2018

Apomixis, Hybridization, And Biodiversity In Ferns: Insights From Genera Phegopteris And Polystichum

Nikisha Patel
University of Vermont

Follow this and additional works at: https://scholarworks.uvm.edu/graddis

Part of the Systems Biology Commons

Recommended Citation
Graduate College Dissertations and Theses. 833.
https://scholarworks.uvm.edu/graddis/833

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
ABSTRACT

Apomixis is an evolutionarily important phenomenon across plant lineages. The interaction of apomixis with hybridization and polyploidy can lead to complex patterns of reticulation, complicating efforts to reconstruct evolutionary history in groups where apomixis is common. Ferns, in particular, are rich in apomictic species, notably in centers of species diversity like East Asia. Eastern North America too is home to a number of apomictic species. We investigated the East Asian ferns in *Polystichum* sections *Xiphopolystichum* and *Duropolystichum* (Dryopteridaceae) in order to elucidate the evolutionary and biogeographic history of seven apomictic species in the group: *Polystichum tsus-simense*, *P. xiphophyllum*, *P. sinotsus-simense*, *P. pseudoxiphophyllum*, *P. mayebarae*, *P. rigens*, and *P. neolobatum*. In addition, we examined the evolutionary origin of an undescribed apomictic cytotype of North American genus *Phegopteris* (Thelypteridaceae). The datasets comprised phylogenetic inference based on three nuclear and three plastid markers, analysis of mixed nucleotide signals from chromatograms generated from Sanger sequencing of nuclear markers, ploidy estimates based on flow cytometry data and spore length measurements, morphometric analysis of representative specimens collected in southwest China and nearby regions, and climatic niche models. By interpreting these multiple lines of evidence synthetically, we have discerned multiple highly reticulate complexes of polyploid lineages derived largely from diploid sexual progenitors. Our findings highlight the importance of understanding the role of apomictic reproduction in the context of species diversity, an understanding central to similar future inquiry into the diversity of East Asian and North American ferns.
# TABLE OF CONTENTS

## CHAPTER 1: COMPREHENSIVE LITERATURE REVIEW .........................................................1

- INTRODUCTION TO ASEXUAL REPRODUCTION ..................................................1
- EVOLUTIONARY SIGNIFICANCE OF APOMIXIS IN FERNS ........................................2
  - POLYPLOIDY, HYBRIDIZATION, AND APOMIXIS ...............................................3
  - APOMIXIS, DISPERAL, AND COLONIZATION ...................................................5
  - THE EVOLUTIONARY TRAJECTORY OF APOMICTIC LINEAGES ................................6
- PATTERNS OF FERN APOMIXIS ..............................................................................7
  - PHYLOGENETIC DISTRIBUTION OF APOMICTIC FERNS ......................................7
- BIOGEOGRAPHY OF APOMICTIC FERNS .............................................................7
  - APOMIXIS IN POLYSTICHUM AND PHEGOPTERIS .............................................9
  - PHEGOPTERIS ......................................................................................................9
  - POLYSTICHUM ....................................................................................................10
- LITERATURE CITED ...............................................................................................14

## CHAPTER 2: BIOGEOGRAPHIC DISJUNCTION AND EVOLUTION IN PHEGOPTERIS ..........22

- ABSTRACT .............................................................................................................22
- INTRODUCTION .....................................................................................................22
- MATERIALS AND METHODS ..................................................................................25
  - TAXONOMIC SAMPLING ...................................................................................25
  - SPORE MEASUREMENTS ....................................................................................28
  - DNA EXTRACTION, AMPLIFICATION, AND SEQUENCING .......................................28
  - SEQUENCING .....................................................................................................30
  - MIXED NUCLEOTIDE SIGNALS AND RETICULATION ........................................34
  - TIME CALIBRATION ..............................................................................................34
- RESULTS ................................................................................................................35
  - SPORE MEASUREMENTS AND PLOIDY ............................................................35
  - PLASTID BASED PHYLOGENETIC RELATIONSHIPS ............................................38
  - SINGLE MOLECULE NUCLEAR PHYLOGENETIC RELATIONSHIPS ......................40
  - SUMMED NUCLEOTIDE SIGNALS AND RETICULATION .......................................41
  - DIVERGENCE TIME ESTIMATES .........................................................................41
- DISCUSSION ...........................................................................................................46
  - ORIGIN OF THE UNDESCRIBED TETRAPLOID ..................................................46
  - DIVERGENCE TIME ESTIMATES AND BIOGEOGRAPHIC DISJUNCTION ..................47
  - MULTIPLE APPROACHES TO PARING ALLELIC INHERITANCE .............................49
- LITERATURE CITED ...............................................................................................51

## CHAPTER 3: BIODIVERSITY AND APOMIXIS: INSIGHTS FROM EAST-ASIAN HOLLY FERNS WITHIN POLYSTICHUM SECTION XIPHOPOLYSTICHUM ..................................................56

- ABSTRACT .............................................................................................................56
- INTRODUCTION .....................................................................................................57
- METHODS ..............................................................................................................64
  - TAXON SAMPLING AND SPECIES DELIMITATION .............................................64
  - FLOW CYTOMETRY ...............................................................................................65
COMPREHENSIVE LITERATURE CITED .................................................................................. 144
APPENDICES ................................................................................................................... 159
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Distribution of <em>Phegopteris</em></td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Plastid Phylogeny of <em>Phegopteris</em></td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>ITS phylogeny of <em>Phegopteris</em> from vector cloning</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td><em>gapCp</em> phylogeny of <em>Phegopteris</em></td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>Time calibrated <em>Phegopteris</em> phylogeny</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Hypothesized reticulation in <em>Phegopteris</em></td>
<td>48</td>
</tr>
<tr>
<td>7</td>
<td>Mean spore length and genome size in <em>Xiphopolystichum</em></td>
<td>73</td>
</tr>
<tr>
<td>8</td>
<td>Plastid Phylogeny of <em>Xiphopolystichum</em></td>
<td>77</td>
</tr>
<tr>
<td>9</td>
<td>Nuclear phylogeny of <em>Xiphopolystichum</em></td>
<td>78</td>
</tr>
<tr>
<td>10</td>
<td>Pinnae illustration in <em>Xiphopolystichum</em></td>
<td>79</td>
</tr>
<tr>
<td>11</td>
<td>Morphological characters in <em>Xiphopolystichum</em></td>
<td>80</td>
</tr>
<tr>
<td>12</td>
<td>Proposed reticulate relationships in <em>Xiphopolystichum</em></td>
<td>83</td>
</tr>
<tr>
<td>13</td>
<td>Plastid phylogeny representing <em>Duropolystichum</em></td>
<td>113</td>
</tr>
<tr>
<td>14</td>
<td>Morphological groups in <em>Duropolystichum</em></td>
<td>117</td>
</tr>
<tr>
<td>15</td>
<td>Biogeographic distribution of <em>Duropolystichum</em></td>
<td>119</td>
</tr>
<tr>
<td>16</td>
<td>Time calibrated phylogeny of <em>Duropolystichum</em></td>
<td>122</td>
</tr>
<tr>
<td>17</td>
<td>Proposed reticulate relationships in <em>Duropolystichum</em></td>
<td>127</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. List of primers used to amplify *Phegopteris* .................................................................32
Table 2. Spore measurements in *Phegopteris* ........................................................................37
Table 3. Mixed chromatographic signals in *Phegopteris* .......................................................42
Table 4. List of primers used to amplify *Xiphopolystichum* ...................................................68
Table 5. Mixed chromatographic signals in *Xiphopolystichum* ............................................76
Table 6. List of primers used to amplify *Duropolystichum* ..................................................105
Table 7. Spore measurements in *Duropolystichum* ...............................................................112
Table 8. Mixed chromatographic signals in *Duropolystichum* .............................................116
Table 9. Schoener’s D values across *Duropolystichum* .........................................................120
Table 10. Worldclim variables of high permutation importance in *Duropolystichum* ..........131
CHAPTER 1: COMPREHENSIVE LITERATURE REVIEW

INTRODUCTION TO ASEXUAL REPRODUCTION

Asexual reproduction is a phenomenon known from virtually all taxonomic groups, which brings with it numerous evolutionary benefits as well as drawbacks (Butlin 2002). Asexual reproduction offers the benefits of allowing for preservation of beneficial genes and gene combinations across generations (Barton and Charlesworth 1998; Otto 2009; Butlin 2002), as well as the ability to disperse and colonize new habitats without the constraint of obligate fertilization by another individual (Baker 1955; Frey and Kurschner 2011). In contrast, asexually reproducing lineages generally do not carry the benefit of genetic diversity generated by recombination afforded to sexually reproducing lineages, which are often hypothesized to be able to adapt to changing conditions more quickly and be less susceptible to the proliferation of deleterious mutations in a population (Jaenike 1978; Otto and Nuismer 2004).

The mechanisms of asexual reproduction among various taxonomic groups are highly variable, and each form of asexuality may confer specific benefits and drawbacks to a species that may vary depending on changing selection pressures (Hamilton et al. 1990; Judson 1995). Asexuality is common among plants, and may be particularly strategic in colonization for a group of organisms that are sessile (Pannell and Barrett 1998). Mechanisms of asexual reproduction among plants can be roughly divided into two forms: vegetative reproduction and apomictic reproductive (Asker and Jerling 1992). While in vegetative reproduction, independent organisms are generated from the somatic tissues of a progenitor plant, in apomictic reproduction, progeny develop spontaneously from sex
cells, free of fertilization (Asker and Jerling 1992). Though each form of reproduction has evolved numerous times across the history of plant evolution, the origin and trajectory of apomictically reproducing lineages is often of greater interest to evolutionary biology in that lineages that reproduce in this manner, most often do so obligately. Additionally, apomictically reproducing lineages typically occur via evolutionary transition from sexuality, rather than as an augmentation to sexual reproduction as with vegetative reproduction (Whitton et al. 2008; Silvertown 2008).

**EVOLUTIONARY SIGNIFICANCE OF APOMIXIS IN FERNS**

Apomictic reproduction occurs in roughly 1% of all angiosperm species and is relatively well understood given the importance of apomixis in preserving agriculturally beneficial traits (Asker and Jerling 1992; Spillane et al. 2001). In ferns, apomixis is less well understood although approximately 10% of pteridophytes are apomictic (Wagner 1984; Gastony and Windham 1989). Apomixis among ferns is characterized by differences in sporogenesis and life cycle as compared to sexually reproducing ferns. In sexual fern life cycles, sporophytic plants produce spores through four mitotic divisions of an archesporial cell, followed by one meiotic division, resulting in 64 haploid spores. The resulting haploid spores disperse and develop into independent haploid gametophytes, which develop archegonia and antheridia (male and female reproductive structures, respectively). Fertilization may then occur intragametophytically or intergametophytically. Fertilization triggers the development of a new diploid sporophyte (Bell 1959; Klekowski 1973). The apomictic life cycle diverges from the sexual life cycle in terms of the nature of sporogenesis, particularly in terms of the number of successful mitotic and meiotic cell divisions leading to mature spores. There are three schemes of apomixis representing three
possible variants on this series of cell divisions (Manton 1950; Evans 1964; Braithwaite 1964). The result of all three aberrations from sexual sporogenesis, is 32 unreduced spores, characteristic of nearly all apomicts.

Apomixis in ferns is critical to speciation and evolution in two major ways. (1) First, apomixis interacts with the phenomena polyploidy and hybridization in several ways with complex implications for diversification (Otto and Whitton, 2000; Soltis et al., 2004; Grusz, 2016). (2) Second, as with apomixis in angiosperms and some animals, fern apomixis has the potential to facilitate colonization (de Groot 2012). Though apomictic plant lineages are sometimes thought to be “short-lived” and “evolutionary dead-ends” (van Dijk et al. 2003; Beck et al. 2011), their central importance to reticulate complexes and colonization makes them critical to the long-term success and diversification of the larger taxonomic groups to which they belong, nonetheless.

POLYPLOIDY, HYBRIDIZATION, AND APOMIXIS

Polyploidy is estimated to occur in as much as 95% of all ferns species (Grant 1981). They may arise either through autoployploidy (chromosomal doubling) or allopolyploidy, which typically follows a hybridization event. Many of these polyploid species can reproduce sexually, but allopolyploids with an “unbalanced” genome, that is a genome with proportionally greater contribution from one progenitor than the other, may face difficulty in syngamy as a result of a lack of homology among chromosomal pairings during meiosis (Manton 1950). Apomictic reproduction is one means of overcoming this barrier to reproduction among allopolyploids (Cosendai et al. 2011). Although most apomicts are unbalanced triploid lineages (Wagner and Wagner, 1980; Asker and Jerling, 1992), some autotetraploid fern lineages are apomictic, such as Pellaea glabella Mett. ex
Kuhn and *P. occidentalis* Rydb. (Gastony, 1988), and some diploid apomicts are known, including *Dryopteris wallichiana* (Spreng.) Hyl., *Cheilanthes leucopoda* Link, and *Pteris cretica* L. (Verma and Khullar, 1965; Knobloch, 1967; Fraser-Jenkins, 2007).

Many recent investigations have taken a systematics approach to parsing the evolutionary history of groups in which apomixis, polyploidy, and hybridization are prominent. Grusz (2009) demonstrated that *Myriopteris wootonii* (Maxon) Grusz and Windham and *M. yavapensis* (T. Reeves ex Windham) Grusz and Windham, not only share a pair of diploid progenitors, but are also likely both unbalanced allopolyploids. Dyer et al. (2012) finds a similar pattern in the *Asplenium monanthes* L. complex, where four unbalanced allopolyploids all share a sexual diploid progenitor. In the *Dryopteris varia* L. complex, Hori et al. (2014), find evidence for a unique hybridization event leading to each of several closely related sexual diploid lineages. Similar findings on allopolyploid apomicts in the genera *Pellaea*, (Gastony 1988), *Cornopteris* (Park and Kato 2003), and *Phegopteris* (Driscoll and Barrington 2003) have contributed to a growing body of research on the involvement of apomixis in the reticulate evolution of ferns. In addition to hybridization, reticulate complexes are further complicated by multiple hybrid origins resulting from repeated hybridization events as well as the presence of multiple cytotypes, evident in the *Asplenium monanthes* and *Dryopteris varia* complexes. The clear importance of apomixis to diversification in many species-rich and widespread fern groups demonstrates its important as an evolutionary phenomenon in need of further study in order to understand extant fern diversity.
Apomictic ferns as well as intragametophytically selfing ferns are often better colonizers than their sexually reproducing, outcrossing relatives; they readily proliferate into new habitats previously unoccupied by related species (Wubs et al. 2010; de Groot et al. 2012). Apomictic and selfing ferns may be better colonizers largely because they are not constrained by the need for the dispersal of at least two individuals for sexual reproduction (Baker 1955; Longhurst 1955; Randle et al. 2009). In addition, in both apomictic and self-compatible ferns, the benefits of genetic diversity are maintained through meiosis and therefore recombination during sporogenesis (Klekowski 1973; Lloyd and Davis 1993; Schneller et al. 1998).

Though both apomictic and self-compatible ferns likely exceed their sexual relatives in colonization ability for the reasons described, apomictic ferns possess some characteristics that have the potential to afford them advantages over self-compatible ferns. Given that nearly all apomictic ferns are polyploids (Wagner and Wagner 1980; Liu et al. 2012), they dually benefit from the ability to disperse and establish potentially with a single spore, as well as the lower risk of inbreeding depression concomitant with polyploidy, by virtue of possessing multiple genomes (Masuyama 1979; Soltis and Soltis 2000). Additionally, polyploidy may confer certain physiological benefits, including increased cold and drought tolerance (te Beest et al. 2011). In addition, where gametophytes of self-compatible ferns are able to fertilize by virtue of swimming sperm traveling from antheridia to archegonia, apomictic ferns obviate the necessity of moisture altogether by not requiring fertilization at all (Haufler et al. 2016). This is of obvious benefit to apomictic ferns in dispersal, and recent investigations have suggested that apomictic ferns tend to occur in
drier, colder regions than their sexually reproducing relatives (Liu et al. 2012; Tanaka et al. 2014). Taken together, these features of apomictic ferns are likely key to their persistence and proliferation.

THE EVOLUTIONARY TRAJECTORY OF APOMICTIC LINEAGES

A lack of relative genetic diversity within apomictic lineages has long been theorized to limit the evolutionary potential of these species. Lynch (1993) theorizes that asexual plant lineages would inevitably face a mutational meltdown as a result of accumulation of deleterious mutations. Accordingly, extant asexual lineages are most often relegated to the “tips” of phylogenetic trees (Schwander and Crespi 2009; Beck et al. 2011). However, although a given apomictic lineage may only occupy a short branch in a phylogenetic tree, it may still have enormous evolutionary implications by acting as a catalyst to speciation. The tendency of apomictic lineages to be polyploid products of hybridization, as well as to participate in hybridization as paternal progenitors, means that while apomictic lineages themselves may be short-lived, they play an important role in preserving and proliferating the genomes of related species. In addition, their ability to colonize new habitats potentially uninhabitable to their sexual relatives, makes them important not only to the phylogenetic history of the groups to which they belong, but also to their biogeography and ecology. Hence, while apomictic species themselves may lead short evolutionary lives, they can have enormous implications for the evolutionary trajectory of the larger complexes to which they belong.
PATTERNS OF FERN APOMIXIS

PHYLOGENETIC DISTRIBUTION OF APOMICTIC FERNS
The distribution of most apomictic lineages in only a few families is often attributed to species richness. The Pteridaceae and Dryopteridaceae, for instance, are the two most species-rich families of ferns, and the simple availability of many species with the ability to hybridize may increase the likelihood of occurrence of apomictic lineages (Liu et al. 2012; Grusz 2016). Indeed, within each of these families, there is a strong correlation between the occurrence of apomixis and polyploidy, particularly odd-numbered ploidies (triploid and pentaploid) suggesting that these lineages have allopolyploid origins (Liu et al. 2012; Chao et al. 2012). Among the Pteridaceae and the Dryopteridaceae, apomictic lineages are largely concentrated in a few recently divergent clades (<8 mya) (Liu et al. 2012; Guo and Liu 2013).

BIOGEOGRAPHY OF APOMICTIC FERNS
The majority of these apomict-rich clades are native or endemic to East Asia, and many have theorized a relationship between uplift of the Qinghai Tibetan Plateau (QTP), the diversification of these groups, and the evolution of apomixis (Liu et al. 2006; Liu et al. 2012; Huang et al. 2012; Guo and Liu 2013; Tanaka et al. 2014). The rise of the Himalaya has altered the climate of surrounding regions, partially by impacting rainfall regimes (Zheng et al. 2000; Yao et al. 2011), forming arid desert regions in central Asia and wet tropical zones in southern Asia (Zhisheng et al. 2001). Some have cited the ability of apomicts to reproduce without water, which is required for fertilization in sexual species, as a major advantage in proliferating across region dominated by a monsoon climate regime and therefore prolonged annual periods of drought (Tanaka et al. 2014; Haufler et al. 2016).
While geologically dynamic regions like the Himalaya facilitate speciation and diversity through their spatio-temporal heterogeneity, they also foster speciation through secondary contact, allowing for hybridization (Zhou et al. 2017), a phenomenon closely linked to the evolution of apomixis. The Himalaya and neighboring Hengduan Mountains, have been subject to significant climatic fluctuations during the Quaternary, including alternating glacial and interglacial periods (Owen 2008). Many plant species, as a result, have had constantly expanding and contracting distributions allowing for secondary contact among related species that previously experienced long intervening periods of geographic and ecological-niche isolation (Hewitt 2000). Such hybridization events may result in hybrid swarms and therefore continuous morphology, making species delineation challenging (Stebbins 1959; Gao et al. 2015). Many apomictic fern species with origins in the Himalaya are in fact part of hybrid complexes that include numerous cytotypes and morphotypes, confounding attempts at taxonomic categorization. The *Pteris cretica* complex in the Western Himalaya, for instance, includes diploid and triploid apomictic cytotypes that are morphologically indistinguishable at the macro level (Verma and Khullar 1965), and more recent studies of the group in Asia suggest that multiple apomictic lineages have arisen from a series of recent hybridizations with closely related taxa, yielding nearly continuous morphologies in the complex (Jaruwattanaphan 2013). Similarly, the *Lepisorus clathratus* complex, distributed across the QTP and in the Hengduan Mountains, includes multiple apomictic cytotypes, with evidence of hybrid origins, and continuous morphological variation among haplotypes in the group (Wang et al. 2012).
APOMIXIS IN POLYSTICHUM AND PHEGOPTERIS

PHEGOPTERIS
The fern genus Phegopteris (Thelypteridaceae) exhibits a distribution disjunct between northeastern North America and East Asia, and includes multiple apomictic, polyploid lineages. Hence, illuminating the evolutionary and biogeographic history of the group presents a significant challenge and opportunity for elucidating the relationship between a history vicariance or dispersal and the evolution of apomixis in ferns. The three species circumscribed in the genus Phegopteris are P. connectilis (Michaux) Watt, P. decursivipinnata (van Hall) Fee, and P. hexagonoptera (Michaux) Fée. Phegopteris connectilis, the most widespread of the three, has a circumboreal distribution, reaching its southernmost latitude approximately in the North American mid-Atlantic. This lineage is a known triploid apomict, though a diploid race is known from high elevations in Japan (Matsumoto 1982). Phegopteris hexagonoptera, a sexual diploid, overlaps in distribution with P. connectilis, with a distribution that spans Eastern North America. Phegopteris decursivipinnata is distributed across Japan and Eastern China. This species includes a diploid, triploid, and tetraploid cytotype. In addition to these described species, a fourth, undescribed lineage occurs in Northern Vermont and Southern Quebec (Mulligan et al. 1972). Though morphologically very similar to P. connectilis, this undescribed lineage is distinct in terms of frond shape. Two studies have examined this lineage, and though both suggest it is likely an allopolyploid hybrid, findings on its progenitors are conflicting. Mulligan et al. (1972) found that this lineage is a tetraploid and an apomict, and they further suggest, based on morphological analysis, that it is a hybrid between P. hexagonoptera and P. connectilis. However, Driscoll et al. (2003) find, using isozyme as well as quantitative
morphological analysis, that although *P. connectilis* is a likely progenitor to the 
undescribed lineage, *P. hexagonoptera* is not.

Here we utilize molecular phylogenetic techniques including, molecular cloning of 
a nuclear marker, as well as next generation sequencing, with the objective of isolating 
single molecule reads to illuminate hybrid-parent relationships among potential 
allopolyploids. We hypothesize that the undescribed tetraploid apomict in northeastern 
North America is a product of hybridization between triploid apomict *Phegopteris 
connectilis* and the diploid Japanese cytoype of *P. connectilis*. Additionally, we utilize time 
calibration of a plastid phylogeny of the group in order to elucidate the potential 
relationship between the evolutionary history of the group with vicariance resulting from 
the separation of North America and Asia.

**POLYSTICHUM**

The monophyletic *Polystichum* Roth sect. *Xiphopolystichum* Daigobo of 
*Polystichum* (Dryopteridaceae) (Daigobo et al., 1972; Zhang 1996; Lu et al., 2007; Li et 
al., 2008), is comprised of both section *Xiphopolystichum* sensu stricto (s.s.) (Kung et al., 
2001) and section *Duropolystichum* Fraser-Jenkins (Zhang and Barrington, 2013). 
Virtually all *Xiphopolystichum sensu lato* (s.l.) species are native or endemic to East and 
Southeast Asia, and most are abundant in the Himalaya and Hengduan mountains (Zhang 
and Barrington, 2013). Several species within section *Xiphopolystichum* s.l. have been 
identified previously as triploid apomicts.

**XIPHOPOLYSTICHUM**

Apomicts identified within *Xiphopolystichum s.s.* are *Polystichum tsus-simense* 
(Hook.) J. Sm, *P. xiphophyllum* (Baker) Diels, *P. mayebarae* Tagawa, and *P. sinotsus-
*simense* Ching & Z. Y. Liu (Daigobo 1973; Gibby 1985). They exhibit the morphologically complexity common to apomictic complexes. They have an array of leaf dissections between once and twice pinnate, most typically once pinnate-pinnatifid. Other species in the group may be apomictic on the basis of morphological similarity to known apomicts. *Polystichum pseudoxiphophyllum* Ching ex H. S. Kung is highly morphologically similar to *P. xiphophyllum*, though cytological and reproductive data are lacking for this taxon. Given that hybrid ferns often exhibit a morphology intermediate between their parents (e.g., Barrington, 1986), we hypothesize that in *Xiphopolystichum* s.s., apomictic lineages with a once pinnate-pinnatifid laminar dissection are likely to have originated as hybrids of a once-pinnate progenitor and a twice-pinnate progenitor (Wagner, 1983). *Polystichum herbaceum* Ching & Z. Y. Liu and *P. revolutum* P.S. Wang are proposed to be sexually reproducing diploids (Gibby 1985) belonging to *P*. sect. *Xiphopolystichum* s.s. *Polystichum revolutum* is once pinnate and *P. herbaceum* is fully twice pinnate. Hence the pair is a reasonable set of progenitors.

We establish phylogenetic relationships among species currently circumscribed in *Xiphopolystichum* s.s. to elucidate the evolutionary history of apomictic lineages in the group and to delimit species using a combination of criteria appropriate for apomicts. In *Xiphopolystichum* s.s., we contend that delineating species is best based on the phylogenetic, phenetic, and evolutionary species criteria considered together. Our dataset lends itself to evaluating each of these criteria both independently and synthetically toward defining species.

We explore the contribution of the once-pinnate *Polystichum revolutum* to the ancestry of apomictic lineages in section *Xiphopolystichum* s.s. through a study of nuclear markers.
in the light of ploidy levels, breeding systems, and morphology. Specifically, we test the hypothesis that *P. revolutum* is the once-pinnate progenitor that contributed the less-dissected morphology to the apomicts. Further, we address three major questions regarding the genomic composition of the apomicts in relation to their morphology. Are the genomes balanced? Do the identities and proportions of the contributed genomes relate to the variation in lamina dissection that is prominent in the complex? Do these genetic profiles yield insights into the delimitation of taxonomic species in the section?

**DUROPOLYSTICHUM**

Like *Xiphopolystichum* s.s., *Duropolystichum* includes known apomicts: *Polystichum neolobatum* and *P. rigens*. Understanding *Duropolystichum* is confounded by morphological complexity and a lack of phylogenetic resolution. While *Xiphopolystichum* s.s. is monophyletic, *Duropolystichum* has so far only been defined, in a molecular phylogenetic context, as paraphyletic (Le Pechon et al. 2017; Patel et al. 2017). *Duropolystichum* has some synapomorphic morphological characters, including thick, leathery lamina, spinules on the pinna margins, and large, brown ovate or lanceolate scales spread across the rachis (Kung et al. 2001; Zhang and Barrington 2013; Le Pechon 2016). However, nearly continuous morphological variation within *Duropolystichum* has led to taxonomic confusion. Species in *Duropolystichum* are routinely misidentified, and synonymy of certain taxa is disputed among experts among various treatments of the group (Fraser Jenkins 1985; 1991; 1997; Zhang and Barrington 2013). In particular, the known apomict *P. neolobatum* is highly morphologically variable (Fraser-Jenkins 1997).

*Duropolystichum* is most abundant and species-rich in the Himalaya and Hengduan mountaints, and study of *Duropolystichum* using a molecular phylogenetic and
morphometric approach offers an opportunity to understand apomixis, hybridization, and polyploidy in relationship to speciation in Qinghai Tibetan Plateau (QTP) and surrounding montane regions. Here, we analyze both plastid and nuclear sequence datasets to resolve relationships among species currently circumscribed in the most recent treatment of Duropolystichum (Zhang and Barrington 2013), and to test for reticulate evolution. The molecular analysis is grounded in a morphometric analysis to define morphological boundaries across the named species in the section. Molecular dating and niche modeling is utilized to offer historical geologic and current climatic context. Based on strong evidence from our previous work, documenting reticulation in Xiphopolystichum s.s., we test four non-exclusive hypotheses: (1) highly morphologically similar species, such as P. stimulans, P. cyclolobum, and P. rhomboideum comprise single species, (2) the apomictic species P. rigens and P. neolobatum are allopolyploid apomicts, (3) P. neolobatum has multiple origins, and (4) diversification of Duropolystichum coincides with the most recent Himalayan/Hengduan uplift events concomitant with intensification of the East Asian monsoon approximately 5 million years ago (mya).
LITERATURE CITED


BARRINGTON, D.S. 1986. The morphology and cytology of *Polystichum x potteri* hybr. nov. (= *P. acrostichoides* x *P. braunii*). *Rhodora* 297–313.


GAO, Y.-D., A.J. HARRIS, and X.-J. He. 2015. Morphological and ecological divergence of Lilium and Nomocharis within the Hengduan Mountains and Qinghai-Tibetan Plateau may result from habitat specialization and hybridization. BMC evolutionary biology 15: 147.


CHAPTER 2: BIOGEOGRAPHIC DISJUNCTION AND EVOLUTION IN

PHEGOPTERIS

ABSTRACT

Phegopteris C. Presl., like many other Eastern North American genera, exhibits a distribution disjunct between North America and East Asia. The evolutionary and biogeographic history of the group is further complicated by the presence of apomictic polyploid lineages, suggesting a reticulate history, which has yet to be investigated using a molecular phylogenetic approach. Here, we utilize phylogenetic analysis of plastid markers as well as single molecule reads generated using vector cloning and next-generation sequencing, in order to parse the evolutionary origin of an undescribed apomictic lineage native to Vermont. We integrate inference from time calibration of a plastid phylogeny in order to elucidate the potential role of the separation of North America and Asia in the origins of the undescribed tetraploid. Multiple nuclear and one plastid dataset reveal the same pattern and suggest that the undescribed tetraploid apomict resulted from hybridization between an diploid sexual Japanese lineage, and a potentially extinct North American triploid apomict.

INTRODUCTION

Several plant genera and species exhibit distributions disjunct between Asia and North America, and some studies have demonstrated that Eastern North America is more floristically similar to Japan than to Western North America (Koyama and Koyano, 1964; Krutzsch, 1989). The floristic similarity between these regions is often attributed to
relictual distribution of species that were widespread before physical separation of the two continents (Guo et al., 1998). However, some disjunct species or taxonomic groups have undergone substantial diversification within Eastern North America and East Asia after continental separation (Lee et al., 1996; Wen, 1999). Though much work on this biogeographic pattern has been focused on angiosperm groups (Sing-Chi et al., 1983; Eyde, 1963), similar patterns of disjunction are evident in some fern groups. Studies in the fern genera _Adiantum, Deparia, Osmunda_, and _Onoclea_ have demonstrated disjunct distributions within species as well as among closely related species, though time calibration and ancestral area reconstruction analyses suggest a combination of dispersal and a variety of complex geologic, climatic, and evolutionary factors shaping distribution in each group, not simply limited to vicariance (Kato, 1993; Iwatsuki, 1994; Wolf et al., 2001).

Divergence and speciation among these disjunct groups and among ferns in general, often involves hybridization, polyploidy, and apomixis. Investigation of the biogeography of _Deparia_ across Northeastern North America and East Asia reveals ploidy level changes in several taxa hypothesized to have subsequently dispersed into Eastern North America (Kuo et al., 2016). Among ferns in general, contact between sexually reproducing lineages before dispersal may lead to the formation of allopolyploid lineages, which subsequently disperse and establish. Indeed, polyploidy is hypothesized to benefit the long distance dispersal and colonization capacity of pteridophytes (Dassler and Farrar, 2001; Kuo et al., 2016). Allopolyploid ferns may be capable of ordinary sexual reproduction as a means of proliferation after dispersal. However, both auto- and allopolyploids may have an unbalanced set of genomes, meaning they may have a genome from at least one parent that
is represented with an odd number of copies. Apomixis is one means of reproducing in spite of the meiotic complications and therefore sterility imposed by this condition (Consendai et al., 2011). Indeed, there is a strong association between polyploidy, both auto- and allopolyploid, and apomixis. Nearly three quarters of apomictic ferns are triploid (Wagner and Wagner, 1980; Asker and Jerling, 1992). For ferns, in contrast to angiosperms, apomixis is defined as a form of asexual reproduction in which spores are still produced via meiosis, thereby maintaining the benefits of dispersal, though fertilization is bypassed.

The fern genus *Phegopteris* C. Presl (Thelypteridaceae) exhibits a distribution disjunct between northeastern North America and East Asia; it includes two apomictic, polyploid lineages. Accordingly, illuminating the evolutionary and biogeographic history of the group presents a significant challenge. The three species circumscribed in the genus *Phegopteris* are *P. connectilis* (Michaux) Watt, *P. decursivipinnata* (van Hall) Fee, and *P. hexagonoptera* (Michaux) Fée. *Phegopteris connectilis*, the most widespread of the three, has a circumboreal distribution, reaching its southernmost latitude in the North American mid-Atlantic region (Figure 1). This lineage is a well-known triploid apomict, though a diploid race is known only from high elevations in Japan (Matsumoto, 1982). The eastern North American *P. hexagonoptera*, a sexual diploid, overlaps in distribution with *P. connectilis* significantly. *Phegopteris decursivipinnata* is found exclusively in Japan and far Eastern China (Figure 1a), and includes a diploid, triploid, and tetraploid cytotypes. In addition to these established and morphologically distinct species, a fourth, undescribed lineage occurs in Northern Vermont and Southern Quebec (Mulligan et al., 1972, Driscoll et al., 2003) (Figure 1b). Though morphologically very similar to *P. connectilis*, this
undescribed lineage is distinct in terms of frond shape. Whereas *P. connectilis* has a tear-drop shaped frond with an acuminate tip, the undescribed tetraploid apomictic has a triangular frond shape. Two studies have examined this lineage, and though both suggest it is likely an allopolyploid hybrid, conclusions on its origins are conflicting. Mulligan et al. (1972) determined that this lineage is tetraploid and apomictic, and they further suggest, based on a morphological analysis, that it is a hybrid between *P. hexagonoptera* and *P. connectilis*. However, Driscoll et al. (2003) found, using isozyme electrophoresis as well as quantitative morphological analysis, that although *P. connectilis* is a likely progenitor to the undescribed lineage, *P. hexagonoptera* is not. Here we utilize molecular phylogenetic techniques including, molecular cloning of a nuclear marker as well as next-generation amplicon sequencing to illuminate hybrid-parent relationships for the allopolyploids. We hypothesize that the undescribed tetraploid apomict in northeastern North America is the product of hybridization between triploid apomict *P. connectilis* and the diploid Japanese cytotype of *P. connectilis*. Additionally, we utilize time calibration of a plastid phylogeny of the group to explore the role of vicariance resulting from the separation of North America and Asia in the evolutionary history of the group.

MATERIALS AND METHODS

TAXONOMIC SAMPLING

Here we have sampled all species in *Phegopteris*. We sampled 18 accessions of *Phegopteris connectilis* including 16 from northeastern North America, one from Northern France, and four accessions from Japan. Additionally, we sampled four accessions of *P. hexagonoptera* from across the Eastern United States, six accessions of the undescribed
tetraploid *Phegopteris* from three sites within Vermont, and one accession of *P. decursivipinnata* from central China (Appendix A, Figure 1a and 1b).
Figure 1. A) Distribution of the three species described in genus *Phegopteris*. B) Collections sites of the undescribed tetraploid *Phegopteris*. 
SPORE MEASUREMENTS
For each species in *Phegopteris*, sporangia and spores from one to two specimens (Appendix A) were mounted in Hoyer’s medium on glass slides and imaged at 100x magnification. The sporangia and spores were imaged in order to count the number of spores per sporangium as well as measure the length of spores. One to two sporangia per specimen were counted: plants with spore counts of 32 or fewer were inferred to be apomictic; those with more than 32 per sporangium were inferred to be sexual. 25 to 30 spores per specimen were measured to calculate mean length and standard deviation for each species. Spore length was measured from the images using ImageJ (Schneider et al., 2012). The external spore membrane (*perispore*) was excluded from measurements. These lengths were used to infer ploidy.

DNA EXTRACTION, AMPLIFICATION, AND SEQUENCING
Total genomic DNA was extracted from fresh (1 g) or silica-dried (0.5 g) leaves using a cetyl trimethyl-ammonium bromide (CTAB) procedure (Doyle and Doyle, 1987) with some modifications. Leaves were ground with a bead-beating machine using glass beads. CTAB buffer was supplemented with polyvinyl pyrrolidone (PVP), and crushed leaf tissue was precipitated in chloroform. Samples were then subjected to washes in 70% and 90% ethanol and re-suspended in Tris-EDTA buffer. The plastid DNA sequences *psba-trnH* and *trnS-rps4* spacer region were PCR-amplified under standard conditions using previously published primers, with modifications (Table 1). The marker *trnS-rps4* was amplified using an initial denaturation step of 94 °C for 3 min, followed by 30 cycles of 94 °C for 45 s, 53 °C for 45 s, and 72 °C for 2 min, and a final extension at 72 °C for 10 min. Amplification of *psba-trnH* followed similar conditions except that the initial and final
steps were done for 10 min, the cycles were done 35 times with 58 °C annealing for 1 min, and extension at 72 °C for 1 min. Resulting PCR products were cleaned using ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA), and sequenced on an ABI PRISM 3730x automated sequencer (Beckman Coulter Genomics, Danvers, MA, USA). Each plastid marker was sequenced in both the forward and reverse direction using the amplification primers.

Nuclear marker PgiC (Exons 14-16) (Table 1) was amplified using an initial denaturation step of 94 °C for 3 min, followed by 30 cycles of 94 °C for 1 minute 60 °C for 30 s, and 72 °C for 2 min, and a final extension at 72 °C for 10 min. Resulting PCR products were cleaned using ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA), and sequenced as above.

Nuclear marker ITS (introns 11-17) (Table 1) was amplified using an initial denaturation step of 94 °C for 3 min, followed by 30 cycles of 94 °C for 1 minute 58 °C for 45 s, and 72 °C for 2 min, and a final extension at 72 °C for 10 min. The PCR products were visualized on an agarose gel, then excised and purified using the Prep-Ease gel extraction kit (Affymetrix, Santa Clara, California). Purified fragments from *Phegopteris hexagonoptera, P. connectilis*, and the undescribed tetraploid cytotype were cloned using the PGEM vector cloning kit (Promega, Fitchburg, Wisconsin) following the manufacturer's instructions. Target fragments were amplified from purified plasmid DNA from 4–6 isolated colonies and sequenced as above.

Nuclear marker gapCp (exons 8-11) (Table 1) was amplified using barcoded forward and reverse primers, using an initial denaturation step of 94 °C for 3 min, followed by 30 cycles of 94 °C for 30 seconds, 65 °C for 45 s, and 72 °C for 2 min, and a final extension at 72 °C
for 10 min. Amplicons were prepared for next-generation sequencing using the Pacific Biosciences sequencing platform (Pacific Biosciences, Menlo Park, CA, USA) according to the protocol of Rothfels et al. (2017). PCR products were run on agarose gels and imaged. Relative concentration of each PCR product was estimated using a five-point scale from very weak to very strong. These products were multiplexed in various volumes proportional to concentration in order equalize total nanograms of DNA from each product in the final library. The multiplexed PCR products were purified using Ampure XP magnetic beads (Beckman Coulter Genomics, Danvers, MA, USA). The library was sequenced at the Arizona Genomics Institute.

SEQUENCING

Phylogenetic analysis was performed on three datasets, the two plastid markers (psba-trnH and trnS-rps4) and each of two of the nuclear markers (ITS and gapCp) separately, resulting in three phylogenies. Reads of the two plastid markers, as well as single-molecule reads of ITS generated from vector cloning, were prepared for phylogenetic analysis using the following approach. Consensus sequences were generated from assemblies of forward and reverse reads of both markers, and aligned using MUSCLE (Edgar, 2004) with minor manual adjustments. Assembly, consensus generation, and alignment were implemented in Geneious version 9.0 (Kearse et al., 2012). Indels were coded simply as single characters with binary states (simple gap coding; Simmons and Ochoterena, 2000). The resulting indel data were appended to the end of sequences for use in all subsequent phylogenetic inference analysis. Plastid sequences were concatenated into one plastid data set for further analysis.
Raw sequences generated from sequencing on the PacBio platform for nuclear marker \textit{gapCp} were processed using the PURC pipeline, created by Rothfels et al. (2017), for parsing reticulate relationships using single-molecule reads of nuclear sequences. Reads were demultiplexed according to barcodes in order to identify accessions.
<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>psba</em>-</td>
<td>F 5’GGTTCAAGTCCCTCTATCCC3’ R 5’ATTTGAACCTGGTGACACGAG3’</td>
<td>Schneider et al. (2012)</td>
</tr>
<tr>
<td>trnH</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>rps4</em>-</td>
<td>F 5’TTCACGAGGTTCTGAATCCCTC3’ R 5’GAGTATATTACTCCGCAGGAAG3’</td>
<td>McHenry and Barrington (2014)</td>
</tr>
<tr>
<td>trnS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITS</td>
<td>F 5’TCCCTCCGTATTTGATATGC3’ R 5’ACGAATTCATGGTCCGGGTAAAG3’</td>
<td>Topik et al. (2005)</td>
</tr>
<tr>
<td>gapCp</td>
<td>F 5’CYACMAACTGCCTTGRCRCTCTTGC3’ 5’GTATCCCCACTCRRTATCATACC3’</td>
<td>Rothfels et al. (2017)</td>
</tr>
<tr>
<td>PgiC</td>
<td>F 5’TTTGCTCCTCATTCAAC 3’ R 5’CTTAGTATGGAAAGCATAATGGGAGG3’</td>
<td>Lyons et al. 2017; Koenemann et al. 2011</td>
</tr>
</tbody>
</table>

**Table 1.** Primers used, modified from given references.
Chimeric sequences and reads shorter than 500 bp were removed. In order to cluster reads into alleles, we chose to build consensus sequences from reads with greater than 95% similarity as a parameter in the PURC pipeline. The minimum retained cluster size was ten reads. Subsequently, consensus sequences were aligned using MUSCLE (Edgar, 2004) with minor manual adjustments. Assembly, consensus generation, and alignment were implemented in Geneious version 9.0 (Kearse et al., 2012). Indels were coded simply as single characters with binary states (simple gap coding; Simmons and Ochoterena, 2000). The resulting indel data were appended to the end of sequences for use in all subsequent model testing and phylogenetic analysis.

BI analysis was applied to the plastid dataset as well as the cloned ITS dataset and the next-generation sequencing gapCp dataset, each using MrBayes version 3.2.6 (Ronquist et al., 2012) on the CIPRES Science Gateway server (Miller et al., 2010), generating three phylogenetic trees. The BI analysis of plastid data included sequences from species of several closely related genera Pseudophegopteris, Macrothelypteris, and Thelypteris. The BI analysis of ITS is rooted only with Phegopteris decursivipinnata, and the BI analysis of gapCp is rooted with one species of related genus Thelypteris. For BI analysis of plastid data as well as each nuclear marker separately, the optimal evolutionary models for each marker was discerned from jModeltest 2 (Darriba et al., 2012) using the Akaike Information Criterion (AIC). For the plastid dataset, MrBayes was run for 6 million generations with trees sampled every 1000 generations. The first 500,000 trees were discarded as burn-in iterations, the remainder were used to generate a 50% majority-rule
consensus tree. For the ITS and \textit{gapCp} datasets, MrBayes was run for 10 million generations, with trees sampled every 1000 generations. The first 600,000 trees were discarded as burn-in iterations, and the remainder were used to generate 50\% majority-rule consensus trees. For all datasets, posterior probabilities were obtained from MrBayes, and the phylogenetic tree including branch lengths was visualized using FigTree Version 1.4 (Rambaut and Drummond, 2008).

**MIXED NUCLEOTIDE SIGNALS AND RETICULATION**
The aligned chromatograms of the original trimmed forward and reverse sequences for the nuclear marker \textit{pgiC} were examined for multiple nucleotide peaks at each position. Following Tate et al., (2006), Jorgensen and Barrington (2017), and Lyons et al., (2017), these multiple peaks were taken as evidence of alleles inherited from hybrids progenitors, retrieved for the marker in question, summed in the single chromatogram generated from direct Sanger sequencing.

**TIME CALIBRATION**
Divergence times were estimated for \textit{Phegopteris} using the plastid dataset (including the markers \textit{psba-trnH} and \textit{rps4-trnS}) by applying a Bayesian approach using the program BEAST version 2.4.7 (Drummond and Rambaut, 2007) with the relaxed phylogenetic method of Drummond et al. (2006). Data were partitioned by plastid marker region, and the best fit model as determined by jModeltest2 was applied to each partition (Darriba et al., 2012). Secondary time calibrations were applied in BEAUti v. 2.4.6, since fossils are not available for \textit{Phegopteris}. The most recent common ancestor of outgroup \textit{Macrothelypteris torresiana} is estimated to have diverged no less than 68.5 mya, and the most recent common ancestor of \textit{Phegopteris} is estimated to have diverged no less than
45.9 mya (Schuettpelz and Pryer, 2009). A relaxed lognormal clock was applied to the node constraints. A birth-death speciation prior was used with a gamma model of rate variation. The analysis was run for 10 million generations with sampling every 1000 generations. Log files were examined in Tracer v1.5 (Rambaut and Drummond, 2007) to ensure appropriate sampling. Trees were summarized in TreeAnnotator v1.6.2 (Drummond and Rambaut, 2007). Trees with node-age estimates were visualized in figtree (Rambaut, 2008) with the 95% highest posterior density (HPD) intervals.

RESULTS

SPORE MEASUREMENTS AND PLOIDY
Findings from measurement of spore length reveal that spores of the undescribed tetraploid *Phegopteris* have a mean spore length approximately 1/3 longer than triploid apomictic accessions of *P. connectilis*. In addition, sporangia in the undescribed tetraploid contain, on average, 30 spores per sporangium (Table 2). These findings are consistent with previous results of chromosome squashes suggesting that this undescribed lineage is a tetraploid and an apomict (Mulligan et al., 1972; Driscoll et al., 2003). Spores of *P. connectilis*, a well-known triploid apomict (Mulligan et al., 1972), are shorter in average length than the undescribed tetraploid among accessions sampled both of North America and France, and sampled sporangia contained 20-30 spores (Table 2). Japanese accessions of *P. connectilis* collected from Japan were shorter than North American and European *P. connectilis*, consistent with previous findings suggesting that it is a diploid (Matsumoto, 1982). Sporangia from all Japanese *P. connectilis* accessions contained greater than 50
spores (Table 2). Accordingly, our findings corroborate previous findings that *P. connectilis* is a triploid apomict across most of its circumboreal range, but includes a diploid sexual cytotype in Japan. *P. decursivipinnata* accessions all had a spore length similar to the diploid *P. connectilis* (Table 2) and also sporangia sampled had 58-60 spores, suggesting that our sampled *P. decursivipinnata* is the diploid sexual cytotype.
<table>
<thead>
<tr>
<th>Species</th>
<th>Mean Spore Length</th>
<th>Spore Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Phegopteris decursivipinnata</em></td>
<td>41.2 (2.4)</td>
<td>60</td>
</tr>
<tr>
<td><em>Phegopteris connectilis</em> (North America)</td>
<td>40.1 (1.3)</td>
<td>21</td>
</tr>
<tr>
<td><em>Phegopteris connectilis</em> (France)</td>
<td>38.6 (1.1)</td>
<td>-</td>
</tr>
<tr>
<td><em>Phegopteris connectilis</em> (Japan)</td>
<td>32.2 (2.1)</td>
<td>61</td>
</tr>
<tr>
<td><em>Phegopteris hexagonoptera</em> (2x)</td>
<td>31.7 (1.4)</td>
<td>58</td>
</tr>
<tr>
<td>Undescribed tetraploid</td>
<td>52.2 (1.1)</td>
<td>30</td>
</tr>
</tbody>
</table>

**Table 2.** Mean spore length for each species in *Phegopteris* and cytotypes. Confidence intervals for
PLASTID BASED PHYLOGENETIC RELATIONSHIPS

BI analysis of plastid sequences reveals a clade that includes all sampled accessions of all species in the genus *Phegopteris* C. Presl (Figure 2). These findings are similar to that of Schneider et al. 2012, who found the genera *Phegopteris* and *Pseudophegopteris* to be monophyletic.

The *Phegopteris* clade comprising three subclades, representing the species *P. hexagonoptera*, *P. decursivipinnata*, and *P. connectilis*. The *P. connectilis* clade includes the Japanese diploid cytotype of *P. connectilis*, as well as the undescribed apomictic tetraploid
Figure 2. Phylogeny based on the combined analysis of plastid markers, *psba-trnH* and *rps4-trnS*. The tree is the 50% majority rule Bayesian Inference (BI) phylogram of the genus *Phegopteris* including select species representing sister genera *Pseudophegopteris*, *Thelypteris*, and *Macrothelypteris*. Posterior probabilities are given at each node. The triploid cytotype of *Phegopteris connectilis* is shown in pink, the diploid cytotype of *P. connectilis* is shown in orange, the undescribed tetraploid is shown in green, *P. hexagonoptera* is shown in yellow, and *P. decursivipinnata* is shown in blue.
SINGLE MOLECULE NUCLEAR PHYLOGENETIC RELATIONSHIPS

BI analysis of single molecule reads generated from vector cloning of purified amplicons of nuclear markers ITS reveal two clades (Figure 3). *Phegopteris hexagonoptera* is unresolved outside of these two clades. One clade includes accessions of only the undescribed tetraploid. The second clade includes two subclades. One includes accessions of the undescribed tetraploid, and the other includes only accessions of the triploid *Phegopteris connectilis*. All accessions of the sexual diploid cytotype of *P. connectilis* are unresolved outside this clade. All sampled accessions of the undescribed tetraploid species are represented in each of these two clades in which it resolves.

Next-generation sequencing of purified amplicons of the nuclear marker *gapCp* yielded 10,016 reads after eliminating chimeric sequences and raw reads lacking barcodes. Generation of consensus sequences constructed from reads with greater than 95% similarity resulted in 29 consensus sequences. Each consensus sequence was constructed from an average of 345 reads. BI analysis of these consensus sequences reveals results similar to BI analysis of the ITS data. Two monophyletic clades correspond to *Phegopteris hexagonoptera* and *P. connectilis* (Figure 4). The *P. connectilis* clade comprises one subclade and an array of unresolved consensus sequences. The one resolved subclade includes only reads representing the undescribed tetraploid *Phegopteris*. The unresolved accessions include reads representing the undescribed tetraploid, the triploid *P. connectilis*, and the diploid *P. connectilis*. Phylogenetic analysis of both ITS and *gapCp* datasets suggest that the undescribed tetraploid is an allopolyploid, the triploid *P. connectilis* is an
autopolyploid, and the diploid *P. connectilis* as well as *P. decrusivipinnata* are homozygous at the sampled loci.

**SUMMED NUCLEOTIDE SIGNALS AND RETICULATION**
Summed nucleotide signals evident in chromatograms of nuclear marker PgiC (Table 3) were encountered only in the undescribed tetraploid, suggesting that it is an allopolyploid. All of the seven two-nucleotide calls included one nucleotide common at that same site in sequences representing the triploid and diploid *Phegopteris connectilis* and one otherwise not encountered in the remaining taxa. In each case, the chromatogram peak for the nucleotide shared with diploid and triploid *P. connectilis* is lower than the signal not shared with any taxon sampled.

**DIVERGENCE TIME ESTIMATES**
The chronogram obtained from divergence-time estimation using the plastid dataset has a topology very similar to the phylogeny generated from BI analysis of the same dataset (Figure 5). Our divergence-time estimate for the clade including both cytotypes of *Phegopteris connectilis* as well as accessions of the undescribed tetraploid is the Eocene-Miocene boundary (33.3 mya). The evolutionary origin of this clade long predates the physical separation of North America and East Asia by severance of the Bering Land Bridge (Milne and Abbott, 2002).
<table>
<thead>
<tr>
<th></th>
<th>96</th>
<th>110</th>
<th>111</th>
<th>222</th>
<th>398</th>
<th>678</th>
<th>694</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Phegopteris connectilis</em> (3x)</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td><em>Phegopteris connectilis</em> (2x)</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>Undescribed</td>
<td>A/T</td>
<td>G/C</td>
<td>C/T</td>
<td>A/G</td>
<td>T/C</td>
<td>T/G</td>
<td>A/G</td>
</tr>
<tr>
<td><em>Phegopteris hexagonoptera</em></td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td><em>Phegopteris decursivipinnata</em></td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

**Table 3.** Base positions in an alignment of chromatograms of PgiC with multiple peaks. For species with double peaks in the chromatograms, the nucleotide giving a stronger signal is shown in bold.
Figure 3. Phylogeny based on analysis of single molecule reads of nuclear marker ITS generated through sequencing of vector cloned amplicons. The tree is the 50% majority rule Bayesian Inference (BI) phylogram of the genus *Phegopteris* including select species representing sister genera *Pseudophegopteris*, *Thelypteris*, and *Macrothelypteris*. Posterior probabilities are given at each node. The triploid cytotype of *Phegopteris connectilis* is shown in pink, the diploid cyptotype of *P. connectilis* is shown in orange, the undescribed tetraploid is shown in green, *P. hexagonoptera* is shown in yellow, and *P. decursivipinnata* is shown in
Figure 4. Phylogeny based on analysis of single molecule reads of nuclear marker gapCp generated through next generation sequencing of amplicons. The tree is the 50% majority rule Bayesian Inference (BI) phylogram of the genus *Phegopteris* including select species representing sister genera *Pseudophegopteris*, *Thelypteris*, and *Macrothelypteris*. Posterior probabilities are given at each node. The triploid cytotype of *Phegopteris connectilis* is shown in pink, the diploid cyptotype of *Phegopteris connectilis* is shown in orange, the undescribed tetraploid is shown in green, *Phegopteris hexagonoptera* is shown in
Figure 5. A time calibrated phylogeny based on the combined analysis of plastid markers, psba-trnH and rps4-trnS. Node ages are given in millions of years at each node. The tree is a phylogram of the genus Phegopteris including select species representing sister genera Pseudophegopteris, Thelypteris, and Macrothelypteris. The triploid cytotype of Phegopteris connectilis is shown in pink, the diploid cytotype of P. connectilis is shown in orange, the undescribed tetraploid is shown in green, P. hexagonoptera is shown in yellow, and P. decursivipinnata is shown in blue.
DISCUSSION

ORIGIN OF THE UNDESCRIBED TETRAPLOID

The undescribed tetraploid apomict appears to have a single maternal progenitor in that it resolves in one clade in the plastid phylogeny (Figure 2). Its similarity to diploid and triploid *Phegopteris connectilis* in the cpDNA phylogeny suggests a shared maternal origin with these two cytotypes. Since apomictic species typically act as paternal progenitors in hybridization scenarios by virtue of typically only producing functional antheridia (Manton, 1950), the sexual diploid cytotype of *P. connectilis*, rather than the apomictic triploid cytotype, is the more reasonable maternal progenitor to the undescribed tetraploid. The summed nucleotide signals in PgiC would suggest that a larger proportion of the undescribed tetraploid’s genome is constituted by its paternal progenitor than by the genome ostensibly shared with the diploid *P. connectilis* (Table 3). This genomic composition is consistent with the diploid sexual Japanese *Phegopteris connectilis*, contributing only one quarter of the allotetraploid genome, acting as a maternal progenitor.

BI analysis of single-molecule reads from nuclear markers ITS and gapCp reveal a shared pattern in which one allele appears to share a heritage with diploid and triploid *Phegopteris connectilis*, whereas the second allele is not highly similar to any species sampled here (Figure 3 and Figure 4). Given that *Phegopteris* is exhaustively taxonomically sampled here, the second progenitor of the undescribed tetraploid species may be extinct. A similar evolutionary history in which an extinct taxon plays an important role in extant genomic diversity has been observed in multiple angiosperm and pteridophyte groups (Jakob and Blattner, 2010; Stein et al., 2010). Given that the most
likely maternal progenitor to the undescribed tetraploid is a sexual diploid, the unsampled paternal progenitor of the undescribed allotetraploid is likely a triploid apomict.

Neither plastid nor nuclear markers analyzed here offer support for the hypothesis of Mulligan et al. (1972), that Phegopteris hexagonoptera is a progenitor to the undescribed allotetraploid tetraploid. Rather, our findings corroborate the findings of Driscoll et al. (2003), who utilized isozyme electrophoresis to parse hybrid origins of allotetraploid. Isozyme loci pgm-2 and pgi-2 suggest allelic inheritance in the undescribed tetraploid from P. connectilis and from a second, unsampled progenitor (Figure 6).

DIVERGENCE TIME ESTIMATES AND BIOGEOGRAPHIC DISJUNCTION
The estimated divergence time for the clade including triploid and diploid Phegopteris connectilis and the undescribed allopolyploid (33.3 mya), is consistent with an evolutionary origin prior to the physical separation of East Asia and Eastern North America (5 mya) (Figure 5). The circumboreal distribution of the triploid P. connectilis is consistent with a global distribution during the Eocene, a time of physical continuity among North America, Europe, and Asia, that subsequently separated. However, dispersal following separation cannot be ruled out.
Figure 6. Hypothesized reticulation scenario for the allopolyploid origins of the undescribed tetraploid *Phegopteris*. Letters represent genomes within each lineages. Dashed lines indicate a reduced genomic contribution, while a solid line represents an unreduced genomic contribution.
Circumboreal plant distributions have been inferred to be relictual global distributions in some angiosperm groups (Wang et al., 2015; Ickert-Bond et al., 2009), though dispersal has also been inferred in some ferns (Jorgensen and Barrington, 2017).

Some angiosperm groups with a relictual distribution disjunct between East Asia and North America have undergone significant subsequent diversification on each continent, leading to multiple phylogenetically and morphologically distinct lineages (Ickert-Bond et al., 2009). For ferns, this history often involves hybridization and polyploidization (Werth and Windham, 1991). However, given that the origin of the clade comprising *Phegopteris connectilis* and the undescribed tetraploid is estimated to have originated approximately 30 million years before the last physical connections between Eastern North American and East Asia was flooded, it is likely that the hybridization event yielding the undescribed tetraploid, now found in New England, occurred well before North American-East Asian vicariance. Nuclear and plastid data suggesting that the sexual diploid cytotype of *P. connectilis*, found exclusively at elevations above 2000 meters in Japan, is a progenitor to the undescribed allotetraploid, found exclusively in northeastern North America, further suggests that a previously continuous distribution for the diploid is likely.

**MULTIPLE APPROACHES TO PARSING ALLELIC INHERITANCE**

Our results include two sequencing-based approaches to obtaining single-molecule reads, as well as direct Sanger Sequencing of a third marker, which yields reads composed of base-pair calls that are a summation of reads. Consideration of both yields insight into the factors to consider in using allelic variation in polyploids to reconstruct hybrid origins. Although *gapCp* and ITS reveal a similar pattern of allelic inheritance, they differ
substantially in terms of degree of sequence divergence among these alleles. As evident in the long branches of the phylogeny based on nuclear marker ITS, these alleles have undergone substantial mutation before the hybridization event leading to the undescribed tetraploid. On the other hand, there is very little sequence divergence between the two alleles of gapCp retrieved from the allotetraploid. These differences are likely attributable to the known high substitution rates in ITS, an untranscribed intergenic spacer, and hence a non-coding region. However, the summed nucleotide signal data from nuclear marker PgiC, as well as isozyme electrophoresis findings from Driscoll and Barrington (2003) reveal a pattern of allelic inheritance pointing to an allopolyploid origin for the undescribed tetraploid. Thus, the use of multiple witnesses to evolutionary history is important in tracing reticulate evolution.

Before widespread use of next-generation sequencing technology, isozyme electrophoresis and vector cloning were utilized in understanding numerous reticulate fern groups (Gastony et al., 1992; Driscoll et al., 2003; Tsutsumi et al., 2011). Our findings here suggest that findings from early investigations of reticulate ferns groups provide a strong foundation upon which next-generation sequencing technology can build.
LITERATURE CITED


Rambaut, A., and A. Drummond. 2008. FigTree: Tree figure drawing tool, version 1.2. *2. Institute of Evolutionary Biology, University of Edinburgh.*


54


CHAPTER 3: BIODIVERSITY AND APOMIXIS: INSIGHTS FROM EAST-ASIAN HOLLY FERNS WITHIN POLYSTICHUM SECTION XIPHOPOLYSTICHUM

ABSTRACT

Apomixis is an evolutionarily important phenomenon across plant lineages. The interaction of apomixis with hybridization and polyploidy can lead to complex patterns of reticulation, complicating efforts to reconstruct evolutionary history in groups where apomixis is common. Ferns, in particular, are rich in apomictic species, notably in centers of species diversity like East Asia. We investigated the East Asian ferns in Polystichum section Xiphopolystichum sensu stricto in order to elucidate the evolutionary history of five apomictic species in the group: Polystichum tsus-simense, P. xiphophyllum, P. sinotsus-simense, P. pseudoxiphophyllum, and P. mayebarae. The datasets comprised phylogenetic inference based on two nuclear and three plastid markers, analysis of mixed nucleotide signals from chromatograms generated from Sanger sequencing of nuclear markers, ploidy estimates based on flow cytometry data and spore length measurements, and morphometric analysis of representative specimens collected in southwest China and nearby regions. By interpreting these multiple lines of evidence synthetically, we report two novel cytotypes of the known apomictic species Polystichum xiphophyllum and P. tsus-simense and provide a complex scenario for reticulation leading to extant species diversity in the section. We conclude that apomictic species diversity in the group has been generated largely from repeated hybridization between two sexual diploid species in the group, Polystichum revolutum and P. herbaceum. Our findings highlight the importance of understanding the
role of apomictic reproduction in the context of species diversity, an understanding central to similar future inquiry into the diversity of East Asian ferns.

INTRODUCTION

Polyploidy, apomixis, and hybridization are common across many plant lineages and are important catalysts for evolution and speciation. For pteridophytes in particular, these three evolutionary processes are often closely linked and, together, are responsible for the complex patterns of reticulate evolution found in an array of fern genera (Soltis et al., 2004; Otto and Whitton, 2000; Grusz, 2016). Thus, understanding sources of diversity in ferns, especially in world centers of diversity such as western China and the tropical Andes, rests on deciphering how polyploidy, hybridization, and apomixis interact to shape patterns of pteridophyte speciation.

Polyploidy, characteristic of up to 95% of fern species (Grant, 1981), may arise via chromosomal doubling (autopolyploidy) or hybridization followed by doubling of chromosomes (allopolyploidy—Manton, 1950; Stebbins, 1950). These hybrids may either enhance or compromise genetic diversity depending on the nature of the hybrid or parental species involved (Barrington et al., 1989). Many fern polyploids, notably allopolyploids, are capable of ordinary sexual reproduction. However, both auto- and allopolyploids may have an unbalanced set of genomes, i.e. a genome from at least one parent is represented with an odd number of copies. The result is that chromosomes in such a genome are left either with multiple homologs with which to pair during meiosis, or none at all. One means of overcoming sterility imposed by an unbalanced genome is apomictic reproduction (Cosendai et al. 2011). There is a strong association between polyploidy, both auto- and
allopolyploid, and apomixis, and nearly three quarters of apomictic ferns are triploid (Wagner and Wagner, 1980; Asker and Jerling, 1992). It should be noted, however, that some autotetraploid fern lineages are apomictic, such as *Pellaea glabella* Mett. ex Kuhn and *P. occidentalis* Rydb. (Gastony, 1988), and some diploid apomicts are known, including *Dryopteris wallichiana* (Spreng.) Hyl., *Cheilanthes leucopoda* Link, and *Pteris cretica* L. (Verma and Khullar, 1965; Knobloch, 1967; Fraser-Jenkins, 2007). In general, apomixis may be adaptive in relation to environmental factors such as aridity or insularity (Haufler et al., 2016), and hence, as a means of reproduction, it may confer an adaptive advantage to the less genetically resilient autopolyploid apomicts.

Unlike vegetative sexual reproduction, apomixis, by including spore production, maintains the benefit of spore dispersal and also allows for the accumulation of genetic variation through recombination of novel mutations (when meiosis persists). In fern apomixis, unreduced spores are produced at the end of meiosis and syngamy does not occur. In the Döpp-Manton (Manton, 1950) scheme of fern apomixis, successful sporogenesis includes three mitotic divisions followed by chromosomal replication but not mitotic division, resulting in a restitution nucleus with each chromosome represented by two newly produced homologues. A final complete meiotic division results in 32 spores—a spore number characteristic of apomictic ferns. Importantly, in the Döpp-Manton scheme of apomixis, gametophytes usually develop only antheridia (the male reproductive structures) not archegonia (the female reproductive structures) typically found in sexually reproducing gametophytes. Therefore Döpp-Manton apomicts are most often paternal progenitors when they are involved in hybridization.
The relationship between apomixis and reticulation makes it important to decipher the relative contribution of sexually reproducing progenitors and their derivative apomicts to the total pattern of diversity in ferns. Reticulate evolutionary histories can be inferred when two or more alleles are sequenced from a single individual, and end up phylogenetically sister to sequences derived from two or more distinct species. Such multi-allele phylogenies can be generated through the use of vector cloning, next-generation sequencing, and/or careful analysis of chromatographic reads generated from direct sequencing. The last technique can specifically reveal double peaks resulting from summation of multiple alleles at a single position in the nucleotide sequence, as would be expected from a heterozygote or polyploid (Zhidkov, 2010; Chang et al., 2012; Lyons et al., 2017). Molecular phylogenies constructed from sequences generated by direct Sanger sequences that do not utilize such analyses may be misleading in taxonomic groups predisposed to hybridization, as phylogenies are strictly bifurcating representations of speciation (Soltis and Soltis, 2009).

Several prominent recent investigations have employed an integrated systematic approach to disentangling the evolutionary history and delimiting species in highly reticulate fern species complexes in which polyploidy, hybridization, and apomixis all play a role in speciation. Grusz (2009) demonstrated that two closely related North American apomicts — *Myriopteris wootonii* (Maxon) Grusz & Windham and *M. yavapensis* (T. Reeves ex Windham) Grusz & Windham — both unbalanced allopolyploids—actually share diploid progenitors. Similarly, in the Aspleniaceae, four unbalanced polyploid apomictic lineages in the *Asplenium monanthos* L. complex likely share a sexual diploid progenitor, but differ in their ploidy and genomic contributions (Dyer et al., 2012). Hori et
al. (2014) used plastid and nuclear sequence data to determine the genomic composition of several apomicts in the East Asian *Dryopteris varia* (L.) Kuntze complex. Nuclear *pgiC* sequences suggested that apomicts in the complex each resulted from unique hybridization events among a group of closely related sexual diploid species. Other notable work in this area has been done in *Pellaea* (Gastony, 1988), *Cornopteris* (Park and Kato, 2003), and *Phegopteris* (Driscoll et al., 2003). Many of these notable studies found strong evidence from nuclear markers of hybridization between apomictic species and their progenitors, resulting in new apomictic lineages. However, elucidation of the patterns of genomic inheritance as well as the identity and relationships of lineages within these complexes can open the door to other more confounding issues in the study of fern apomixis. For example, Dyer et al. (2012) and Hori et al. (2014) both found evidence of multiple hybrid origins and multiple cytotypes, in the *A. monanthes* and the *D. varia* complex, respectively. Findings such as these often complicate taxonomy as well as efforts to consistently apply criteria with the goal of defining species in these complexes.

Species can be defined by criteria relying on the mechanism of speciation, processes or biological forces maintaining cohesion and isolation from other species, and unique molecular or morphological characteristics (Cracraft, 1990; Nixon and Wheeler, 1990; DeQueiroz, 1998). The biological processes inherent in apomicts can cause problems with each of these approaches to characterizing species. Indeed, the species concepts that are currently most widely applied, including the biological, phenetic, and some versions of the phylogenetic species definitions (Mayr, 1942, 1963; Mishler, 1985; Nixon and Wheeler, 1990), fail to recognize apomictic lineages as distinct species. For instance, apomictic ferns, by undergoing primarily or exclusively asexual reproduction, defy the biological
species definition’s tenet that a species remains cohesive through interbreeding (Gastony, 1988, 1989; Grusz, 2009). Similarly, one implication of the tendency of apomictic ferns to arise from multiple origins, or hybridization, is that many potentially evolutionarily and phylogenetically distinct apomictic lineages are taxonomically united as one species on the basis of morphology under the phenetic species concept (Barrington et al., 1989; Takamiya et al., 2001). In contrast, some apomictic species may belong to a single lineage with a single evolutionary origin, but exhibit more intra- to inter-specific morphological variation. The result is accumulating mutations that become fixed in the absence of sexual reproduction within various populations, leading to taxonomic categorization as multiple species as defined by morphology. The phylogenetic species criterion has three variants, each espoused by different authors: (1) “…the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent” (Nixon and Wheeler 1990), (2) “a population or group of populations defined by one or more apomorphous features” (Rosen 1979), and (3) “…that set of organisms between two speciation events, or between one speciation event and one extinction event” (Ridley 1989). Apomictic lineages often meet criteria versions (1) and (2), in that they are often genetically and phylogenetically distinct units, each with unique apomorphies—molecular, morphological, and reproductive (Figures 1 and 2). The third variant of the phylogenetic species criterion is much more difficult to apply to apomictic lineages because they may be the products of, and participants in, repeated hybridization.

Given the shortcomings of applying just one of these definitions to the problem of species delimitation in apomictic ferns and their allies, we support the use of a species concept that incorporates phylogeny, morphology, ecology, and reproduction, as
incorporated in the general lineage concept of de Queiroz (1998, 1999). De Quieroz has advocated for a more inclusive approach to species definition and delimitation, one that accommodates both sexual and asexual species. He argues that the species concept should be considered in the context of lineages, where the discrete taxon is part of a “series of entities forming a single line of direct ancestry and descent” (de Quieroz, 1998). Under his General Lineage Concept, a species is considered a segment of a lineage, in that a species is temporally separated from its ancestors and descendants on an evolutionary time scale. Of central importance to us, this more unified species concept treats other species concepts as criteria for species delimitation, all relevant to delimiting species under the general lineage concept. Each species in an apomictic complex should be defined, at a minimum, by one accepted species criterion.

The monophyletic Polystichum Roth sect. Xiphopolystichum Daigobo of Polystichum (Dryopteridaceae) (Daigobo et al., 1972; Zhang and Kung, 1996; Kung et al., 2001; Lu et al., 2007; Li et al., 2008), known commonly as holly ferns, currently includes both section Xiphopolystichum sensu stricto (s.s.; Kung et al., 2001) and section Duropolystichum Fraser-Jenkins (Zhang and Barrington, 2013). Virtually all Xiphopolystichum species are native to East and Southeast Asia, with most being endemic to the region. Species in this group are most abundant in mountainous regions surrounding the Sichuan Basin (Zhang and Barrington, 2013), including the majority of known apomictic lineages. Several named species within section Xiphopolystichum s.s. have been identified as triploid apomicts: Polystichum tsus-simense (Hook.) J. Sm, P. xiphophyllum (Baker) Diels, P. mayebarae Tagawa, and P. sinotsus-simense Ching & Z. Y. Liu (Daigobo 1973; Gibby 1985), and have an array of leaf dissections between once and twice pinnate,
most typically once pinnate-pinnatifid. *Polystichum pseudoxiphophyllum* Ching ex H. S. Kung is highly morphologically similar to *P. xiphophyllum*, though cytological and reproductive data are lacking for this taxon. Given the tendency for hybrid ferns to exhibit a morphology intermediate between their parents (e.g., Barrington, 1986), we posit that apomictic lineages with a once pinnate-pinnatifid laminar dissection are likely to have originated as hybrids of once-pinnate and twice-pinnate progenitors (Wagner, 1983). *Polystichum herbaceum* Ching & Z. Y. Liu and *P. revolutum* P.S. Wang are proposed to be sexually reproducing diploids (Gibby 1985) belonging to *P. sect. Xiphopolystichum sensu lato* (s.l.). Among these candidate diploid progenitors, only one, *Polystichum revolutum*, is once-pinnate.

Here, we establish phylogenetic relationships among taxa currently circumscribed in *Xiphopolystichum s.s.* to elucidate the evolutionary history of apomictic lineages in the group and to delimit species using a combination of criteria appropriate for apomicts. In *Xiphopolystichum*, we contend that delineating species is best based on the phylogenetic, phenetic, and evolutionary species criteria considered together. Our dataset lends itself to evaluating each of these criteria both independently and synthetically toward defining species.

We explore the contribution of the once-pinnate *Polystichum revolutum* to the ancestry of apomictic lineages in section *Xiphopolystichum* through a study of nuclear markers in the light of ploidy levels, breeding systems, and morphology. Specifically, we test the hypothesis that *P. revolutum* is the once-pinnate progenitor that contributed the less-dissected morphology to the apomicts. Further, we address three major questions regarding the genomic composition of the apomicts in relation to their morphology. Are the genomes
in balance? Do the identities and proportions of the contributed genomes relate to the variation in lamina dissection so prominent in the complex? Do these genetic profiles yield insights into the delimitation of taxonomic species in the section?

METHODS

TAXON SAMPLING AND SPECIES DELIMITATION

We assembled a broad sample of species in Polystichum sect. Xiphopolystichum s.s., as well as select members of P. sect. Xiphopolystichum s.l. as circumscribed in Zhang and Barrington (2013). Xiphopolystichum s.l. includes members of Duropolystichum. Polystichum otophorum, included in Xiphopolystichum s.s. (Zhang and Barrington 2013) is here considered synonymous with P. xiphophyllum given morphological and molecular similarity. The single prominent exception is Polystichum pseudosetosum Ching & Z. Y. Liu, for which we had no material. Sampling of Xiphopolystichum s.s., comprised 32 accessions collected across western China including Chongqing, Guizhou, and Sichuan provinces during two field trips by the authors, one in 2006 and one in 2015. We chose to sample accessions from as wide a geographic and morphological range as possible. In general, three to five accessions for each ingroup species were analyzed to account for morphological variation and potential population variation within species. Herbarium vouchers for all collections were deposited at the Pringle Herbarium (VT), University of Vermont, Burlington, VT, USA, or the herbarium Yunnan University Herbarium (PYU). Choice of Polystichum outgroups was guided by previous phylogenetic analyses (Little and Barrington 2003; Li et al., 2008) and included representation both from the sections of
Polystichum most closely related to Xiphopolystichum, as well as from genera most closely related to Polystichum. A complete list of taxa used in the study including voucher information and GenBank accession numbers is provided in Appendix B.

FLOW CYTOMETRY
Flow cytometry analysis was conducted for 11 accessions in order to estimate ploidy. Each accession included samples of each taxon in Xiphopolystichum s.s. (Appendix B). Approximately one to three grams of tissue of each accession were dried in silica in the field and later stored at 10°C for two to seven months before preparation for flow cytometry. Tissue preparation for flow cytometry followed the protocol of Bainard et al., (2011) with some modifications. Some coarse tissues were subject to an incubation period in digestive enzymes according to Naill and Roberts (2005) prior to staining. For several accessions, two replicates were run on the flow cytometer for each accession on different days, depending on availability of field-collected tissue. Samples were analyzed using a Coulter Epics XL flow cytometer (Beckman Coulter Genomics), equipped with a UV lamp located in the Plant Biology Dept., University of Vermont. Histograms were analyzed using FlowPy (http://flowpy.wikidot.com) software. Fresh tissue of Pisum sativum, cultivated in the laboratory, was used as a genome-size standard, with a known 2C value of 4.88 pg (Bennett and Smith, 1976). For accessions with peaks overlapping with Pisum sativum, findings were cross validated using Hordeum vulgare var. morex, with a 2C genome size of 11.1 pg (Bennet and Smith, 1976). All sample G1 peaks were adjacent but not overlapping with G1 peaks of the standard, allowing for comparison and estimation of sample 2C values using the formula, Sample 2C value (DNA pg) = Reference 2C value x (sample 2C mean peak position/reference 2C mean peak position). The 2C (pg) values of
named *Xiphopolystichum* species were expected to be multiples of the 1C (pg) value of known diploid *Polystichum acrostichoides*, 7.75 pg (Bainard et al., 2011). DNA ploidies were therefore estimated by comparison to these hypothesized values.

**SPORE MEASUREMENTS**

For each species, sporangia and spores from one to two specimens (details, Appendix B) were mounted in Hoyer’s medium on glass slides and imaged at 100x using a compound microscope. The sporangia and spores were imaged in order to count spores per sporangium (to assess reproductive biology) as well as measure the length of spores (as an indication of ploidy). One to two sporangium per specimen was counted: plants with spore counts of 32 or fewer were inferred to be apomictic; those with more than 32 per sporangium were inferred to be sexual. Twenty to 30 spores per specimen were measured to calculate mean length and standard deviation for each species. Spore length was measured from the images using ImageJ (Schneider et al., 2012). The external spore membrane (*perispore*), which is pronounced among the species examined, was excluded from measurements.

**DNA EXTRACTION, AMPLIFICATION, AND SEQUENCING**

Total genomic DNA was extracted from fresh (1 g) or silica-dried (0.5 g) leaves using a cetyl trimethyl-ammonium bromide (CTAB) procedure (Porebski et al., 1997) with some modifications. Leaves were ground with a bead-beating machine using glass beads. CTAB buffer was supplemented with polyvinyl pyrrolidone (PVP), and crushed leaf tissue was precipitated in chloroform. Samples were then subjected to washes in 70% and 90% ethanol and re-suspended in Tris-EDTA buffer. The plastid DNA sequences *rbcL*, *trnL-F* spacer region, and *trnS-rps4* spacer region were PCR amplified under standard
conditions using previously published primers, with modifications for *rbcL* (Taberlet et al., 1991; Little and Barrington 2003; McHenry and Barrington, 2014), as were the nuclear DNA sequences *gapC* (exons 8–11) and *PgiC* (exons 14–16) (Table 4), modified from Schuettpelz et al., (2008), Koenemann et al. (2011), and Lyons et al. (2017). The marker trnS-rps4 was amplified using an initial denaturation step of 94 °C for 3 min, followed by 30 cycles of 94 °C for 45 s, 53 °C for 45 s, and 72 °C for 2 min, and a final extension at 72 °C for 10 min. Amplification of *rbcL*, trnL-F, and the nuclear markers followed similar conditions except that the initial and final steps were done for 10 min, and the cycles were done 35 times with 58 °C annealing for 1 min (plastid) or 55 °C for 30 s (nuclear) and extension at 72 °C for 1 min. Resulting PCR products were cleaned using ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA), and sequenced on an ABI PRISM 3730x automated sequencer (Beckman Coulter Genomics, Danvers, MA, USA). Each plastid and nuclear region was sequenced in both the forward and reverse direction using the amplification primers.
<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>trnL-F</td>
<td>F 5’GGTTCAAGTCCCTCTATCCC’ R 5’ATTTGAACCTGGTGACACGAGn’</td>
<td>Taberlet et al. (1991)</td>
</tr>
<tr>
<td>rps4-trnS</td>
<td>F 5’TTACCGAGGGTTCGAATCCCTC3’ R 5’GAGTATTACTCCCGCAAG3’</td>
<td>McHenry and Barrington (2014)</td>
</tr>
<tr>
<td>rbcL</td>
<td>F 5’TTCATGCGTTGGAGAGATC3’ R 5’GGACTCCACTCCWGCTTC3’</td>
<td>Little and Barrington (2003)</td>
</tr>
<tr>
<td>gapCp</td>
<td>F 5’CCAAGTTCAACTGGTGCT3’ R 5’TGCTWCATCTGCAAGACACC3’</td>
<td>Schuettpelz et al. 2008</td>
</tr>
<tr>
<td>pgIC</td>
<td>F 5’T GTCCTCCTCACATTCAAC 3’ R5’CT TAGTATGGAAAGCAATGGAAG3’</td>
<td>Lyons et al. 2017; Koenemann et al. 2011</td>
</tr>
</tbody>
</table>

**Table 4.** Primers modified from named references
SEQUENCE ANALYSIS
Consensus sequences were generated from assemblies of forward and reverse reads, and aligned using MUSCLE (Edgar, 2004) with minor manual adjustments. Assembly, consensus generation, and alignment were implemented in Geneious version 9.0 (Kearse et al., 2012); for species represented by multiple accessions, consensus sequences were generated for phylogenetic analysis. Indels were coded simply as single characters with binary states (simple gap coding; Simmons and Ochoterena, 2000). The resulting indel data were appended to the end of sequences for use in all subsequent model testing and phylogenetic inference analysis. Initially, data sets for each marker were aligned and analyzed individually using Bayesian Inference (BI) approaches. The tree topologies generated from the individual analyses were inspected for discordance among topologies. Few differences were observed among topologies; consequently, plastid sequences and nuclear sequences were each concatenated into one nuclear and one plastid data set for further analysis.

Bayesian Inference was applied to both the nuclear and plastid datasets using MrBayes version 3.2.6 (Ronquist et al., 2012) on the CIPRES Science Gateway server (Miller et al., 2010). For BI analysis of plastid sequences, the alignment was partitioned by markers and optimal evolutionary models discerned from jModeltest 2 (Guindon et al., 2003; Darriba et al., 2012) using the Akaike Information Criterion (AIC). The first 500,000 trees were discarded as burn-in iterations, the remainder were used to generate a 50% majority-rule consensus tree. Posterior probabilities were obtained from MrBayes, and the phylogenetic tree including branch lengths was visualized using FigTree Version 1.4 (http://tree.bio.ed.ac.uk/software/figtree/). For BI analysis of nuclear sequences the same
procedure implemented for plastid sequences was used with two modifications. MrBayes was run for ten million generations; the first 400,000 trees were discarded as burn-in. ML analysis was implemented in RaxML version 8.2.9 on the CIPRES Science Gateway server (Miller et al. 2010). 1000 search replicates were used with random starting trees. The data were partitioned by marker and the analysis was implemented using a single model discerned from jModeltest (Guindon et al. 2003; Darriba et al. 2012) using the Akaike Information Criterion (AIC).

MIXED NUCLEOTIDE SIGNALS AND RETICULATION
The aligned chromatograms of the original trimmed forward and reverse sequences for the nuclear markers *gapCp* and *pgiC* were examined for multiple nucleotide peaks at each position. Following Tate et al., (2006), Jorgensen and Barrington (2017), and Lyons et al., (2017), these multiple peaks were taken as evidence of different allelic variants retrieved for the marker in question, summed in the single chromatogram generated from direct Sanger sequencing.

MORPHOLOGICAL ANALYSIS
For all 42 accessions representing the breadth of morphological variation in *Xiphopolystichum*, including the 20 for which ploidy was estimated using flow cytometry, five morphological characters were scored. A principal component analysis (PCA) was then conducted using the R package *gg biplot* (http://ggplot2.tidyverse.org/). Each combination of pairs, PC1, PC2, and PC3 was plotted for interpretation. The PCA was used as a heuristic tool for determining the most important morphological characters in distinguishing *Xiphopolystichum* species.
RESULTS

FLOW CYTOMETRY REVEALS MULTIPLE PLOIDIES
Flow cytometry analysis of accessions of *Xiphopolystichum* revealed only diploid accessions of *Polystichum revolutum* and *P. herbaceum*, and only triploid accessions of *P. mayebarae*, *P. sinotsus-simens*, and *P. pseudoxiphophyllum*. On the other hand, *Polystichum tsus-simens* accessions included both triploid and diploid genome sizes and *P. xiphophyllum* accessions included both triploid and tetraploid genome sizes. For each of these named species with multiple ploidies, the more common DNA ploidy was triploid (Figure 7).

SPORE DATA CORROBORATE FLOW CYTOMETRY PLOIDY ESTIMATES
Spore counts and measurements corroborated findings from flow cytometry. Spore counts for named species *Polystichum xiphophyllum*, *P. tsus-simens*, *P. mayebarae*, *P. sinotsus-simens*, and *P. pseudoxiphophyllum* consistently revealed between 20 and 32 spores per sporangium, suggesting Döpp-Manton apomictic reproduction. Spore counts of *Polystichum revolutum* and *Polystichum herbaceum* each yielded more than 45 spores per sporangium, suggesting sexual reproduction. These two species had mean spore lengths of approximately 30 µm, which in the light of the flow-cytometry data we take to be a typical diploid spore length for the section. The mean spore length of most apomictic species was about 40 µm, consistent with their being triploid, as inferred from the flow-cytometry data. However, some accessions of *P. xiphophyllum* had a higher mean spore length, and some accessions of *P. tsus-simens* have a lower mean spore length. These accessions include
those with atypical results in the flow-cytometry data. One *P. xiphophyllum* accession with a larger genome size than other accessions of the species had larger spores; One *P. tsus-simense* accession with a lower genome size than other accessions of the same species had smaller spores (Figure 7).

**PLASTID AND NUCLEAR DATASETS SUPPORT TWO CLADES**

Alignments of chromatographic sequences for *gapCp* and *PgiC* in our dataset revealed seven and nine sites with two nucleotide calls, respectively (Table 5). Within the *Xiphopolystichum* s.s. clade, chromatograms of nuclear sequences with multiple nucleotide calls were retrieved from four species, *Polystichum sinotsus-simense*, *P. xiphophylum* (both 3x and 4x), *P. mayebarae*, and *P. pseudoxiphophyllum*. All of the two-nucleotide calls included two nucleotides common at that same site in other sequences. For the majority of sites with double-nucleotide calls, the signal combines nucleotides shared with sexual diploids *Polystichum revolutum* in Clade A and *P. herbaceum* in Clade B (Figures 8 and 9). However, the relative strength of nucleotide signals combined at these sites varied among the apomictic polyploids *P. xiphophyllum* (3x and 4x), *P. mayebarae*, *P. sinotsus-simense*, and *P. pseudoxiphophyllum*. The diploid species *Polystichum revolutum* and *P. herbaceum* did not have mixed nucleotide signals, nor did triploid apomict *P. tsus-simense*. 
Figure 7. Mean spore length (A) for each species in *Xiphopolystichum* s.s. and cytotype; bars indicate 95% confidence intervals (CI). Mean genome size (B) for each species and cytotype; bars indicate 95% CI.
Polystichum mayebarae, and both cytotypes of P. xiphophyllum belong to Clade A (Figures 1 and 2), along with sexual diploid P. revolutum. All of the two-nucleotide calls in the polyploid apomicts P. mayebarae and P. xiphophyllum (3x and 4x) share one nucleotide with P. revolutum. At most of these sites for Polystichum mayebarae and P. xiphophyllum (3x and 4x), the higher peak is shared with Polystichum revolutum, and the lower peak is shared with P. herbaceum and P. tsus-simense in Clade B. For Polystichum pseudoxiphophyllum, resolved in clade A in the plastid phylogeny and clade B in the nuclear phylogeny, the stronger peak is shared with Polystichum herbaceum and the weaker peak is shared with P. revolutum. Polystichum sinotsus-simense belongs to Clade B along with P. tsus-simense and P. herbaceum. Most of the two-nucleotide calls in the polyploid apomict Polystichum sinotsus-simense share one nucleotide with P. herbaceum. At most of these sites, the higher peak is shared with Polystichum herbaceum.

MORPHOLOGICAL ANALYSIS
PCA based on five leaf measurements revealed that the first two principal components accounted for 74% of the variance across species. A plot of these two PC axes revealed largely overlapping clusters representing P. herbaceum, P. pseudoxiphophyllum, Polystichum revolutum, P. sinotsus-simense, P. tsus-simense, P. xiphophyllum, and. The diploid race of P. tsus-simense and the tetraploid race of P. xiphophyllum are each represented by only two accessions and hence are not represented by clusters. The most important character in defining clusters on the first principal component was number of pinnules on the second pinna (a proxy for level of leaf dissection), whereas clusters were best resolved by the ratio of the length of first and second pinna (a proxy for overall frond shape) on the second principal component. The level of leaf dissection in Xiphopolystichum
s.s. ranges from once-pinnate in *Polystichum revolutum* to fully twice-pinnate in *P. herbaceum* (Figure 10). All other species and cytotypes recognized in the group are intermediate in level of dissection (Figure 11). The mean ratio of length of the basalmost pinna to the second basalmost pinna is highest in *P. revolutum* and lowest in *P. herbaceum*. All other species in the group have a ratio either identical to the two diploids, or intermediate. The value of this ratio is generally lower for species in Clade B (Figures 8 and 9) than in Clade A.
Table 5. Base positions in an alignment of chromatograms of *gapCp* (A) and *PgIC* (B) with multiple peaks. For species with double peaks in the chromatograms, the nucleotide giving a stronger signal is shown in bold.

<table>
<thead>
<tr>
<th></th>
<th>106</th>
<th>127</th>
<th>213</th>
<th>323</th>
<th>331</th>
<th>337</th>
<th>404</th>
<th>467</th>
<th>787</th>
</tr>
</thead>
<tbody>
<tr>
<td>revolutum</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>xiphophyllus</td>
<td>C/T</td>
<td>A/C</td>
<td>C/T</td>
<td>C/T</td>
<td>A/C</td>
<td>G/A</td>
<td>G/T</td>
<td>T/A</td>
<td>T/C</td>
</tr>
<tr>
<td>xiphophyllus</td>
<td>C/T</td>
<td>A/C</td>
<td>C/T</td>
<td>C/T</td>
<td>A/C</td>
<td>G/A</td>
<td>G/T</td>
<td>T/A</td>
<td>T/C</td>
</tr>
<tr>
<td>pseudoxiphi-</td>
<td>A/C</td>
<td>C/T</td>
<td>T/C</td>
<td>A/C</td>
<td>G/A</td>
<td>G/T</td>
<td>T/A</td>
<td>T/C</td>
<td></td>
</tr>
<tr>
<td>mayebaraee</td>
<td>C/T</td>
<td>A/C</td>
<td>C/T</td>
<td>C/T</td>
<td>A/C</td>
<td>C/T</td>
<td>G/T</td>
<td>T/A</td>
<td>T/C</td>
</tr>
<tr>
<td>sinotsussimins</td>
<td>C</td>
<td>C/T</td>
<td>T/C</td>
<td>C/A</td>
<td>G/A</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>tsussimens-</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>herbaceum</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>
Figure 8. Phylogeny of *Polystichum* section *Xiphopolystichum* based on the combined analysis of nuclear markers plastid markers, *trnL*-F, *rps4*-trnS, and *rbcL*. *Xiphopolystichum* s.s.is comprised of clades A and B. The tree is the 50% majority rule Bayesian Inference (BI) phylogram of *Polystichum* section *Xiphopolystichum* based on the combined analysis of BI Posterior Probability/maximum likelihood support values are given at each node supported by both analyses.
Figure 9. Phylogeny of *Polystichum* section *Xiphopolystichum* based on the combined analysis of nuclear markers, *gapCp* and *PgiC*. The tree is the 50% majority-rule phylogram from the Bayesian Inference (BI) analysis. BI Posterior Probability/Maximum Likelihood support values are given at each node supported by both analyses.
Figure 10. Pinna morphology for species resolved in *Xiphopolystichum* sensu stricto, as well as the tetraploid cytotype of *Polystichum xiphophyllum* and diploid cytotype of *P. tsus-simense*, which are morphologically distinct from the more common triploid cyotypes. Letters underneath each pinna diagram indicate the inferred genomic composition.
Figure 11. (a) The number of pinnules on the acroscopic side of the second basalmost pinna for each Xiphopolystichum taxon (a quantitative proxy for level of dissection). Inferred ploidy levels as in figures 4 and 5. (b) The ratio of the first to the second pinna length for each taxon.
DISCUSSION

Considering evidence from phylogeny, ploidy, reproductive mode, and the genome signatures in each hybrid taxon, we have discerned new apomictic cytotypes and established the ways in which each of the two sexual diploid genomes have contributed to evolution of the known apomictic taxa in *Polystichum* section *Xiphopolystichum*.

HYBRIDIZATION HISTORY OF APOMICTS

CLADE A.

Based on both nuclear and plastid phylogenies, considered in the light of ploidy levels, *Xiphopolystichum* s.s.comprises two clades, both of which include one sexual diploid and multiple apomicts. In clade A of the plastid phylogeny (Figure 8), nucleotide polymorphisms for nuclear genes of apomictic triploid and tetraploid *Polystichum xiphophyllum*, triploid *P. mayebarae*, and triploid *P. pseudoxiphophyllum* indicate that they are allopolyploids derived from the clade A diploid *P. revolutum* and clade B diploid *P. herbaceum* (Figure 12). Furthermore, relative peak heights in the chromatograms are consistent with triploid *P. xiphophyllum* and triploid *P. mayebarae* having incorporated two genomes from *P. revolutum* and one from *P. herbaceum*. We also propose that our newly discovered tetraploid cytotype of *P. xiphophyllum* incorporates three *P. revolutum* genomes with one *P. herbaceum* genome. One plausible scenario for the evolution of tetraploid *P. xiphophyllum* is one or more hybridization events between apomictic triploid *P. xiphophyllum* and sexual *P. revolutum* (Figure 12). Conversely, relative peak heights in
chromatograms for triploid apomict *Polystichum pseudoxiphophyllum* suggest that it has inherited one genome from *P. revolutum* and two genomes from *P. herbaceum.*
Figure 12. Proposed scenarios yielding *Xiphopolystichum* s.s. taxa. *R* represents a constituent genome inherited from *Polystichum revolutum*. *H* represents a constituent genome inherited from *Polystichum herbaceum*. Solid lines represent a meiotically reduced genomic contribution; dotted lines linking named species represent an unreduced genomic contribution; Apomicts are in triangles, while sexual species are in squares. Ploidies are indicated on the left side of the figure as 2x (diploid), 3x (triploid), or 4x (tetraploid).
CLADE B
Unlike clade A apomicts, only one allelic form of the nuclear markers was found for triploid apomict *Polystichum tsus-simense* and diploid apomict *Polystichum tsus-simense* in clade B (Figure 8); in both cases the sequence is identical to *P. herbaceum*. Based on these data, we suggest that triploid *Polystichum tsus-simense* is an autopolyploid derived from *P. herbaceum*, whereas diploid *P. tsus-simense* is derived from the triploid apomict via a loss of chromosomal material, as reported for *Osmunda* apomicts by Manton (1950).

Understanding the heritage of triploid apomict *Polystichum sinotsus-simense* presents a greater challenge than the rest. Like apomicts in clade A, this taxon carries nucleotide polymorphisms at some positions exclusive to the *Polystichum revolutum* and *P. herbaceum* genomes, with relative chromatograph peaks in the exons suggesting a larger genomic contribution from the latter parent (Figure 12). However, unlike the remaining apomicts with signals of hybrid origin, nucleotides in the introns are generally homozygous for the *Polystichum herbaceum* allele, possibly caused by back mutations due to higher mutation rates in introns relative to exons, fixation due to natural selection or drift, or deletions/translocation events that are often easily tolerated by polyploids (Shubert and Lysak, 2011). Because the genomic heritage of *P. sinotsus-simense* has proven to be more difficult to understand using Sanger sequencing approach, the lineage is a good candidate for further investigation using next-generation sequencing approaches.

The species pairs *Polystichum mayebarae–P. xiphophyllum* and *P. pseudoxiphophyllum–P. sinotsussimense* present the whole array of problems complicating effective evolutionary categorization. Notably, each pair results from a hybridization event,
but with each lineage arising from the same progenitors having the same reproductive anomaly and the same balance of contributed genomes. Further, in spite of sharing a pair of progenitors with each parent contributing the same proportion of genetic material to these triploid pairs, they are morphologically distinct from each other. Hence these two species pairs present a case of multiple hybrid origins, and one in which the use of multiple species criteria are required to resolve the two lineages as separate species.

**PARENTAL GENOME DOSAGE EFFECTS**

We hypothesized that the leaf division of hybrid apomicts would be intermediate between those of their progenitors based on an additive genetic model; indeed we found this pattern for clade A *Polystichum xiphophyllum* (both 3x and 4x), *P. mayebarae*, *P. sinotsussimense*, and *P. pseudoxiphophyllum*. Level of dissection and the ratio of the length of basalmost pinna to second to basalmost pinna loaded most strongly on the first and second PCA axes. Also consistent with this hypothesis was leaf dissection in *P. tsus-simense*, which nuclear-sequence data suggest is an autopolyploid derived from *P. herbaceum*; *P. tsus-simense* has a similar level of dissection to *P. herbaceum*. The level of dissection among clade B species *Polystichum tsus-simense*, *P. herbaceum*, and *P. sinotsus-simense* was generally greater than that of species in clade A, with the exception of *P. mayebarae*. This finding is in line with our genetic data revealing that the twice-pinnate *P. herbaceum* makes the greatest or only genomic contribution to lineages in clade B (Figure 12). On the other hand, the less-dissected species in clade A have a stronger genetic signal indicating contribution from the once-pinnate *P. revolutum*.

Morphological intermediacy in hybrids is well documented in ferns (Barrington, 1986). All of the allopolyploid apomictic ferns in *Xiphopolystichum* are intermediate in leaf
division between their proposed progenitors. However, the hypothesized genomic composition of *Polystichum mayebarae* in clade A would predict a lower level of dissection than that observed. There are several scenarios that may independently, or in concert, account for this finding. Potentially, the unexpected morphology of *Polystichum mayebarae* could be accounted for by an unsampled twice-pinnate progenitor closely related to the lineages comprising clade A. Alternatively, the unexpectedly high level of leaf division in *Polystichum mayebarae* could be explained by non-additive (i.e. transgressive) behavior of genes regulating leaf dissection (McDade, 1990; Rieseberg, 1995; Hegarty and Hiscock, 2004; Soltis and Soltis, 2009). Finally, reciprocal hybridization should be considered in understanding morphological differences between *P. mayebarae* and *P. xiphophyllum*. Often hybrids with multiple origins have both progenitors serving as the maternal or paternal parent (Stein and Barrington, 1990; Vogel et al., 1998; Sigel et al., 2014; Jorgensen and Barrington, 2017). If *P. xiphophyllum* and *P. mayebarae* do in fact have different maternal and paternal parents among their hypothesized progenitors, *P. revolutum* and *P. herbaceum*, then we would expect each to behave differently in the nuclear and chloroplast phylogenies. In the chloroplast phylogeny, we find that *P. xiphophyllum* is more closely related to *P. revolutum* than *P. mayebarae*, suggesting that the two may have different maternal progenitors. *Polystichum pseudoxiphophyllum* and *Polystichum sino-tsussimense* present a similar problem in that although they have the same genome composition, *Polystichum revolutum* is likely the maternal progenitor of *Polystichum pseudoxiphophyllum*, and the paternal progenitor of *Polystichum sino-tsussimense*, particularly given differences in the resolution of *Polystichum pseudoxiphophyllum* in the nuclear and plastid phylogenies (Figures 8 and 9).
SPECIES CONCEPTS
The named taxa in our *Xiphopolystichum* s.s. study set stand as species when tested using the phylogenetic, phenetic, and evolutionary criteria together. Although events in the evolution of lineages constituting *Xiphopolystichum* and their sequelae (hybridization, multiple origins, and the concomitant morphological complexity) are confounding to the goal of categorizing any lineage in this group as a species, some characteristics of *Xiphopolystichum* apomicts and their relatives are stable and can be used reliably to distinguish evolutionary lineages from each other. Each named species in *Xiphopolystichum* s.s. has one or more distinct alleles for each genetic marker, both plastid and nuclear, sampled in the present study. Accordingly, they are phylogenetically distinct. Similarly, ploidy and reproductive mode (sexual or apomictic) are stable characters that help to delineate groups of individuals as distinct lineages in *Xiphopolystichum*. For instance, all of the individuals sampled that are morphologically identifiable as *Polystichum revolutum* are sexual diploids, as evidenced by flow cytometry and spore counts. Some lineages included in the present study are morphologically indistinct from closely related lineages, but are phylogenetically distinct. *Polystichum mayebarae* is phylogenetically very similar to *P. xiphophyllum*, and we hypothesize that they are products of hybridization between the same progenitors. However, *Polystichum mayebarae* is morphologically distinct and identifiable. Similarly, *Polystichum herbaceum*, *P. sinotsus-simense*, and *P. tsus-simense* are phylogenetically unresolved (Figures 8 and 9), but are morphologically distinct (Figure 10). We found that new apomictic cytotypes of *Polystichum xiphophyllum* and *P. tsus-simense* are morphologically and genetically highly similar to the more common triploid apomictic cytotypes of each.
Although these cytotypes are single evolutionary lineages that maintain their integrity, we have chosen not to designate the various cytotypes of *P. tsus-simense* and *P. xiphophyllum* as distinct species. Recent phylogenetic and systematic works on East Asian ferns involving apomictic complexes have taken a similar approach to ours, combining molecular and morphological data to define more biologically realistic taxonomic groups (Chang et al., 2013; Chen et al., 2014; Hori et al., 2014). In these East Asian apomictic complexes, cytotypes differing in ploidy and reproduction, but that are morphologically and phylogenetically indistinguishable from one another, are included in single species as multiple variants or cytotypes.

**CONCLUSIONS**

Each of our findings on delineation of species in *Polystichum* sect. *Xiphopolystichum* is relevant to understanding species diversity of ferns in China. Species diversity is dependent on how species are defined and, in ferns, these definitions are complicated by the likes of apomixis, polyploidy, and hybridization. Currently, the *Flora of China* includes 34 species in *Polystichum* sect. *Xiphopolystichum* s.l. (Zhang and Barrington, 2013). In large part, we find that the diversity of the plants we sampled for this project is best represented taxonomically as it is in the Flora of China. At least two other apomictic lineages, *P. neolobatum* and *P. rigens*, exist in the broader terrain of *Xiphopolystichum* s.l. Given the potential of apomictic lineages to either inflate or underrepresent species diversity, it will be important to understand these lineages and their potentially reticulate relationships to other species.
LITERATURE CITED

Barrington, D.S., 1986. The morphology and cytology of *Polystichum x potteri* hybr. nov.(= *P. acrostichoides* x *P. braunii*). Rhodora 297–313.
Cracraft, J., 1990. The origin of evolutionary novelties: pattern and process at different hierarchical levels. Evolutionary innovations 21–44.
Fraser-Jenkins, C.R., 2007a The species and subspecies in the Dryopteris affinis group. Fern Gazette 18, 1.


CHAPTER 4: APOMIXIS AND BIOGEOGRAPHY IN *POLYSTICHUM* SECTION *DUROPOLYSTICHUM* IN THE HIMALAYA AND HENGDUAN MOUNTAINS

ABSTRACT

Orogeny in the Himalayan-Hengduan region has been critical in the evolution of numerous plant groups. Mountain building in East Asia shapes both the geologic landscape and the climate regimes of the region. Historical patterns of shifting landscapes may have facilitated secondary contact among previously isolated species, allowing for hybridization. Among ferns in particular, hybridization is a critical phenomenon in diversification, particularly in terms of its interaction with polyploidy and apomixis. The interaction of apomixis with hybridization and polyploidy can lead to complex patterns of reticulation, complicating efforts to reconstruct evolutionary history in groups where apomixis is common. We investigated the East Asian ferns in *Polystichum* section *Duropolystichum* in order to elucidate the evolutionary history of known apomicts in the group, *P. rigens* and *P. neolobatum*, as well as understand the relationship of species in the group with past and current ecological conditions. The datasets comprised phylogenetic inference based on three plastid markers, analysis of mixed nucleotide signals from chromatograms, ploidy estimates based on spore length, morphometric analysis of representative specimens, and niche modeling of sampled species. By interpreting these multiple lines of evidence synthetically we conclude that several species in the group warrant further scrutiny for consideration as species, and that *P. neolobatum* is an allopolyploid that likely comprises of multiple lineages with a high level of morphological
convergence. Our findings highlight the importance of understanding the role of apomictic reproduction in speciation as well as the complexity introduced by recent orogeny.

INTRODUCTION

Plant diversity is often richest in montane regions; both the Andes and the Himalaya are regarded as global biodiversity hotspots. Montane regions foster plant species diversity in a number of ways that influence biological processes. Orogeny, or mountain building, can facilitate vicariance or dispersal events and therefore speciation, not only by acting as a constantly shifting physical barrier or bridge between regions, but also by influencing climatic patterns (Simpson, 1975; Luebert, 2009; Antonelli et al., 2009; McHenry and Barrington, 2014). The Andes, for instance, are thought to serve as bridges between plant populations in Northern and Southern portions of South America, while dividing plants lying to the East and West of the mountain range (Hayes, 2004; Hoorn, 2010). The resulting effects of mountain ranges on climate, particularly rainfall, and therefore differently adapted plant species, are well documented (Troll, 1968; Luebert and Weigend, 2014), with the tropical Andes home to approximately 10,000 plant taxa, 30% of which are endemic (Rafiqpoor, et al. 2005). The rise of the Qinghai Tibetan Plateau (QTP), and the Himalaya mountains, too has facilitated plant diversification in Southeast Asia (Liu et al., 2006; Huang et al., 2012). Like the Andes, the rise of the Himalaya mountains has altered the climate of surrounding regions, mainly by impacting rainfall regimes (Zheng et al., 2000; Yao et al., 2011), by forming arid desert regions in central Asia and wet tropical zones in southern Asia (Zhisheng et al., 2015). Importantly, the Himalaya and Hengduan mountains themselves are incredibly ecologically heterogeneous environments, simply by virtue of
dramatic elevation and concomitant rainfall and temperature gradients, thereby facilitating colonization and speciation in a relatively narrow geographic range (Vetaas and Grytnes, 2002). The wet Eastern Himalaya alone is home to an estimated 9,000 plant species, of which 39% are endemic (Rafiqpoor et al., 2005).

While geologically dynamic regions like the Andes and the Himalaya facilitate speciation and diversity by offering an incredibly heterogeneous and constantly changing physical landscape, allowing for vicariance and dispersal, they also foster speciation through secondary contact, leading to hybridization (Zhou et al., 2017). The Himalaya and neighboring Hengduan Mountains have been subject to significant climatic fluctuations during the Quaternary, including alternating glacial and interglacial periods (Owen, 2008). Many plant species, as a result, have had constantly shifting distributions allowing for secondary contact among related species that previously experienced long intervening periods of geographic and ecological niche isolation (Hewitt, 2000).

Hybrids can either enhance or diminish genetic diversity depending on the nature of the species involved (Barrington et al., 1990). Although hybrids can be sterile, in plants, polyploidization and asexual reproduction can often overcome this sterility and allow hybrid lineages to persist as “separately evolving metapopulation lineages” (de Quieroz, 2007) or species (Pala and Coelho, 2005; Schranz, 2005; Robertson, 2010). For ferns in particular, asexual reproduction, polyploidy, and hybridization are important forces in evolution; it is estimated that 31% of fern speciation events have a concomitant ploidy level change (Otto and Whitton, 2000; Soltis et al., 2004). Up to 95% of fern species are polyploid (Grant, 1981), via one of two possible paths: chromosomal doubling (autopolyploidy) or hybridization followed by doubling of chromosomes
(allopolyploidy—Manton, 1950; Stebbins, 1950). Many fern polyploids, including allopolyploids, are capable of sexual reproduction. However, both auto- and allopolyploids may have genomes that are considered “unbalanced,” meaning that a genome from at least one parent is represented with an odd number of copies. One means of overcoming sterility imposed by an unbalanced genome is apomixis (Consendai et al., 2011). Although most apomicts are unbalanced triploid lineages (Wagner and Wagner, 1980; Asker and Jerling, 1992), some autotetraploid apomictic ferns are known, such as *Pellaea glabella* Mett. ex Kuhn and *P. occidentalis* Rydb. (Gastony, 1988). Additionally, some apomictic diploids have been identified, including *Dryopteris wallichiana* (Spreng.) Hyl., *Cheilanthes leucopoda* Link, and *Pteris cretica* L. (Verma and Khullar, 1965; Knobloch, 1967; Fraser-Jenkins, 2007). In ferns, apomixis is a form of asexual reproduction in which spores are still produced via meiosis, thereby maintaining the benefits of dispersal, though fertilization is bypassed. There is a strong association between polyploidy, both auto- and allopolyploid, and apomixis; nearly three quarters of apomictic ferns are triploid (Wagner and Wagner, 1980; Asker and Jerling, 1992).

Southeast Asia is home to the greatest apomictic fern species richness in the world (Liu et al. 2012), and the majority of these lineages are thought to have originated during and after major climatic shifts resulting in the development of the monsoon, or heterogeneous rainfall regimes across Asia (Liu et al., 200; Horandl, 2009; Wen et al., 2014). A number of apomictic fern species with origins in the Himalaya are in fact part of hybrid complexes that include numerous cytotypes and clusters, confounding attempts at taxonomic categorization. The *Pteris cretica* L. complex in the Western Himalaya, for instance, includes diploid and triploid apomictic cytotypes that are morphologically
indistinguishable at the macro level (Verma and Khullar, 1965), and more recent studies of the group in Asia suggest that multiple apomictic lineages have arisen from a series of recent hybridizations with closely related taxa, yielding nearly continuous morphologies among lineages in the complex (Jaruwattanaphan, 2013). Similarly, the *Lepisorus clathratus* Ching complex, distributed across the QTP and in the Hengduan Mountains, includes multiple apomictic cytotypes, with evidence of hybrid origins and continuous morphological variation among haplotypes in the group (Wang et al., 2012). In spite of the continuous or nearly continuous morphological variation in these complexes, the ecological heterogeneity of the QTP region can help to elucidate evolutionary history of individual lineages by defining species niches (Poudel et al., 2012). Some morphologically complex reticulate networks of ferns native to the QTP and Hengduan mountains include lineages distinguishable by their ecological niches, which are often defined by elevation in particular (Wang et al., 2011).

The third largest fern genus in the world, *Polystichum* Roth, has its center of diversity in East Asia (Zhang and Barrington, 2013; Le Pechon, 2016), and is relatively apomict rich, particularly in the Himalaya and the Hengduan mountains (Liu et al., 2012). The genus includes thirteen sections; current infrageneric classifications are largely based on morphology (Zhang and Barrington, 2013; LePechon, 2016). However, estimates of the number of species in the group vary widely ranging from 200 to 500 (Barrington, 1990; Mabberly, 1997; Le Pechon et al., 2016). Some of the difficulty in estimating *Polystichum* diversity is attributable to ongoing species discovery in the Himalaya and montane Western China (Zhang and He, 2009; He and Zhang, 2011). These species are often considered cryptic in that they are morphologically difficult to distinguish from closely related taxa,
which may be attributable to significant polyploidy, hybridization and apomixis (Wagner, 1973; Barrington, 1990; Little and Barrington, 2003; McHenry and Barrington, 2014).

*Polystichum* section *Xiphopolystichum* Daigobo *sensu lato* (s.l.) is the most apomict-rich section in the genus; it includes 34 species distributed largely in Western China, the Western Himalaya, and Nepal, with some species native also to Japan and Bhutan, and one species endemic to Hawaii. This section includes six known apomictic species, four of which belong to the *Xiphopolystichum sensu stricto* (s.s.) clade, and are part of a complex network of reticulation resulting from repeated hybridization events among two diploid species (Patel et al., 2017). Species in *Xiphopolystichum* outside the s.s. clade were previously circumscribed on the basis of morphology, as section *Duropolystichum* Fraser-Jenkins (previously *Scleropolystichum* Daigobo in the classification of Kung (2001); the group includes two known apomictic species, *Polystichum neolobatum* Nakai and *P. rigens* Tagawa. *Duropolystichum* has proven a difficult group to understand, both in a morphological and molecular-phylogenetic context. While *Xiphopolystichum* s.s. is monophyletic, *Duropolystichum* has so far only been defined as paraphyletic in a molecular phylogenetic context (Le Pechon et al., 2016; Patel et al., 2017). Morphologically, *Duropolystichum* has some synapomorphic characters, including thick, leathery lamina, spinules on the pinna margins, and large, brown ovate or lanceolate scales spread along the rachis (Kung et al., 2001; Zhang and Barrington, 2013; Le Pechon, 2016). However, within the group, nearly continuous morphological variation has led to significant taxonomic confusion. Species in *Duropolystichum* are routinely misidentified, and synonymy of certain species is disputed among experts in various treatments of the section (Fraser-Jenkins, 1985; 1991; 1997; Zhang and Barrington, 2013).
In particular, the known apomict *P. neolobatum* is highly morphologically variable (Fraser-Jenkins, 1997). Study of *Duropolystichum* using a molecular phylogenetic and morphometric approach offers an opportunity to understand apomixis, hybridization, and polyploidy as they relate to speciation and evolutionary radiation in a global biodiversity hotspot, the Himalaya.

Here we analyze both plastid and nuclear sequences to resolve relationships among species currently circumscribed in the group in its most recent treatment (Zhang and Barrington, 2013) and to test for reticulate evolution. The molecular analysis is grounded in a morphometric analysis to define morphological boundaries across the named species in the section. Molecular dating and niche modeling is utilized to offer historical geologic and current climatic context.

Based on strong evidence from our previous work, documenting reticulation in *Xiphopolystichum* s.s., we proceeded with three hypotheses about *Duropolystichum* in mind:

1. The apomictic species *Polystichum rigens* and *P. neolobatum* are polyploid apomicts.

2. *Polystichum neolobatum* likely has multiple origins given significant intraspecific morphological variation.

3. Highly morphologically similar species, such as *Polystichum acanthophyllum* (Franchet) Christ, *P, cyclolobum* C. Christensen, and *P. rhomboideum* Ching, are in fact single species.
4. Diversification of *Duropolystichum* coincides with the time frame of the most recent Himalayan and Hengduan uplift events and concomitant intensification of the East Asian monsoon, occurring approximately 5 million years ago (mya).

METHODS

TAXON SAMPLING AND SPECIES DELIMITATION

We assembled a representative sample of species in *Xiphopolystichum* s.l. as circumscribed in Zhang and Barrington (2013). Collections were made during two trips by the authors to Western China in 2006 and 2015, and one trip to Nepal in 2005. Samples were augmented by herbarium specimens by various collectors (Appendix C), all belonging to collections at the Pringle Herbarium (VT) or the Yunnan University Herbarium (PYU). These tissues were used for subsequent extraction, amplification, and sequencing for phylogenetic analysis. The phylogenetic analysis was augmented by existing sequences for species representative of the genus *Polystichum*, taken from GenBank (Appendix A). Sampling for phylogenetic analysis included 48 accessions representing 28 species within *Polystichum*, 37 accessions representing 17 species within *Xiphopolystichum* s.l., and 31 accessions representing 11 species within *Duropolystichum* (Appendix C). For morphometric and niche modeling analyses the study set was augmented by digital vouchers from the GBIF (gbif.org) database (Appendix C).
SPORE MEASUREMENTS

For each species, sporangia and spores from one to two specimens (Appendix C) were mounted in Hoyer’s medium on glass slides and imaged at 100x using a compound microscope. The sporangia and spores were imaged in order to count spores per sporangium (to assess reproductive biology) and to measure the length of the spores (as an indication of ploidy). One to two sporangia per specimen was counted: plants with spore counts of 32 or fewer were inferred to be apomictic; those with more than 32 per sporangium were inferred to be sexual. Twenty to 30 spores per specimen were measured to calculate mean length and standard deviation for each species. Spore length was measured from the images using ImageJ (Schneider et al., 2012). The external spore membrane (perispore), which is pronounced among the species examined, was excluded from measurements.

DNA EXTRACTION, AMPLIFICATION, AND SEQUENCING

Total genomic DNA was extracted from fresh (1 g) or silica-dried (0.5 g) leaves using a cetyl trimethyl-ammonium bromide (CTAB) procedure (Doyle and Doyle, 1987) with some modifications. Leaves were ground with a bead-beating machine using glass beads. CTAB buffer was supplemented with polyvinyl pyrrolidone (PVP), and crushed leaf tissue was precipitated in chloroform. Samples were then subjected to washes in 70% and 90% ethanol and re-suspended in Tris-EDTA buffer. The plastid DNA sequences rbcL, trnL-F spacer region, and trnS-rps4 spacer region were PCR amplified under standard conditions using previously published primers, with some modifications (Table 6). The marker trnS-rps4 was amplified using an initial denaturation step of 94 °C for 3 min,
followed by 30 cycles of 94 °C for 45 s, 53 °C for 45 s, and 72 °C for 2 min, and a final extension at 72 °C for 10 min. Amplification of rbcL and trnL-F followed similar conditions except that the initial and final steps were done for 10 min, and the cycles were done 35 times with 58 °C annealing for 1 min (plastid) or 55 °C for 30 s, and extension at 72 °C for 1 min. The nuclear marker ApPEFP_C was amplified under standard conditions using PCR conditions taken from Rothfels et al. (2017). Resulting PCR products were cleaned using ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA), and sequenced on an ABI PRISM 3730x automated sequencer (Beckman Coulter Genomics, Danvers, MA, USA). Each plastid and nuclear region was sequenced in both the forward and reverse direction using the amplification primers.

SEQUENCE ANALYSIS

All three plastid markers, rbcL, trnL-F spacer region, and trnS-rps4, were used in phylogenetic analysis. Consensus sequences were generated from assemblies of forward and reverse reads of all three markers, and aligned using MUSCLE (Edgar, 2004) with minor manual adjustments. Assembly, consensus generation, and alignment were implemented in Geneious version 9.0 (Kearse et al., 2012); for species represented by multiple accessions, consensus sequences were generated for phylogenetic analysis. Indels were coded simply as single characters with binary states (simple gap coding; Simmons and Ochoterena, 2000). The resulting indel data were appended to the end of sequences for use in all subsequent phylogenetic inference analysis.

Bayesian Inference was applied to the plastid dataset using MrBayes version 3.2.6 (Ronquist et al., 2012) on the CIPRES Science Gateway server (Miller et al., 2010).
For BI analysis of plastid sequences, the alignment was partitioned by markers and optimal evolutionary models discerned from jModeltest 2 (Darriba et al., 2012) using the Akaike Information Criterion (AIC). MrBayes was run for 6 million generations with trees sampled every 1000 generations. The first 500,000 trees were discarded as burn-in; the remainder were used to generate a 50% majority-rule consensus tree. Posterior probabilities were obtained from MrBayes, and the phylogenetic tree including branch lengths was visualized using FigTree Version 1.4 (Rambaut and Drummond, 2008).
<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>trnL-F</em></td>
<td>F 5’GGTTCAAGTCCCTCTATCCC’ R 5’ATTTGAACCTGGTACACGAGn’</td>
<td>Taberlet et al. (1991)</td>
</tr>
<tr>
<td><em>rps4-trnS</em></td>
<td>F 5’TACCGAGGGTTCAATCCCTC3’ R 5’GAGTATTACTCCCGAAAG3’</td>
<td>McHenry and Barrington (2014)</td>
</tr>
<tr>
<td><em>rbcL</em></td>
<td>F 5’TTCATGCGTTGGAGAGGATC3’ R 5’GGACTCCACTTACWAGCTTC3’</td>
<td>Little and Barrington (2003)</td>
</tr>
<tr>
<td><em>ApPEFP_C</em></td>
<td>F 5’GGACCTGGSCTYGCTGARGAGT3’ R 5’GCAACRTGAGCAGCYGGTTTCRGGG3’</td>
<td>Rothfels et al. (2017)</td>
</tr>
</tbody>
</table>

Table 6. Primers used, modified from given references.
BEAST ANALYSIS

Divergence times were estimated via a Bayesian approach using the program BEAST version 2.4.7 (Drummond and Rambaut, 2007) with the relaxed phylogenetic method of Drummond et al. (2006). Data were partitioned by plastid region, and the best fit model as determined by jModeltest2 was applied to each partition (Darriba et al., 2012). Fossils are not available for Polystichum, so secondary time calibrations were applied to the plastid dataset in BEAUti version 2.4.6. The most recent common ancestor of outgroup Arachniodes denticulata is estimated to have diverged no less than 67.8 mya, and the most recent common ancestor of Polystichum is estimated to have diverged no less than 30.8 mya (Schuettpelz and Pryer, 2009). A relaxed lognormal clock was applied to the node constraints. A birth-death speciation prior was used with a gamma model of rate variation. The analysis was run for 10 million generations with sampling every 1000 generations. Log files were inspected in Tracer v1.5 (Rambaut et al., 2014) to ensure an appropriate level of sampling. Trees were summarized in TreeAnnotator v1.6.2 (Drummond and Rambaut, 2007). Trees with node age estimates were visualized in figtree (Rambaut and Drummond, 2008) with the 95% highest posterior density (HPD) intervals.

MIXED NUCLEOTIDE SIGNALS

The aligned chromatograms of the original trimmed forward and reverse sequences for the nuclear marker ApPEFP_C were examined for multiple nucleotide peaks at each position. Following Tate et al., (2006), Jorgensen and Barrington (2017), and Lyons et al., (2017), these multiple peaks were taken as evidence of different allelic variants retrieved
for the marker in question, summed in the single chromatogram generated from direct Sanger sequencing.

MORPHOMETRIC ANALYSIS

We measured seven quantitative morphological characters for 70 accessions representing 11 species in *Duropolytichum*, as well as 10 morphological characters for 40 accessions of *Polystichum neolobatum* (Tables S1 and S2). Each dataset was analyzed independently to characterize morphology among *Duropolytichum* as well as among accessions of *P. neolobatum*, which exhibit significant intraspecific morphological variation. These characters were measured using scanned images within the software ImageJ (Schneider et al., 2012). Each dataset was analyzed using a pipeline modified from Cadena et al. (2017). For each dataset, we used the R package clustvarsel (Fraley et al., 2014) to select the set of principal components most useful for defining groups without *a priori* group definitions. We used the R package McCluster (Scrucca et al., 2016) to find the best fit normal mixture model (NMM) based on Bayesian Information Criterion (BIC). For the full *Duropolytichum* dataset, we implemented the model fitting in McCluster in two iterations. One iteration assumes a minimum of one group, to allow for the possibility that there is no clustering, and a maximum of 70 groups, to allow for the possibility that each sampled accession is an individual group. The second iteration allows for a minimum of one group and a maximum of 11 groups, to allow for the possibility that groupings will correspond to the number of taxonomic species. The groups defined by McCluster were plotted against the current taxonomic treatment for the sampled species of *Duropolytichum*. For the *P. neolobatum* dataset we implemented model fitting in
McCluster in one iteration, assuming a minimum of one group, to allow for the possibility that there is are morphological definition into clusters, and a maximum of 40 groups, to allow for the possibility that each sampled accession is best supported as an individual group.

ECOLOGICAL NICHE MODELING AND BIOGEOGRAPHY

Decimal GPS coordinates for each sample used in phylogenetic analysis were combined with coordinates of digital vouchers (Appendix A) downloaded from The Global Biodiversity Information Facility (GBIF.org). In order to approximate climatic conditions, we used data layers representing all 19 bioclimatic variables from the WorldClim database at a 30 arcsec resolution (WorldClim 1.4, [www.worldclim.org](http://www.worldclim.org), Hijmans et al., 2005). In addition, we included the variables ‘Solar Radiation,’ and ‘Altitude,’ from the WorldClim database (WorldClim 2, [www.worldclim.org](http://www.worldclim.org)). All 21 variables were extracted for a region within the longitudinal range 70 °E to 150 °E, and the latitudinal range 15 °E to 60 °E. We used Pearson’s correlation coefficient to detect high levels of autocorrelation among our chosen variables, which led to our eliminating variables with coefficients higher than 0.75.

Using variables selected after testing for collinearity, ecological niche models (ENM) for the 11 sampled species in *Duropolystichum* as well as for the clusters of *Polystichum neolobatum* identified were constructed using Maxent v. 3.3.3 implemented in R v3.4 (Philips et al., 2004). In order to optimize models created for each species, the R package ‘ENMeval’ v0.1.1 ([(http://cran.r-project.org/web/packages/ENMeval; Muscarella et al., 2014)](http://cran.r-project.org/web/packages/ENMeval; Muscarella et al., 2014)) was used to adjust two settings, the regularization multiplier (RM) and the feature class (FC). The feature classes used were L (‘linear’), H (‘hinge’, i.e.
modelling piecewise linear responses to the environmental variables), LQ (‘linear–quadratic’) and LQH (a combination of all three). The regularization multipliers used were 1, 2, and 3. Twelve possible models result. The mean ORMin (omission rate based on the minimum training presence logistic threshold) and AUC were calculated using ENMeval. The model settings that minimized the ORMin and maximized the AUC were used in subsequent analyses. All analyses were conducted using the a priori partition ‘checkerboard1’ in ENMeval in order to account for potential spatial autocorrelation in our distribution data. The ‘permutation importance’ values in the Maxent output were used to evaluate the most important variables for the optimal model for each species.

We evaluated niche divergence between selected species pairs using the metric Schoener’s D (Schoener, 1968), as determined by the R packages ‘ENMTools’ (Warren et al., 2010), with a value of 0 representing no niche overlap and a value of 1 representing full niche overlap. In addition, we plotted the geographic distribution of each of intraspecific morphological cluster of Polystichum neolobatum in two iterations: once using only occurrences sampled in the present study for phylogenetic analysis, and once using the full dataset of occurrences including records taken from GBIF, and which therefore lack accompanying molecular phylogenetic data.

RESULTS

SPORE DATA SUGGEST TRIPLOID AND DIPLOIDS

Spore counts for named species Polystichum neolobatum, P. rigens, and P. cyclolobum consistently revealed between 21 and 32 spores per sporangium, suggesting
Döpp-Manton apomictic reproduction. The mean spore length for each accession of each species ranged from 38 to 42, suggesting that they are triploids (Patel et al., 2017). Spore counts for all accessions of species *P. acanthophyllum*, *P. mehrae* Fraser-Jenkins & Khullar, *P. stimulans* Kunze ex Mettenius, *P. squarrosum* Don, *P. rhomboideum*, *P. hillebrandii* Carruth, and *P. integripinnulum* Ching each yielded more than 48 spores per sporangium, suggesting sexual reproduction. The mean spore length for each accessions of each of these species ranged from 30 to 36, suggesting that these accessions are diploid (Table 7).

**PHYLOGENETIC ANALYSIS**

In the phylogeny constructed from plastid markers, *Duropolystichum* and *Xiphopolystichum* are resolved as two well-supported clades (Figure 13). *Duropolystichum* includes five well-supported clades. A-E (Figure 13). Posterior probabilities are low for most clades more recently divergent than the most recent common ancestors of each lettered clade. Sequences suggest that accessions resolved within each clade have highly similar plastid genomes. In addition, there is low support for uniting clades A, B, and C as a monophyletic group.

Clade A includes accessions representing species *Polystichum acanthophyllum*, *P. mehrae*, and *P. squarrosum* (PP=1). (Zhang and Barrington, 2013). Clade B includes *P. meiguense*, *P. stimulans*, *P. cyclolobum*, and *P. rhomboideum* (PP=1). These species are small, with once-pinnate to once-pinnate pinnatifid fronds ranging from 12-30 cm in length (Zhang and Barrington, 2013).
Clade C (PP=.99) includes accessions representing *Polystichum neolobatum*, a triploid apomict, and *P. hillebrandii*. *Polystichum hillebrandii* is a Hawaiian endemic (Driscoll and Barrington, 2007; Zhang and Barrington, 2013). Clade D (PP=1) comprises the remaining *P. neolobatum* accessions as well as *P. integripinnulum*. Clade E (PP=1) is comprised exclusively of accessions representing *P. rigens*, also a triploid apomict.
Table 7. Spore counts as well as mean length of spores for each sampled species of *Duropolystichum*, with confidence intervals given in parentheses. Estimated ploidy and reproductive mode based on spore lengths and counts, respectively, are also given.

<table>
<thead>
<tr>
<th>Species</th>
<th>Spore Count</th>
<th>Mean Spore Length (µm)</th>
<th>Estimated Ploidy</th>
<th>Reproductive Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. neolobatum</td>
<td>29</td>
<td>43.7 (1.2)</td>
<td>3x</td>
<td>Apomictic</td>
</tr>
<tr>
<td>P. rigens</td>
<td>30</td>
<td>41.6 (1)</td>
<td>3x</td>
<td>Apomictic</td>
</tr>
<tr>
<td>P. cyclolobum</td>
<td>46</td>
<td>43.2 (2.2)</td>
<td>3x</td>
<td>Apomictic</td>
</tr>
<tr>
<td>P. stimulans</td>
<td>56</td>
<td>32.1 (2.8)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
<tr>
<td>P. mehrae</td>
<td>55</td>
<td>30.1 (1.1)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
<tr>
<td>P. acaanthophyllum</td>
<td>55</td>
<td>36.1 (.9)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
<tr>
<td>P. rhomboideum</td>
<td>60</td>
<td>29.2 (1.9)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
<tr>
<td>P. squarrosom</td>
<td>58</td>
<td>33.7 (3.3)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
<tr>
<td>P. hillebrandii</td>
<td>61</td>
<td>32.4 (.5)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
<tr>
<td>P. integripinnulum</td>
<td>56</td>
<td>30.06 (1.6)</td>
<td>2x</td>
<td>Sexual</td>
</tr>
</tbody>
</table>
Figure 13. 50% majority rule Bayesian Inference phylogram of *Polystichum* section *Xiphopolystichum* s.l. based on the combined analysis of plastid markers, *trnL*-F, *rps4-trnS*, and *rbcL*. 
MIXED NUCLEOTIDE SIGNALS AND RETICULATION

Alignments of chromatographic sequences for ApPEFP_C in our dataset revealed six sites with two nucleotide calls (Table 8). Within the Duropolystichum clade, chromatograms of nuclear sequences with multiple nucleotide calls were retrieved only from Polystichum neolobatum. All of the two-nucleotide calls included two nucleotides common at that same site in other sequences. For the majority of sites with double-nucleotide calls, the signal combines one nucleotide shared with P. acanthophyllum, P. stimulans, P. integripinnulum, P. cyclolobum, and P. squarrosum. The second nucleotide signal at these six sites is shared with P. hillebrandii and P. rigens (Table 8). However, the relative strength of nucleotide signals combined at these sites varied among the accessions of P. neolobatum.

Accessions of Polystichum neolobatum with a stronger nucleotide signal shared with Polystichum hillebrandii and P. rigens resolve in Clade C. Accessions of P. neolobatum with stronger signals shared with P. acanthophyllum, P. stimulans, P. integripinnulum, P. cyclolobum, P. squarrosum resolve in both clade D and clade C. This suggests that the multiple origins of P. neolobatum may result from hybridization of multiple pairs of progenitors.

MORPHOMETRIC ANALYSIS SUGGESTS TWO GROUPS

Both iterations of the McCluster analysis of the full Duropolystichum dataset, one allowing for between one and 11 groups, and the second defining between one and 70 groups, yielded the same result. In each case the division of all accessions into two groups...
was best supported, with BIC values of -698.1 and -699.0 for each analysis, respectively. Accessibility representing taxa *Polystichum acanthophyllum*, *P. mehrae*, *P. cyclolobum*, *P. meiguense*, *P. stimulans*, and *P. rhomboideum* all fall squarely into group 2, whereas *P. squarrosum*, *P. hillebrandii*, *P. integripinnulum*, and *P. rigens* belong only to group 1 (Figure 14). In contrast, *P. neolobatum* has represented in both groups, consistent with its being found in two major clades of the *Duropolystichum* cpDNA phylogeny (Figure 13). The variables with highest loading in principal components used to define clusters are frond length and the ratio of the length of the basalmost pinna to the second basalmost pinna. Hence the overall frond size and shape best defined groups.
Table 8. Base positions in an alignment of chromatograms of ApPEFP_C with multiple peaks. For species with double peaks in the chromatograms, the nucleotide giving a stronger signal is shown in bold. Where some accessions reveal a different pattern than others, the species is listed twice to show both patterns.
Figure 14. McCluster morphometric analysis for *Duropolystichum* dataset. Bar graphs show number of specimens with the given taxon name that belong to each McCluster defined group.
In the follow up analysis of only *Polystichum neolobatum*, three groups were best supported—with BIC scores of -791 and -790, respectively. The variables with the highest loading in the principal components used to define these clusters are pinnule shape and level of dissection. Cluster 1 is defined by a low level of dissection and rounded pinnules, cluster 2 is defined by intermediate dissection and rhombic pinnules, and cluster 3 is defined by a high level of dissection and rhombic pinnules.

**BIOGEOGRAPHY AND ECOLOGICAL NICHE MODELING**

Plotting the distribution of each of the three morphological clusters of *Polystichum neolobatum* using the full dataset, including GBIF digital vouchers not sampled phylogenetically, reveals a geographic divide between cluster 1 and clusters 2 and 3. Cluster 1 occurs in the Himalaya, whereas clusters 2 and 3 occur within the Hengduan mountains and somewhat into Yunnan and Central China (Figure 15). Plotting the three clusters using only occurrences for accessions sampled for the present study with molecular phylogenetic data reveals a similar pattern. In terms of ecological niche modeling, the highest degree of niche overlap was between accessions of types two and three, with a Schoener’s D value of 0.8 (Table 9). The environmental variable with the highest permutation importance in characterizing niche models for *P. neolobatum* cluster 1 was rainfall periodicity; for clusters two and three it was elevation (Table 9).
Figure 15. Distribution of three *P. neolobatum* morphotypes. Blue points represent cluster 1, red points represent cluster 2, and purple points represent cluster 3. Distribution of proposed progenitors *Polystichum stimulans, P. squarrosum, P. rigens, P. integripinnulum*. Elevation in meters is shown and labeled according to color, along with latitude and longitude.
<table>
<thead>
<tr>
<th></th>
<th>neolobatum1</th>
<th>neolobatum2</th>
<th>neolobatum3</th>
<th>acanthophyllum</th>
<th>stimulans</th>
<th>mehrae</th>
<th>cyclolobum</th>
<th>squarrosum</th>
<th>hillebrandii</th>
<th>rigens</th>
<th>rhomboideum</th>
</tr>
</thead>
<tbody>
<tr>
<td>neolobatum1</td>
<td>0.68</td>
<td>0.5</td>
<td>0.78</td>
<td><strong>0.88</strong></td>
<td>0.77</td>
<td>0.5</td>
<td>0.4</td>
<td>0.8</td>
<td>0.83</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>neolobatum2</td>
<td>0.8</td>
<td>0.6</td>
<td>0.53</td>
<td>0.89</td>
<td>0.5</td>
<td>0.8</td>
<td><strong>0.95</strong></td>
<td>0.4</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>neolobatum3</td>
<td>0.57</td>
<td>0.46</td>
<td>0.77</td>
<td>0.6</td>
<td><strong>0.9</strong></td>
<td>0.7</td>
<td>0.66</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acanthophyllum</td>
<td></td>
<td></td>
<td></td>
<td>0.66</td>
<td>0.81</td>
<td>0.49</td>
<td>0.84</td>
<td>0.51</td>
<td>0.33</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>stimulans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77</td>
<td>0.94</td>
<td>0.46</td>
<td>0.23</td>
<td>0.67</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>mehrae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td>0.2</td>
<td>0.43</td>
<td>0.77</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>cyclolobum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.11</td>
<td>0.49</td>
<td>0.49</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>squarrosum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.91</td>
<td>0.2</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>hillebrandii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.91</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>rigens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>0.53</strong></td>
</tr>
<tr>
<td>rhomboideum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 9.** Schoener’s D values comparing *P. neolobatum* clusters given in column one, with each other morphotype as well as hypothesized progenitors given in row one. The hypothesized progenitors with which *Polystichum neolobatum* clusters 1, 2, and 3 have the highest level of niche overlap have Schoener’s D values shown in bold.
Polystichum rigens and P. integripinnulum are all geographically distributed in the Hengduan mountains, overlapping significantly with P. neolobatum clusters 2 and 3. Polystichum squarrosum and P. stimulans occur in both the Hengduan and Himalaya mountains, overlapping with P. neolobatum clusters 1, 2, and 3. In terms of ecological niche modeling, Polystichum neolobatum cluster 1 has the highest level of niche overlap with P. stimulans (Schoener’s D = 0.88), P. neolobatum cluster two has the highest level of niche overlap with P. hillebrandii (Schoener’s D = 0.95), and P. neolobatum cluster 3 has the highest level of niche overlap with P. squarrosum (Schoener’s D = 0.9) (Table 9).

DIVERSION TIME ESTIMATES

The dated cpDNA phylogeny suggests a divergence time for the origin Xiphopolystichum s.l. in the mid Miocene, with a mean estimate of 10.99 mya (Figure 16). The next event is the divergence of Duropolystichum from Xiphopolystichum s.s., with a mean estimate of 9.88 mya. Within Duropolystichum the ancestor of clades A, B, and C, and of clades D and E, are estimated to be 9.3 mya and 7.1 mya, respectively. The origins of Xiphopolystichum s.s. and Duropolystichum coincide with major uplift events in the QTP (Wang et al., 2009) Diversification of extant lineages within each lettered clade is estimated to occur in the Pliocene or Pleistocene, between 3.4 mya and 0.5 mya, suggesting a correlation between speciation in Duropolystichum and Hengduan uplift as well as the most recent intensification of the East Asian monsoon as well as a period of intense climatic fluctuation in terms of temperature and precipitation (Wang et al., 2011).
Figure 16. Time calibrated phylogeny for *Duropolystichum*. Estimated divergence time is given at each node. Major geologic events occurring in the Qinghai Tibetan Plateau are labeled below the chronogram, and major climatic events are labeled above the chronogram.
DISCUSSION

SPECIES DELIMITATION EVOLUTIONARY RELATIONSHIPS

The morphological similarities between Duropolystichum sensu Kung et al. (2001) and Xiphopolystichum s.s. have recently led to the two being included in the larger section Xiphopolystichum s.l. (Zhang and Barrington 2013). Additionally, the most recent phylogenetic work in Polystichum finds Duropolystichum paraphyletic relative to Xiphopolystichum (LePechon et al., 2017). Here we resolve Duropolystichum as a monophyletic group on the basis of three plastid markers (Figure 13). Xiphopolystichum s.s. is a monophyletic clade.

In our inquiry into diversity and evolution in Duropolystichum, we have used a species concept that incorporates phylogeny, morphology, reproduction, and ecology as incorporated in the general lineage concept of de Queiroz (1998, 1999). Under his General Lineage Concept, a species is considered a segment of a lineage, in that a species is temporally separated from its ancestors and descendants on an evolutionary time scale. Of central importance to us, this more unified species concept treats other species concepts as criteria for species delimitation, all relevant to delimiting species under the general lineage concept. Each species in an apomictic complex should be defined, at a minimum, by one accepted species criterion.
PHYLOGENETIC AND MORPHOLOGICAL SPECIES CRITERIA

Although there are five well-supported sub-clades of *Duropolystichum* resolved in the plastid phylogeny, McCluster analysis defined only two morphological groups. A lack of phylogenetic resolution and nearly continuous morphological variation is not uncommon in taxonomic groups where apomixis, polyploidy, and hybridization are prevalent—especially in groups native or endemic to montane regions like the Himalaya and the Andes (Hori et al., 2014; Hughes and Atchison, 2015).

Accordingly, in *Duropolystichum*, taxa resolved in the same sub-clade, hence possessing a highly similar or identical plastid genome, as well as in the same McCluster-defined morphological group, warrant further scrutiny for consideration as species. For example, *Polystichum rhomboideum, P. cyclolobum, P. stimulans*, and *P. meiguense* belong to clade B (Figure 13) and both belong to the same McCluster defined group (Figure 14). Similarly, *P. mehrae* and *P. acanthophyllum*, both resolved in clade A (Figure 13), both belong to the same McCluster-defined group (Figure 14), and both appear to be diploid and sexual (Table 7). In contrast, although *P. squarrosum* is also resolved in clade A, it is morphologically distinct from *P. acanthophyllum* and *P. mehrae* and belongs to a different McCluster-defined morphological group (Figure 3). *Polystichum squarrosum* has strongly rhombic pinnules, elongate pinnae, and is 2-3 times larger than *P. acanthophyllum* and *P. mehrae* on average.

*Polystichum neolobatum* presents a challenge in defining species using its resolution in the plastid phylogeny coupled with morphology. Accessions of *P. neolobatum* resolve in
both clades C and D (Figure 13), and each clade includes accessions representing multiple morphological clusters of *P. neolobatum* (Figure 17). In first considering plastid phylogenetic relationships, resolution of *P. neolobatum* in clades C and D suggests at least two independent contributors of plastid genomes to lineages defined as *P. neolobatum*. Accessions of *P. neolobatum* in clade C have a plastid genome highly similar to *P. hillebrandii* (Figure 13), a sexual diploid endemic to Hawaii. *Polystichum hillebrandii* or an ancestor occupying Asia prior to its dispersal (Driscoll and Barrington 2003), is a likely progenitor to *P. neolobatum* in clade C (Figure 17). *Polystichum integripinnulum* is a hypothesized progenitor to accessions of *P. neolobatum* in clade D, on the basis of similarity in plastid genomes. In terms of morphological clusters, clade D includes representatives of *P. neolobatum* clusters 2 and 3, and clade C includes representatives of all three clusters. This suggests that cluster 2 and 3 each includes more than one lineage that is convergently morphologically similar, meaning that different progenitors have contributed to lineages belonging to each morphological cluster.

Based on the morphologies of each cluster and the characters most important in defining them, we posit three hypotheses as to the species acting as the second progenitors of lineages in each *P. neolobatum* cluster (Figure 17), in addition to *P. hillebrandii* and *P. integripinnulum* which are inferred as progenitors from the plastid phylogeny (Figure 1).

1) *Polystichum stimulans* in cpDNA clade B, small with rounded pinnules, is a reasonable hypothetical progenitor to lineages of *P. neolobatum* cluster 1.

2) Cluster 2 of *Polystichum neolobatum* in clade D and C has the same genomic composition including *P. integripinnulum* and *P. hillebrandii* as progenitors, but given their divergent positions in the plastid phylogeny, the reciprocal parentage.
3) Cluster 3 of *Polystichum neolobatum* in both clades D and C has morphology consistent with either *P. rigens* or *P. squarrosum* as a paternal contributor (Figure 17).
Figure 17. Hypothesized reticulate relationship among named species in *Duropolystichum*. Letters indicate genome dosage. Male and female progenitors are indicated with Linnaean symbols. Clades are colored according to clade color designations in Figure 1.
ALLELIC VARIATION CRITERION

Polystichum neolobatum is the only lineage with mixed nucleotide signals for nuclear marker ApPEFP_C (Table 8). Apparent homozygosity in the other identified apomicts, P. rigens and P. cyclolobum, suggest that they are autopolyploids. Other species in Duropolysichum sampled are diploid and sexual, and hence simply homozygous at the sampled locus.

Patterns of heterozygosity in Polystichum neolobatum lend support to multiple hybrid origins corroborating findings from the plastid phylogenetic relationships (Figure 1) and intraspecific morphological variation (Figure 17). Considering summed nucleotides in chromatographic sequence data; most accessions of P. neolobatum within clade C have a higher chromatographic nucleotide signal matching that in members of clades A, B, and D. These clades include P. squarrosum, P. stimulans, and P. integripinnulum, which are P. neolobatum progenitors hypothesized on the basis of morphology from each of these clades, respectively. Only one accession of P. neolobatum in clade C has nucleotide signals consistent with stronger genomic contribution from progenitor P. hillebrandii, and therefore a lower genomic contribution from its other progenitor, hypothesized to be P. squarrosum on the basis of morphology (Figure 17).

For Polystichum neolobatum accessions in Clade D, mixed nucleotide signals suggest a higher genome dosage from P. integripinnulum, one hypothesized progenitor (Figure 13). Nucleotide signals are also consistent with P. hillebrandii or P. rigens contributing one-third of the triploid genome of P. neolobatum accessions in clade D, as hypothesized on the basis of morphology (Figure 17).
DISTRIBUTION AND ECOLOGICAL NICHE

The majority of species in Duropolystichum are distributed across both the Himalaya and Hengduan mountains (Zhang and Barrington 2013), with some exceptions: Polystichum rigens, and P. integripinnulum are recorded only in the Hengduan and farther east into central China (Figure 15). However, most species in Duropolystichum with strongly overlapping geographic distributions surrounding the QTP occupy distinct ecological niches defined either by elevation or patterns of precipitation (Tables 9 and 10).

Polystichum neolobatum cluster 1 appears to be a distinct lineage with a single origin that is not morphologically convergent with other lineages (Figure 17). Its identity as a discrete lineage is further supported by its exclusive occurrence in the Himalaya, geographically isolated from clusters 2 and 3 in the Hengduan mountains (Figure 15). In addition, seasonality of precipitation has the highest permutation importance for P. neolobatum cluster 1, whereas the ecological niches of clusters 2 and 3 are most defined by elevation. Polystichum neolobatum clusters 2 and 3 exhibit a greater degree of niche overlap with each other than either exhibits with cluster 1 (Table 17). Hence, ecological differences between these clusters also lend support to our hypothesis that P. neolobatum cluster 1 has an evolutionary history unique from clusters 2 and 3.

However, efforts to use niche modeling as a tool to understand reticulation are complicated in Duropolystichum by the fact that Polystichum neolobatum cluster 3 likely includes at least three lineages that are morphologically convergent. Recorded occurrences of P. neolobatum cluster 3 occupy a niche very similar to P. squarrosum, perhaps suggesting that most collected accessions of this cluster originate with P. hillebrandii and P. squarrosum as progenitors.
The geographic distribution and ecological niche of proposed progenitors to *Polystichum neolobatum* also lend support to the proposed reticulate history (Figure 17). Each cluster has the highest level of niche overlap, quantified with Schoener’s D, with at least one of its proposed progenitors (Table 17). In addition, *P. squarrosum*, *P. rigens*, and *P. integripinnulum*, proposed progenitors of *P. neolobatum* clusters 2 and 3 in clade C, are also distributed largely in the Hengduan.

**SPECIES UNDER THE GENERAL LINEAGE CONCEPT**

Species in *Duropolystichum* that are phylogenetically, morphologically, reproductively, cytologically, biogeographically, and ecologically similar should be subject to further scrutiny as distinct species. *Polystichum cyclolobum* and *P. rhomboideum*, *P. stimulans*, and *P. meiguense* are highly phylogenetically (Figure 13 and Table 8) and morphologically (Figure 14) similar. However, they are geographically distinct in that *Polystichum stimulans* occupies a much broader geographic region than the other three. In addition, they are ecologically distinct (Table 9), and *Polystichum cyclolobum* is an apomictic triploid (Table 7), whereas the other species are sexual diploids. Synonymy of *P. rhomboideum* and *P. cyclolobum* has been proposed before (Fraser-Jenkins 1997). Further investigation of these two taxa with denser sampling is required to determine if the two lineages are cytotypes of a single species.

*Polystichum rigens*, an autopolyploid triploid apomict, appears have a single origin (Figure 13 and Table 7). It is morphologically similar to *Polystichum squarrosum* and *P. integripinnulum*, two sexual diploids, and hence may share an origin with a progenitor of these species also distributed in the Hengduan Mountains.
<table>
<thead>
<tr>
<th>Species</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. neolobatum</em>1</td>
<td>Seasonality of Precipitation</td>
</tr>
<tr>
<td><em>P. neolobatum</em>2</td>
<td>Elevation</td>
</tr>
<tr>
<td><em>P. neolobatum</em>3</td>
<td>Elevation</td>
</tr>
<tr>
<td><em>P. acanthophyllum</em></td>
<td>Precipitation of the Wettest Quarter</td>
</tr>
<tr>
<td><em>P. stimulans</em></td>
<td>Elevation</td>
</tr>
<tr>
<td><em>P. mehrae</em></td>
<td>Precipitation of the Wettest Quarter</td>
</tr>
<tr>
<td><em>P. cyclolobum</em></td>
<td>Precipitation of the Driest Quarter</td>
</tr>
<tr>
<td><em>P. squarrosum</em></td>
<td>Elevation</td>
</tr>
<tr>
<td><em>P. hillebrandii</em></td>
<td>Precipitation of the Wettest Quarter</td>
</tr>
<tr>
<td><em>P. rigens</em></td>
<td>Temperature of the Wettest Quarter</td>
</tr>
<tr>
<td><em>P. rhomboideum</em></td>
<td>Precipitation of the Wettest Quarter</td>
</tr>
</tbody>
</table>

Table 10. Worldclim variables with highest level of permutation importance in characterizing niches using Maxent for all sampled species in *Duropolystichum*
*Polystichum neolobatum* includes multiple lineages, arising from multiple origins and pairs of progenitors. Accessions in cluster 1 likely arise from the only unique hybridization event, which we hypothesize involves *P. hillebrandii* and *P. stimulans* as progenitors. It’s identity as a unique lineage with origins independent from the other two clusters is further supported by its geographic and ecological isolation from clusters 2 and 3. Although each cluster is a triploid apomict, other criteria for species definition, taken together, strongly suggest that *P. neolobatum* cluster 1 is a unique lineage that should be considered a species. Clusters 2 and 3 each appear to include two lineages that are morphologically convergent and hence require phylogenetic inference for distinction. Cluster 2, however, is morphologically distinguishable from cluster 3, though the two have a similar geographic distribution and ecological niche in the Hengduan mountains. Clusters 2 and 3 hence should each be further divided into two unique lineages.

**DIVERGENCE TIME ESTIMATES**

Orogeny has likely played an important role in the evolution of *Polystichum*, both in the Andes and the Himalaya (Li et al., 2004; McHenry and Barrington, 2013). The estimated time of origin for the genus *Polystichum* is approximately 30 mya (Schuettpelz and Pryer, 2009), coinciding with the first collision of the Indian plate with Eurasia (Zheng et al., 2000). *Duropolystichum* lineages, most of which are native to the Himalaya or Hengduan mountains (Zhang and Barrington, 2013), appear to have evolved in concert with critical geologic and climatic shifts in the QTP and surrounding mountains during the late Miocene (Figure 16), a pattern of evolution common in numerous groups native to the
QTP (Wang et al., 2009; Jabour and Renner, 2012; Favre et al., 2014). The most intense period of uplift in the Himalaya, which likely brought the QTP to an elevation comparable to that of the present day, is estimated to have occurred in the mid Miocene (10 mya) (Axelrod, 1997; Mulch and Chamberlain, 2006), coinciding with the origin of the *Xiphopolystichum* s.l. clade (10.99 mya) (Figure 16).

The most intense phase of uplift in the Hengduan mountains is estimated to have occurred 3.4 mya (Fubin 1992), coinciding with much of the diversification of extant lineages in *Duropolystichum* (Figure 16). Studies of both angiosperm and pteridophyte diversification in the Hengduan mountains have inferred a correlation between evolutionary radiations in the Hengduan mountains and climatic shifts concomitant with mountain building in the region. Particularly, interglacial periods occurring at high elevations may allow for range expansion, while glaciation may facilitate allopatric speciation (Wang et al., 2011; Li et al., 2011).

Repeated hybridization leading to allopolyploid lineages may be linked to historic orogeny and climatic shifts. The divergence time estimates for the accessions representing clusters 2 and 3 range from 1.61 to 1.89 mya (Figure 16). This coincides with a major interglacial period occurring at high elevations in the Hengduan (1.7 mya), which could allow for secondary contact among previously isolated populations.

**CONCLUSIONS**

Reticulate evolution is a critical part of the evolutionary history of numerous fern groups that presents a significant challenge to parsing and identifying lineages of single
origins. This is especially complicated by reticulation in landscapes influenced by recent orogeny, where numerous lineages occupy a narrow geographic range. Considering multiple species criteria, including ecological niche, is critical to understanding these lineages and their origins. The findings here, suggesting that one allopolyploid species, *Polystichum neolobatum*, may actually comprise multiple lineages with unique origins, exemplify the complexity of systematics and taxonomy of ferns in mountains surrounding the QTP. Future systematics based investigations of *Polystichum* in East Asia will benefit from incorporating a consideration of ecological niche in defining lineages.
LITERATURE CITED


GAO, Y.-D., A.J. HARRIS, and X.-J. HE. 2015. Morphological and ecological divergence of Lilium and Nomocharis within the Hengduan Mountains and Qinghai-Tibetan Plateau may result from habitat specialization and hybridization. BMC evolutionary biology 15: 147.


RAMBAUT, A., and A. DRUMMOND. 2008. FigTree: Tree figure drawing tool, version 1.2. 2. Institute of Evolutionary Biology, University of Edinburgh.


COMPREHENSIVE LITERATURE CITED


Barrington, D. S. (1986). The morphology and cytology of *Polystichum x potteri* hybr. nov. (= *P. acrostichoides* x *P. braunii*). *Rhodora, 297*–313.


Gao, Y.-D., Harris, A. J., & He, X.-J. (2015). Morphological and ecological divergence of Lilium and Nomocharis within the Hengduan Mountains and Qinghai-Tibetan Plateau may result from habitat specialization and hybridization. *BMC Evolutionary Biology, 15*(1), 147.


Rambaut, A., & Drummond, A. (2008). FigTree: Tree figure drawing tool, version 1.2. 2. *Institute of Evolutionary Biology, University of Edinburgh.*


APPENDICES

Appendix A: Voucher information for each species given as: Accession number (locality), herbarium code (data sets). Datasets are coded as follows: rps4-trnS – R, psba-trnH-B, PgiC – P, gapcp – G, I -ITS, Spore Length – L, Spore Counts – C. For genbank sequences, genbank accessions numbers are given as (R=XXX)

Phegopteris decursivipinnata: WA614 (Japan), VT (R, B, T, P, I, L, C) Phegopteris connectilis: 305(Vermont), VT (R, B, P, G); 306(Vermont), VT (R, B, P, G, L); 311(France), VT (R, B, P, G, I, L, C); 310(Vermont), VT (R, B, P, G, I, L); 309(Vermont), VT (R, B, I); 313(Vermont), VT (R, B, G, L, C); 307(Vermont), VT (R, B, L); 316(Vermont), VT (R, B, P, L); 308(Vermont) (R, B, L, C); 318(Vermont), VT (R, B); 303(Vermont), VT (R, B, P, I); 304(Vermont), VT (R, B); 315(Vermont), VT (R, B, P, L); 303(Vermont), VT (I, P); 319(France), VT (I); 401(Japan), VT (R, B, P, G, I L, C); 402A(Japan), VT (I, L, C); 403A(Japan), VT (I, L, C) Undescribed tetraploid: 312(Vermont), VT (R, B, P, G, I, L, C); 302(Vermont), VT (R, B, P, G, I, L); 415(Vermont), VT (R, B, I, L, C); 314(Vermont), VT (R, B, P, I); 405(Vermont), (I)

Phegopteris hexagonopteria: 501(Alabama), VT (R, B, P, I, L, C); 502 (Alabama), VT (R, B, P, I, L, C); 500 (Maine), VT (R , B, P, I); 421 (Maine), VT (G, C, P, L); 526 (Vermont), VT (G, L) (Macrothelypteris torresiana: (R=AF425172.1, B= AB575674.1)

Thelypteris viridifrons: (B= AB575679.1) Pseudophegopteris pyrorachis: (R=JX874906) Pseuophegopteris levingei (R = JX874915.1, B = HQ890391.1)

Thelypteris noveboracensis (G= KF553803.1).

**Arachniodes denticulata:** ES AJ 56, VT (P); 12322, VT (R=KX768068); 10457, VT (T); 10551, VT (B)

**Dryopteris goldiana:** 10431-G60, VT (P); 10508, VT (R, B=AF537228.1)

**Cyrtomium yunnanense:** 1922-N17 (T= DQ202418.1) Phanerophlebia nobilis: 10404, VT (R=EU03178, B=, T) Polystichum lepidocaulon: 10534, VT (B=AF537224, T)

**Polystichum lonchitis:** 10413, VT (R, B, T); 12402, Z (G=KX866669.1) Polystichum latilepis: 11909, PYU (R, B, T=DQ202428, P) Polystichum neolobatum: 10523, VT (R, B=AF537252); 11944, PYU (P) Polystichum acanthophyllum: 11900, PYU (R, B, T); 12281, VT (P); 11768, VT (G) Polystichum squarrosum: 11656, VT (R, B=EF177339, T, G) Polystichum cycloclon: 11656DL28, PYU (R, B, T); 12287DL43, PYU (G)

**Polystichum stimulans:** 11907K48, PYU (R, B); 10375, VT (P, G) Polystichum hillebrandii: 10274, VT (R, B, T, G) Polystichum sinotsussimense: 11904, PYU (R, B=KC878857, T=DQ150416, P, G); SY62811, VT (P, G, F); SY62812, VT (L, C)

**Polystichum herbaceum:** 11905, PYU (R, B, T=DQ150405.1, P, G); HX62714, VT (F, L); HX6273, VT (L); JO6287, VT (R) Polystichum rigens: R92, VT (R, P) Polystichum xiphophyllum: 11903, PYU (R=-, B=T=DQ150421); 11728, VT (P, G); DU62311, VT (F, C); 622-10, VT (L); 622EX, VT (L); EM 12617, VT (C) Polystichum xiphophyllum (4x): 62223, VT (P, L, C, F); DU 6237, VT (P, F) Polystichum tsussimense: 11906, PYU (R, B, T=DQ150419); 11794, VT (P, G); 12353, PYU (P); JO6287, VT (R); HX6272, VT (C, L); SH6261, VT (L); DU 623-3, VT (F) Polystichum tsussimense (2X): DU6232, VT (R, P, L, C); JO6283, VT (L, F) Polystichum mayebarae: 11887, VT (B, T=DQ150408.1, F); 11790, VT (P, G); Polystichum pseudoxiphophyllum: 707, CIB (R=KU244858, T=KU244943.1); CQ6259, VT (P, F, C, L) Polystichum revolutum: 11727, VT (R, B, T, P, G); EM1617, VT (P, C, L); EM4617, VT (F); EM6617, VT (F)
Appendix C. Voucher information given as: Accession number, herbarium code (data sets). Datasets are coded as follows: rps4-trnS – R, rbcL- B, trnL-f – T, ApPEFP_C- A, Spore Length – L, Spore Counts – C, Morphometric Analysis- M. For species in which GBIF digital vouchers were used in morphometric analysis, the catalog numbers of the vouchers are given as a list in brackets.