina.*

The low sequence variation in *W. scopulina* is in sharp contrast to the much greater variation in the southern sister species *W. mollis* and *W. montevidensis*. This might be a result of the Pleistocene glaciations in a similar way as is hypothesized for the circumboreal species, see below.

***Woodsia mollis.* —** Chromosome counts of *W. mollis* have found it to be diploid (n = 38) or triploid (n = 2n = 114) (MDW unpublished). The large-spored specimens have 32 spores per sporangium, as opposed to 64 in the diploid specimens, indicating that the large spored specimen might represent unreduced triploids with an apomictic asexual life cycle (Manton, 1950). This is the only species of *Woodsia* with a spore count of 32 (Ma, 1987; MDW unpublished).

A presumably haploid gametophyte grown from spores of a diploid specimen (F139) contained only one *pgiC* allele, compared to the two alleles of the sporophyte, indicating that *pgiC* sequences

of *W. mollis* sometimes may be alleles and not polyploid homoeologues. However, a gametophyte grown from spores of a polyploid specimen (F143) had three *pgiC* sequences identical to the three sequences of the sporophyte (presumably two alleles and two homoeologues), which is in line with an apomictic life cycle where the gametophyte not is haploid, but unreduced triploid.

The samples of *W. mollis* form two very distinct lineages in both the chloroplast and nuclear phylogenies. The molecular distance between these lineages is as large as the distance to the closest related species, *W. scopulina* and *W. montevidensis*. Spore measurements (Electr. Suppl. Table S2) and allelic counts of nuclear *pgiC* indicate that there are diploid (i.e., F139, F144, F249, F261) and polyploid (i.e., F18, F141, F143) specimens in both lineages. The results of the analyses of the chloroplast dataset show that that polyploid specimens occur in both of the two different lineages, which indicate multiple polyploidization. Furthermore, as the polyploids have matK haplotypes that are identical or almost identical to sequences of diploids, this polyploidization is probably recent. The indusium of *W. mollis* varies from commonly globose to plate-like (Mickel & Windham, 2004). However the shape of the indusium is not correlated with the two lineages of *W. mollis*. In the same lineage, specimens with globose or plate-like indusia have been found intermixed.

Woodsia canescens is sometimes treated as distinct from *W. mollis,* mainly due to its plate-like indusia, in contrast to the globose indusia in *W. mollis*, and also the characteristics of the segment margins (Mickel & Windham, 2004). As noted above the characteristics of the indusium is varying in both of the lineages and we also find nuclear *pgiC* sequences of specimens identified as *W. canescens* nested among sequences of *W. mollis*.

***Woodsia montevidensis.* —** *Woodsia montevidensis* has a disjunct distribution in Africa and South America (Fig. 11) and Brown (1964) suggested that this disjunction could be the result of migration from the "Himalayan" area of Gondwanaland into today's Africa and South America. In both the chloroplast and *pgiC* phylogenies (Figs. 1, 2) the African samples are nested among the South American ones indicating that the direction of dispersal is from South America to Africa. The closest relatives of *W. montevidensis* are *W. mollis* and *W. scopulina,* which are distributed in Central and North America, indicating that *Woodsia* migrated into South America from the north.

All African samples of *W. montevidensis* have the same *matK* haplotype, indicating that they might stem from a single colonization event. The *pgiC* sequences of the African samples form two monophyletic groups (F120a versus F120b and F46a in Fig. 2) that either could be the result of a polyploidization event or simply divergent alleles from a diploid specimen dispersed from South America. This *pgiC* sequence pattern neither supports nor refutes a single colonization hypothesis. The molecular distance between the African and South American samples is no greater than the variation within the South American samples. The dated chloroplast phylogeny indicates that the dispersal from South America occurred sometime within the last 1.5 Ma (Fig. 3).

In Africa *W. montevidensis* occurs in montane areas in South Africa, Lesotho and Zimbabwe and is also found in one isolated location in Angola and one on Madagascar. The synonym *W. burgessiana* Gerr. ex Hook. & Baker, or *W. montevidensis* var. *burgessiana* (Gerr. ex Hook. & Baker) Schelpe, is described from South Africa, whereas *W. angolensis* Schelpe is based on material from Angola.

In South America it ranges from Colombia to Argentina, Uruguay and eastern Brazil, and there are also records from the Galapagos Islands (CFD Galapagos Species Checklists, 2013) and the Island of

Hispaniola (Haiti and Dominican Republic).

There are no chromosome counts available for *W. montevidensis,* but spore sizes of different specimens and the number of nuclear *pgiC* alleles found indicate that there are both diploid and polyploid specimens in our sample. All specimens, except F96, have two *pgiC* alleles. F96 instead has three alleles and relatively large spores, indicating polyploidy. There are also African specimens, such as F123 and AL811, which may be polyploid as indicated by their large spore size, but there are also some specimens, such as F120, that have an intermediate spore size and are impossible to assign to cytotype.

American clade 2: *Woodsia obtusa, W. oregana, W. phillipsii, W. neomexicana, W. plummerae, W. cochisensis, W. mexicana*

Three diploid taxa, *Woodsia obtusa* subsp. *occidentalis*, *W. oregana* subsp. *oregana* and *W. phillipsii*, are found in this clade along with several auto- and allopolyploids, such as *W. oregana* subsp. *cathcartiana, W. obtusa* subsp. *obtusa, W. plummerae, W. mexicana, W. cochisensis* and *W. neomexicana*, and numerous hybrids. In the chloroplast phylogeny the analyzed samples group into three distinct subclades representing the three diploid taxa (Fig. 1). Both the *W. obtusa* clade and the *W. phillipsii/mexicana*-complex can be further subdivided into two distinct chloroplast clades each.

In the *pgiC* phylogeny (Fig. 2) *W. oregana* subsp. *oregana* has sequences in two clades (Clade oregana 1 and 2) that are clearly separated from each other in the phylogeny and thus not each others closest relatives. The clades represent diploid alleles and not gene copies as verified by gametophyte sequencing (see below under *W. oregana*). *W. obtusa* also has two clades of divergent *pgiC* alleles, but in this case they are sister clades with two matching distinct chloroplast lineages.

The *W. phillipsii/mexicana*-complex is comprised of four clades of *pgiC* alleles (Clade phillipsii, Clade mexicana 2, Clade neomexicana 1 and Clade neomexicana 2). In Clade phillipsii the diploid *W. phillipsii* specimens are found, whereas the three other clades are represented by polyploid species only. However, we have found *W. neomexicana* to be an autopolyploid of an unsampled diploid. The spore size also indicate that specimen F217 (W. cf. "coloradoensis") is diploid and could be a representative of the missing diploid.

The numerous distinct *pgiC* clades are correlated with only small variations in the matK haplotypes. This could be the result of ancient hybridization events where fewer chloroplast lineages have been retained. The multiple divergent *pgiC* lineages connected with one chloroplast could be the result of ancient homoploid or polyploid hybridization. The fact that the various allo- and autopolyploid taxa have matK haplotypes or nuclear alleles that are identical or very similar to the diploid parents suggests that the polyploidizations are recent.

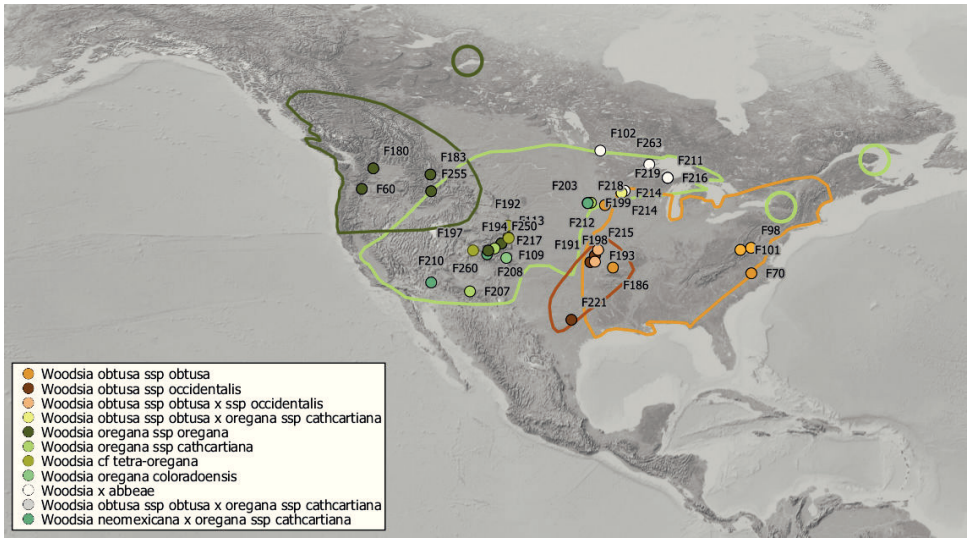

Fig. 12. Distributions of taxa in American clade 2 with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

Woodsia obtusa. — In all phylogenies *W. obtusa* is found to be monophyletic. Although the species includes both diploid (subsp. *occidentalis*) and autotetraploid (subsp. *obtusa*) taxa, it is not involved in extensive allopolyploid speciation or hybridization with other species, the only instances being the triploid hybrid *W.* × *kansana* (*W. obtusa* subsp. *occidentalis* x *W. oregana* subsp. *cathcartiana*) and the triploid hybrid *W. obtusa* subsp. *obtusa* × *W. obtusa* subsp. *occidentalis.* Apart from the ploidy level the two subspecies also have slightly divergent morphology and geographic and ecological ranges (Windham, 1993). They both have an eastern North American distribution (Fig. 12).

The sequenced specimens group into two distinct clades in both the chloroplast and nuclear *pgiC* phylogenies, one consisting of only subsp. *obtusa* and the other of both subsp. *obtusa* and subsp. *occidentalis.* The tetraploid subsp. *obtusa* has probably originated at least twice, since *matK* haplotypes representing both clades were found. When tetraploid subsp. *obtusa* has the chloroplast haplotype of subsp. *occidentalis* their sequences are identical, indicating that the tetraploidization is a recent event. If so, a cryptic diploid lineage might exist with the chloroplast and nuclear sequences in the second clade where we have only found tetraploid *W. obtusa* subsp. *obtusa*. Small-spored *W. obtusa* specimens seen in the southern Appalachian Mountains (Windham, pers. obs.) could possibly represent this cryptic/overseen diploid.

Our phylogenies confirm previous observations (Windham, 1993) that the two subspecies hybridize in areas with overlapping distributions, resulting in sterile triploid specimens.

In the Great Lakes region where both *W. obtusa* subsp. *obtusa* and *W. oregana* subsp. *cathcartiana* occur, we have identified the sterile tetraploid *W.* × *kansana* (specimen F214 in Figs. 1, 2) which we confirm is a hybrid between these two tetraploid taxa, as earlier postulated (Brooks,

1982; Windham, 1993).

***Woodsia oregana.* —** *W. oregana*, with its diploid subsp. *oregana* (n = 38) and tetraploid subsp. *cathcartiana* (n = 76), has been noted as a source of taxonomic difficulties due to its variability and promiscuity (Windham, 1993). One reason for these difficulties is that *W. oregana* subsp. *cathcartiana* is not an autopolyploid of diploid subsp. *oregana* as suggested by Brown (1964), but an allopolyploid between *W. oregana* subsp. *oregana* and a species in the *W. phillipsii/mexicana*-complex. In subsp. *cathcartiana* some specimens (F109, F203) were found to have the subsp. *oregana* chloroplast haplotype and others (F210F210, F207) the *W. neomexicana* chloroplast haplotype, indicating that this polyploidization has occurred multiple times. As there are no differences in the chloroplast sequences between subsp. *oregana* and subsp. *cathcartiana* this seems to be recent. The repeated polyploidizations have apparently given rise to a morphological variation ranging all the way from typical *W. oregana* subsp. *oregana* to species in the *W. phillipsii/mexicana*-complex. This would explain, for example, why certain specimens of *W. plummerae* (a polyploid in the *W. mexicana*-complex) are very difficult to separate from *W. oregana* subsp. *cathcartiana* (Windham, 1993).

In the nuclear *pgiC* phylogeny the sequences of diploid *W. oregana* subsp. *oregana* are found in two very divergent clades. To verify that the divergent sequences are alleles and not gene copies we have sequenced a haploid gametophyte that contained only one of the two alleles found in the parent sporophyte. Another indication for this is provided by the specimen F248 that has *pgiC* alleles in one of the clades only. This pattern may be the result of a homoploid hybridization, where the *pgiC* sequences from both the parent species have been preserved as alleles and one of the chloroplast lineages has gone extinct.

Our results are congruent with earlier suggestions that *W. oregana* subsp. *cathcartiana* forms sterile tetraploid hybrids with *W. neomexicana* and triploid hybrids with diploid *W. phillipsii* where they come in contact (Windham, 1993).

W. × *abbeae* is an extreme hybrid between *W. oregana* subsp. *cathcartiana* and *W. ilvensis* found in the Great Lakes region where the parent species meet (see further under *W. ilvensis*).

Woodsia phillipsii/mexicana* complex: *W. phillipsii, W. neomexicana, W. plummerae, W. mexicana, W. cochisensis

All these species, *W. phillipsii, W. neomexicana, W. plummerae, W. mexicana, and W. cochisensis* fall into one clade in the chloroplast phylogeny (Fig. 1). This clade can be further subdivided into two subclades, where *W. phillipsii* (the only diploid found in this complex) is represented in one of the subclades, whereas in the other the autopolyploid *W. neomexicana* is represented but no diploids are found. These two chloroplast subclades are each compatible with two distinct nuclear *pgiC* clades, making this complex contain two chloroplast clades and four nuclear *pgiC* clades. There is no strong support for the relationships between the *pgiC* clades. Hybridization and loss of chloroplast haplotypes (or the existence of unsampled diploids) are needed to explain this pattern of only one diploid taxon and four distinct clades of nuclear *pgiC* alleles. The number of auto- and allopolyploids and the several known hybrids in this complex of lineages indicate that they have had a reticulate history that we today cannot fully resolve.

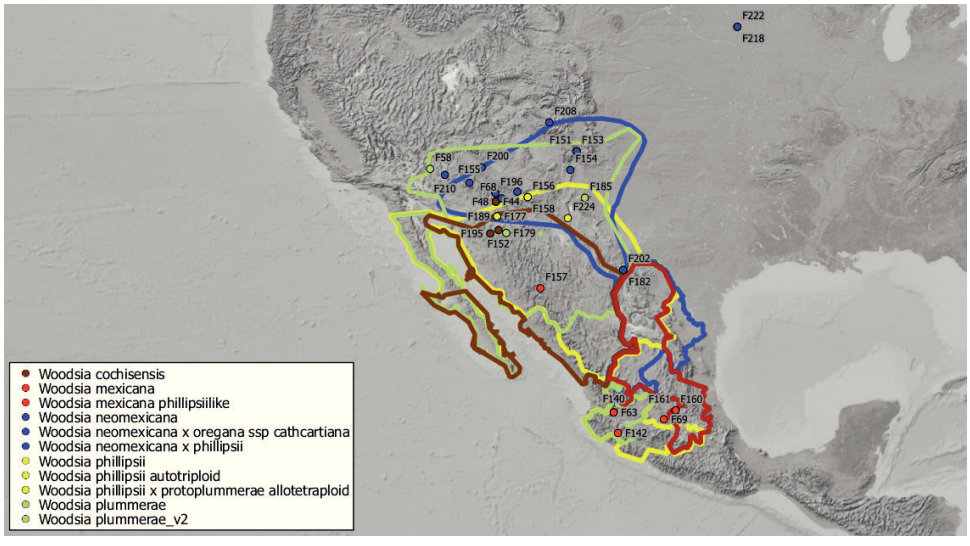

Fig. 13. Distributions of taxa within the *W. phillipsii/mexicana* complex with collection sites for specimens indicated. Range boundaries compiled from published revisions, floras and checklists (see text for references).

Wooodsia phillipsii. — *Woodsia phillipsii* was identified as *W. mexicana* until 1993 when, based on morphological and cytological studies, this diploid (n = 38) was separated from the otherwise tetraploid taxa in the complex (Windham, 1993). All samples of *W. phillipsii* sequenced fall into one of the two subclades in the chloroplast phylogeny and into one of the four clades in the *pgiC* phylogeny in the *W. phillipsii/mexicana* complex.

Woodsia neomexicana. — *Woodsia neomexicana* is a tetraploid of which all our samples fall into the second chloroplast subclade (the one without *W. phillipsii*) in the phylogeny. The *pgiC* sequences fall into two of the four clades (clade neomexicana 1 & 2) in the *W. phillipsii/mexicana* complex. Some specimens have sequences in one of the clades only (F154, F155, F68, F200, F202, F196, F206) and some in both *pgiC* clades (F151, F153, F207). Our data does not clarify whether this second clade of *pgiC* sequences represents another diploid lineage of which the chloroplast was not sampled, or if it represents a divergent *pgiC* allele of the same *W. neomexicana*-lineage. Since we have no indication of an unsampled chloroplast we interpret *W. neomexicana* as an autopolyploid of one of the chloroplast lineages in the *W. phillipsii/mexicana* complex. *W. neomexicana* was previously thought to be derived from hybridization between *W. oregana* and *W. phillipsii* because of the intermediate dissection of the frond (Windham, 1993), but our analyses do not reveal any close relationship with *W. oregana*. There is also a specimen (F208) of the hybrid *W. neomexicana* ×*W. oregana* subsp. *cathcartiana* that has sequences from both of the *W. neomexicana pgiC* clades, as well as from the *W. oregana* subsp. *oregana* clade.

Woodsia plummerae. — *Woodsia plummerae* is tetraploid with *pgiC* sequences that fall both into the *W. neomexicana* clade of *pgiC*-sequences and into the clade represented by the diploid *W. phillipsii*. The analysis of the chloroplast dataset places most of the *W. plummerae* specimens in the

W. neomexicana subclade, but we have also found a specimen (F156) that has a chloroplast haplotype in the *W. phillipsii* subclade. We therefore propose that *W. plummerae* is an allopolyploid of *W. phillipsii* and the diploid progenitor of the autopolyploid *W. neomexicana. W. plummerae* has sequences that are identical or very similar to those of *W. neomexicana* and *W. phillipsii,* suggesting that this allopolyploidization is recent.

Woodsia mexicana* and *W. cochisensis. — All our samples of *Woodsia mexicana* and *W. cochisensis* are most likely tetraploids, as indicated by spore size, nuclear sequences and chromosome counts. They all have the *W. phillipsii* lineage chloroplast haplotype. They have *pgiC* sequences that fall into the *W. phillipsii* clade, but there is also a second *pgiC* copy from a clade of *pgiC* sequences that is exclusive to these *W. mexicana* and *W. cochisensis* samples. We cannot say if this *W. mexicana* clade of *pgiC* sequences represents another diploid lineage of which a chloroplast has not been sampled or if it represents a divergent *pgiC* lineage with the same chloroplast as the *W. phillipsii*-lineage. This makes us interpret *W. mexicana* and *W. cochisensis* as allopolyploids of the *W. phillipsii*-lineage and the *W. mexicana*-lineage. *W. cochisensis* was described in Windham (1993) based on specimens that had previously been identified as *W. plummerae* or *W. mexicana*. The main morphological differences between *W. mexicana* and *W. cochisensis* are the shape of the indusia and the number of glands on the leaves (Mickel & Windham, 2004). However, we find these characters to vary substantially between sequence-wise closely related specimens.

Species not seen and not included

Woodsdia guizhouensis P. S. Wang, Q. Luo & Li Bing Zhang, Novon. 22: 191. 2012. According to the description this species is very similar to *W. andersonii* in general leaf morphology, but differs in its slightly larger indusia that more resemble those of *W. rosthorniana* (Luo & Zhang, 2012).

Woodsia oblonga Ching & S. H. Wu, Fl. Tsinling. 2: 221. 1974. Said in the protologue to be closely related to *W. macrochlaena* but with cup-shaped indusia.

Woodsia pilosa Ching, Fl. Tsinling. 2: 222. 1974. Said in the protologue to be closely related to *W. polystichoides* but with cup-shaped indusia.

Woodsia shensiensis Ching, Sinensia. 3: 141. 1932. In the original publication this is said to be nearest to *W. hancockii,* but much larger and with stipe, rachis and undersurface of leaves densely clothed in long fine flexible hairs.

Woodsia sinica Ching, Sinensia. 3: 145. 1932. This is stated to be closely related to *W. polystichoides* in the original publication and Brown (1964) suggests that it may be a small shade form of either *W. macrochlaena* or *W. polystichoides*.

Cheilanthopsis kangdingensis (H. S. Kung, Li Bing Zhang & X. S. Guo) Shmakov in Chandra & Srivastava, Pteridol. New Millennium. 64. 2003. Closely related to *W. elongata* according to Luo & Zhang (2012).

GENERAL DISCUSSION

High frequency of polyploids and hybrids in *Woodsia* and closely related families

Out of the 36 taxa of *Woodsia* included in the analyses (homoploid hybrids excluded), 22 (61%) are polyploids, 14 (39%) are diploid, and four taxa have both cytotypes (Table 3). Of the 22 polyploids, 11 (50%) are autopolyploids and 11 are allopolyploids. This frequency of polyploids is substantially higher than the infrageneric polyploid frequencies that have been reported for leptosporangiate ferns in general (33%) or vascular plants (34.5%) (Wood & al., 2009).

In the expanded analysis where Woodsiaceae is included together with the nine most closely related families (the Eupolypods II clade), 346 (54%) of the species are estimated to be polyploid and 291 (46%) diploid (Electr. Suppl. Table S1).

One reason for this big discrepancy between our results and those of Wood et al. (2009) is that they did not count polyploids when there was also a diploid count for the same species. We treat such cases as two separate taxa, one diploid and one polyploid, in line with Soltis & al. (2007), who argue that reproductive isolation could be just as efficient in morphologically distinct as in morphologically cryptic polyploids.

Although we recognize the different cytotypes as largely reproductively isolated entities, we prefer leaving the cases of unnamed cytotypes as they are. In *Woodsia* the autopolyploids and their diploid progenitors have either been treated as subspecies (*W. obtusa*, *W. scopulina*), as distinct species (*W. intermedia/W. polystichoides*), or have not been formally named (*W. mollis*, *W. montevidensis*, *W. elongata*, *W. manchuriensis*). We believe that an informal reference to the ploidy level in species with multiple cytotypes is a better option than formal naming.

Six probable homoploid hybrids have been detected in our phylogenies (Table 3, Figs. 1–2) and at least another four homoploid hybrids have been reported in the literature (*W.* × *maxonii*, *W. scopulina* subsp. *scopulina* × subsp. *laurentiana*, *W. plummerae* × *W. phillipsii*, *W. alpina* × *W. ilvensis*).

Speciation through polyploidization common in *Woodsia* and related fern families

Species formation through polyploidization has been shown to be frequent among vascular plants (Wood & al., 2009), who estimated a speciation frequency by polyploidization of 34% for leptosporangiate ferns in general and 15% for flowering plants. In *Woodsia* we estimate that 54% of all speciation events are connected to a shift in ploidy level (Electr. Suppl.: Fig S3.; Table S3), i.e. it is considerably higher than in leptosporangiate ferns in general. In the expanded analysis with all 10 families in the Eupolypods II clade the polyploid speciation rate is 46% (118 polyploid speciations/258 speciation events), again higher than in leptosporangiate ferns in general.

A major reason for the discrepancy between our results and those of Wood & al. (2009) is that they did not count species with both diploid and polyploid counts as a speciation event. If our data is reanalyzed without counting autopolyploids as speciation events, the polyploid speciation rate is 51% in *Woodsia* and 38% in the whole Eupolypods II clade. However, we argue that taxa with multiple cytotypes should be counted as separate species and the forming of a new cytotype as a speciation event, as the genetic circumscription of taxonomically defined auto- or allopolyploids is arbitrary. Another reason for the discrepancy is that Wood & al. (2009) only counted infrageneric shifts (as we do with *Woodsia*), whereas we counted all shifts for the much larger Eupolypods II

phylogeny, this effect should on the other hand push the results in the opposite direction since polyploids have higher extinction rate (see below), and there will be more speciation events only involving diploids because we include more distantly related taxa and therefore an older most recent common ancestor.

Multiple formations or introgression?

Three of the allopolyploid taxa of *Woodsia*, *W. alpina*, *W. oregana* subsp. *cathcartiana* and *W. plummerae*, include specimens with *matK* chloroplast haplotypes from both parent species (reciprocal formation). Another two "autopolyploid" species, *W. obtusa* subsp. *obtusa* and *W. mollis*, have specimens with different chloroplast haplotypes that represent different diploid chloroplast lineages within the species.

These non-monophyletic polyploid lineages could either be the result of multiple formations of the polyploid species or, alternatively, the result of gene transfer via introgression between the diploid parents and the polyploid. With our data we cannot discriminate between the two processes, but either way the genetic variation has increased in the polyploid species compared to a single polyploidization event, as has been shown in several cases previously (Thórsson & al., 2001; Slotte & al., 2008; Oberle & al., 2012). Although the genetic variation within the polyploid species has increased compared to that of a single founder event, there could still be compatibility problems between lineages that arise from different events due to varying patterns of gene silencing and gene expression in different linages (Werth & Windham, 1991; Soltis & al., 2004).

High extinction rate of polyploid species - few long surviving polyploid lineages

It has been shown that polyploids in general diversify at lower rates and have a higher extinction rate than diploids (Mayrose & al., 2011). This implies that most polyploid species should be young and continuously replaced by new ones formed by the diploid stock.

In *Woodsia*, 15 out of 22 polyploids have formed relatively recently (less than 2 Ma), and the DNA regions analysed here show that the DNA sequences retrieved from the polyploids are identical or very similar to those of the diploid parent taxa. Seven of the polyploid *Woodsia* species could be older, 2–20 Ma (Fig. 3; Table 3). These estimates of the age of the polyploids are a maximum upper limit, calculated as the time when the polyploid lineage intersects with an existing diploid lineage. Due to extinction or the existence of unsampled diploids there could be more recent ancestors of the polyploid lineages, the age of specific polyploidizations could be greatly overestimated.

The few old polyploid lineages detected in *Woodsia* are congruent with the findings of Mayrose & al. (2011), who showed that polyploid lineages of both seedless vascular plants and seed plants have a higher extinction rate than diploid lineages. Asexual polyploid fern lineages in the genus *Astrolepis* Benham have also been shown to be short lived (Beck & al., 2011).

Old surviving polyploid lineages are also rare in the expanded Eupolypods II analysis. No polyploid lineage older than 50 Ma seems to exist, whereas there are 20 diploid lineages of this age. There are only 19 polyploid lineages possibly older than 25 Ma, compared to 56 diploid lineages (Electr. Suppl.: Fig S4.; Table S4).

Numerous hypotheses have been put forward to explain why polyploid lineages have a lower fitness than diploid lineages, see reviews by Soltis & al. (2004), Comai (2005). Two major factors mentioned are reduction in outcrossing capabilities due to population incompatibilities when

different populations might silence different duplicated gene copies (Werth & Windham, 1991), and changes in gene expression (Comai, 2005). There may also be advantages with polyploidy, not the least in the short run, in allopolyploids for example through fixed heterozygosity that will reduce the risk for accumulating recessive mutations (Soltis & Soltis, 2000), and through increased individual variation due to the existence of multiple gene copies that will reduce the risk of suffering from deleterious mutations, particularly important during the haploid gametophyte generation (Comai, 2005).

Woodsia and most ferns have high base chromosome numbers (n = 33–41 in *Woodsia*), and this has given rise to the hypothesis that all or most ferns are ancient polyploids, but there is very little concrete evidence (Duncan & Smith, 1978).

Although ferns have high base numbers, the base numbers seem very stable during the last 100 MA and there is no indication of ancient polyploidization within *Woodsia* or the Eupolypods II during this time (Electr. Suppl. Fig S4). In *Woodsia* the base numbers of extant diploid taxa are 33, 38, 39 and 41, and all polyploids have multiples of these. When reconstructing the ancestral chromosome states the most parsimonious base number of the most common recent ancestor is x = 41, with three independent reductions to 33, 38 and 39 (Electr. Suppl. Fig S4). In the parsimony base number reconstruction of the whole Eupolypods II clade, the ancient chromosome base number is 40–42, and also in this expanded analysis there are no indications of an ancient polyploidization since the most common recent ancestor of the clade (Electr. Suppl. Fig S4).

If any of the current diploid lineages were the result of an ancient polyploidization after the most common recent ancestor of the Eupolypods II clade, we would probably see a substantial shift in base number in that lineage, unless almost all duplicated chromosomes had been lost through aneuploidy. Polyhaploidy, the instantaneous reduction of a tetraploid genome to a diploid, is not commonly accepted as a mechanism for chromosome reduction in polyploids, instead a gradual loss of duplicated chromosomes (aneuploidy) over time is the preferred explanation (Mayrose & al., 2011).

Although it might be intuitive to think that it would be possible to spot older polyploidizations through some intermediate chromosome numbers, this might not always be the case as illustrated by the angiosperm family Fabaceae, where a proposed WGD (whole genome duplication) event ca. 50 Ma is not clearly detected in extant chromosome numbers (Doyle, 2012). A possible way to shed light on ancient polyploidizations not detectable by chromosome counts of extant species, would be to analyze the K-values of duplicated genes to detect if many duplicated genes have a similar age (Cui & al., 2006). More than 50 fern transcriptomes that could be used for such an analysis are available via the 1KP-project (OneKP, 2013)

Conclusion about polyploidy and hybridization in *Woodsia* and related ferns

Polyploidization is probably the most common mode of speciation in *Woodsia* and the eupolypods II clade, and although polyploids have a higher extinction rate than diploids, they do sometimes survive and may acquire beneficial mutations and make good use of some of the duplicated genes. Their survival does not have to be directly attributed to the fact that they are polyploid, there might be other factors that make some of them survive, but since polyploids are common it is likely that at least some survive in the long run.

Our results support the idea that although polyploidizations are common among plants and are

one of the most important speciation processes, due to the higher extinction rate polyploids rarely succeed to persist in the long term (Mayrose & al., 2011). Polyploids are occasionally successful over a long period of time but this might just be because there are so many of them, and when they go extinct they are continuously being replaced by new ones from the more long-lived diploid lineages.

Low sequence variability within the northern taxa

In *Woodsia ilvensis* 16 specimens were sequenced from all parts of its circumboreal distribution (Figs. 4,7), but in all five chloroplast regions used no sequence variation was found apart from a single SNP (single nucleotide polymorphism) in specimens from North-East Asia. All sequences of the nuclear markers *pgiC* and *RPA2* were identical. Similarly, low genetic variation is also characteristic for the other two circumboreal species, *W. glabella* and *W. alpina*. This is in contrast to the larger interspecific variation found in most of the species outside the circumboreal region, such as *W. mollis*, *W. montevidensis* and *W. obtusa*.

With the reservation that our genetic markers might be evolving to slowly for describing the scenario correct, the low sequence variation in the circumboreal species is in line with a "Leading Edge" colonization scenario (Hewitt, 1996). Rapid latitudinal colonization after the Pleistocene glaciations could lead to lower genetic variation due to repeated bottlenecks, where the initial founder populations in turn recolonize the gradually emergent land and at the same time become a barrier for the successful dispersal of the populations left behind (Hewitt, 2004; Ehrich & al., 2007). In the case of the circumboreal *Woodsia* species, the low genetic variation is not only along the latitudinal axis but also along the longitudinal one, indicating either low species variation already before recolonization or a dominant longitudinal colonization from one of the refugial populations.

Summary of geographic patterns in *Woodsia*

Sister species in *Woodsia* have almost exclusively neighboring or overlapping distribution areas (Figs. 4–13), indicating that although *Woodsia* spores are dispersed by wind, long distance dispersal with establishment is rare. Apart from this general distribution pattern there are some interesting disjunctions and deviating distributions.

There are three species with a similar circumboreal distribution, the diploids *W. ilvensis*, *W. glabella* and the allopolyploid of these species, *W. alpina.* The distribution of *W. alpina* could be a result of multiple polyploidization events over the distribution area of the parent species, or few polyploidization events with dispersal. The circumboreal species *W. ilvensis* and *W. glabella* do not have an obvious sister-relationship in the phylogeny and their adaptations to the circumboreal habitat probably is of independent origin.

Although these circumboreal species have a distribution that includes North America, it is not one of their recent ancestors that established the "American clade" of *Woodsia* species (Fig. 1). The most recent ancestor of the American clade is instead shared with the Asian *"Cheilanthopsis* clade". Since the American clade (Figs. 1–2) is monophyletic and nested within Asian species, a single introduction to America 19 Ma (Fig. 3) explains the biogeographic pattern of this clade.

One species in the American clade, *W. montevidensis*, has its main distribution in South America (including Hispaniola in the Caribbean and the Galapagos Islands), but also has made a relatively recent leap to Southern Africa and Madagascar. So, after its origin in Asia, the "American &

Cheilanthopsis clade" has made a remarkable global round-the-world trip, first to North America, then to South America, and via Southern Africa to Madagascar andalmost back to Asia again.

ACKNOWLEDGEMENTS

This work was supported by grants from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas) to P.K. (2006-429 and 2010-585). The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project p2009050.

LITERATURE CITED

Barrington, D. S., Paris, C. A., & Ranker, T. A. 1986. Systematic Inferences from Spore and Stomate Size in the Ferns. *American Fern Journal* 76:149-159. http://dx.doi.org/10.2307/1547723

Beck, J. B., Windham, M. D., & Pryer, K. M. 2011. Do asexual polyploid lineages lead short evolutionary lives? A case study from the fern genus *Astrolepis. Evolution* 65:3217.

Brooks, R. E. 1982. A New *Woodsia* Hybrid from Kansas. *American Fern Journal* 72:79-84. http://dx.doi.org/10.2307/1546600

Brown, D. F. M. 1964. *A monographic study of the fern genus Woodsia* (Vol. H. 16). J. Cramer

Brysting, A. K., Mathiesen, C., & Marcussen, T.. 2011. Challenges in polyploid phylogenetic reconstruction: A case story from the arctic-alpine *Cerastium alpinum* complex. *Taxon* 60:333-347.

CFD Galapagos Species Checklists. 2013. CDF Galapagos Species Checklists. Retrieved January 14, 2014, from http://www.darwinfoundation.org/datazone/checklists/

Ching, R. C. 1932. The Studies of Chinese Ferns VIII. In *Sinensia* (Vol. 3, pp. 131-156).

Ching, R. C. 1945. The Studies of Cinese ferns XXXII. *Lingnan Science Journal* 21.

Comai, L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* 6:836-846. http://dx.doi.org/10.1038/nrg1711

Crouch, N., Burrows, J., Klopper, R., & more, & 0. 2012. *Ferns of Southern Africa: A Comprehensive Guide*. Cape Town : London: Random House Struik.

Cui, L., Wall, P. K., Leebens-Mack, J. H., Lindsay, B. G., Soltis, D. E., Doyle, J. J., … dePamphilis, C. W. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research* 16:738-749. http://dx.doi.org/10.1101/gr.4825606

Doyle, J. J. 2012. Polyploidy in Legumes. In P. Soltis & D. E. Soltis (Eds.), *Polyploidy and Genome Evolution*. Springer.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29:1969-1973. http://dx.doi.org/10.1093/molbev/mss075

Duncan, T., & Smith, A. R. 1978. Primary Basic Chromosome Numbers in Ferns: Facts or Fantasies? *Systematic Botany* 3:105-114. http://dx.doi.org/10.2307/2418535

Dyer, R. J., Savolainen, V., & Schneider, H. 2012. Apomixis and reticulate evolution in the *Asplenium monanthes* fern complex. *Annals of Botany* 110:1515-1529.

http://dx.doi.org/10.1093/aob/mcs202

Ehrich, D., Gaudeul, M., Assefa, A., Koch, M. A., Mummenhoff, K., Nemomissa, S., … Brochmann, C. 2007. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* 16:2542–2559. http://dx.doi.org/10.1111/j.1365-294X.2007.03299.x

Feng, Y., Oh, S.-H., Manos, P. S., & Zomlefer, W. B. 2005. Phylogeny and Historical Biogeography of the Genus *Platanus* as Inferred From Nuclear and Chloroplast DNA. *Systematic Botany* 30:786-799. http://dx.doi.org/10.1600/036364405775097851

FloraGREIF Team. 2013. FloraGREIF - Virtual Flora of Mongolia. Retrieved May 11, 2014, from http://greif.uni-greifswald.de/floragreif/

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58:247-276. http://dx.doi.org/10.1006/bijl.1996.0035

Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 359:183-195; discussion 195. http://dx.doi.org/10.1098/rstb.2003.1388

Hultén, E., 1894-1981, & Fries, M. 1986. Atlas of north European vascular plants north of the Tropic of Cancer. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300649971

Juslén, A., Väre, H., & Wikström, N. 2011. Relationships and evolutionary origins of polyploid *Dryopteris* (Dryopteridaceae) from Europe inferred using nuclear pgiC and plastid trnL-F sequence data. *Taxon* 60:1284-1294.

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30:3059-3066.

Kurita, S. 1965. Chromosome numbers and systematic position of the genus *Woodsia*. *Journ. Jap. Bot.* 40:358–362.

Larsson, A. 2010. abioscript. http://ormbunkar.se/phylogeny/abioscripts/.

Larsson, A. 2013. SplitTrace. http://ormbunkar.se/phylogeny/splittrace

Larsson, A. 2014. AliView: a fast and lightweight alignment viewer and editor for large data sets. *Bioinformatics* :btu531. http://dx.doi.org/10.1093/bioinformatics/btu531

Löve, Á., Löve, D., & Pichi Sermolli, R. E. G. 1977. *Cytotaxonomical atlas of the Pteridophyta*. Vaduz.

Löve, Á., Löve, D., & Sermolli, R. E. G. P. 1977. *Cytotaxonomical Atlas of the Pteridophyta*. J. Cramer.

Li, H. 1975. *Flora of Taiwan. Vol. 1, Pteridophyta and Gymnospermae*. Taipei.

Luo, Q., & Zhang, L.-B. 2012. *Woodsia guizhouensis* (Woodsiaceae), a New Species from a Limestone Area in Guizhou, China. *Novon: A Journal for Botanical Nomenclature* 22:191-195. http://dx.doi.org/10.3417/2012011

Maddison, W. P, & D.R. Maddison. 2011. *Mesquite: a modular system for evolutionary analysis*. http://mesquiteproject.org

Manton, I. 1950. *Problems of cytology and evolution in the Pteridophyta*. Cambridge: Cambridge U.P.

Manton, I. 1950. *Problems of Cytology and Evolution in the Pteridophyta*. Cambridge University

Press.
Marticorena, C., & Rodríguez, R. 1995. *Flora de Chile. Vol. 1, Pteridophyta - Gymnospermae*. Concepción: Universidad de Concepción.
Ma, Y.-L. 1985. Cytology and Taxonomy in Woodsiaceae. *Fern Gazette* 13:17-24.
Ma, Y.-L. 1987. Studies on the Chinese Woodsiaceae (3)—Phylogeny and Speciation. *Acta Phytotaxonomica Sinica* 25:299-306.
Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., & Otto, S. P. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333:1257–1257.
Mehra, P. N., & Singh, H. P. 1955. Cytology of Cyatheaceae, *Woodsia* and Marattiaceae. *Current Science* 24:425.
Mickel, J., & Windham, M. D. 2004. *Woodsia*. In *The pteridophytes of Mexico* (pp. 686-692). Bronx, N.Y.: New York Botanical Garden Press.
Mitui, K. 1965. Chromosome studies on Japanese ferns (1). *Journal of Japanese Botany* 40:117-124.
Mitui, K. 1966. Chromosome studies on Japanese ferns (2). *Journal of Japanese Botany* 41:61-63.
Nylander, J. A. A. 2009. *MrAic*. Program distributed by the author. http://www.abc.se/~nylander/mraic/mraic.html
Nylander, J. A. A., Wilgenbusch, J. C., Warren, D. L., & Swofford, D. L. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics (Oxford, England)* 24:581-583. http://dx.doi.org/10.1093/bioinformatics/btm388
Oberle, B., Montgomery, R. A., Beck, J. B., & Esselman, E. J. 2012. A morphologically intergrading population facilitates plastid introgression from diploid to tetraploid *Dodecatheon* (Primulaceae). *Botanical Journal of the Linnean Society* 168:91–100. http://dx.doi.org/10.1111/j.1095-8339.2011.01191.x
Ohwi, J., Meyer, F. G., & Walker, E. H. 1965. *Flora of Japan : a combined, much revised and extended translation by the author of his Flora of Japan (1953) and Flora of Japan - Pteridophyta (1957)*. Washington: Smithsonian Institution.
OneKP. 2013. 1KP. Retrieved January 31, 2014, from http://okekp.com
Pilotti, M., Brunetti, A., Tizzani, L., & Marani, O. 2009. *Platanus* × *acerifolia* genotypes surviving to inoculation with *Ceratocystis platani* (the agent of canker stain): first screening and molecular characterization. *Euphytica* 169:1-17. http://dx.doi.org/10.1007/s10681-009-9884-9
Popp, M., & Oxelman, B. 2004. Evolution of a RNA Polymerase Gene Family in *Silene* (Caryophyllaceae)—Incomplete Concerted Evolution and Topological Congruence Among Paralogues. *Systematic Biology* 53:914-932. http://dx.doi.org/10.1080/10635150490888840
Rambaut, A. 2012. *Figtree, version 1.4.0*. http://tree.bio.ed.ac.uk/software/figtree/
Rambaut, A., & Drummond AJ, A. J. 2007. *Tracer, version 1.4*. http://beast.bio.ed.ac.uk/Tracer
Rautenberg, A., Sloan, D. B., Aldén, V., & Oxelman, B. 2012. Phylogenetic Relationships of *Silene multinervia* and *Silene* Section *Conoimorpha* (Caryophyllaceae). *Systematic Botany* 37:226-237. http://dx.doi.org/10.1600/036364412X616792
Ronquist, F., & Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574. http://dx.doi.org/10.1093/bioinformatics/btg180

Rothfels, C. J., A. K. Johnson, P. H. Hovenkamp, D. L. Swofford, H. C. Roskam, C. R. Fraser-Jenkins, M. D. Windham, and K. M. Pryer. 2014. Natural hybridization between parental lineages that diverged from each other approximately 60 million years ago. *American Naturalist*. In review.

Rothfels, C. J., Larsson, A., Kuo, L. Y., Korall, P., Chiou, W. L., & Pryer, K. M. 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Systematic biology* 61:490–509.

Rothfels, C. J., Larsson, A., Li, F.-W., Sigel, E. M., Huiet, L., Burge, D. O., … Pryer, K. M. 2013. Transcriptome-Mining for Single-Copy Nuclear Markers in Ferns. *PLoS ONE* 8. http://dx.doi.org/10.1371/journal.pone.0076957

Rothfels, C. J., Sundue, M. A., Kuo, L.-Y., Larsson, A., Kato, M., Schuettpelz, E., & Pryer, K. M. 2012. A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales). *Taxon* 61:515-533.

Scheen, A.-C., Pfeil, B. E., Petri, A., & Heidari, N. 2012. Use of allele-specific sequencing primers is an efficient alternative to PCR subcloning of low-copy nuclear genes. *Molecular Ecology Resources* 12:128. http://dx.doi.org/10.1111/j.1755-0998.2011.03070.x

Schmakov, A. 2001. Synopsis of the ferns of Russia. *Turczaninowia* 4:36–72.

Schuettpelz, E., & Pryer, K. M. 2009. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences of the United States* 106.

Shimura, Y., & Matsumo, S. 1975. Study on the chromosomes of *Woodsia intermedia* Tagawa. *Journal of Geobotany; or the Hokuriku Journal of Botany* 23:2-4.

Shmakov, A. 2003. Review of the family Woodsiaceae (Diels) Herter of Eurasia. In *Pteridology in the new millennium.* (pp. 49–64). Dordrecht: Kluwer Academic Publishers.

Slotte, T., Huang, H., Lascoux, M., & Ceplitis, A. 2008. Polyploid speciation did not confer instant reproductive isolation in *Capsella* (Brassicaceae). *Molecular biology and evolution* 25:1472-1481. http://dx.doi.org/10.1093/molbev/msn092

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. http://dx.doi.org/10.2307/25065646

Soltis, D. E., & Rieseberg, L. H. 1986. Autopolyploidy in *Tolmiea menziesii* (Saxifragaceae): Genetic Insights from Enzyme Electrophoresis. *American Journal of Botany* 73:310-318. http://dx.doi.org/10.2307/2444186

Soltis, D. E., Soltis, P. S., Schemske, D. W., Hancock, J. F., Thompson, J. N., Husband, B. C., & Judd, W. S. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* 56:13-30.

Soltis, D. E., Soltis, P. S., & Tate, J. A. 2004. Advances in the study of polyploidy since Plant speciation. *New Phytologist* 161:173-191. http://dx.doi.org/10.1046/j.1469-8137.2003.00948.x

Soltis, P. S., & Soltis, D. E. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences* 97:7051-7057. http://dx.doi.org/10.1073/pnas.97.13.7051

Staden, R., Beal, K. F., & Bonfield, J. K. 2000. The Staden package, 1998. *Methods in molecular biology (Clifton, N.J.)* 132:115.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690. http://dx.doi.org/10.1093/bioinformatics/btl446

Sukumaran, J., & Holder, M. T. 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26:1569-1571. http://dx.doi.org/10.1093/bioinformatics/btq228

Tagawa, M. 1938. Spicilegium Pteridographiae Asiae Orientalis 16. *Acta phytotaxonomica et geobotanica* 7:184-191.

Thórsson, Æ. T., Salmela, E., & Anamthawat-Jónsson, K. 2001. Morphological, Cytogenetic, and Molecular Evidence for Introgressive Hybridization in Birch. *Journal of Heredity* 92:404-408. http://dx.doi.org/10.1093/jhered/92.5.404

The Inkscape Team. 2013. Inkscape. http://sourceforge.net/projects/inkscape/

The Plant List. 2013. The Plant List. Retrieved January 25, 2014, from http://www.theplantlist.org/

Tutin, T. G. (Ed.). 1993. Woodsia. In *Flora Europaea. Vol. 1, Psilotaceae to Platanaceae* (pp. 25-26). Cambridge: Cambrigde Univ. Press.

Vogel, J. C., Russell, S. J., Rumsey, F. J., Barrett, J. A., & Gibby, M.. 1998. Evidence for Maternal Transmission of Chloroplast DNA in the Genus *Asplenium* (Aspleniaceae, Pteridophyta). *Botanica Acta* 111:247-249. http://dx.doi.org/10.1111/j.1438-8677.1998.tb00704.x

Wagner, F. S. 1987. Evidence for the Origin of the Hybrid Cliff Fern, *Woodsia* x *Abbeae* (Aspleniaceae: Athyrioideae). *Systematic Botany* 12:116-124.

Werth, C. R., & Windham, M. D. 1991. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. *The American Naturalist* 137:515-526.

Wilgenbusch, J. C., Warren, D. L., & Swofford, D. L. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Windham, M. D. 1993. *Woodsia*. In *Flora of North America Editorial Committee (eds.)* (Vol. 2, pp. 270-280). New York: Oxford University Press.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences* :pnas.0811575106. http://dx.doi.org/10.1073/pnas.0811575106

Yoon, C.-S., Glawe, A., & Shaw, P. D. 1991. A Method for Rapid Small-Scale Preparation of Fungal DNA. *Mycologia* 83:835-838.

Zhang, G. M., Kato, M., & Shmakov, A. 2013. Woodsiaceae. In Z. Y. Wu, P. H. Raven, & D. Y. Hong (Eds.), *Flora of China* (Vol. 2-3, pp. 397-404). Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. Retrieved from http://www.efloras.org

Appendix 1. Voucher data, presented in following order: Id number in this study, Taxon, Voucher, Herbarium, Provenance, number of *pgiC* sequence copies found, GenBank accession numbers for *pgiC*, *matK*, *RPA2*, *rbcL*, *atpA*, *atpB*, *trnGR*. Missing data indicated by "-"

ID	*Taxon*	Country	State	Voucher	Herbarium	*pgiC*	*pgiC* copies	*matK*	*RPA2*	*rbcL*	*atpA*	*atpB*	*trnGR*
F1	*W. elongata*	Nepal	Maharigaon	O. Polunin, W.R. Sykes & L.H.J. Williams	UPS		-	X	X	X	X	X	X
F3	*W. ilvensis*						1			X	X	X	X
F4	*W. alpina*	Sweden		PO Karis	UPS		2	X		X	X	X	X
F6	*W. macrochlaena*						1	X		X	X	X	X
F11	*W. polystichoides*						1			X	X	X	X
F13	*Athyrium filix-femina*			A. Larsson 5	UPS		2			X	X	X	X
F14	*Thelypteris palustris*	Sweden	Uppsala	A. Larsson 7	UPS		-			X	X	X	X
F18	*W. mollis*	Sweden	Uppsala	A. Larsson 11	UPS		3	X		X	X	X	X
F19	*W. fragilis*	Mexico	Hidalgo	A. Larsson 103, with J. Beck, R. Dyer, A. Reyes Garcia, C. J. Rothfels, P. Rothfels, G. Yatskievych	UPS		3	X		X	X	X	X
F21	*W. ilvensis*			A. Larsson 6	DUKE		1	X		X	X	X	X
F22	*W. alpina*	Sweden	Uppsala	Anders Larsson 9	UPS		2	X	X	X	X	X	X
F28	*W. montevidensis*	Ecuador	Pichincha Province	E Schuettpelz 997 (with M Sundue)	DUKE, QCA, NY, QCNE		2	X		X	X	X	X
F31	*W. plummerae*	USA	Arizona	J. Metzgar 152 (with A. Grusz, N. Nagalingum, K. Pryer and M. Windham)	DUKE		2			X	X	X	X
F44	*W. plummerae*	USA	Arizona	E. Schuettpelz 1235A	DUKE		2			X	X	X	X
F45	*W. andersonii*	China	Yunnan	AGS Expedition to China (1994) 1432	E 00161648		2	X	X	X	X	X	X
F46	*W. montevidensis syn burgessiana*	South Africa	Mokhotlong District	Nüsser, M. 4384	E 00293495		1 missing probably 1	X	X	X	X	X	X
F47	*W. manchuriensis*	Japan	Tokyo	Nakaike, T.; Sakakibara, M. & Ishizuka, M.	E 00294459		-	X	X	X	X	X	X
F48	*W. cochisensis*	USA	Arizona	J. Metzgar 138 (with A. Grusz, N. Nagalingum, K. Pryer and M. Windham)	DUKE		2	X		X	X	X	X
F49	*W. andersonii*	China	Qinghai	Sino-American-British Yushu Expedition (1996) 2601	E 00067169		2 maybe missing 1	X	X	X	X	X	X
F50	*W. subcordata*	China	Nei Mongol	Boufford et al. 25932	DUKE (dupe at Harvard)		2	X	X	X	X	X	X
F53	*W. manchuriensis*	Japan	Tokyo	Nakaike, T.; Sakakibara, M. & Ishizuka, M.	E 00294459			X		X	X	X	X
F54	*W. rosthorniana*	China	Sichuan	Harry Smith 2415	UPS 06/020-63		-	X		X	X	X	X
F55	*W. hancockii syn gracilima*	China	Shanxi	Harry Smith 7072	UPS 06/020-47		-	X		X	X	X	
F56	*W. andersonii*	China	Qinghai	Sino-British Qinghai Expedition 1997. 907	E00136514		3	X		X	X	X	
F58	*W. plummerae*	USA	Nevada	Beck 1099 (with MD Windham, CJ Rothfels, A Grusz, K Pryer)	DUKE		2			X	X	X	X
F60	*W. oregana ssp oregana*	USA	Oregon.	Windham 3665	DUKE 400180		2	X		X	X	X	X
F63	*W. mexicana sometimes phillipsiilike*	Mexico	Jalisco	A. Larsson 119	UPS		2						
F65	*W. scopulina ssp scopulina*	USA	Idaho	S.F.Smith 162	DUKE		2						
F66	*W. canescens*	Mexico	Oaxaca	A. Larsson 145	UPS		2	X	X	X	X	X	X
F68	*W. neomexicana*	USA	Arizona	Windham 3719, Grusz, Wessinger	DUKE 400252		1				X		
F69	*W. mexicana sometimes phillipsiilike*	Mexico	Guanajuato	C.J.Rothfels 3083, with P.Rothfels, J.Beck, G.Yatskievych, A.Larsson, R.Dyer, A.J.Reyes Garcia	MEXU, DUKE, UPS, MO		2	X	X	X	X	X	X
F70	*W. obtusa ssp obtusa*	USA	North Carolina	C.J.Rothfels 2839	DUKE		3	X	X	X	X	X	X
F71	*W. intermedia*	Plant in cultivation at Juniper Level Botanic Garden, North Carolina		Schuettpelz 391	DUKE		1	X	X	X	X	X	X
F72	*W. cycloloba*	China	Xikang	Harry Smith 11664	UPS		-	X	X	X	X	X	X
F73	*W. cyclolobalike hybrid*	China	Xizang(Tibet)	D. E. Boufford, S. L. Kelley, R. H. Ree & S. K. Wu. 29615	HUH 298062		2	X	X	X	X	X	
F74	*W. rosthorniana*	China	Sichuan	D. E. Boufford, B. Bartholomew, W. Y. Chen, M. J. Donoghue, R. H. Ree, H. Sun & S. K. Wu. 28573	HUH 298056		2	X	X	X	X	X	X
F75	*W. andersonii*	China	Sichuan	D. E. Boufford, J. H. Chen, K. Fujikawa, S. L. Kelley, R. H. Ree, H. Sun, J. P. Yue, D. C. Zhang & Y. H. Zhang. 34183	HUH 289202		2 maybe missing 1	X	X	X	X	X	X
F76	*W. glabella*	China	Yunnan	D. E. Boufford, B. Bartholomew, W. Y. Chen, M. J. Donoghue, R. H. Ree, H. Sun & S. K. Wu. 29211	HUH 298057		1	X	X	X	X	X	X
F77	*W. ilvensis*	Russia	Karelia	Bengt Jonsell 8133	UPS V-146810		1	X		X	X	X	X
F79	*W. ilvensis*	Canada	Ontario	Inga Hedberg 85146	UPS V-215-738		-			X	X	X	X
F80	*W. cycloloba*	China	Yunnan	D. E. Boufford, B. Bartholomew, W. Y. Chen, M. J. Donoghue, R. H. Ree, H. Sun & S. K. Wu. 29210	HUH 298059		2	X		X	X	X	X
F84	*W. glabella*	Norway	Troms	A. Larsson 291	UPS		1			X	X	X	X
F86	*W. glabella*	Norway	Troms	A. Larsson 296	UPS		1						
F87	*W. ilvensis*	Norway	Finnmark	A. Larsson 319	UPS		1	X		X	X	X	X
F88	*W. glabella*	Norway	Finnmark	A. Larsson 321	UPS		1	X	X	X	X	X	X
F89	*W. alpina*	Norway	Finnmark	A. Larsson 323	UPS		2	X		X	X	X	X
F90	*W. manchuriensis*	Japan		Kuo 995	TAIF		2	X		X	X	X	X

F91	*W. indusiosa*	China	Yunnan	Kuo 1330	TAIF	2	X	X				X
F92	*W. elongata*	China	Yunnan	Kuo 1339	TAIF	1	X	X	X	X	X	X
F93	*W. andersonii*	China	Yunnan	Kuo 1495	TAIF	3 maybe more	X	X		X		X
F94	*W. rosthorniana*	China	Sichuan	Kuo 1547	TAIF	2	X	X				
F95	*W. alpina*	France	Provence-Alpes-Côte-d'Azur				X	X				
F96	*W. montevidensis brasil*	Brasil	Rio de Janeiro	Prado, Schwartsburd, Schuettpeltz, Yatskievych 2105	DUKE	3	X					
F97	*W. ilvensis*	USA	New York	Sigel 2010-63	DUKE		X					
F98	*W. obtusa*	USA	Virginia	A.L.Grusz 167	DUKE	1	X					
F99	*W. ilvensis*	Canada	Ontario	C.J.Rothfels 3941, with P.Rothfels	DUKE	-	X	X			X	X
F100	*W. ilvensis*	Canada	Ontario	C.J.Rothfels 3940, with P.Rothfels	DUKE	-	X					
F101	*W. obtusa*	USA	Virginia	C.J.Rothfels 3924, with S.Zylinski	DUKE	3	X					
F102	*W. x abbeae*	Canada	Ontario	M.J. Oldham, S. Brinker, W.D. Bakowsky & Lorraine 35556	MICH, DUKE	3	X	X				
F103	*W. alpina*	Canada	Ontario	M.J. Oldham & W.D. Bakowsky 25799	OAC, MJO, DUKE		X	X				
F104	*W. montevidensis*	Ecuador	Pichincha Province	C.J.Rothfels 3537 with M.McHenry, P.Rothfels	QCA, UPS, DUKE	1	X					
F105	*W. montevidensis*	Ecuador	Loja Province	C.J.Rothfels 3670 with P.Rothfels, E.Sigel, M.McHenry	QCA, DUKE	2	X					
F106	*W. montevidensis*	Ecuador	Imbabura Province	C.J.Rothfels 3553 with M.McHenry, P.Rothfels	QCA, UPS, DUKE	2	X					
F107	*W. montevidensis*	Ecuador	Pichincha Province	C.J.Rothfels 3680 with P.Rothfels, E.Sigel, M.McHenry	QCA, DUKE, UPS	2	X					
F108	*W. montevidensis*	Ecuador	Pichincha Province	C.J.Rothfels 3722 with P.Rothfels, E.Sigel, M.McHenry	QCA, DUKE, UPS		X					
F109	*W. oregana ssp cathcartiana*	USA	Colorado	S.F.Smith 109B	DUKE	2	X					
F111	*W. scopulina ssp scopulina*	USA	Colorado	S.F.Smith 109A	DUKE	-	X	X				
F112	*W. scopulina ssp scopulina*	USA	Colorado	S.F.Smith 124	DUKE	1	X					
F113	*W. cf tetra-oregana*	USA	Colorado	S.F.Smith 118	DUKE	-	X	X				
F114	*W. fragilis*	Georgia	Mtiuleti	Schmidt, H.H.; Merello, M.; Eristavi, M. 2890	E 00283852	2	X					
F115	*W. andersonii*	China	Qinghai	Sino-British Qinghai Expedition 1997. 1138	E 00136509	3	X	X				
F116	*W. polystichoides*	China	Henan	Boufford, D.E.; Cheng, H.X.; Wang, Z.Y.; Ying, T.S. & Zhang, X.C. 26448	E 00068125		X					
F118	*W. lanosa*	China	Yunnan	AGS Expedition to China (1994). 1212	E 00161647	3	X					
F119	*W. lanosa*	China	Yunnan	AGS Expedition to China (1994). 1369	BOL 44992	3	X	X				
F120	*W. angolensis syn montevidensis*	South Africa	Transvaal	J. E. Burrows 3067	BOL 44987	2	X	X	X	X	X	X
F122	*W. montevidensis syn burgessiana*	South Africa	Natal	O.M. Hilliard, B.L. Burtt. 14131	BOL 44984		X					
F123	*W. montevidensis syn burgessiana*	South Africa	Northern Cape	J.J. Blom	BOL 44983	-	X					
F124	*W. montevidensis syn burgessiana*	Lesotho		F K Hoener 1427	BOL 44988	-	X					
F125	*W. montevidensis syn burgessiana*	South Africa	Natal	J. E. Burrows 1422	HUH 298061	-	X					
F126	*W. andersonii*	China	Yunnan	D. E. Boufford, B. Bartholomew, W. Y. Chen, M. J. Donoghue, R. H. Ree, H. Sun & S. K. Wu. 29289	HUH 298058		X					
F127	*W. andersonii*	China	Qinghai	T. N. Ho, B. Bartholomew, M. F. Watson & M. G. Gilbert 2601	HUH 298060	3	X					
F129	*W. glabella*	USA	Alaska	Carolyn L. Parker,Bruce A. Bennett 13162	ALA 150090		X					
F130	*W. glabella*	USA	Alaska	Carolyn L. Parker 15657	ALA 138195	-	X					
F131	*W. ilvensis*	USA	Alaska	Carolyn L. Parker,Reidar Elven,Nikki Guldager 13786	ALA 115415		X					
F132	*W. ilvensis*	Russia	Chukotka	Carolyn L. Parker 4408	ALA 115399	-	X					
F133	*W. ilvensis*	Russia	Chukotka	Carolyn L. Parker 4386	ALA 151097	-	X					
F134	*W. ilvensis*	USA	Alaska	Korem Bosworth,Jesse Grunblatt KSB01-7	ALA 165876		X					
F135	*W. alpina*	USA	Alaska	Alan R. Batten 08-51	ALA 167536	-	X					
F136	*W. ilvensis*	USA	Alaska	Brad Krieckhaus 729	ALA 140640	-	X					
F137	*W. alpina*	USA	Alaska	Carolyn L. Parker,Reidar Elven,Bruce A. Bennett,Nikki Guldager 12735	ALA 156127		X					
F139	*W. mollis*	Mexico	Guanajuato	A. Larsson 113	UPS	2	X					
F140	*W. mexicana sometimes phillipsiilike*	Mexico	Jalisco	A. Larsson 120	UPS	2	X	X				
F141	*W. mollis*	Mexico	Nayarit	A. Larsson 125	UPS	3	X	X				
F142	*W. mexicana*	Mexico	Jalisco	A. Larsson 136	UPS	-	X	X			X	
F143	*W. mollis*	Mexico	Guerrero	A. Larsson 139	UPS	3	X					
F144	*W. mollis*	Mexico	Jalisco	A. Larsson 121	UPS	1	X					
F146	*W. canescens*	Mexico	Oaxaca	A. Larsson 145	UPS	2	X					
F147	*W. alpina*	Norway	Nordland	A. Larsson 279	UPS	-	X					
F148	*W. alpina*	Sweden	Västerbotten	A. Larsson 232	UPS	-	X					
F149	*W. alpina*	Sweden	Västerbotten	A. Larsson 233	UPS	-	X					
F150	*W. polystichoides*	Taiwan	Nantou County	E Schuettpelz 1066A (with M Windham and H Schneider)	DUKE 398635	1	X					
F151	*W. cf neomexicana det cf oregana*	USA	New Mexico	S.F.Smith 141	DUKE	2	X					
F152	*W. cochisensis*	USA	Arizona	E. Schuettpelz 492 (with J. Metzgar, K. Pryer, H. Schneider, and M. Windham)	DUKE 391426	2	X	X				

F153	*W. cf neomexicana det cf oregana*	USA	New Mexico	S.F.Smith 140	DUKE	-	X						
F154	*W. neomexicana*	USA	New Mexico	S.F.Smith s.n.	UPS	1	X						
F155	*W. neomexicana*	USA	Arizona	C.J.Rothfels 2574 with A.L.Grusz	DUKE	1	X	X					
F156	*W. plummerae_v2 det phillipsii*	USA	New Mexico	Worthington 35857b	DUKE	2	X						
F157	*W. mexicana phillipsiilike*	Mexico	Chihuahua	D.O.Burge 1249	DUKE	3	X	X					
F158	*W. cf phillipsii*	USA	New Mexico	Worthington 35857a	DUKE	2	X	X			X	X	
F159	*W. scopulina ssp scopulina*	USA	Idaho	S.F.Smith 163	DUKE	2	X						
F160	*W. mexicana phillipsiilike*	Mexico	Querétaro	C.J.Rothfels 3456, with P.Rothfels, J.Beck, R.Dyer	MEXU, DUKE. UPS	2	X						
F161	*W. mexicana sometimes phillipsiilike*	Mexico	Querétaro	C.J.Rothfels 3448, with P.Rothfels, J.Beck, R.Dyer	MEXU, DUKE. UPS	2	X						
F164	*W. alpina*	Austria	Tirol	Andreas Tribsch, 31.8.2001	WU 4365	-							
F165	*W. ilvensis*	Russia	Altai	Andreas Tribsh & Franz Essel 10270	WU 10270	-	X						
F166	*W. ilvensis*	Greenland	Ilulissat county	Kristine Westergaard 29b	WU 4766	-	X	X					
F167	*W. glabella*	Russia	Taimyrsky	Peter Schönswetter & Andreas Tribsch, August 5, 2004	WU 4854	-	X						
F168	*W. macrospora*	China	Gansu	J. F. Rock 13712	S	1	X						
F169	*W. indusiosa*	China	Yunnan	Maire 6487	S	2	X						
F170	*W. subcordata*	Japan	Hondo	M. Furuse	S	-	X						
F173	*W. hancockii*	China	Sichuan	Kuo 2299	TAIF	1	X						
F174	*W. cinnamomealike*	China	Sichuan	Kuo 2260	TAIF	1	X	X					
F175	*W. andersonii*	China	Sichuan	Kuo 2261	TAIF	1 maybe missing some allele	X	X					
F176	*W. okamotoi*	Taiwan		Kuo ?	TAIF	2	X						
F177	*W. phillipsii autotriploid*	USA	Arizona	Windham & Haufler 643b	MDW		X	X	X	X	X	X	
F178	*W. scopulina ssp laurentiana*	USA	Montana	Windham et al. 812	MDW	2	X						
F179	*W. plummerae*	USA	Arizona	Windham 463	MDW	2	X						
F180	*W. oregana ssp oregana*	USA	Washington	Windham & Paris 825	MDW	2	X						
F181	*W. scopulina ssp laurentiana*	USA	Colorado	Windham & Haufler 668	MDW	2	X						
F182	*W. phillipsii*	USA	Texas	Windham & Yatskievych 750	MDW	2	X						
F183	*W. oregana ssp oregana*	USA	Montana	Windham et al. 810	MDW	2	X						
F184	*W. scopulina ssp laurentiana*	USA	Oregon	Windham et al. 833	MDW	2	X						
F185	*W. plummerae*	USA	New Mexico	Windham & Haufler 602	MDW	2	X						
F186	*W. obtusa ssp obtusa*	USA	Missouri	Windham & Czech 675	MDW	-	X						
F187	*W. neomexicana det oregana ssp cathcartiana*	USA	Nevada	Windham 423	ASC?; MDW	1	X						
F188	*W. scopulina ssp scopulina*	USA	Montana	Windham et al. 815	MDW	2	X						
F189	*W. phillipsii*	USA	Arizona	Windham & Haufler 643a	MDW	1	X						
F190	*W. scopulina ssp scopulina*	USA	Washington	Windham et al. 818	MDW	2	X						
F191	*W. obtusa ssp occidentalis*	USA	Kansas	Windham & Windham 681	MDW	2	X						
F192	*W. cf tetra-oregana det ssp cathcartiana*	USA	Wyoming	Windham et al. 843	MDW	1	X						
F193	*W. obtusa ssp occidentalis*	USA	Kansas	Windham & Windham 683	MDW	2	X						
F194	*W. oregana ssp cathcartiana*	USA	Colorado	Windham & Haufler 667	MDW	3	X						
F195	*W. cochisensis*	USA	Arizona	Windham 60	ASC; MDW	1	X						
F196	*W. neomexicana*	USA	Arizona	Windham & Yatskievych 788	MDW	1	X						
F197	*W. cf tetra-oregana det ssp cathcartiana*	USA	Utah	Windham & Czech 582	MDW	1	X						
F198	*W. obtusa ssp obtusa x ssp occidentalis*	USA	Kansas	Windham & Windham 684	MDW	4	X						
F199	*W. oregana ssp cathcartiana*	USA	Minnesota	Windham & Haufler 696	MDW	1	X						
F200	*W. neomexicana*	USA	Arizona	Windham & Czech 591	MDW	1	X						
F202	*W. neomexicana*	USA	Texas	Windham & Yatskievych 758	MDW	1	X						
F203	*W. oregana ssp cathcartiana*	USA	South Dakota	Windham & Haufler 697	MDW	2	X						
F204	*W. ilvensis*	USA	Wisconsin	Windham & Ranker 880	MDW	1	X						
F205	*W. cf neomexicana det oregana ssp cathcartiana*	Canada	Ontario	Windham & Ranker 866	MDW	2	X						
F206	*W. neomexicana*	USA	Arizona	Windham & Yatskievych 792	MDW	1	X	X					
F207	*W. oregana ssp cathcartiana*	USA	Arizona	Windham & Yatskievych 793a	MDW	1	X						
F208	*W. neomexicana x oregana ssp cathcartiana*	USA	Colorado	Windham & Windham 896	COLO; UT	3	X						
F210	*W. neomexicana x oregana ssp cathcartiana*	USA	Arizona	Windham & Haufler 648	MDW	2	X						

F211	*W. oregana ssp cathcartiana*	USA	Michigan	Windham & Ranker 873	MDW	1 missing probably 1 or 2	X	
F212	*W. obtusa ssp obtusa*	USA	Iowa	Windham et al. 692	MDW	1	X	
F213	*W. neomexicana x phillipsii*	USA	Arizona	Windham & Yatskievych 789	MDW	2	X	X
F214	*W. obtusa ssp obtusa x oregana ssp cathcartiana (W. × kansana)*	USA	Wisconsin	Windham & Ranker 885	MDW	2	X	
F215	*W. obtusa ssp obtusa x ssp occidentalis*	USA	Kansas	Windham 702	MDW	-	X	
F216	*W. x abbeae*	USA	Michigan	Windham & Ranker 875	MDW	2	X	
F217	*W. cf. coloradoensis*	USA	Colorado	Smith 148	DUKE 401392	1 maybe missing 1 but very weak	X	
F218	*W. neomexicana x oregana ssp cathcartiana*	USA	South Dakota	Windham & Haufler 698	MDW	1 missing 1 or 2	X	X
F219	*W. x abbeae*	USA	Wisconsin	Windham & Ranker 886	MDW	-	X	
F221	*W. obtusa ssp occidentalis*	USA	Texas	Yatskievych & Gastony 86-343	DUKE 396264; IND?	2	X	
F222	*W. neomexicana*	USA	South Dakota	Windham & Haufler 699	MDW	2	X	X
F223	*W. ilvensis*	USA	Michigan	Windham & Ranker 874	MDW	-	X	
F224	*W. phillipsii x protoplummerae allotetraploid*	USA	New Mexico	Worthington 35554	MDW	2	X	
F227	*W. fragilis*	Georgia	Mtskheta-Mtia neti	G.M. Schneeweiww, A. Tribsch, M. Staudinger & P. Schönswetter	WU 4783	2	X	
F247	*W. mollis*	Mexico	Distrito Federal	C.J.Rothfels 2989	DUKE	1 missing 1 or maybe 2		
F248	*W. oregana ssp oregana*	U.S.A.	Colorado	S.F.Smith 515	DUKE	1	X	
F249	*W. mollis*	Mexico	Querétaro	C.J.Rothfels 3458	DUKE	-	X	X
F250	*W. oregana ssp oregana*	U.S.A.	Colorado	S.F.Smith 516	DUKE	1	X	
F251	*W. polystichoides*	Taiwan	Nantou	F.W.Li 1443	DUKE	1	X	
F252	*W. mollis*	Mexico	Querétaro	C.J.Rothfels 3457	DUKE	1 missing 1	X	
F253	*W. scopulina ssp scopulina*	U.S.A.	Idaho	C.J.Rothfels 4187	DUKE	2	X	
F254	*W. mollis*	Mexico	Oaxaca	C.J.Rothfels 3286	DUKE	-	X	
F255	*W. oregana ssp oregana*	U.S.A.	Idaho	C.J.Rothfels 4189	DUKE	-	X	
F256	*W. mollis*	Mexico	Querétaro	C.J.Rothfels 3424	DUKE	1 missing probably 1	X	
F257	*W. scopulina ssp scopulina*	U.S.A.	Idaho	C.J.Rothfels 4190	DUKE	2	X	
F258	*W. mollis*	Mexico	Municipio Ocuilan	C.J.Rothfels 3406	DUKE	2	X	
F259	*W. scopulina ssp scopulina*	U.S.A.	Colorado	C.J.Rothfels 4194	DUKE	2	X	
F260	*W. oregana*	Mexico	Oaxaca	S.F.Smith 548	DUKE		X	
F261	*W. mollis*	Mexico	Oaxaca	C.J.Rothfels 3292	UPS	1	X	X
F263	*W. x abbeae*	USA	Minnesota	D. Schimpf 164	Min 456572	2	X	X
F265	*W. scopulina ssp laurentiana*	USA	Minnesota	Lynden B. Gerdes 5758	MIN 914833	2	X	X
F266	*W. scopulina ssp laurentiana*	USA	Minnesota	Lynden B. Gerdes 5784	MIN 914830	2	X	X
F267	*W. scopulina ssp laurentiana*	United States	Minnesota	Doris Lawson Gerdes 753	MIN 917163	2	X	X
F269	*Woodsia yunnan-okamotoi*	China	Yunnan	Kuo 3325	UPS	1		
F270	*Woodsia andersonii*	China	Yunnan	Kuo 3322	UPS		X	
F271	*Woodsia alpina*	Italy	Trentino-Alto Adige	AL863	UPS	2	X	
F272	*Woodsia glabella*	Italy	Trentino-Alto Adige	AL867	UPS	1	X	
F274	*Woodsia ilvensis*	Italy	Trentino-Alto Adige	AL871B	UPS	1	X	
F275	*Woodsia ilvensis*	Austria	Steiermark	AL850:1-3	UPS	1	X	
F277	*Woodsia mollis* gametophyte	Mexico	Guanajuato	AL113	UPS	1		
F285	*Woodsia oregana ssp oregana* gametophyte	USA	Oregon	Windham3665	UPS	1		
F287	*Woodsia glabella*	Italy	Trentino-Alto Adige	AL870:1	UPS	1	X	

Fig. S1. MrBayes consensus tree of nuclear RPA2 sequences. Names ending with a,b,c or d indicate specimen with multiple sequences, x indicates specimen with multiple sequences where only one has been retrieved. Diploid taxa are not colored. Colors indicate polyploid or hybrid taxa. Numbers at nodes are Bayesian posterior probabilities (PP) followed by RAxML Maximum Likelihood bootstrap support. Branches are collapsed at 0.5 PP. Scale bar represents average number of substitutions per site.

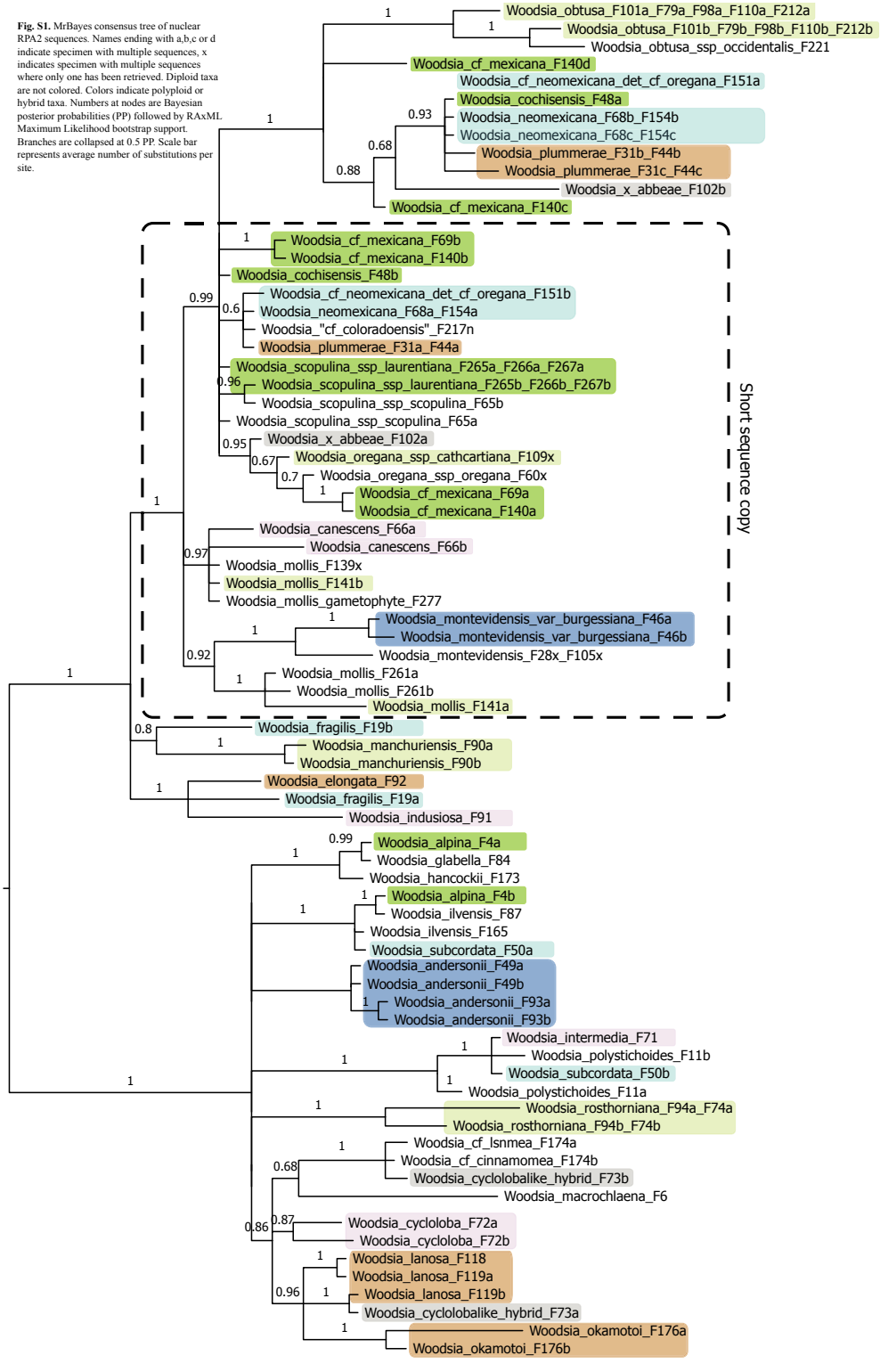

Fig. S2. Polyploid speciation events and most parsimonious reconstruction of ancestral chromosome base numbers drawn on maximum clade credibility chronogram from BEAST analysis of the extended Eupolypods II chloroplast dataset. Green color indicates diploid lineage, red = polyploid, yellow = taxa with both diploid and polyploid cytotypes, grey = taxa not included in calculation due to lack of chromosome counts. X denotes a speciation event that includes a shift in ploidy level. Reconstructed chromosome base numbers are given above branches.

Table S1A. Eupolypods II ploidy level count per genus and family including subspecific taxa.
Different polyploid ploidy levels are not counted separately.

Family	Genus	total_counted	diploid_count	polyploid_count	polyploid_ratio
Aspleniaceae	*Asplenium*	183	55	128	0,6994535519
Aspleniaceae	*Ceterach*	5	1	4	0,8
Aspleniaceae	*Hymenasplenium*	18	8	10	0,5555555556
Aspleniaceae	*Phyllitis*	2	2	0	0
Aspleniaceae	*Pleurosorus*	3	1	2	0,6666666667
Aspleniaceae Result		***211***	***67***	***144***	0,682464455
Athyriaceae	*Anisocampium*	2	1	1	0,5
Athyriaceae	*Deparia*	9	3	6	0,6666666667
Athyriaceae	*Acystopteris*	4	2	2	0,5
Athyriaceae	*Allantodia*	4	2	2	0,5
Athyriaceae	*Athyrium*	75	35	40	0,5333333333
Athyriaceae	*Cornopteris*	7	2	5	0,7142857143
Athyriaceae	*Diplazium*	80	30	50	0,625
Athyriaceae	*Dryoathyrium*	3	2	1	0,3333333333
Athyriaceae	*Lunathyrium*	8	4	4	0,5
Athyriaceae	*Pseudocystopteris*	1	1	0	0
Athyriaceae Result		***193***	***82***	***111***	0,5751295337
Blechnaceae	*Blechnum*	41	26	15	0,3658536585
Blechnaceae	*Brainea*	1	1	0	0
Blechnaceae	*Doodia*	8	4	4	0,5
Blechnaceae	*Pteridoblechnum*	1	1	0	0
Blechnaceae	*Sadleria*	1	1	0	0
Blechnaceae	*Salpichlaena*	1	1	0	0
Blechnaceae	*Woodwardia*	6	5	1	0,1666666667
Blechnaceae Result		***59***	***39***	***20***	0,3389830508
Onocleaceae	*Struthiopteris*	2	2	0	0
Onocleaceae	*Matteuccia*	3	3	0	0
Onocleaceae	*Onoclea*	3	3	0	0
Onocleaceae	*Onocleopsis*	1	1	0	0
Onocleaceae Result		***9***	***9***	***0***	0
Thelypteridaceae	*Abacopteris*	8	4	4	0,5
Thelypteridaceae	*Amauropelta*	2	2	0	0
Thelypteridaceae	*Ampelopteris*	1	1	0	0
Thelypteridaceae	*Amphineuron*	3	1	2	0,6666666667
Thelypteridaceae	*Christella*	13	7	6	0,4615384615
Thelypteridaceae	*Cyclosorus*	22	10	12	0,5454545455
Thelypteridaceae	*Glaphyropteridopsis*	1	1	0	0
Thelypteridaceae	*Goniopteris*	6	3	3	0,5
Thelypteridaceae	*Lastrea*	6	3	3	0,5
Thelypteridaceae	*Leptogramma*	4	3	1	0,25
Thelypteridaceae	*Macrothelypteris*	3	1	2	0,6666666667
Thelypteridaceae	*Meniscium*	2	1	1	0,5
Thelypteridaceae	*Metathelypteris*	2	1	1	0,5
Thelypteridaceae	*Oreopteris*	2	2	0	0
Thelypteridaceae	*Parathelypteris*	4	1	3	0,75
Thelypteridaceae	*Phegopteris*	7	2	5	0,7142857143
Thelypteridaceae	*Pneumatopteris*	2	1	1	0,5
Thelypteridaceae	*Pronephrium*	12	7	5	0,4166666667
Thelypteridaceae	*Pseudocyclosorus*	6	5	1	0,1666666667
Thelypteridaceae	*Pseudophegopteris*	9	5	4	0,4444444444
Thelypteridaceae	*Sphaerostephanos*	3	2	1	0,3333333333
Thelypteridaceae	*Stegnogramma*	4	2	2	0,5
Thelypteridaceae	*Thelypteris*	60	35	25	0,4166666667
Thelypteridaceae	*Trigonospora*	4	1	3	0,75
Thelypteridaceae Result		***186***	***101***	***85***	0,4569892473
Woodsiaceae	*Woodsia*	18	5	13	0,7222222222
Woodsiaceae Result		***18***	***5***	***13***	0,7222222222
Cystopteridaceae	*Cystopteris*	**18**	**5**	**13**	0,7222222222
Diplaziopsidaceae	*Diplaziopsis*	**1**	**1**	**0**	0
Cystopteridaceae	*Gymnocarpium*	**12**	**7**	**5**	0,4166666667
Grand Total		***707***	***316***	***391***	0,5530410184

Table S1B. Eupolypods II ploidy level count per genus and family excluding subspecific taxa.
Different polyploid ploidy levels are not counted separately.

Family	Genus	total	diploid_count	polyploid_count	polyploid_ratio
Aspleniaceae	*Asplenium*	155	47	108	0,6967741935
Aspleniaceae	*Ceterach*	3	1	2	0,6666666667
Aspleniaceae	*Hymenasplenium*	18	8	10	0,5555555556
Aspleniaceae	*Phyllitis*	1	1	0	0
Aspleniaceae	*Pleurosorus*	3	1	2	0,6666666667
Aspleniaceae Result		***180***	***58***	***122***	0,6777777778
Athyriaceae	*Anisocampium*	2	1	1	0,5
Athyriaceae	*Deparia*	8	3	5	0,625
Athyriaceae	*Acystopteris*	4	2	2	0,5
Athyriaceae	*Allantodia*	4	2	2	0,5
Athyriaceae	*Athyrium*	69	30	39	0,5652173913
Athyriaceae	*Cornopteris*	5	2	3	0,6
Athyriaceae	*Diplazium*	69	27	42	0,6086956522
Athyriaceae	*Dryoathyrium*	3	2	1	0,3333333333
Athyriaceae	*Lunathyrium*	8	4	4	0,5
Athyriaceae	*Pseudocystopteris*	1	1	0	0
Athyriaceae Result		***173***	***74***	***99***	0,5722543353
Blechnaceae	*Blechnum*	39	24	15	0,3846153846
Blechnaceae	*Brainea*	1	1	0	0
Blechnaceae	*Doodia*	8	4	4	0,5
Blechnaceae	*Pteridoblechnum*	1	1	0	0
Blechnaceae	*Sadleria*	1	1	0	0
Blechnaceae	*Salpichlaena*	1	1	0	0
Blechnaceae	*Woodwardia*	6	5	1	0,1666666667
Blechnaceae Result		***57***	***37***	***20***	0,350877193
Onocleaceae	*Matteuccia*	2	2	0	0
Onocleaceae	*Onoclea*	3	3	0	0
Onocleaceae	*Onocleopsis*	1	1	0	0
Blechnaceae	*Struthiopteris*	2	2	0	0
Onocleaceae Result		***6***	***6***	***0***	0
Thelypteridaceae	*Abacopteris*	***7***	***4***	***3***	0,4285714286
Thelypteridaceae	*Amauropelta*	2	2	0	0
Thelypteridaceae	*Ampelopteris*	1	1	0	0
Thelypteridaceae	*Amphineuron*	3	1	2	0,6666666667
Thelypteridaceae	*Christella*	13	7	6	0,4615384615
Thelypteridaceae	*Cyclosorus*	18	9	9	0,5
Thelypteridaceae	*Glaphyropteridopsis*	1	1	0	0
Thelypteridaceae	*Goniopteris*	6	3	3	0,5
Thelypteridaceae	*Lastrea*	6	3	3	0,5
Thelypteridaceae	*Leptogramma*	4	3	1	0,25
Thelypteridaceae	*Macrothelypteris*	3	1	2	0,6666666667
Thelypteridaceae	*Meniscium*	2	1	1	0,5
Thelypteridaceae	*Metathelypteris*	2	1	1	0,5
Thelypteridaceae	*Oreopteris*	2	2	0	0
Thelypteridaceae	*Parathelypteris*	4	1	3	0,75
Thelypteridaceae	*Phegopteris*	7	2	5	0,7142857143
Thelypteridaceae	*Pneumatopteris*	2	1	1	0,5
Thelypteridaceae	*Pronephrium*	10	6	4	0,4
Thelypteridaceae	*Pseudocyclosorus*	6	5	1	0,1666666667
Thelypteridaceae	*Pseudophegopteris*	8	4	4	0,5
Thelypteridaceae	*Sphaerostephanos*	3	2	1	0,3333333333
Thelypteridaceae	*Stegnogramma*	4	2	2	0,5
Thelypteridaceae	*Thelypteris*	58	34	24	0,4137931034
Thelypteridaceae	*Trigonospora*	4	1	3	0,75
Thelypteridaceae Result		***176***	***97***	***79***	0,4488636364
Woodsiaceae	*Woodsia*	18	5	13	0,7222222222
Woodsiaceae Result		***18***	***5***	***13***	0,7222222222
Cystopteridaceae	*Cystopteris*	**13**	**5**	**8**	0,6153846154
Diplaziopsidaceae	*Diplaziopsis*	**1**	**1**	**0**	0
Cystopteridaceae	*Gymnocarpium*	**11**	**6**	**5**	0,4545454545
Grand Total		***637***	***291***	***346***	0,5431711146

Table S2. Spore size and spore number per sporangium.

ID	Species	Sporemeasure(μM)	Number of spores per sporangium
F18	*Woodsia mollis*	65	32(<32)
F19	*Woodsia fragilis*	56	
F46	*Woodsia montevidensis* var. *burgessiana*	60.1	64(>32)
F60	*Woodsia oregana* subsp *oregana*	42.4	
F63	*Woodsia* cf. *mexicana*	48	
F66	*Woodsia canescens*	sterile	
F69	*Woodsia* cf. *mexicana*	sterile	
F94	*Woodsia rosthorniana*	50	64(>32)
F96	*Woodsia montevidensis brasil*	49.5	
F104	*Woodsia montevidensis*	45.5	
F106	*Woodsia montevidensis*	45,0	
F107	*Woodsia montevidensis*	51.0	
F108	*Woodsia montevidensis*	51.2	
F139	*Woodsia mollis*	46.5	64(>32)
F140	*Woodsia* cf. *mexicana*	48	
F141	*Woodsia mollis*	62	32(<32)
F142	*Woodsia mexicana*	48	
F143	*Woodsia mollis*	64	32(<32)
F144	*Woodsia mollis*	42.9	
F146	*Woodsia canescens*	42	64(>32)
F157	*Woodsia mexicana phillipsiilike*	56	
F160	*Woodsia mexicana phillipsiilike*	49	
F168	*Woodsia macrospora*	59,6	
F174	*Woodsia cf. cinnamomea*	49.4	
F175	*Woodsia andersonii*	64	
F178	*Woodsia scopulina* subsp *laurentiana*	52,4	64
F179	*Woodsia plummerae*	45,6	64
F181	*Woodsia scopulina* subsp *laurentiana*	54,0	64
F183	*Woodsia oregana* subsp *oregana*	38.9	64
F185	*Woodsia plummerae*	46,3	64
F187	*Woodsia neomexicana* det. *oregana* subsp *cathcartiana*	47,8	
F192	*Woodsia* cf. *tetra-oregana det subsp cathcartiana*	45,8	64
F217	*Woodsia* cf. *"oregana coloradoensis"*	small	
F227	*Woodsia fragilis*	56,5	
F247	*Woodsia mollis*	60	--
F249	*Woodsia mollis*	42.9	64(>32)
F250	*Woodsia oregana* subsp *oregana*	40,4	
F253	*Woodsia scopulina* subsp *scopulina*	39,9	
F256	*Woodsia mollis*	51,6	64(>32)
F257	*Woodsia scopulina* subsp *scopulina*	48,3	
F258	*Woodsia mollis*	49,0	64(>32)
F261	*Woodsia mollis*	42	
F266	*Woodsia scopulina* subsp *laurentiana*	56	
F269	*Woodsia yunnan-okamotoilike*	71	
F270	*Woodsia andersonii*	72	64(>32)

Table S3. Calculations and frequencies of polyploid and diploid speciation events and ages in *Woodsi*

Polyploidization speciation events within *Woodsia*	
(This is with 2 version of mollis but no canescens)	
Total number of speciation events	39
Total number of polyploid speciation events	20
Taxa with both diploid and polyploid cytotypes	2
Diploid taxa	14
Polyploid taxa	28
Taxa/names with diploid only cytotype	12
Taxa/names with polyploid only cytotype	26
Total number of taxa/names included in analysis	40
Frequency polyploid speciation	0,5128205128
Frequency polyploid speciation (counting double cytotype taxa)	0,5365853659
Polyploid taxa percent (including double cytotype taxa twice)	0,6666666667

Table S4. Calculations and frequencies of polyploid and diploid speciation events and ages in Eupolypods II clade.

Full dataset is on other sheet

Polyploidization speciation events within Eupolypods II clade	
Including 29 Woodsia species (from superEu2_w_woodsia_tree 2014-01-22)	
Total number of speciation events	250
Total number of polyploid speciation events	94
Taxa/names with both diploid and polyploid cytotypes	39
Diploid taxa/names	123
Polyploid taxa/names	167
Taxa/names with diploid only cytotype	84
Taxa/names with polyploid only cytotype	128
Total number of taxa/names included in analysis	251
Frequency polyploid speciation	0,376
Frequency polyploid speciation (counting double cytotype taxa twice)	0,4602076125
Polyploid taxa percent (including double cytotype taxa twice)	0,575862069