8-2021

Insights on Reticulate Evolution in Ferns, with Special Emphasis on the Genus Ceratopteris

Sylvia P. Kinosian
Utah State University

INSIGHTS ON RETICULATE EVOLUTION IN FERNS, WITH SPECIAL EMPHASIS ON THE GENUS *CERATOPTERIS*

by

Sylvia P. Kinosian

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Ecology

Approved:

Zachariah Gompert, Ph.D. Major Professor	Paul G. Wolf, Ph.D. Committee Member
William D. Pearse, Ph.D. Committee Member	Karen Mock, Ph.D Committee Member
Karen Kaphiem, Ph.D Committee Member	Michael Sundue, Ph.D. Committee Member

D. Richard Cutler, Ph.D.
Interim Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2021

ABSTRACT

Insights on reticulate evolution in ferns, with special emphasis on the genus *Ceratopteris*

by

Sylvia P. Kinosian, Doctor of Philosophy

Utah State University, 2021

Major Professor: Zachariah Gompert, Ph.D.
Department: Biology

The history of life is often viewed as a bifurcating tree; however, in reality it is more like a tangled hedgerow. Many groups of organisms are known to have such a reticulate evolutionary history, but it is particularly common in ferns and lycophytes (also known as pteridophytes). This dissertation investigates the role of polyploidy and hybridization in pteridophyte evolution, with a special emphasis on the model fern *Ceratopteris*. Population genomic analyses are used to understand hybridization across large evolutionary distances in *Dryopteris*; in addition, evolutionary network and hybridization detection analyses infer cryptic species and reticulate evolution in *Ceratopteris*. Finally, a close look is taken at the model fern species *C. richardii*, also known as the C-fern. *Ceratopteris richardii* has been an important part of pteridophyte research since the 1980s, but no wild specimens of the C-fern have been included in an evolutionary context. This work presents the first evidence that *C. richardii* is a hybrid species, or potentially a mix of morphotypes derived from different pairs of species, all under the same name. These findings are important to consider for future research using *C. richardii* and any of its genomic resources.

(197 pages)

PUBLIC ABSTRACT

Insights on reticulate evolution in ferns, with special emphasis on the genus *Ceratopteris*

Sylvia P. Kinosian

The history of life is often viewed as a evenly branching tree; however, in reality it is more like a tangled hedgerow. Many groups of organisms are known to have such a net-like or reticulate evolutionary history, but it is particularly common in ferns and lycophytes (also known as pteridophytes). This dissertation investigates how net-like evolution affects different groups of ferns, with a special emphasis on the model species C-fern (*Ceratopteris richardii*, also called the antler or water sprite fern). Genomic data are utilized to understand hybridization, cryptic species and reticulate evolution in two groups of ferns. The C-fern is shown to be a potential hybrid species, which has important implications for future research using this model organism.

To my grandmothers, Ann Hinman Lilley and Dr. Mary Jane Kinosian.

ACKNOWLEDGMENTS

I am grateful to advisors my Paul Wolf and Will Pearse. Paul was instrumental in the development of my dissertation, and Will provided immeasurable help actualizing each chapter.

I am indebted to Dave Barrington and Cathy Paris for introducing me to the world of ferns. The opportunity to do research as an undergrad in the Barrington Lab changed my career trajectory, and has lead my to some of the most memorable experiences of my life.

Thank you to parents, Margaret and Henry, for introducing me to the natural world, and supporting me unconditionally throughout my academic endeavors.

Thank you to Jacob Suissa for many thoughtful discussions and endless encouragement. I'm looking forward to many more years of collaboration and tropical field work.

To my friends in Logan, I couldn't have done this without you. All of you made life so much more fun with many awesome trips, big and small. Whenever I needed break, an excuse to go outside, or just space to vent, someone was always there.

And finally, I am so grateful to Ian Burbidge, has been the most supportive partner as I figured out how to be a scientist.

Sylvia P. Kinosian

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Evolution is often depicted as a dichotomously branching tree, but in reality the history of life is much more complicated. Lineages split through a variety of evolutionary processes, but can rejoin later on via horizontal gene transfer (Wickell & Li, 2020), hybridization (Adjie, Masuyama, Ishikawa, & Watano, 2007; Nelson et al., 2020), or range expansion (Dauphin, Mossion, Wirth, Sandoz, & Grant, 2017; Pflugbeil, Affenzeller, Tribsch, & Comes, 2021; Zhou et al., 2018). Known as reticulate (net-like) evolution, this has been documented in many groups of organisms (Hipp et al., 2020; Ma et al., 2018; Rieseberg, 1991; Rothfels, 2021; Rutherford, Wilson, Rossetto, & Bonser, 2016; Schley et al., 2020; Sigel, Windham, Huiet, Yatskievych, & Pryer, 2011; Vargas, Ortiz, & Simpson, 2017). There has been a special emphasis on the study of reticulate evolution in spore-dispersed vascular plants (Barrington, 2020; Barrington, Haufler, & Werth, 1989). This group comprises ferns and lycophytes, also known as pteridophytes; it is a paraphyletic group is united by a life history of alternating generations of independent sporophytes and gameotphyes (PPG I, 2016).

The purpose of this dissertation is to build upon previous research on reticulate evolution in pteridophytes using novel methodologies and analyses. In particular, the model fern genus *Ceratopteris* is examined to understand the reticulate species relationships among several species, and the potential hybrid history of the model species *C. richardii*.

1.1 History of reticulate evolution in ferns

The study of reticulate evolution in ferns dates back to the mid-twentieth century, with work on pteridophyte cytology (Manton, 1950; Manton & Sledge, 1954). Variable chromosome numbers (cytotypes) and polyploid hybrid plants are common in ferns (Barrington et al., 1989; Manton, 1950), and it was through their investigation that researchers began to unravel some of the complex relationships among pteridophytes (Conant & Cooper-Driver,

1980; Manton & Sledge, 1954; R. M. Tryon & Tryon, 1982; W. H. Wagner, 1954, 1969, 1983; W. H. Wagner Jr., 1970). In the 1980s and 90s, isozyme banding patterns were used to detect hybrid and polyploid ferns, as well as provide the first genetic evidence for reticulate species complexes (Barrington et al., 1989; Haufler, Windham, & Rabe, 1995; Paris, Wagner, & Wagner, 1989; Werth, Guttman, & Eshbaugh, 1985a, 1985b). Around the turn of the century, Sanger sequencing of plastid and nuclear markers provided a new way to investigate larger species complexes (Grusz, Windham, & Pryer, 2009; Sánchez-Baracaldo, 2004; Schneider et al., 2013; Sigel, Windham, & Pryer, 2014), as well as creating the first phylogenies for ferns as a whole (Pryer et al., 2001, 2004; Smith et al., 2006). More recently, next generation sequencing has once again changed the way that researchers can look at reticulate species complexes. These massive datasets can consist of dozens (to hundreds) of nuclear genes (Breinholt et al., 2021; Burbano et al., 2010; Weitemier et al., 2014; Wolf et al., 2018), or thousands of single nucleotide polymorphisms (SNPs; Andrews, Good, Miller, Luikart, and Hohenlohe (2016); Davey and Blaxter (2010); Wolf et al. (2019)). In addition, there are now several fern genomes available (Li et al., 2018; Marchant et al., 2019), increasing fern genomic resources and potential for future work untangling to complex pteridophyte evolutionary tree.

1.2 Behind the research: people of reticulate evolution in pteridology

All of this work on reticulate evolution in ferns has been driven by dedicated scientists, working to understand the intricacies of the pteridophyte evolutionary tree. Work on reticulate evolution in ferns began with Irene Manton, who pioneered the chromosome "squash" technique, and published chromosome counts for many fern species (Manton, 1950; Preston, 1990). Warren "Herb" Wagner presented the first reticulate scheme for relationships of North American *Aspenium* ferns in 1954 (W. H. Wagner, 1954). This was later confirmed by isozyme work by Charles Werth (Werth et al., 1985a, 1985b). Alice Tryon published numerous works on fern spore characters (A. F. Tryon & Lugardon, 1991; A. F. Tryon & Moran, 1997; R. M. Tryon & Tryon, 1982), which continue to be helpful resources when studying polyploid plants. Barrington *et al.* (1986) correlated fern spore size to ploidy level

(Barrington, Paris, & Ranker, 1986), a trait that can be very helpful when initially studying a polyploid species complex (e.g. (Suissa et al., 2020)).

More recently, Kathleen Pryer and Michael Windham, along with many of their students, have made great strides in understanding the difficult relationships within the Pteridaceae family, and reconstructing the fern evolutionary tree as a whole (Grusz et al., 2009; Kao, Pryer, Freund, Windham, & Rothfels, 2019; Pryer et al., 2001, 2004; Sigel et al., 2011). Amanda Grusz and colleagues have greatly increased our understanding of apogamy[1] in ferns, and its ecological and evolutionary role in reticulate species complexes (Grusz, 2016; Grusz et al., 2021, 2009). David Barrington and collaborators have focused on difficult reticulate relationships in *Polysitchum* ferns (Barrington, 1990; Driscoll & Barrington, 2007; Jorgensen & Barrington, 2017; Perrie, Brownsey, Lockhart, & Large, 2003; Stein & Barrington, 1990), as well as apogamy and the role of multiple ploidy levels in reticulate species complexes (N. Patel, 2018; N. R. Patel, Fawcett, & Gilman, 2019). Paul Wolf's lab and collaborators have utilized RADseq to understand the origin of polyploids in *Pteridium* (Wolf et al., 2019); they also produced the first set of genes for targeted enrichment sequencing in fern (Wolf et al., 2018). Work by the labs of Fay-Wei Li and Carl Rothfels is now paving the way for targeted enrichment data processing methods to understand how polyploid genomes are affected by reticulate evolution (Kao et al., 2019; Rothfels, 2021; Rothfels, Pryer, & Li, 2017). In addition, the Genealogy of Flagellate Plants project is sequencing large targeted enrichment datasets for many groups of plants with flagellate sperm (non-angiosperm land plants), which will greatly improve the resolution of relationships across this large group.

1.3 History of the model fern genus *Ceratopteris*

A majority of this dissertation focuses on the model fern genus *Ceratopteris*. The species *C. richardii* has been instrumental in many areas of fern research (Aragon-Raygoza, Vasco, Blilou, Herrera-Estrella, & Cruz-Ramirez, 2020; Bui et al., 2017; Chasan, 1992; Cordle, Irish, & Cheng, 2007; Hickok, Warne, & Fribourg, 1995; Hickok, Warne, & Slocum,

[1] Apogamy is a type of asexual reproduction where no gametes (egg or sperm) are produced, and a new sprophyte grows directly from gametophytic tissue, with no change in ploidy (Grusz et al., 2009)

1987; Marchant et al., 2019; Nakazato, Jung, Housworth, Rieseberg, & Gastony, 2006; Vasco, Moran, & Ambrose, 2013; Vasco et al., 2016), but very little about its evolutionary history is known (see chapters 4-6).

1.3.1 *Ceratopteris* as a model organism

The fern species *Ceratopteris richardii* was developed as a model organism in the 1980s and 90s because of its fast life cycle and easy cultivation (Hickok et al., 1995). It was later selected for whole genome sequencing because it has a relatively small genome size for homosporous ferns of about 11 gigabases, a low base chromosome number ($n = 39$), and genetic diploidy (Marchant et al., 2019; Sessa et al., 2014).

While there has been a great deal of work in a laboratory setting on *Ceratopteris richardii* (e.g., (Cordle et al., 2007; Hickok et al., 1987; Marchant et al., 2019; Nakazato et al., 2006), the taxonomy of the genus is still greatly understudied. What we do know, however, is that polyploidy and hybridization are rampant within the genus (Adjie et al., 2007; Masuyama & Watano, 2010). *Ceratopteris richardii*, *C. cornuta*, and *C. pteridoides* are diploid ($n = 39$) (Hickok et al., 1987; Lloyd, 1974), while the remainder of the species in this genus are tetraploid ($n = 77, 78$) (Lloyd, 1974; Masuyama & Watano, 2010). Species are known to hybridize readily (Hickok et al., 1987; Masuyama & Watano, 2005), and hybrid plants are incredibly difficult to identify morphologically from their progenitors (Hickok et al., 1987; Lloyd, 1974; Masuyama & Watano, 1994).

Some concerns about the choice of *C. richardii* are that it only has 16 spores per sporangium (the typical diploid number is 64) (Hickok et al., 1987; Lloyd, 1974), it is also known to hybridize with *C. pteridoides*, and there has been speculation that *C. richardii* is itself a hybrid (Lloyd, 1974). The original collection of *C. richardii* for laboratory studies was from Cuba (*Killip 44595*, GH), and is the basis of the *Hnn* strain, the most commonly used experimental lineage of *Ceratopteris richardii*; however, *Hnn* has been hybridized with other strains for certain studies (Hickok et al., 1987), making the origin of current experimental plants all the more confusing.

1.3.2 Taxonomic history

The genus *Ceratopteris* is a group of aquatic, homosporous[2] ferns that are found throughout the world's tropics and subtropics. *Ceratopteris* was first described in 1821 by Brongniart based on the taxa *Acrostichum siliquosum* and *A. thalictroides*, originally named by Linneaus (Benedict, 1909; Linnaeus, 1764). Brongniart named one species, *C. thalictroides* (Brongniart, 1821) in 1821, and then *C. richardii* in 1823 (Brongniart, 1823). The species *C. cornuta* was described by Leprieur in 1830. In 1825, Hooker described the genus under the name *Parkeria*, including one species: *P. pteridoides* (Hooker, 1825). This species was revised in 1907 as *Ceratopteris pteridoides* (Underwood, 1907). It was not until 1909 that a monograph of the genus was published, which recognizes two additional species: *C. lockhartii* (synonym of *C. pteridoides*) and *C. deltoidea* (synonym of *C. richardii*) (Benedict, 1909). Benedict remarks, however, that this revision is incomplete due to the lack of samples available at the time (Benedict, 1909). Only plants from the New World were examined, which is problematic considering that the original description of the plants by Linneaus was based on collections from the Old World (Linnaeus, 1764).

The next and most recent comprehensive examination of the genus was by Lloyd (1974). This monograph of the genus included a much greater number of accessions, and recognized four species with revisions on Benedict's work: *Ceratopteris thalictroides*, *C. cornuta*, *C. pteridoides*, and *C. richardii* Lloyd (1974). Lloyd used a number of morphological characters to distinguish these species, but the ones most relied upon are spore number per sporangium (16 or 32) and number of annulus cells (0 - 70), stipe width (narrow to inflated, and growth habit (floating or rooted) (Lloyd, 1972). Benedict, in contrast, based his classification mainly on leaf shape which, although important in some cases, is exceedingly variable in the group, and can cause confusion if the developmental stage of the plant is unknown (Benedict, 1909).

In the past two decades, several authors have added to Lloyd's 1974 monograph. Masuyama and Watano (2010) published a revision to the genus based on molecular work

[2]Homospory is the condition of producing one type of spore, that develops into bisexual gametophytes. All seed-bearing plants, (and some pteridophytes) and heterosporous, in that they produce male and female spores.

examining cryptic species within *Ceratopteris thalictroides*. They proposed the addition of three new species and two new varieties: *C. gaudichaudii* var. *gaudichaudii*, *C. gaudichaudii* var. *vulgaris*, *C. oblongiloba*, and *C. froesii*. These distinctions were made based on the results of three preceding studies that used allozymes and cross-breeding experiments (Masuyama, Yatabe, Murakami, & Watano, 2002), chromosome counts (Masuyama & Watano, 2005), and morphology from wild and cultivated plants (Masuyama, 2008) to detect cryptic species within *C. thalictroides*. This work done by Masuyama and colleagues is of particular importance to the continued study of *Ceratopteris* because it indicates that there is a great deal of genetic variation within morphologically similar individuals in the genus. It also exemplifies the importance of multiple and varying data sets to accurately detect cryptic species in particular difficult complexes, such as *Ceratopteris*.

Finally, in 2020 a new species was described from the Hainan province in China. *Ceratopteris shingii* is endemic to this island province, and is the only species in the genus to have creeping rhizomes (Zhang, Yu, Shao, Wang, & Yan, 2020).

1.4 Dissertation summary

A solid taxonomic background is the basis for any research in evolutionary biology. In the case of *Ceratopteris richardii*, our understanding of its evolutionary history has lagged far behind other areas of research. This is partly because *Ceratopteris*, like many ferns, has a highly reticulate phylogeny, which makes reconstructing this history challenging (Adjie et al., 2007). In the following chapters, reticulate evolution in ferns is examined, with a special focus on the evolutionary history of *Ceratopteris*, and the model species *C. richardii*:

Chapter 1, *Mothers of Pteridology*, describes some of the contribution made by early female pteridologists, throughout the nineteenth and twentieth centuries. Some of these women were instrumental in the study of hybridization and polyploidy in ferns, including Irene Manton, Alice Tryon, and Florence Wagner. It can be easy to pass over the people behind such important scientific works. This chapter takes time to recognize the women behind some of the most important pteridological discoveries in the past 200 years, without which this dissertation would not be possible.

Chapter 2, *Using RAD Data to Confirm Parentage of Polyploids in a Reticulate Complex of Ferns*, uses next generation sequencing data to take a new look at the species relationships of North American wood ferns (*Dryopteris*). This work builds upon previous work on North American *Dryopteris* (Sessa, Zimmer, & Givnish, 2012a, 2012b), and utilizes genomic data to investigate genomic composition of the species in this complex. The main objective was to see if population structure analysis could detect an extinct progenitor of several species.

Chapter 3, *Cryptic diversity in the model fern genus Ceratopteris (Pteridaceae)* takes a different look at species relationships using RADseq data. There are several known cryptic species within the genus *Ceratopteris*, all described from the "highly polymorphic" *C. thalictroides* (Adjie et al., 2007; Lloyd, 1974; Masuyama & Watano, 2010). This study finds support for previously named cryptic species, and detected a new cryptic species from the New World.

Chapter 4, *There and Back Again: Reticulate evolution in Ceratopteris*, explores the evolutionary history of the cryptic species within the genus. It builds upon research by Adjie et al. (2007), and reconstructs the reticulate evolutionary relationships within the genus, including known natural and experimental hybrids.

This work concludes with Chapter 5, *The evolutionary history of the C-fern (Ceratopteris richardii)*, which dives into the complex history of the model species *C. richardii*. In his 1974 monograph, Lloyd notes that *C. richardii* may be a hybrid species, although this has not been investigated in any subsequent studies (Lloyd, 1974). Most phylogenies of *Ceratopteris* use only the lab strain Hnn as their representative of *C. richardii* (Adjie et al., 2007; Zhang et al., 2020). In this chapter, non-model material of *C. richardii* is included in a phylogenetic context for the first time. Population structure and split network analyses place all accessions of *C. richardii* in different clades, with the two model lineages coming out within the species *C. cornuta*. There is also significant evidence of hybridization within several specimens of *C. richardii*. These findings bring in to question to taxonomic status of *C. richardii*, and how that might change the future of work on this model fern.

References

Adjie, B., Masuyama, S., Ishikawa, H., & Watano, Y. (2007). Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. *Journal of Plant Research*, *120*(1), 129–138.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92.

Aragon-Raygoza, A., Vasco, A., Blilou, I., Herrera-Estrella, L., & Cruz-Ramirez, A. (2020). *Development and cell cycle dynamics of the root apical meristem in the fern Ceratopteris richardii.*

Barrington, D. S. (1990). Hybridization and allopolyploidy in central american *Polystichum*: Cytological and isozyme documentation. *Annals of the Missouri Botanical Garden*, *77*(2), 297–305.

Barrington, D. S. (2020). The biogeography of polyploid ferns across space and time. *American Fern Journal*, *110*(4), 233–254.

Barrington, D. S., Haufler, C. H., & Werth, C. R. (1989). Hybridization, reticulation, and species concepts in the ferns. *American Fern Journal*.

Barrington, D. S., Paris, C. A., & Ranker, T. A. (1986). Systematic inferences from spore and stomate size in the ferns. *American Fern Journal*, *76*(3), 149–159.

Benedict, R. C. (1909). The genus *Ceratopteris*: A preliminary revision. *Bulletin of the Torrey Botanical Club*, *36*, 463-476.

Breinholt, J. W., Carey, S. B., Tiley, G. P., Davis, E. C., Endara, L., McDaniel, S. F., ... Burleigh, J. G. (2021). A target enrichment probe set for resolving the flagellate land plant tree of life. *Applications in Plant Science*, *9*(1), e11406.

Brongniart, A. T. (1821). Description d'un nouveau genre de Fougere, nomme *Ceratopteris*. *Bulletin des Sciences, par la Société Philomatique*, *3*(8), 184–187.

Brongniart, A. T. (1823). *Dictionnaire classique d'histoire naturelle.*

Bui, L. T., Pandzic, D., Youngstrom, C. E., Wallace, S., Irish, E. E., Szövényi, P., & Cheng,

C.-L. (2017). A fern AINTEGUMENTA gene mirrors BABY BOOM in promoting apogamy in *Ceratopteris richardii*. *Plant Journal*, *90*(1), 122–132.

Burbano, H. A., Hodges, E., Green, R. E., Briggs, A. W., Krause, J., Meyer, M., ... Pääbo, S. (2010). Targeted investigation of the neandertal genome by array-based sequence capture. *Science*, *328*(5979), 723–725.

Chasan, R. (1992). *Ceratopteris*: A model plant for the 90s. *The Plant Cell*, *4*(2), 113-115.

Conant, D. S., & Cooper-Driver, G. (1980). Autogamous allohomoploidy in *Alsophila* and *Nephelea* (Cyatheaceae): A new hypothesis for speciation in homoploid homosporous ferns. *American Journal of Botany*, *67*(9), 1269–1288.

Cordle, A. R., Irish, E. E., & Cheng, C. (2007). Apogamy induction in *Ceratopteris richardii*. *International Journal of Plant Science*, *168*(4), 361–369.

Dauphin, B., Mossion, V., Wirth, S., Sandoz, F., & Grant, J. (2017). Recent postglacial recolonization from in situ survival and intense gene flow within natural populations of the alpine fern *Botrychium lunaria* (Ophioglossaceae) revealed by ddRAD sequencing. *Monsieur Benjamin Dauphin*, 149.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: Next-generation population genetics. *Briefings in Functional Genomics*, *9*(5-6), 416–423.

Driscoll, H. E., & Barrington, D. S. (2007). Origin of hawaiian polystichum (dryopteridaceae) in the context of a world phylogeny. *American Journal of Botany*, *94*(8), 1413–1424.

Grusz, A. L. (2016). A current perspective on apomixis in ferns. *Journal of Sytematics Evolution*, *54*(6), 656–665.

Grusz, A. L., Windham, M. D., Picard, K. T., Pryer, K. M., Schuettpelz, E., & Haufler, C. H. (2021). A drought-driven model for the evolution of obligate apomixis in ferns: evidence from pellaeids (pteridaceae). *American Journal of Botany*.

Grusz, A. L., Windham, M. D., & Pryer, K. M. (2009). Deciphering the origins of apomictic polyploids in the *Cheilanthes yavapensis* complex (Pteridaceae). *American Journal of Botany*, *96*(9), 1636–1645.

Haufler, C. H., Windham, M. D., & Rabe, E. W. (1995). Reticulate evolution in the *Polypodium vulgare* complex. *Systematic Botany*, *20*(2), 89–109.

Hickok, L. G., Warne, T. R., & Fribourg, R. S. (1995). The biology of the fern *Ceratopteris* and its use as a model system. *International Journal of Plant Science*, *156*(3), 332–345.

Hickok, L. G., Warne, T. R., & Slocum, M. K. (1987). *Ceratopteris richardii*: Applications for experimental plant biology. *American Journal of Botany*, *74*(8), 1304-1316.

Hipp, A. L., Manos, P. S., Hahn, M., Avishai, M., Bodénès, C., Cavender-Bares, J., ... Valencia-Avalos, S. (2020). Genomic landscape of the global oak phylogeny. *New Phytologist*, *226*(4), 1198–1212.

Hooker, W. J. (1825). *Exotic flora*.

Jorgensen, S. A., & Barrington, D. S. (2017). Two beringian origins for the allotetraploid fern *Polystichum braunii* (Dryopteridaceae). *Systematic Botany*, *42*(1), 6–16.

Kao, T.-T., Pryer, K. M., Freund, F. D., Windham, M. D., & Rothfels, C. J. (2019). Low-copy nuclear sequence data confirm complex patterns of farina evolution in notholaenid ferns (Pteridaceae). *Molecular Phylogenetics and Evolution*, *138*, 139–155.

Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., de Vries, J., Delaux, P.-M., ... Pryer, K. M. (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nat Plants*, *4*(7), 460–472.

Linnaeus, C. v. (1764). *Species plantarum: Volume 2*. Imprensis Laurentii Salvii.

Lloyd, R. M. (1972). Species delination in the genus *Ceratopteris* (Parkeriaceae). *American Journal of Botany*, *59*, 676.

Lloyd, R. M. (1974). Systematics of the genus *ceratopteris* Brong.. (Parkeriaceae) II. taxonomy. *Brittonia*, *26*(2), 139–160.

Ma, Z.-Y., Wen, J., Tian, J.-P., Jamal, A., Chen, L.-Q., & Liu, X.-Q. (2018). Testing reticulate evolution of four *Vitis* species from east asia using restriction-site associated DNA sequencing. *Journal of Systematics and Evolution*, *56*(4), 331–339.

Manton, I. (1950). *Problems of cytology in the pteridophyta*. Cambridge University Press.

Manton, I., & Sledge, W. A. (1954). Observations on the cytology and taxonomy of the pteridophyte flora of ceylon. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *238*(654), 127–185.

Marchant, D. B., Sessa, E. B., Wolf, P. G., Heo, K., Barbazuk, W. B., Soltis, P. S., & Soltis, D. E. (2019). The C-Fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. *Scientific Reports*, *9*(1), 18181.

Masuyama, S. (2008). Cryptic species in the fern *ceratopteris thalictroides* (Parkeriaceae). III. Referential diagnostic characters of three cryptic species. *Journal of Plant Research*, *12*(3), 279-286.

Masuyama, S., & Watano, Y. (1994). Hybrid sterility between two isozymic types of the fern *Ceratopteris thalictroides* in japan. *Journal of Plant Research*, *107*(3), 269–274.

Masuyama, S., & Watano, Y. (2005). Cryptic species in the fern *ceratopteris thalictroides* (L.) Brong.. (parkeriaceae). II. Cytological characteristics of three cryptic species. *Acta phytotaxonomica et geobotanica*, *56*(3), 231-240.

Masuyama, S., & Watano, Y. (2010). Cryptic species in the fern *ceratopteris thalictroides* (L.) Brong.. Parkeriaceae. IV. Taxonomic revision. *Acta. Phytotax. Geobot.*, *61*(2), 75-86.

Masuyama, S., Yatabe, Y., Murakami, N., & Watano, Y. (2002). Cryptic species in the fern *ceratopteris thalictroides* (L.) Brong.. (parkeriaceae). I. Molecular analyses and crossing tests. *Journal of Plant Research*, *115*(1118), 87-97.

Nakazato, T., Jung, M.-K., Housworth, E. A., Rieseberg, L. H., & Gastony, G. J. (2006). Genetic Map-Based analysis of genome structure in the homosporous fern *Ceratopteris richardii*. *Genetics*, *173*(3), 1585–1597.

Nelson, T. C., Stathos, A. M., Vanderpool, D. D., Finseth, F. R., Yuan, Y.-W., & Fishman, L. (2020). Ancient and recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model monkeyflower radiation (*Mimulus* section *Erythranthe*). *bioRxiv*, 2020.09.08.287151.

Paris, C. A., Wagner, F. S., & Wagner, W. H. (1989). Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. *American Fern Journal*, *79*(2), 46–54.

Patel, N. (2018). *Apomixis, hybridization, and biodiversity in ferns: Insights from genera Phegopteris and Polystichum* (Unpublished doctoral dissertation). University of Vermont.

Patel, N. R., Fawcett, S., & Gilman, A. V. (2019). *Phegopteris excelsior* (Thelypteridaceae): A new species of North American tetraploid beech fern. *Novon*.

Perrie, L. R., Brownsey, P. J., Lockhart, P. J., & Large, M. F. (2003). Evidence for an allopolyploid complex in new zealand *Polystichum* (Dryopteridaceae). *New Zealand Journal of Botany*, *41*(2), 189–215.

Pflugbeil, G., Affenzeller, M., Tribsch, A., & Comes, H. P. (2021). Primary hybrid zone formation in *Tephroseris helenitis* (Asteraceae), following postglacial range expansion along the central Northern Alps. *Molecular Ecology*.

PPG I. (2016). A community-derived classification for extant lycophytes and ferns: PPG I. *Jnl of Sytematics Evolution*, *54*(6), 563–603.

Preston, R. D. (1990). Irene manton. 17 april 1904-13 may 1988. *Biogr. Mem. Fellows R. Soc.*, *35*, 249–261.

Pryer, K. M., Schneider, H., Smith, A. R., Cranfill, R., Wolf, P. G., Hunt, J. S., & Sipes, S. D. (2001). Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature*, *409*(6820), 618–622.

Pryer, K. M., Schuettpelz, E., Wolf, P. G., Schneider, H., Smith, A. R., & Cranfill, R. (2004). Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. *American Journal of Botany*, *91*(10), 1582–1598.

Rieseberg, L. H. (1991). Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *American Journal of Botany*, *78*(9), 1218–1237.

Rothfels, C. J. (2021). Polyploid phylogenetics. *New Phytologist*.

Rothfels, C. J., Pryer, K. M., & Li, F.-W. (2017). Next-generation polyploid phylogenetics: rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *New Phytologist*, *213*(1), 413–429.

Rutherford, S., Wilson, P. G., Rossetto, M., & Bonser, S. P. (2016). Phylogenomics of the green ash eucalypts (myrtaceae): a tale of reticulate evolution and misidentification. *Aust. Systematic Botany*, *28*(6), 326–354.

Sánchez-Baracaldo, P. (2004). Phylogenetics and biogeography of the neotropical fern genera *Jamesonia* and *Eriosorus* (Pteridaceae). *American Journal of Botany*, *91*(2), 274–284.

Schley, R. J., Pennington, R. T., Pérez-Escobar, O. A., Helmstetter, A. J., de la Estrella, M., Larridon, I., ... Klitgård, B. (2020). Introgression across evolutionary scales suggests reticulation contributes to Amazonian tree diversity. *Molecular Ecology*.

Schneider, H., Navarro-Gomez, A., Russell, S. J., Ansell, S., Grundmann, M., & Vogel, J. (2013). Exploring the utility of three nuclear regions to reconstruct reticulate evolution in the fern genus asplenium. *Journal of Systematics and Evolution*, *51*(2), 142–153.

Sessa, E. B., Banks, J. A., Barker, M. S., Der, J. P., Duffy, A. M., Graham, S. W., ... Wolf, P. G. (2014). Between two fern genomes. *Gigascience*, *3*, 15.

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012a). Phylogeny, divergence times, and historical biogeography of new world *Dryopteris* (Dryopteridaceae). *American Journal of Botany*, *99*(4), 730–750.

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012b). Unraveling reticulate evolution in north american *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology*, *12*, 104.

Sigel, E. M., Windham, M. D., Huiet, L., Yatskievych, G., & Pryer, K. M. (2011). Species relationships and farina evolution in the cheilanthoid fern genus *Argyrochosma* (pteridaceae). *Systematic Botany*, *36*(3), 554–564.

Sigel, E. M., Windham, M. D., & Pryer, K. M. (2014). Evidence for reciprocal origins in *Polypodium hesperium* (Polypodiaceae): a fern model system for investigating how

multiple origins shape allopolyploid genomes. *American Journal of Botany*, *101*(9), 1476–1485.

Smith, A. R., Pryer, K. M., Schuettpelz, E., Korall, P., Schneider, H., & Wolf, P. G. (2006). A classification for extant ferns. *Taxon*, *55*(3), 705–731.

Stein, D. B., & Barrington, D. S. (1990). Recurring hybrid formation in a population of *Polystichum* x *potteri*: Evidence from chloroplast DNA comparisons. *Annals of the Missouri Botanical Garden*, *77*(2), 334–339.

Suissa, J. S., Kinosian, S. P., Schafran, P. W., Bolin, J. F., Carl Taylor, W., & Zimmer, E. A. (2020). *Revealing the evolutionary history of a reticulate polyploid complex in the genus Isoëtes.*

Tryon, A. F., & Lugardon, B. (1991). *Spores of the pteridophyta: Surface, wall structure, and diversity based on electron microscope studies.* Springer Science & Business Media.

Tryon, A. F., & Moran, R. C. (1997). *The ferns and allied plants of new england.* Center for Biological Conservation, Massachusetts Audubon Society.

Tryon, R. M., & Tryon, A. F. (1982). *Ferns and allied plants: With special reference to tropical America.* Springer Science & Business Media.

Underwood, L. M. (1907). The names of some of our native ferns. *Torreya*, *7*(10), 193–198.

Vargas, O. M., Ortiz, E. M., & Simpson, B. B. (2017). Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-andean diversification (Asteraceae: Astereae: Diplostephium). *New Phytologist*, *214*(4), 1736–1750.

Vasco, A., Moran, R. C., & Ambrose, B. A. (2013). The evolution, morphology, and development of fern leaves. *Frontiers in Plant Science*, *4*, 345.

Vasco, A., Smalls, T. L., Graham, S. W., Cooper, E. D., Wong, G. K.-S., Stevenson, D. W., ... Ambrose, B. A. (2016). Challenging the paradigms of leaf evolution: Class III HD-Zips in ferns and lycophytes. *New Phytologist*, *212*(3), 745–758.

Wagner, W. H. (1954). Reticulate evolution in the appalachian aspleniums. *Evolution*, *8*(2), 103–118.

Wagner, W. H. (1969). The role and taxonomic treatment of hybrids. *Bioscience*, *19*(9), 785–795.

Wagner, W. H. (1983). Reticulistics: the recognition of hybrids and their role in cladistics and classification. *Advances in cladistics*, *2*, 63–79.

Wagner, W. H., Jr. (1970). Biosystematics and evolutionary noise. *Taxon*, *19*(2), 146–151.

Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., & Liston, A. (2014). Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Science*, *2*(9).

Werth, C. R., Guttman, S. I., & Eshbaugh, W. H. (1985a). Electrophoretic evidence of reticulate evolution in the Appalachian *Asplenium* complex. *Systematic Botany*, *10*(2), 184–192.

Werth, C. R., Guttman, S. I., & Eshbaugh, W. H. (1985b). Recurring origins of allopolyploid species in *Asplenium*. *Science*, *228*(4700), 731–733.

Wickell, D. A., & Li, F.-W. (2020). On the evolutionary significance of horizontal gene transfers in plants. *New Phytologist*, *225*(1), 113–117.

Wolf, P. G., Robison, T. A., Johnson, M. G., Sundue, M. A., Testo, W. L., & Rothfels, C. J. (2018). Target sequence capture of nuclear-encoded genes for phylogenetic analysis in ferns. *Applications in Plant Science*, *6*(5), e01148.

Wolf, P. G., Rowe, C. A., Kinosian, S. P., Der, J. P., Lockhart, P. J., Shepherd, L. D., ... Thomson, J. A. (2019). Worldwide relationships in the fern genus *Pteridium* (bracken) based on nuclear genome markers. *American Journal of Botany*, *106*(10), 1365–1376.

Zhang, R., Yu, J.-H., Shao, W., Wang, W.-Q., & Yan, Y.-H. (2020). *Ceratopteris shingii*, a new species of ceratopteris with creeping rhizomes from hainan, china. *Phytotaxa*, *449*(1), 23–30.

Zhou, W., Ji, X., Obata, S., Pais, A., Dong, Y., Peet, R., & Xiang, Q.-Y. J. (2018). Resolving relationships and phylogeographic history of the *Nyssa sylvatica* complex using data from RAD-seq and species distribution modeling. *Molecular Phylogenetics*

and Evolution, *126*, 1–16.

CHAPTER 2

Mothers of Pteridology

Published in the American Fern Journal[1]

2.1 Abstract

Women have long been underrepresented in science, technology, engineering, and mathematics (STEM), although there are certain fields within this umbrella which show less of a disparity — the biological sciences being one example. Specifically within biology, pteridology has a rich history of female contributors involved since its inception. In this review I strive to highlight some of the foremost female pteridologists including Elizabeth Knight Britton, Alma Stokey, Irene Manton, Alice Tryon, Barbara Hoshizaki, and Florence Wagner. This is not an exhaustive list, but rather an insight into the strong maternal lineage of the fern community. While the field would not be the same without the contributions of many male pteridologists, in this piece I want to emphasize the important work that some of these founding women have made. Some of the research conducted by early female pteridologists was largely undermined by their time and circumstance; here I bring their lives and works to the foreground. Furthermore, I hope this paper inspires young botanists to enter my unique and historically rich field.

[1]Authors: Sylvia Kinosian and Jacob Suissa

2.2 Introduction

> My pine lot is truly a fern-land, as I discovered to my surprise not long ago when my eyes were opened and I became a fern lover... When the list in my pine lot had reached sixteen, I climbed my neighbor's fences and wandered farther afield.
>
> — Edith Scamman (1923)

In 1893 The Linnean Fern Bulletin–later becoming the American Fern Society–unfurled as a small chapter of the Agassiz Association for the Study of Nature (Clute, 1902,4; Winslow, 1919). The founding members of the society included Mr. Willard N. Clute, Mr. Henry C. Cowles, Mrs. T. D. Dershimer, Mr. Reuben M. Strong, Mr. James A. Graves, and Mrs. A. D. Dean, with the goal to "promote the study of ferns by correspondence, the exchange of specimens, the publication of knowledge thus obtained, [and] the promotion of field trips for ferns," (Clute, 1943). Pteridology was starting to grow, and the initiation of a centralized society helped amalgamate the field. Since its inception, the society had a strong female presence. In-fact, Mrs. T.D. Dershimer who helped initiate the society, was the first secretary, later becoming the second elected Vice President in 1894 (Clute, 1894). Many female authors were included in the first volume of the journal, such as Nellie Mirick (Vol. 1 No. 1), Kate D. Spalding (Vol. 1 No. 3), Elizabeth G. Britton (Vol. 1 No. 4), and Emily H. Terry (Vol. 1 No. 6). The prevalence of women early in the history of a scientific society was uncommon, but not restricted to pteridology.

Other journals in the biological sciences also have a long history of female involvement. Recently, Bronstein & Bolnick (2018) published an article in The American Naturalist (Am. Nat.), reviewing female authorship in the journal since its beginnings in 1867 (Fig. 2.1). They found that from 1867 to 1970, the proportion of female authorship was variable, ranging from around 0.03 to 0.2. Since the 1970s, however, there has been a steady increase in the number of female authors publishing in the Am. Nat. with nearly half of all papers in the journal having at least one female author today (Bronstein & Bolnick, 2018). I found a similar pattern in The American Fern Journal (AFJ) to that in the Am. Nat. (Fig. 2.1),

with the proportion of female authorship being relatively low from 1910 to 1970, however the AFJ had slightly less variability than the Am. Nat. in female authorship through this time period. Since the 1970s, female authorship in the AFJ has increased consistently as well. A notable difference is that the proportion of female authors in the Am. Nat. is nearly twice that of the AFJ today. It will be intriguing to see how the number of female authors in the AFJ, and other biology-centered journals, changes in years to come. I hope to see a further increase in female contributions to the AFJ, and hope this review will inspire young female pteridologists to pursue a career in my field.

2.3 Historical women in pteridology: Premise

The field of pteridology had a strong base of foundational women early in its history. Both directly or indirectly, these pioneering women have influenced my field and the people within it. In addition, the diversity of male and female contributions has been important for the growth and development of the field. In the following pages, I highlight some founding mothers of pteridology. This is not an exhaustive list, but rather an insight into the strong maternal lineage of the fern community. Additionally, as some of these pterdiologist's work was undermined by their time and circumstance, I wish to bring their lives and works to the foreground. These women, as exemplified by the opening quote from Edith Scamman (Scamman, 1923), were unconditional lovers of not just ferns, but of the natural world and scientific exploration. Their enthusiasm and ingenuity contributed to some of the most important pteridological discoveries to date, while it also nurtured and inspired subsequent works by women and men alike.

2.4 Historical women in pteridology: Biographies

2.4.1 Margaretta Riley (1804 - 1899)

Margaretta "Meta" Riley is regarded as the first female pteridologist, having contributed much to the field in its earliest days. She was born in 1804 in Nottingham, England to Richard and Margaretta Hooper. Meta's father worked in the hosier and cotton

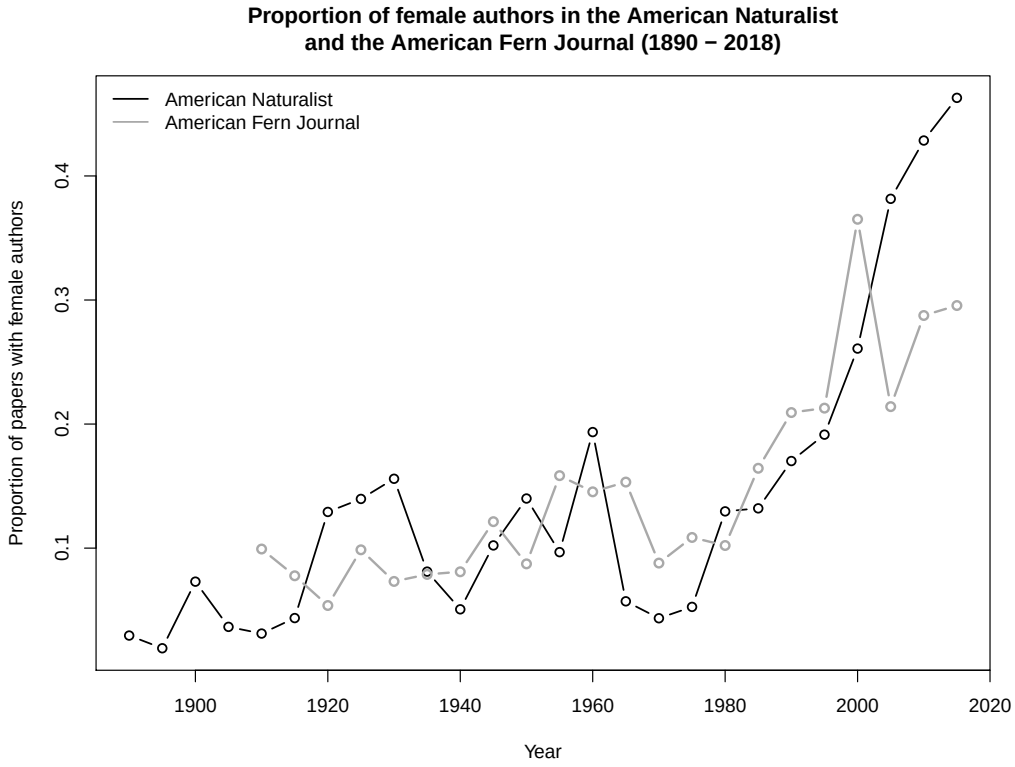


Fig. 2.1: Proportion of female authors from the American Fern Journal (AFJ; light gray line) of papers (collected by SPK), and proportion of papers with female authors from the American Naturalist (AN; black line; data from Bronstein & Bolnick (2018), measured in cumulative 5-year increments. The female authorship in the early days of the AFJ is greater than the female authorship in that of the AN, but current female authorship in the AFJ is lower than that in the AN.

business, and the family was fairly well-off (Creese, 2000). Meta Hooper married John Riley in 1826 and moved to the nearby village of Papplewick. The couple shared a passion for the collection, cultivation, and classification of ferns, and would go on to contribute greatly to the understanding of the British fern flora (Ellis, 2004).

Although Meta worked closely with her husband, she was not always credited as an author on their final works, so her full contribution to pteridology is unknown (Ellis, 2004). Despite this, records do show that Meta made several advancements in the growing field. During the late 1830s, she donated a complete pressed collection of Britain's fern flora to the Botanical Society of London. She also published on the classification of *Cystopteris*

and *Thelypteris* in England, and, along with her husband John, published a comprehensive review of the cultivation of native English ferns (Allen, 1978). In 1838, John Riley was elected to the Botanical Society of London, following a paper on fern hybridization; Meta was elected to join the following year, an unusual honor for women at the time (Allen, 1978; Ellis, 2004).

Meta's contributions to the botanical world ended after John's death in 1846 (Creese, 2000). After his passing, she donated their extensive collection of dried and living specimens to other botanists and museums, but remained a member of the Botanical Society of London. In lieu of botany, Meta took up interest in gardening, watercolor painting, poetry, history, philosophy, and politics. She died of bronchitis in 1899 at the age of 95 (Ellis, 2004). Her name and impact persist in broader venues than pteridology, as a large crater on Venus was named after her in the late 20th century: The Riley Crater (NASA/JPL, 1996).

2.4.2 Sarah "Sadie" Frances Price (1849-1903)

Kentucky's most well regarded botanists of the 19th century was the amateur naturalist Sarah "Sadie" Price. Sadie, the daughter of Maria and Alexander Price, was born in Evansville, Indiana. The family moved to Bowling Green, Kentucky when Sadie was a young child; it was here where she spent 12 years with her older brother Frederick and sister Mary. The beginning of the Civil War in 1861 proved difficult for the Price family; soon after the war began, Frederick was enlisted in the Union and the rest of the family left their work and moved back to Indiana (Lovell, 1959). During this time Sadie was educated at St. Agnes Hall. In the following decade both Sadie's brother and her parents passed away in 1873 and 1875, respectively. At the age of 26, Price was bedridden with unexplained back pain, but her passion for the natural world trumped her mortal woes and despite her ailments she would teach art courses in watercolor painting. Her students would bring plants and other wildlife to her bedside as items for illustration (Green & Rollins, 2005). After traveling to visit physicians in Philadelphia to treat her back pains, she was healthy and botanizing by 1880.

Following her return to good health, Price began a vigorous career in natural history.

On top of thousands of botanical illustrations, she published 24 botanical manuscripts, and 40 scientific papers (Green & Rollins, 2005), including works on the ferns of Kentucky, rare species of *Asplenium*, and on cave ferns (Price, 1893, 1901,0). In 1897 Sadie published The Fern-Collector's Handbook and Herbarium. This was an 80-page treatise on ferns of the northern United States including detailed descriptions of over 70 species. As a skilled artist, Sadie has detailed illustrations for each of these species in the handbook, her attention to detail can be captured in these lovely illustrations (Fig. 2.2). Additionally, she included figures on the putative characters that unite some of the common species in the northern U.S. which is a boon to any flora. On top of illustrations and keys, in the preface of her handbook Ms. Price includes descriptions of how to properly preserve and mount herbarium specimens for the layperson.

While Sadie was the premier botanist of Kentucky and impactful in early pteridology, her work and contributions may have been broadly overlooked because of John Williamson's Flora of Kentucky, published two decades prior (1878). The Flora of Kentucky was the first state fern flora in the United States which, in retrospect, seemed to overshadow subsequent work in the state. While his work was foundational, Sadie nonetheless expanded on it to include ferns of the broader northern United States (Price, 1901). She also corrected some of Williamson's work, specifically on the genus Ophioglossum and its prevalence in North America (Price, 1897).

Not only did Ms. Price directly work and publish on ferns, but her enthusiasm for these organisms is expressed throughout her other publications. She advocated for the conservation of ferns; she was greatly concerned about some of the more "beautiful forms" and how they were being destroyed in the eastern mountains (Lovell, 1959). Sadie's work was a boon to the study of ferns in Kentucky during a depauperate period of pteridology in the state. The significance of her work was not only appreciated by practicing botanists, but also by lay plant lovers. In a review in the Fern Bulletin (v. 5 page 36) of Sadie's fern handbook it was said that her "volume [is] designed to assist the beginner in fern study... one who knows nothing about botany, may identify any fern in the region mentioned, by

Fig. 2.2: llustrations by Sadie Price of A. *Asplenium pinnatifidum*, and B. *Botrychium virginianum*. Photo credit: Missouri Botanic Garden Archive

merely turning the pages until [they] come to the illustration that matches the specimen." This review is not only an ode to her artistic abilities, but to her desire to engage lay folk in the study and appreciation of ferns.

In addition to her publications, Sadie's contributions include 2,912 herbarium sheets as well as about 965 watercolor illustrations (Green & Rollins, 2005). In 1903, Sadie passed away relatively young at the age of 54. She left behind many works still unpublished, some of which were picked up and published posthumously by her sister Mary Price (Lovell, 1951,5). Mary sold Sadie's herbarium specimens the Missouri Botanical Gardens where they are now housed. In one of her several obituaries Reverend Frank Thomas regarded her as "a true high-priestess of nature," (Green & Rollins, 2005); she had a strong influence on the botanical community, and her work lives on through her thousands of botanical illustrations and important publications.

2.4.3 Elizabeth Gertrude Knight Britton (1858 - 1934)

Elizabeth Knight was born in New York City on January 9th, 1858 (Fig. 2.3). As is true with many of the pioneering women in pteridology, Ms. Britton did not receive a graduate degree and never held a formal salaried job. Despite these setbacks she published over 340 manuscripts and was a figurehead in the botanical community (Barnhart, 1935). She spent much of her early life in Matanzas, Cuba where her grandfather owned a sugar plantation and factory (Barnhart, 1935). This time was spent with her sisters and father naturalizing the island's ecosystem; these experiences must have been strongly influential in her botanical interests as she continued the study of natural history throughout her life.

As Elizabeth grew older, she spent much of her time with her grandmother in New York where she attended private school. She went on to enroll and graduate from Normal (now Hunter) College in 1875 at the age of 17. In 1885, Elizabeth married Nathaniel Britton who worked as a geologist at Columbia University. Nathaniel was interested in plants, but fully transitioned to studying botany following his marriage to Elizabeth (Creese, 2000). Elizabeth was a vivacious individual whose personality was infections. Her overall accomplishments in botany landed her a spot as one of the American Men of Science, now aptly American Men and Women of Science (Creese, 2000). While she was primarily known for being a figurehead in bryology, she had a passion for ferns and published on them throughout her career.

She contributed to pteridology in her treatment of Ophioglossaceae (Britton, 1897), Schizaeaceae (Britton, 1901; Britton & Taylor, 1901), Hymenophyllaceae (Britton, 1902), and Pteridaceae (Britton & Taylor, 1902). One of her most significant findings in pteridology was her documentation of the relationship between fungal hyphae and *Schizaea* gametophytes (Britton, 1901). She noted that the fungus forms a symbiosis with the gametophyte and then with the developing sporophyte. This was an interesting finding at the time because mycorrhizal association with *Schizaea* gametophytes were never observed before. Elizabeth made significant contributions to pteridology in its early days; she was influential in many ways from her publications to her philanthropy.

Fig. 2.3: Photo of Elizabeth Knight Britton peering into a microscope. Photo credit: Archives of The New York Botanical Garden.

In total she published 16 papers on ferns including an account of one of the only collected individuals of *Schizaea pusilla* in Nova Scotia, this specimen was sent to Asa Gray of Harvard University (Britton, 1896; Howe, 1934). She continued with her fern studies publishing revisions of North American *Ophioglossum* (Britton, 1897), and capped it off with final expeditions to Cuba (Britton, 1911) and Bombay (Britton, 1923). Her broad impact

to botany as a whole was tremendous. She was the first woman charter member of The Botanical Society of America and founded The American Bryological Society. Additionally, she was paramount in the establishment of The New York Botanic Garden (NYBG). Her husband Nathaniel was the first director of the NYBG (Rudolph, 1982), and Elizabeth was prominent in many aspects of this position. She was integral in the genesis and progression of the NYBG, assisting Nathaniel with the responsibilities of director in chief for 30 years. Her joint work with her husband was impressive but it did not overshadow her own work as a botanist. Elizabeth was also a conservationist (Howe, 1934), and published many papers on the need for protecting plants (including ferns) which aided in the passing of plant conservation laws in North America. Both during her career and posthumously she was recognized widely as one of the most influential and productive women in botany.

Elizabeth passed away February 25th, 1934 and a lovely memorial was written in her honor in the Journal of The New York Botanical Garden (Howe, 1934). Britton's philanthropic, conservation, and scientific contributions echo through the field of plant biology as a whole.

2.4.4 Alma Gracey Stokey (1877 - 1968)

When reading through literature on fern gametophytes the name Stokey is hard to miss. Alma Gracey Stokey, the mother of gametophyte morphology and anatomy, paved the way for focusing on this unique stage of the fern life cycle in the 1900s. One of the great pteridologists of the 20th century, Stokey was born on June 17th, 1877 in Canton, Ohio to a large family of five children. Before entering college, Stokey taught at a public high school, demonstrating her passion for education and knowledge at a young age (Atkinson, 1968). Soon after her time as a teacher, she went on to receive a B.A. at Oberlin College in 1904. Stokey stayed for an additional 2 years as a research assistant in botany with Dr. Frederick Grover. She then went on to become a PhD student at the University of Chicago and received her doctorate in 3 years. Although Dr. Stokey is widely known for her work on gametophytes, one of her first papers published as a PhD student at the University of Chicago was on the anatomy of *Isoëtes* (Stokey, 1909). Progressing to full

professor at Holyoke College in Massachusetts by 1916, she focussed heavily on anatomy and morphology of fern gametophytes and used their traits to try and understand the evolutionary relationships between fern lineages.

Although there is a deep literature on gametophyte morphology, her detailed approach to working with these organisms provided new perspective and insight into this stage of the fern life cycle. Stokey made significant contributions on the gametophytes of many lineages of ferns including the Gleicheniaceae, Blechnaceae, Thelypteridaceae, Polypodiaceae, Dryopteridaceae and Hymenophyllaceae (Atkinson & Stokey, 1973; Stokey, 1957,4,5; Stokey & Atkinson, 1952,5). Her broad comparative studies including many representative species within each lineage allowed her to draw taxonomic and phylogenetic conclusions within these groups. Growing gametophytes is not a trivial task, sometimes it can take a few years for spores to germinate and mature (Stokey, 1942), thus the amount of work invested into Alma's studies was immense as documented by her numerous publications. Two of her most well cited papers include the "Comparative morphology of the gametophyte of homosporous ferns" (Atkinson & Stokey, 1964) and "The Contributions by the gametophyte to the classification of homosporous ferns" (Stokey, 1951).

Alongside Stokey was Lenette R. Atkinson, who published with her constantly. Atkinson was another prominent female pteridologists and I would be remiss to discuss gametophytes or Alma Stokey without mentioning Lenette Atkinson. Dr. Stokey and Atkinson made aware to the botanical community that there are specific morphological, developmental, and anatomical character traits of the fern gametophyte that are taxonomically and phylogenetically informative. Some of these characters that Stokey elucidated as taxonomically important traits include thallus symmetry, hair formation, rhizoid form, and antheridia morphology. Stokey's work on gametophyte morphology, development, and anatomy has permeated the field and lead to more detailed work on this life stage of ferns ranging from taxonomic treatments to physiology.

Alma's commitment to the pteridology was recognized by The American Fern Society in 1953, when they named her the 10th Honorary Member and the first woman to be bestowed

with this honor (Atkinson, 1968). Additionally, later in her career her undergraduate alma mater granted her an honorary doctorate degree, recognizing her foundational contributions to the field of pteridology and botany as a whole. In addition to Dr. Stokey's focus on scientific endeavors, her outreach for women in science was demonstrated through her position at the historically all women's college of Mount Holyoke and through her assistance with initiating the department of botany at the Women's Christian College Madras, now the Women's Christian College, Chennai, India (Atkinson, 1968). Stokey will be long-remembered for her impassioned and groundbreaking work on fern gametophytes, and her persistent dedication to the advancement of pteridology.

2.4.5 Norma E. Pfeiffer (1878 - 1978)

Norma E. Pfeiffer lived to the ripe old age of 100 years. In her time, she contributed greatly to pteridology, and in particular to the study of *Isoëtes*. A resident of the state of Chicago, Norma received her bachelor's and doctorate from the University of Chicago. She received her doctorate following her discovery of a new plant species *Thismia americana*, and in 1913, she was the youngest person to receive a PhD from the university. Norma taught at the University of North Dakota until 1924 when she then moved to Yonkers to work at the Boyce Thompson Institute for Plant Research, which is now a part of Cornell University (The New York Times, 12 September 1989).

While she may be most well known for her discovery and breeding of lilies, Norma worked on many things pteridological such as the gametophytes of Pteris and *Ophioglossum* (Pfeiffer, 1912,1), but her foundational contribution to my field is on the genus *Isoetes*. The first monograph of *Isoetes* published in 1883 by Motelay & Vendryes (1882) contains beautiful illustrations and detailed species descriptions; Pfeiffer's new edition was more detailed, broader in scope, and includes dichotomous keys that still prove important in morphological identification of species in the complex genus. Her monograph includes a historical perspective, introduction to the genus, species descriptions, and many spore images (Pfeiffer, 1922). Norma passed away in 1978 from complications due to a brain tumor. To date her 1922 monograph has been cited over 200 times and is still an important reference for

Isoetologists.

2.4.6 Irene Manton (1904 - 1988)

An important aspect of fern biology is the rampant polyploidy that exists across nearly all major fern lineages. Among polyploid ferns, odd-numbered polyploids can have a difficult time reproducing sexually, and so adopt an apomictic life history (Grusz, 2016). Irene Manton, the mother of chromosome counts, was integral in my understanding of these two common but complex fern traits.

Irene was born on April 17th, 1904 to George and Milana Manton in the London suburb of Kensington. Her parents ensured that Irene and her older sister Sidnie received a good education, sending them to private school, taking them on many outdoor trips, and introducing them to drawing, painting, and music (Leadbeater, 2004). They also encouraged both their daughters' interest in biology; George Manton was the first to introduce Irene to a microscope, and Milana educated the girls on natural history (Preston, 1990). Irene and Sidnie were good students, they passed their college entrance exams with honors and both received scholarships (Leadbeater, 2004). Irene attended Girton College, Cambridge and it was there that her interest in chromosomes began (Preston, 1990). Some years later, after her graduate work, Irene had a postdoc in Sweden and then moved to the University of Manchester for an Assistant Lectureship position in 1929. There she worked with W. H. Lang, who inspired her work on pteridophytes. In 1946, Irene took a position as the Chair of Botany at Leeds where she would remain for the remainder of her career (Leedale, 1989).

When Irene first began working with chromosomes, they were counted by staining the material of interest and then examining serial cross sections. Since this method was quite time-consuming and not very accurate, Irene developed the "squash" method commonly used today (Preston, 1990). It was through this technique that she was able to produce the large volume of chromosome counts and cytological data needed for her influential book Problems of cytology and evolution in the Pteridophyta (Manton, 1950). This volume is also where Irene presents her scheme for apogamy in ferns, now referred to as the Dopp-Manton scheme (Cordle, Bui, Irish, & Cheng, 2010; Grusz, 2016). Irene's chromosome work also

helped pteridologists understand the importance of genus-level base chromosome numbers in distinguishing different genera, namely through her work on *Athyrium* ($n = 40$) and *Diplazium* ($n = 41$) (Manton & Sledge, 1954).

Later in her career, Irene began to focus more on electron microscopy, with algae as her primary study organisms. While she made many important contributions to phycology (Manton & Parke, 1960; Preston, 1990), she always humbly claimed to be not a phycologist, but rather a simple botanist interested in plant cytology Leedale (1989). After a brief illness in 1988, Irene passed away at the age of 84. She continued her scientific work up until a few weeks before her death (Preston, 1990). In an obituary for Irene, phycologist Gordon Leedale wrote that the world had "...lost a scientist of genius and many of us lost a generous and steadfast friend," Leedale (1989).

2.4.7 Alice Faber Tryon (1920 - 2009)

Regarded as one of the 20th century's eminent pteridologists, Alice Tryon made significant contributions in the areas of fern spore morphology, taxonomy, reproductive biology, and biogeography (Fig. 2.4). She, alongside her husband Rolla M. Tryon Jr. (1916 - 2001), mentored and inspired a generation of pteridologists. Their lab at Harvard produced many important publications and initiated the careers of some of today's leading pteridologists (Gastony, Barrington, & Conant, 2002).

Alice Elizabeth Faber was born August 2nd, 1920 in Milwaukee, Wisconsin. She attended Milwaukee State Teachers College (now University of Wisconsin, Milwaukee) in 1941 and taught public school for several years before returning to her studies at the University of Wisconsin, Madison. It was there that she met Rolla Tryon, whom she married in 1945 (Gastony, Barrington, & Conant, 2009). As Rolla's student, Alice completed her master's degree on the spore characteristics of *Selaginella* (Tyron, 1949). Alice began her doctoral degree on the genus Pellaea (Tryon, 1957; Tryon & Britton, 1958) at Madison, but finished her studies at Washington University in 1952 after she and Rolla moved to the Missouri Botanical Garden. A few years later, after a stint at the University of California at Berkeley, Alice and Rolla moved to Harvard University, both taking positions at the Grey Herbarium

Fig. 2.4: Photo of Alice F. Tryon in the field, Brazil. Photo Credit: Mettenius - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/wiki/File:Alicetryoni_brazil.jpg

in 1958 (Gastony et al., 2009).

Through her career, Alice made numerous foundational contributions to the field of pteridology, primarily on the subjects of fern taxonomy and spores. Shortly after arriving at Harvard, Alice produced monographs of the closely related Andean genera *Jamesonia* (Tryon, 1962) and *Eriosorus* (Tryon, 1970). She hypothesized that "*Jamesonia* is derived from one or more elements in *Eriosorus*," (Tryon, 1970). Subsequent molecular analysis

has provided support for Alice's hypothesis. Studies have shown the two genera to form one monophyletic clade; as traditionally defined, *Jamesonia* is polyphyletic and *Eriosorus* is paraphyletic (Sanchez-Baracaldo, 2004).

Alice is perhaps best remembered for her work on pteridophyte spores, a topic she returned to several times throughout her career. Her book, Spores of the Pteridophyta (Tryon & Lugardon, 1991) is a detailed volume on the spore features of ferns from a diversity of lineages, and includes 2,792 scanning electron microscope (SEM) images. The use of SEM for examining fern spores was pioneered by Alice during her time at Harvard (Gastony et al., 2009). She also included spore images in the book The Ferns and Allied Plants of New England (Tryon & Moran, 1997), an unusual but powerful tool to include in such a volume.

While Alice had many impactful solo projects, collaboration with her husband Rolla was a constant throughout her career. One of their most notable publications is Ferns and Allied Plants with Special Reference to Tropical America (Tryon & Tryon, 1982), a book still used today not only for identification of ferns, but to provide working taxonomic hypotheses for molecular studies (Gastony et al., 2009). The Tryons also initiated many lecture series including the annual Systematics Symposium at the Missouri Botanical Gardens, the New England Fern Conference, and the Tryon Lecture Series at the University of South Florida after their retirement (Gastony et al., 2009). While in Florida and technically retired, the Tryons still continued to publish and engage with their community until Rolla's passing in 2001 (Gastony et al., 2002). In her last years, Alice made several generous gifts to the Field Museum in Chicago, the New England Botanical Club, the University of Vermont, and the University of Wisconsin, Madison. Alice was cared for by friends and surrounded by mementos of a happy life with Rolla when she passed away peacefully in her apartment on March 29th, 2009 (Gastony et al., 2009).

Without the contributions of Alice and Rolla Tryon, pteridology would be in a very different place today. We would lack a large body of knowledge generated by their lab, as well as some of the leading fern researchers and mentors to the next generation of pteridologists.

2.4.8 Barbara Joe Hoshizaki (1928 - 2012)

Barbara Joe Hoshizaki was a world expert in fern horticulture. Her book, the Fern Growers' Manual, is the preeminent volume on the subject. The second edition of the book, co-authored with Dr. Robbin C. Moran, contains information on about 700 species of ferns from 124 genera (Hoshizaki & Moran, 2001). Throughout her life, Barbara contributed greatly to not only my knowledge of fern horticulture, but taxonomy and popular interest in ferns as well. She published numerous papers on the cultivation of ferns (Hoshizaki, 1970a,8; Hoshizaki & Wilson, 1999; Yansura & Hoshizaki, 2012), fern hybrids (Hoshizaki, 1975,9), and systematics and identification of ferns — notably of *Platycerium* (Hoshizaki, 1970b,7; Hoshizaki & Price, 1990).

Barbara's interest in pteridophytes began when she attended the University of California, Los Angeles in 1951. There she received her bachelor's degree and met Mildred Mathias, who would become Barbara's mentor and prime advocate for her study of ferns (Moran, 2012). Mildred herself was an incredibly influential botanist in her own right: an Apiaceae expert and important founding member of the Organization for Tropical Studies (Gibson, 2017).

In 1954, Barbara received her master's degree from the University of California, LA (UCLA) and became a professor of biology at the City College of Los Angeles, where she would remain for 28 years. In addition to teaching, Barbara was the Curator of Ferns at the UCLA Herbarium; she served as the president of the American Fern Society, Southern California Horticultural Institute, and Los Angeles International Fern Society; and was the vice-president for the Pacific Horticultural Foundation (Moran, 2012).

In 1967, Barbara spent six weeks on the first Organization for Tropical Studies Ferns and Lycophytes course (Moran, 2012). This experience was so meaningful to her that after her passing in 2012, her family established the Barbara Joe Hoshizaki Memorial Scholarship to aid students in attending OTS courses. I was fortunate enough to receive this scholarship for the 50th anniversary of the Tropical Ferns and Lycophytes course in 2017, an opportunity that I am extremely grateful for. Through her numerous publications and support

of organizations like OTS, Barbara's memory and contributions to the fern community live on.

2.4.9 Florence Wagner (1919 - 2020)

Florence Wagner (née Signaigo) was born in 1919. In the 1940s she attended graduate school at the University of California, Berkeley. It was there that she met Warren "Herb" Wagner. They married in 1948, beginning a long and fruitful career of collaboration (Farrar, 2002). Florence was trained as a phycologist, and remarked that she would happily study ferns as long as Herb kept her away from the coasts (Christopher Haufler, personal communication). Florence and Herb helped increase pteridologist's interest in hybrids, publishing review and novel empirical papers on the subject (Wagner, 1992; Wagner & Wagner, 1979)

In the obituary written for Herb Wagner in the American Fern Journal, Don Farrar shares a touching anecdote about the Wagners:

"The excitement of field trips with Herb and Florence remain highlights of my graduate years at Michigan, as do warm memories of holidays at the Wagner home, cutting out snowflakes or whatever project Florence had designed for their 'extended family' of graduate students. This nurturing of an Ozark farm boy a long way from home made a difference. It was a personal gift. Yet I know it was only one of many such personal gifts, bestowed on many others as well, by Herb and Florence. For those gifts I all say thanks!" — Farrar (2002)

2.5 Conclusion

The women highlighted here contributed broadly to my field and have inspired many generations of past and current pteridologists. Their work spans the spectrum of biology from anatomy to systematics and cytology to horticulture. Through my review of their work I hope to exemplify not only their impact on my field, but also their lives and achievements as people and scientists.

In addition to the female pteridologists I have covered in detail, there are many women whose contributions were impactful, but who were not detailed in this short review. A

handful of these women include: Doris Löve, who published the greatly influential volume, "A Cytotaxonomic atlas of the Pteridophyta" (Kaersvang, Weber, & Ives, 2000; Love, Love, & Pichi Sermolli, 1977) with her husband Áskell, as well as having a second career as a translator for 12 languages (Kaersvang et al., 2000; Love et al., 1977); Diana Stein, one of the first people to work on fern chloroplasts, investigating their structure, component genes, and evolutionary signal (Palmer & Stein, 1982; Stein & Barrington, 1990; Stein, Palmer, & Thompson, 1986); Gillian Cooper-Driver, who conducted work on secondary compounds in bracken (*Pteridium*) and their variation in the context of phenology, herbivory, and taxonomy (Cooper-Driver, Finch, Swain, & Bernays, 1977; Cooper-Driver & Swain, 1976,7); Aura Star, who researched frond exudates and allelopathy in *Pityrogramma* (Star, 1980; Star & Mabry, 1971), and Margaret Slosson, whose book "How Ferns Grow" (Slosson, 1906) was one of the first on pteridophyte development, and her work on fern hybridization was the first empirical evidence of such contributed by an American researcher (Clute, 1902; Slosson, 1902).

The history of early women in pteridology is quite unique in STEM and this is reflected in the current state of the field. Today there are numerous labs run by prominent female pteridologists all over the world, all contributing regularly to my growing knowledge of ferns. With this piece I bring to light the foundational women, the work they contributed to fern biology, and the fruitful collaborations they had with their peers; I hope that they continue to inspire young scientists to enter the unique field of pteridology.

References

Allen, D. E. (1978). The first woman pteridologist. *British Pteridolgical Society Bulletin*, *1*(6), 247–249.

Atkinson, L. R. (1968). Alma Gracey Stokey (1877-1968). *American Fern Journal*, *58*(4), 145–152.

Atkinson, L. R., & Stokey, A. G. (1964). Comparative morphology of the gametophyte of homosporous ferns. *Phytomorphology*.

Atkinson, L. R., & Stokey, A. G. (1973). The gametophyte of some Jamaican thelypteroid ferns. *Botanical Journal of the Linnean Society*, *66*(1), 23–36.

Barnhart, J. H. (1935). The published work of Elizabeth Gertrude Britton. *Bulletin of the Torrey Botanical Club*, *62*(1), 1–17.

Britton, E. G. (1896). How i found *Schizaea pusilla*. *Linnaean Fern Bulletinetin*, *4*, 17–19.

Britton, E. G. (1897). A revision of the North American species of *Ophioglossum*. *Bulletin of the Torrey Botanical Club*, *24*(12), 545–559.

Britton, E. G. (1901). The curly grass, *Schizaea pusilla*. *The Plant World*, *4*(4), 72–73.

Britton, E. G. (1902). *Trichomanes radicans*. *Bulletin of the Torrey Botanical Club*, *29*(7), 475–477.

Britton, E. G. (1911). Fern collecting in Cuba. *American Fern Journal*, *1*(4), 75–77.

Britton, E. G. (1923). The ferns of Bombay.

Britton, E. G., & Taylor, A. (1901). Life history of *Schizaea pusilla*. *Bulletin of the Torrey Botanical Club*, (pp. 1–19).

Britton, E. G., & Taylor, A. (1902). The life history of *Vittaria lineata*. *Memoirs of the Torrey Botanical Club*, *8*(3), 185–211.

Bronstein, J. L., & Bolnick, D. I. (2018). "her joyous enthusiasm for her Life-Work…": Early women authors in the American Naturalist. *American Naturalist*.

Clute, W. N. (1894). Nature and the ground-nut. *The Asa Gray Bulletin*, (pp. 3–4).

Clute, W. N. (1902). *Botrychium ternatum* and *obliquum*. *Fern Bulletin*, *10*, 76–77.

Clute, W. N. (1943). Early days of the American fern society. *American Fern Journal*, *33*(1), 1–8.

Cooper-Driver, G., Finch, S., Swain, T., & Bernays, E. (1977). Seasonal variation in secondary plant compounds in relation to the palatability of *Pteridium aquilinum*. *Biochemical Systematics and Ecology*, *5*(3), 177–183.

Cooper-Driver, G. A., & Swain, T. (1976). Cyanogenic polymorphism in bracken in relation to herbivore predation. *Nature*, *260*(5552), 604–604.

Cordle, A. R., Bui, L. T., Irish, E. E., & Cheng, C.-L. (2010). *Laboratory-Induced Apogamy and Apospory in Ceratopteris richardii*. New York, NY: Springer New York.

Creese, M. R. S. (2000). *Ladies in the Laboratory? American and British Women in Science, 1800-1900: A Survey of Their Contributions to Research*. Scarecrow Press.

Ellis, M. (2004). Margaretta "Meta" Riley, a local philanthropist.

Farrar, D. R. (2002). Obituary: Warren H. Wagner, Jr. (1920-2000). *American Fern Journal*, *92*(2), 39–49.

Gastony, G. J., Barrington, D. S., & Conant, D. S. (2002). Obituary: Rolla Milton Tryon, Jr. (1916-2001). *American Fern Journal*, *92*(1), 1–9.

Gastony, G. J., Barrington, D. S., & Conant, D. S. (2009). Obituary: Alice Faber Tryon (1920–2009). *American Fern Journal*, *99*(4), 231–235.

Gibson, A. C. (2017). Mildred E. Mathias. Mildred E. Mathias Botanical Garden. botgard.ucla.edu/mildred-e-mathias/. Accessed November 2, 2018.

Green, N., & Rollins, J. (2005). Women's Work Portraits of 12 Scientific Illustrators from the 17th to 21st Century. womenswork.lindahall.org/price1.shtml. Retrieved December 17, 2018.

Grusz, A. L. (2016). A current perspective on apomixis in ferns. *Journal of Sytematics Evolution*, *54*(6), 656–665.

Hoshizaki, B. J. (1970a). The genus *Adiantum* in cultivation Polypodiaceae. *Baileya*, *17*(3), 97–144.

Hoshizaki, B. J. (1970b). The rhizome scales of *Platycerium*. *American Fern Journal*, *60*(4), 144–160.

Hoshizaki, B. J. (1972). Morphology and phylogeny of *Platycerium* species. *Biotropica*, (pp. 93–117).

Hoshizaki, B. J. (1975). A staghorn fern (*Platycerium*) hybrid. *American Fern Journal*, *65*(4), 99–101.

Hoshizaki, B. J. (1981). The fern genus *Davallia* in cultivation (Davalliaceae). *Baileya*.

Hoshizaki, B. J. (1991). An" intergeneric" hybrid: *Aglaomorpha* x *Drynaria*. *American Fern Journal*, (pp. 37–43).

Hoshizaki, B. J., & Moran, R. C. (2001). *Fern grower's manual*, vol. 604. Timber Press Portland.

Hoshizaki, B. J., & Price, M. G. (1990). *Platycerium* update. *American Fern Journal*, *80*(2), 53–69.

Hoshizaki, B. J., & Wilson, K. A. (1999). The cultivated species of the fern genus *Dryopteris* in the united states. *American Fern Journal*, (pp. 1–98).

Howe, M. A. (1934). Elizabeth Gertrude Britton. *Journal of the New York Botanical Gard*, *35*, 97–103.

Kaersvang, L. L., Weber, W. A., & Ives, J. D. (2000). In Memoriam: Doris Löve, 1918-2000. *Arct. Antarct. Alp. Res.*, *32*(3), 360–362.

Leadbeater, B. (2004). Irene Manton: a biography. *The Linnean (Special issue)*, *5*, 1–94.

Leedale, G. (1989). Irene Manton, FRS, 1904–1988. *British Phycological Journal*, *24*(2), 103–109.

Love, A., Love, D., & Pichi Sermolli, R. E. G. (1977). *Cytotaxonomical atlas of the Pteridophyta*. J. Cramer.

Lovell, H. B. (1951). Bibliography of Sarah F. Price, Kentucky Naturalist. *Transactions of the Kentucky Academy of Sciences*, *13*, 121–128.

Lovell, H. B. (1959). Biographical sketch of Sadie Price. *The Kentucky Warbler*, *35*, 19–37.

Manton, I. (1950). *Problems of cytology in the Pteridophyta*. Cambridge University Press.

Manton, I., & Parke, M. (1960). Further observations on small green flagellates with special reference to possible relatives of *Chromulina pusilla* butcher. *Journal of the Marine Biological Association of the United Kingdom*, *39*(2), 275–298.

Manton, I., & Sledge, W. A. (1954). Observations on the cytology and taxonomy of the pteridophyte flora of Ceylon. *Philosophical Transactions of the Royal Society B*, *238*(654), 127–185.

Moran, R. C. (2012). The life of barbara joe hoshizaki (1928–2012). *American Fern Journal*, *102*(4), 252–255.

Motelay, L., & Vendryes, A. (1882). Monographie des isoeteae. *Actes de la Societe Linneenne, de Bordeaux*.

NASA/JPL (1996). Venus - oblique view of crater riley. www.photojournal.jpl.nasa.gov/catalog/PIA00266. Accessed: 2019-03-09.

Palmer, J. D., & Stein, D. B. (1982). Chloroplast DNA from the fern *Osmunda cinnamomea*: physical organization, gene localization and comparison to angiosperm. *Current Genetics*, *5*(3), 165–170.

Pfeiffer, N. E. (1912). Abnormalities in prothallia of *Pteris longifolia*. *Botanical Gazzette*, *53*(5), 436–438.

Pfeiffer, N. E. (1916). The prothallia of *Ophioglossum vulgatum*. *Botanical Gazzette*, *61*(6), 518–522.

Pfeiffer, N. E. (1922). Monograph of the Isoetaceae. *Annals of the Missouri Botanical Garden*, *9*(2), 79–233.

Preston, R. D. (1990). Irene Manton. 17 april 1904-13 may 1988. *Biographical Memoirs of Fellows of the Royal Society*, *35*, 249–261.

Price, S. F. (1893). A rare fern. *Garden and Forest*, *6*, 99–100.

Price, S. F. (1897). *The Fern-Collector's Handbook*. Henry Holt and Company.

Price, S. F. (1901). Notes from western kentucky. *The Plant World*, *4*(8), 143–144.

Price, S. F. (1904). Some ferns of the cave region of Stone County, MO. *Fern Bulletin*, *12*, 72–77.

Rudolph, E. D. (1982). Women in nineteenth century American botany; a generally unrecognized constituency. *American Journal of Botany.*, *69*(8), 1346–1355.

Sanchez-Baracaldo, P. (2004). Phylogenetics and biogeography of the neotropical fern genera *Jamesonia* and *Eriosorus* (Pteridaceae). *American Journal of Botany.*, *91*(2), 274–284.

Scamman, E. (1923). The ferns in my pine lot. *The Maine Naturalist*, *3*, 18–30.

Slosson, M. (1902). The origin of *Asplenium ebenoides*. *Bulletin of the Torrey Botanical Club*, *29*(8), 487–495.

Slosson, M. (1906). *How Ferns Grow*. H. Holt.

Star, A. E. (1980). Frond exudate flavonoids as allelopathic agents in *Pityrogramma*. *Bulletin Torrey Botanical Club*, *107*(2), 146–153.

Star, A. E., & Mabry, T. J. (1971). Flavonoid frond exudates from two Jamaican ferns, *Pityrogramma tartarea* and *P. calomelanos*. *Phytochemistry*, *10*(11), 2817–2818.

Stein, D. B., & Barrington, D. S. (1990). Recurring hybrid formation in a population of *Polystichum x potteri*: Evidence from chloroplast DNA comparisons. *Annals of the Missouri Botanical Garden*, *77*(2), 334–339.

Stein, D. B., Palmer, J. D., & Thompson, W. F. (1986). Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda*. *Current Genetics*, *10*(11), 835–841.

Stokey, A. (1957). The gametophyte of some American species of *Elaphoglossum* and *Rhipidopteris*. *Phytomorphology*, *7*, 275–292.

Stokey, A. G. (1909). The anatomy of *Isoetes*. *Botanical Gazzette*, *47*(4), 311–335.

Stokey, A. G. (1942). Gametophytes of *Marattia sambucina* and *Macroglossum smithii*. *Botanical Gazzette*, *103*(3), 559–569.

Stokey, A. G. (1948). Reproductive structures of the gametophytes of hymenophyllum and *Trichomanes*. *Botanical Gazzette*, *109*(4), 363–380.

Stokey, A. G. (1950). The gametophyte of the Gleicheniaceae. *Bulletin of the Torrey Botanical Club*, (pp. 323–339).

Stokey, A. G. (1951). The contribution by the gametophyte to classification of the homosporous ferns. *Phytomorphology*, *1*(1-2), 39–58.

Stokey, A. G., & Atkinson, L. R. (1952). The gametophyte of *Blechnum spicant* (l.) wither. and *B. buchtienii* rosenst. *Phytomorphology*, *2*, 9–15.

Stokey, A. G., & Atkinson, L. R. (1958). The gametophyte of the Grammitidaceae. *Phytomorphology*, *8*(3-4), 391–403.

Times, T. N. Y. (1989). Norma Pfeiffer, expert on lilies, dies at 100. *The New York Times*, (p. 10).

Tryon, A. F. (1957). A revision of the fern genus *Pellaea* section *Pellaea*. *Annals of the Missouri Botanical Garden*, *44*(2), 125–193.

Tryon, A. F. (1962). A monograph of the fern genus *Jamesonia*. *Contributions from the Gray Herbarium of Harvard University*, *191*, 109–197.

Tryon, A. F. (1970). A monograph of the fern genus *Eriosorus*. *Contributions from the Gray Herbarium of Harvard University*, *200*, 54–174.

Tryon, A. F., & Britton, D. M. (1958). Cytotaxonomic studies on the fern genus *Pellaea*. *Evolution*, *12*(2), 137–145.

Tryon, A. F., & Lugardon, B. (1991). *Spores of the Pteridophyta: Surface, Wall Structure, and Diversity Based on Electron Microscope Studies*. Springer Science & Business Media.

Tryon, A. F., & Moran, R. C. (1997). *The Ferns and Allied Plants of New England*. Center for Biological Conservation, Massachusetts Audubon Society.

Tryon, R. M., & Tryon, A. F. (1982). *Ferns and Allied Plants: With Special Reference to Tropical America*. Springer Science & Business Media.

Tyron, A. F. (1949). Spores of the genus *Selaginella* in North America North of Mexico. *Annals of the Missouri Botanical Garden*, *36*(4), 413–431.

Wagner, F. S. (1992). Cytological problems in lycopodium sens. lat. *Annals of the Missouri Botanical Garden*, (pp. 718–729).

Wagner, W. H., Jr, & Wagner, F. S. (1979). Polyploidy in pteridophytes. *Basic Life Sciences*, *13*, 199–214.

Winslow, E. J. (1919). Early days of the American Fern Society. *American Fern Journal*, *9*(2), 33–38.

Yansura, D. G., & Hoshizaki, B. J. (2012). The tree fern highland lace is a cultivar of *Sphaeropteris cooperi*. *American Fern Journal*, *102*(1), 69–77.

CHAPTER 3

Using RAD data to confirm parentage of polyploids in a reticulate complex of ferns

Published in the American Fern Journal[1]

3.1 Abstract

Reticulate evolution, in which phylogenetic relationships are not strictly bifurcating (tree-like), is a common feature of fern evolution. Ferns are prone to hybridization and whole genome duplication, two processes that can make untangling phylogenetic relationships among species challenging. Next-generation sequencing technologies have greatly increased the amount of data available for analyzing various aspects of evolutionary history, and here I test the ability of one next-generation sequencing approach to identify the progenitors of allopolyploids. I produced and analyzed double-digest restriction-site-associated DNA (ddRAD) sequences from six species of North American *Dryopteris*, including two allopolyploids and their respective diploid parents. The relationships of these species have been confidently established in previous studies, and the goal was to determine the extent to which RAD data are capable of identifying these known relationships. Analyses of the genetic structure in the samples reliably separated the diploids from one other, but in general each polyploid sample resembled one or the other of its progenitors, or had genetic variation unassignable to either parent. None of the polyploid samples had unambiguous genetic contributions from both known parents, as I had expected. These results may have been influenced by small overall sample size, different numbers of samples from the two diploid parents in each pair, and the large divergence times between the diploids. These are all potentially important issues to consider when designing similar studies, and the results therefore have useful implications for researchers interested in using a RAD approach to study polyploid complexes.

[1]Authors: Sylvia Kinosian, Weston Testo, Sally Chambers, and Emily Sessa

3.2 Introduction

Fern enthusiasts, amateur and professional alike, are captivated by these plants for a multitude of reasons. For researchers interested in evolutionary processes, ferns are a particularly fascinating lineage to study because of their propensity for polyploidy and hybridization. These two processes, which are particularly common in ferns compared to other land plants (Wood et al., 2009), have the potential to influence many aspects of evolution, particularly those related to genome size, structure, and complexity, as well as phylogenetic relationships (Soltis, Visger, & Soltis, 2014). Hybridization occurs when members of two distinct evolutionary lineages interbreed and produce offspring. Polyploidy, or whole genome duplication, is a complete doubling of the genome that results in offspring with at least twice the chromosomes and genetic content of their progenitors. These processes can occur independently or in synchrony, producing organisms known as allopolyploids that are the product of both genome doubling and hybridization.

The non-bifurcating phylogenies that result from these reticulate evolutionary processes require extra effort to decipher. The traditional workhorse of plant phylogenetics, the chloroplast genome, is maternally inherited in most plants, including in ferns (Gastony & Yatskievych, 1992; Vogel, Russell, Rumsey, Barrett, & Gibby, 1998), and can therefore identify only one parent of a putative hybrid or allopolyploid. Identifying the second, paternal parent, requires information from biparentally inherited nuclear markers. For the last two decades, obtaining data from these markers has relied on painstaking and time-consuming laboratory procedures to isolate each homeologous copy (using an Escherichia coli vector that replicates via cloning), so that each can be sequenced independently. Genes such as gapCp (Schuettpelz, Grusz, Windham, & Pryer, 2008), pgiC (Ishikawa, Watano, Kano, Ito, & Kurita, 2002), and many others (Rothfels, Li, et al., 2015) have been analyzed in this way and produced the first DNA-sequencing based confirmations of parentage in many fern polyploid complexes (e.g., in *Dryopteris* Adans. (Sessa, Zimmer, & Givnish, 2012b), *Polystichum* Roth (Jorgensen & Barrington, 2017), various Pteridaceae genera (Beck, Windham, Yatskievych, & Pryer, 2010; Grusz, Windham, & Pryer, 2009), and many

others). However, the rise of next-generation sequencing (NGS) technologies has inspired a natural desire to use these powerful and data-rich approaches as an alternative to labor-intensive gene-by-gene cloning and sequencing for analyzing polyploid complexes. Despite this, few studies have used NGS approaches to resolve reticulate evolutionary histories in ferns (but see Rothfels, Pryer, and Li (2017)).

A principal limitation to the use of NGS data in studies of polyploids has been the challenge of correctly assembling homeologous copies, especially when sequence data are generated on platforms with relatively short read lengths. Overcoming this and related bioinformatic challenges will be a critical step before widespread use of NGS approaches in the study of polyploid complexes can become feasible. An additional problem for ferns is their large genomes, which have made them less tractable than other plant lineages as study groups for NGS methods (for example, ferns were the last major lineage of land plants to have a reference nuclear genome sequenced, due in part to their massive genomes (F.-W. Li et al., 2018; Sessa et al., 2014).

"Reduced representation" sequencing methods seek to minimize the complexity of genome assembly by sequencing only a subset of the complete genome (Andrews, Good, Miller, Luikart, & Hohenlohe, 2016; Rowe, Renaut, & Guggisberg, 2011). While the cost of next-generation sequencing has decreased steadily, assembling the millions of short (typically 100-150 base pairs) sequencing reads produced by these approaches remains an immense challenge. For many study systems, sequencing and assembling an entire genome remains out of reach, due either to financial limitations, assembly issues, or a combination of the two (e.g., difficult-to-assemble genomes benefit from long-read sequencing, which is more expensive than short-read sequencing). Whole genome sequencing may also be unnecessary for addressing questions of interest, such as identifying the parents of a polyploid species, a query for which sequence data from only one or a handful or markers is typically sufficient. In groups like ferns, where whole-genome sequencing is still impractical for the average researcher, reduced representation strategies are ideal for capitalizing on next-generation sequencing approaches while using resources efficiently.

In the present study, I evaluated the utility of one reduced representation, next-generation sequencing approach – restriction-site-associated DNA sequencing (known as RAD or RADseq) – for identifying the progenitors of polyploid ferns. RAD and the related genotyping-by-sequencing (GBS) are both approaches that utilize the restriction enzyme cut sites that occur naturally across the genome (Andrews et al., 2016). Whole genomic DNA is first digested with restriction enzymes that cut the DNA only at specific sequences that are unique to each enzyme; these cut sites typically occur thousands of times throughout the genome, at different frequencies for different enzymes. The resulting DNA fragments will span a range of sizes, and those in the ideal range for next-generation sequencing can be selected at a later step in library preparation. By sequencing only fragments that are adjacent to restriction enzyme cut sites, RAD targets a non-random portion of the genome and therefore increases the likelihood of sequencing homologous regions across samples. This is especially important for organisms with large genomes, like ferns, where approaches such as genome skimming or "shotgun" sequencing (which theoretically sequence a truly random subset of the genome) are less likely to capture homologous sequences from different samples and species. Because SNPs are identified from individual (unassembled) reads in the RAD data analysis pipeline, this method does not suffer from the issues associated with assembling homeologous copies from short-read data. However, while RAD and other reduced-representation approaches attempt to deal with the issue of large genome sizes, the analytical complications introduced by polyploidy still haunt these approaches, since the programs available for analyzing the datasets typically operate under the assumption that included taxa are diploid, a fundamental assumption (with numerous implications for expected copy numbers and locus behavior, among other things) that is not readily altered to accommodate data from known polyploids. For example, two of the most commonly used pipelines for processing GBS and RADseq data, Stacks (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) and TASSEL (Bradbury et al., 2007)) assume genotypes are diploid and do not permit allele frequencies to deviate from those expected in diploids,

typically treating as noise information that may be in fact be the signal of polyploid genotypes.

I applied a double-digest RAD approach (ddRAD; so called because two restriction enzymes are used) to a small dataset for a polyploid complex in which relationships are known with confidence: North American *Dryopteris*. These ferns have been the focus of previous study by the research group (Sessa & Givnish, 2014; Sessa, Zhang, Väre, & Juslén, 2015; Sessa, Zimmer, & Givnish, 2012a; Sessa et al., 2012b; Sessa, Zimmer, & Givnish, 2012c; Testo, Watkins, & Barrington, 2015), and the North American complex as a whole includes one extinct and four extant diploids, four allotetraploids, and one allohexaploid (Fig. 3.1). The parents of the allopolyploids were first hypothesized based on morphology and cytological observations (Walker, 1955, 1959, 1961, 1962, 1969), and later confirmed using sequences of *gapCp* and *pgiC* ((Sessa et al., 2012b)). Here I focused on two of the allotetraploids, *D. campyloptera* (Kunze) Clarkson and *D. celsa* (W. Palmer) Knowlt., Palmer and Pollard, and their respective diploid parents: *D. intermedia* (Willd.) A. Gray and *D. expansa* (C. Presl) Fraser-Jenk. and Jermy for *D. campyloptera*, and *D. ludoviciana* (Kunze) Small and D. goldiana (Hook. ex Goldie) A. Gray for *D. celsa* (Fig. 3.1).

My goal was to determine whether analysis of ddRAD sequences from a small sampling of these six species could recover evidence of their known polyploid-progenitor relationships by correctly grouping sequences from the two allotetraploids with their respective parent taxa, ideally with each polyploid showing evidence of equal genetic contributions from the two progenitors. The six species I selected are ideal for this study because all diploid progenitors are extant and can be sampled, and the diploids are sufficiently diverged from one another ((Sessa et al., 2012a)) that sequences from the polyploids can theoretically be assigned unambiguously to each of the four diploids.

3.3 Materials and Methods

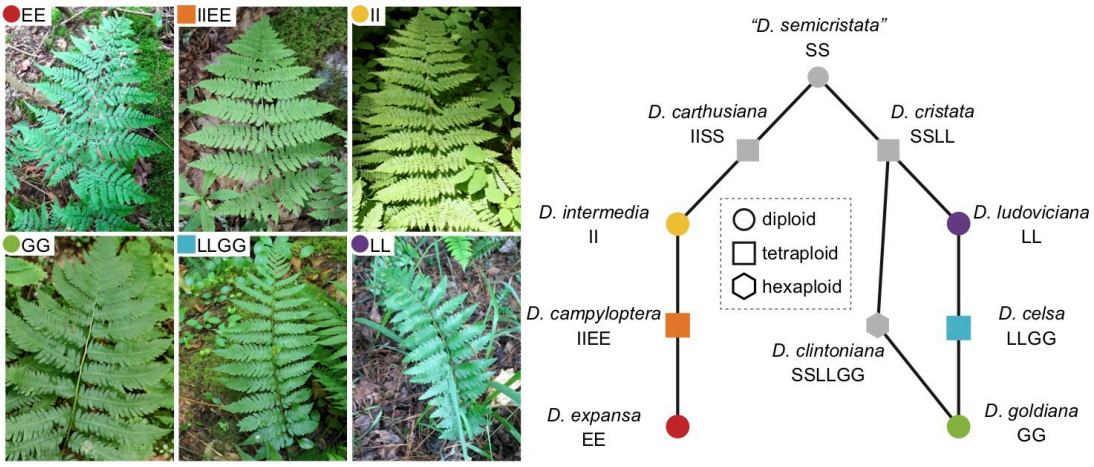


Fig. 3.1: Relationships of diploid and polyploid species in the North American *Dryopteris reticulate* complex (based on (Sessa et al., 2012a, 2012b, 2012c)). Letters below species names refer to their genomic designations. Pictures of all species are included at left (all photos by EB Sessa).

3.3.1 Taxon sampling and DNA extraction

I included sixteen samples representing six *Dryopteris* species, with all but two species represented by multiple accessions (Table 3.1). I extracted total genomic DNA using a DNeasy Plant Mini Kit (Qiagen, Valencia, California, USA) following the manufacturer's protocol.

3.3.2 ddRAD library preparation and sequencing

I followed the ddRAD library construction protocol established by (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012), with a few modifications. Because of the large genome sizes of the focal taxa (average in *Dryopteris* diploids is 1C = 7.63 pg; (Bainard, Henry, Bainard, & Newmaster, 2011), I replicated each sample three times during library preparation. To obtain equal numbers of reads for all individuals, I standardized DNA quantity prior to library preparation.

I used two enzymes, MseI and EcoRI – a frequent cutter and an infrequent cutter, respectively (referring to the distribution of the enzymes' cut sites across the genome) – to digest 6 microliters of genomic DNA from each sample. I then ligated enzyme-specific, double-stranded adapters (8–14 base pairs in length) to the digested DNA fragments, with

Table 3.1: Sample ID and voucher information for samples included in this study.

	Genus	Ploidy	Sample ID	U.S. state	Voucher
1	*Dryopteris campyloptera*	4x	E648	VA	EBS 9722006 (FLAS)
2	*Dryopteris campyloptera*	4x	E650	VA	EBS 9722005 (FLAS)
3	*Dryopteris campyloptera*	4x	E655	VA	EBS 9722004 (FLAS)
4	*Dryopteris campyloptera*	4x	E649	VA	EBS 9722016 (FLAS)
5	*Dryopteris celsa*	4x	E657	GA	EBS 27 (WIS)
6	*Dryopteris celsa*	4x	E658	SC	EBS 49 (WIS)
7	*Dryopteris celsa*	4x	E661	LA	Price 94-2 (NY)
8	*Dryopteris celsa*	4x	E664	MO	3479307 (MO)
9	*Dryopteris expansa*	2x	E734	AK	5710532 (MO)
10	*Dryopteris goldiana*	2x	E700	NY	EBS 65A (WIS)
11	*Dryopteris goldiana*	2x	E701	NY	EBS 72 (WIS)
12	*Dryopteris goldiana*	2x	E715	NY	EBS 12 (WIS)
13	*Dryopteris intermedia*	2x	E713	WV	EBS 9722021 (FLAS)
14	*Dryopteris intermedia*	2x	E714	VA	EBS 9722010 (FLAS)
15	*Dryopteris intermedia*	2x	E736	MO	5198765 (MO)
16	*Dryopteris intermedia*	2x	E739	SC	EBS 48 (WIS)

the EcoRI adapter containing a unique barcode specific to each sample and replicate. I ensured successful ligation of the adapters to the digested DNA by inspecting PCR products via gel electrophoresis visualization. I then pooled the restriction ligation product from each of the successful libraries and cleaned this pooled product using a QIAquick PCR Purification Kit (Qiagen, Valencia, California, USA). The pooled product was then run on an Elf Pippin Bioanalyzer (Sage Science, Massachusetts, USA) to select genomic fragments ranging from 350 to 700 bp (size selection service provided by the University of Florida Interdisciplinary Center for Biotechnology Research; UF ICBR). I checked the success of the fragment size selection via gel electrophoresis and analysis on a TapeStation 2200 Automated Electrophoresis (Agilent, California, USA) system.

I performed a final round of PCR to anneal the Illumina sequencing primers to the digested DNA fragments, working with 1 microliter at a time of digested, size-selected, pooled DNA. To reduce opportunity for PCR errors I performed eighteen separate reactions and then combined the resulting PCR products for sequencing. I visualized a subset of the pooled product via gel electrophoresis to check for amplification success, and cleaned the

remaining product using a QIAquick PCR Purification Kit. Pooled, clean PCR product was submitted to the UF ICBR where it was cleaned further using Ampure beads to remove unincorporated adapters before being sequenced on an Illumina NextSeq500 platform, generating 2 × 150 bp reads. A 10% phiX spike was included during sequencing as an internal control.

3.3.3 Data cleaning

I processed the raw Illumina reads using the Process Radtags pipeline in Stacks (Catchen et al., 2013), retaining all reads with a quality score above 20 and splitting the raw reads by sample and replicate based on the unique EcoRI barcode, which was subsequently trimmed along with the cut site. This resulted in roughly 277 million reads. I then used the FAST-X Trimmer from the FAST-X Toolkit (hannonlab.cshl.edu/fastx_toolkit/) to remove the MseI cut site from each secondary read and to trim the last four bases from the 3' end of the primary reads to make all reads 127 bp in length. Cleaned, trimmed forward and reverse reads for each individual were paired used PEAR v. 0.9.2 (Zhang, Kobert, Flouri, & Stamatakis, 2014), and all reads from each of the three replicates per sample pooled together. Unfiltered, demultiplexed sequences have been deposited in NCBI GenBank Sequence Read Archive (PRJNA542715).

3.3.4 Data analysis

Processing the raw Illumina data occurred in three stages: creation of a pseudo-reference genome, alignment of reads and variant calling, and admixture analysis. An R Markdown file describing this pipeline and all custom scripts, is located at github.com/sylviakinosian/dryopteris_gbs.

Creation of the pseudo-reference genome—Because there is no reference genome available for *Dryopteris*, I constructed a pseudo-reference using the sequences from the diploid taxa. I chose to use only the diploid taxa because they contain the majority of the sequence variation present in the tetraploids (which are known to be hybrids between the diploids), but harbor none of the possibly divergent sequences that may be present in the

tetraploid species. I also performed the analyses separately on each tetraploid clade: *Dryopteris celsa* and its progenitors, and *D. campyloptera* and its progenitors. These two clades are separated by almost 40 million years of evolution (Sessa et al., 2012b).

I created two pseudo-references (one for each clade) by first clustering highly similar sequences for each diploid taxon separately using VSEARCH v 2.4.2 (Rognes, Flouri, Nichols, Quince, & Mahe, 2016). Clustering was done at 92% similarity to create centroids for further clustering. I then clustered at 84% similarity and removed all sequences that collapsed at this stage, to exclude paralogs. I next combined the two diploid taxa from a given clade (either *D. intermedia* and *D. expansa*, or *D. ludoviciana* and *D. goldiana*) and clustered again using VSEARCH at 84% similarity; the resulting sequences were then used as the two pseudo-reference genomes.

Alignment of reads and variant calling—Before calling variants for all of the included species, I first had to index the pseudo-reference genome from the previous step, which identifies sequence position points for the alignment. This was done using the INDEX function of BWA v. 0.7.10 (H. Li & Durbin, 2009). Next I used PicardTools v. 2.9.0 (Institute, 2019) to create a sequence dictionary, and the INDEX function of SAMTOOLS v. 1.5 (H. Li et al., 2009) was then used to create a FASTA index file. I used BWA ALN and SAMSE to align all individual reads to the appropriate pseudo-reference. Next, I used the SAMTOOLS functions VIEW, SORT, and INDEX to prepare all individual reads for variant calling. To call variants, I used two different methods. First, I used the GATK HaplotypeCaller v. 3.8.0 (McKenna et al., 2010) to call variants on all samples as diploids. I then used VCFTOOLS v. 0.1.15 (Danecek et al., 2011) to filter the resulting VCF file. For the second approach, I again used the GATK HaplotypeCaller, but called variants separately on the diploids and tetraploids (HaplotypeCaller allows the user to specify any ploidy). I then used a custom Python script to filter the resulting VCF files, as VCFTOOLS does not support polyploids. Finally, I used custom Perl (v. 5.15.3, www.perl.org) scripts to find the intersection of SNPs from the diploids and tetraploids.

Admixture analysis—I used the population genetics program ENTROPY v. 1.2 (Gompert et al., 2014), which is very similar to the popular program STRUCTURE (Pritchard, Stephens, & Donnelly, 2000). While both are Bayesian, model-based approaches to population genetics, a key difference between them is that STRUCTURE assumes that individual genotypes are known a priori, whereas ENTROPY does not. ENTROPY calculates genotype likelihoods from raw sequence data and quality estimates; these genotype likelihoods are then used as an input for the model. STRUCTURE calculates the likelihoods repeatedly at each MCMC step from prior genotype assignments (Gompert et al., 2014).

The first step of the ENTROPY analysis was to convert the VCF variant file to a Genotype Likelihood (GL) file format using a custom Perl script. A second Perl script was used to convert the GL file to a matrix for input to R v. 3.5.1 (R Core Team, 2016). I used the R package ADEGENET v. 2.1.1 (Jombart, 2008) to perform a discriminant analysis of principal components (DAPC) to find the most likely source population for each individual. This analysis is similar to that performed by ENTROPY, but is less complex, and generates starting values that can be used to seed ENTROPY, which helps eliminate label swapping and allows the MCMC analysis to converge more quickly. I followed the DAPC protocol of Jombart and Collins (available at: adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf).

I ran ENTROPY for $K = 2$, 3, 4, and 5 with 2 chains in each analysis, and I examined the results obtained with and without the DAPC starting values, with high and low burn-in, and with various numbers of iterations. I found that the various values of these parameters did very little to alter the results, and so decided to run the final analysis using the starting values and with a large number of iterations (65,000) and a high burn-in (15,000). I used the program ESTPOST v. 1.2. (Gompert et al., 2014) to extract admixture proportions for each individual, and then used a custom R script and functions to visualize the ENTROPY output.

Finally, I performed a principal component analysis on the ENTROPY output. I used the program ESTPOST to extract genotype probabilities for $K = 2 - 5$ and then read those data into R. I averaged the genotype probabilities for all K values, and then performed a

PCA using the R function prcomp. All of the R code used to analyze and visualize the data is available as an Rmarkdown file on Github (github.com/sylviakinosian/dryopteris_gbs).

3.4 Results

The pseudo-reference genomes constructed for the two clades differed in the relative evenness of contributions from the two sets of diploids: for the *intermedia - expansa - campyloptera* clade, 837,895 and 64,631 contigs were retrieved from *D. intermedia* and *D. expansa*, respectively. In the *ludoviciana - goldiana - celsa* clade, the contigs were a bit more evenly divided, with 440,468 retrieved from *D. goldiana* and 132,480 from *D. ludoviciana*.

The first analysis, with variants called on all samples as diploids, retained 1288 SNPs for the *intermedia - expansa - campyloptera* clade and 2122 SNPs for the *ludoviciana - goldiana - celsa* clade. The second analysis, which called variants for diploids and tetraploids separately, retained 3964 SNPs for the *intermedia - expansa - campyloptera* clade, and 4419 SNPs for the *ludoviciana - goldiana - celsa* clade. I performed the ENTROPY analysis on both sets of SNPs, and although calling variants separately on diploids and tetraploids obviously increased the total number of SNPs retained for both clades, I did not see a marked difference in the ENTROPY results between the two variant calling routines. I used the set of SNPs called only as diploids in the final analyses reported here.

ENTROPY analyses of the genotype data generally separate the diploid taxa from one another, but for both clades, the polyploid species contain genomic contributions primarily from one parent or the other (in the case of *Dryopteris* celsa), or they contain a substantial fraction from one parent as well as additional fractions that are not attributed to the second parent (in the case of *D. campyloptera*, Fig. 3.2). For the *D. campyloptera* samples, the diploid *D. intermedia* was the dominant contributor, with some small contributions from the other diploid parent, *D. expansa*, and additional fractions unassigned to either diploid.

The principal component analysis yielded a similar pattern to the ENTROPY analysis. In the *D. expansa - intermedia - campyloptera* clade, the *D. campyloptera* individuals clustered most closely with the *D. intermedia* individuals (Fig. 3.3a). In the *D. ludoviciana - goldiana - celsa* clade, *D. celsa* clustered with both progenitor species, although more

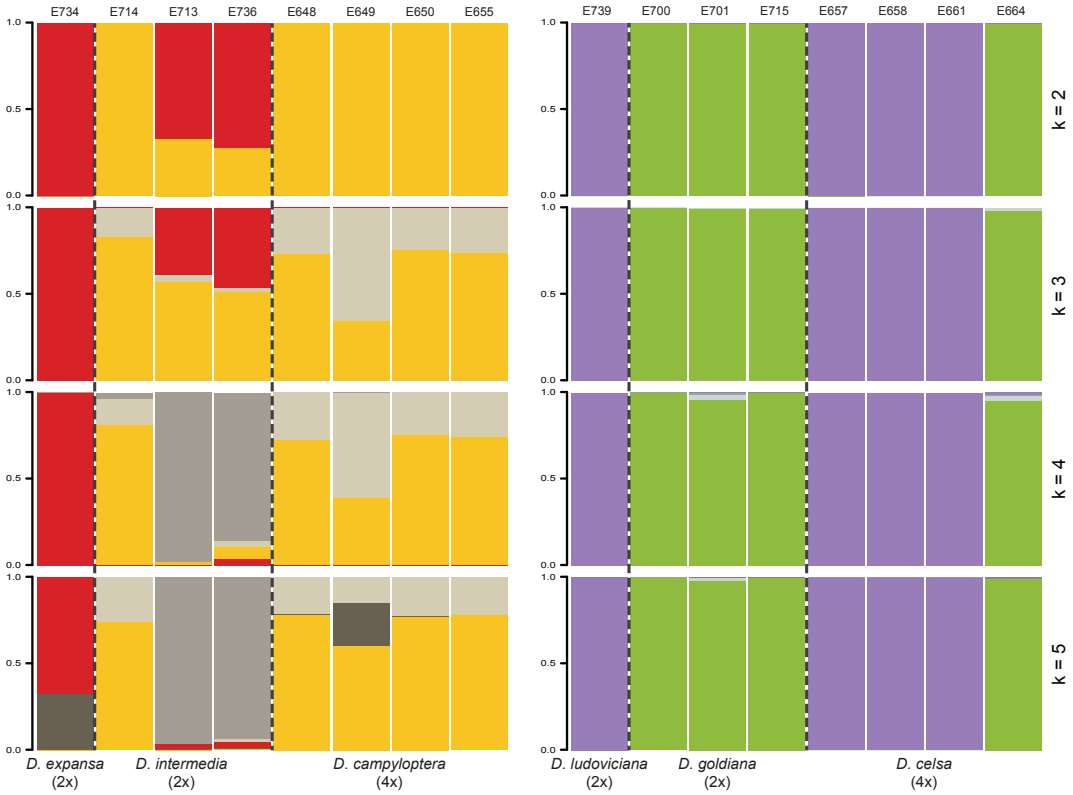


Fig. 3.2: Results of ENTROPY analysis of genotype data for two polyploid clades in North American *Dryopteris*. For each clade, I investigated admixture proportions between the species of each clade with $K = 2 - 5$ ancestral source populations. My expectation, given that I know the parentage of the tetraploids in both clades, would have been to see equal representation of the two diploid progenitors in the genetic makeup of each tetraploid sample. In contrast to that expectation, I found that for the *D. campyloptera* clade, at all K values, *D. intermedia* was the greatest contributor to the genome of the tetraploid hybrid *D. campyloptera*, and in the *D. celsa* clade, while the two diploid progenitors were both found to contribute to the tetraploid samples, it was always as an overwhelming contribution towards one parent or the other, in each of the samples.

individuals clustered with *D. ludoviciana* individuals than with *D. goldiana* (Fig. 3.3b).

3.5 Discussion

3.5.1 *Dryopteris campyloptera* clade

For all values of K tested, SNPs associated with the diploid *Dryopteris intermedia* dominate the genetic complement of the tetraploid samples (Fig. 3.2, 3.3). *Dryopteris*

expansa, the other diploid parent, has a distinct genotype that rarely appears in any of the tetraploids. I had an unequal number of samples from the two diploid parents in this group: three from *D. intermedia* and one from *D. expansa*. When constructing the reference genome for this clade, this unevenness would have resulted in a majority of the sequences belonging to the *D. intermedia* genome, as I saw – there were nearly 13 times as many contigs from *D. intermedia* in the pseudo-reference genome as there were from *D. expansa*. Consequently, most of the SNPs called were from this progenitor. Some SNPs associated with *D. expansa* do occur in each of the *D. campyloptera* samples at $K = 3$, 4, and 5 (visible as extremely thin red lines at the tops of the stacked bars for those K values in Fig. 3.2), but these are extremely minor contributions to the tetraploid genome samples. At $K = 5$, a more substantial *D. expansa* contribution occurs in one of the tetraploids (sample E649), further supporting the hypothesis that only a small number of *D. expansa* sequences made it into the reference and subsequent SNP calling.

An additional genetic component found in the *D. campyloptera* samples at $K = 3$ – 5 (light grey sections in Figure 3.2) may be due to sequence divergence and accumulation of mutations in cut sites that would have occurred subsequent to the formation of the polyploid species, whose earliest inferred age of formation is 4.6 million years ago (Sessa et al., 2012b). That is an adequate amount of time for numerous mutations to occur that would alter the location and/or frequency of cut sites, and this would potentially produce fragments in the polyploids with lengths and sequence compositions not found in either parent species. Considering this, *D. campyloptera* sample E649 perhaps best represents the a priori expectations for a successful result from this study, as it includes sets of SNPs found in both parents as well as an additional component not found in either diploid.

I also found evidence of genetic diversity in one of the diploid parents in this clade, *Dryopteris intermedia*. The three samples of *D. intermedia* differ considerably from one another in both the ENTROPY admixture analysis and the PCA (Fig. 3.2, 3.3), and at $K = 2$ a large portion of SNPs in two of the *D. intermedia* samples are associated with the *D. expansa* sample. At higher K values, the three *D. intermedia* samples become much more

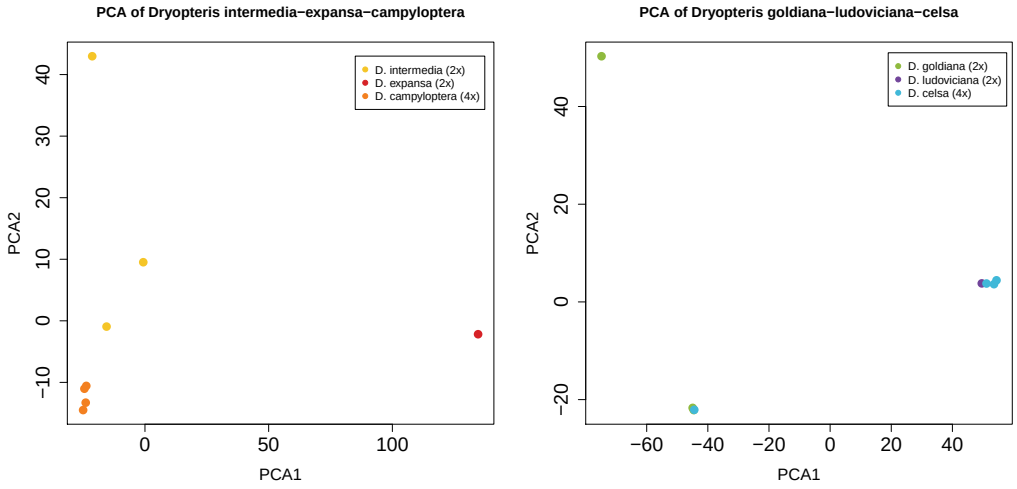


Fig. 3.3: Principal component analyses for the two clades. Legends indicate species names and number of samples. A: In the *D. expansa* - *intermedia* - *campyloptera* clade, the tetraploid *D. campyloptera* clusters most closely with *D. intermedia*, as in the admixture analysis (Fig. 2). There is also reasonable genetic variation in the progenitor *D. intermedia*. B: In the *D. ludoviciana* - *goldiana* - *celsa* clade, the tetraploid *D. celsa* clusters with both progenitors. Interestingly one progenitor, *D. goldiana*, shows more variation here than in the admixture plot (Fig. 3.2). Clusters of points in the lower left-hand corners of both plots have been jittered slightly to make the points distinguishable.

distinct from *D. expansa*; at $K = 4$ and 5, two of the *D. intermedia* samples resemble each other closely, while the third is distinct and resembles the *D. campyloptera* samples. These results suggest that there is substantial genetic diversity within and between populations of *Dryopteris intermedia*, which is unsurprising given its wide geographic range (Montgomery & Wagner, 1993). Both *D. intermedia* and *D. expansa* diverged from their closest diploid relatives in the Miocene (ca. 10–15 mya; (Sessa et al., 2012a)), and it therefore seems likely that I might discover additional genetic variation in *D. expansa* as well, if I sampled additional populations across its broad range.

3.5.2 *Dryopteris celsa* clade

For this clade, I also had three samples from one parent (*D. goldiana*) and one from the other (*D. ludoviciana*), but there was less apparent sampling bias of the under-represented parent in the pseudo-reference assembly. The balance between the number of contigs these

species contributed to the reference was more equal than in the *D. campyloptera* clade: after clustering, 440,468 contigs were retrieved from the *D. goldiana* individuals and 132,480 from *D. ludoviciana*. In all of the analyses of this clade ($K = 2 - 5$), *D. goldiana* and *D. ludoviciana* are assigned to separate populations; three of the *D. celsa* individuals are assigned to the "*ludoviciana*" population, and one is assigned to the "*goldiana*" population. There is less apparent bias in these assignments than there was in the *D. campyloptera* clade, in that both parental genotypes are present in the tetraploids. However, the results are clearly at odds with the expectations, which would have been for each of the tetraploid samples to show clear evidence of genetic contributions from (at least) two sources, rather than being dominated by only one.

In this clade I also saw some evidence of genetic variation in one progenitor species, *Dryopteris goldiana*. This was most evident from the PCA (Fig. **??**), where one of the *D. goldiana* individuals clustered very differently from the other two. *Dryopteris goldiana* has a broad geographic range, but is regionally rare and locally abundant across its distribution, and is found only in rich woods and ravines (Montgomery & Wagner, 1993). This could potentially isolate populations from one another, accounting for the diversity observed in the PCA results. *Dryopteris ludoviciana* has a much narrower range than any of the other three progenitor species in this study (Montgomery & Wagner, 1993), and further sampling would be required to reveal whether it has a similar amount of genetic variation as the other progenitor in this particularly clade.

3.5.3 RAD data and analysis of polyploid complexes

The goal of this research was to assess the utility of RAD sequence data for identifying progenitors in a polyploid complex. The dataset, which consisted of a small number of samples from a group with known polyploids and no reference genome, is typical for what might be available to researchers interested in polyploid ferns, and thus serves as a test case for these types of analyses. Based on the results discussed above, the success was mixed, and interpretation of the results was greatly facilitated by knowing in advance the progenitors of the allopolyploid species. I suspect that studies attempting to use a RAD-based approach to

determine relationships in a polyploid complex where progenitors were not known would face substantial challenges and likely obtain at least somewhat ambiguous results. Nonetheless, for researchers interested in these methods despite their potential shortcomings, there are several aspects of the study, including features of the species complex itself and of the experimental design, that are informative and which I discuss below.

As mentioned earlier in the discussion, for both tetraploid complexes there was a disparity in the sampling of the parents: both groups had three samples from one parent and one sample from the other. When building the pseudo-reference genomes for each clade, I attempted to balance the number of contigs contributing to the reference from both diploid species equally. This was done using a Perl script to find the intersection of contigs from two species after the final clustering step. However, at each locus that matched the search criteria, I used the consensus sequence created by VSEARCH. In many cases, these consensus sequences were built from clusters of contigs that had a greater representation of one diploid over the other, making the consensus somewhat biased toward one parent. There are a handful of programmatic ways to remedy this using custom scripts, but perhaps the most practical and effective would be to increase the sample size and better balance sampling from the two progenitors. This would not only help create consensus sequences that are built of similar numbers of contigs from each species, but would also increase the size of the pseudo-reference and consequently the number of SNPs that could be called.

RAD approaches were originally developed to address questions about population genetics (Andrews et al., 2016; Rowe et al., 2011), and are therefore most informative across relatively short evolutionary distances, typically for individuals whose maximum divergence is between 5 and 10 mya. GBS approaches have been successful at analyzing hybrid complexes that are more recently diverged than the *Dryopteris* system, for example in *Juglans* (Zhao et al., 2018). In the present study, the two clades of *Dryopteris* are separated by about 40 million years of evolution, and even within the hybrid clades, roughly 15–25 million years have elapsed since the divergence of the diploid progenitors (Sessa et al., 2012a, 2012b). Mutations can start to accumulate in enzyme cut sites that can result in non-homologous

fragments being cut and amplified across species, which becomes more likely the longer ago the species diverged (Eaton, Spriggs, Park, & Donoghue, 2017). This is perhaps part of the reason that the results did not reflect particularly well the known relationships between the tetraploid *Dryopteris* species and their progenitors.

To investigate the time scales at which RAD approaches are most effective, Eaton et al. (2017) examined RADSeq data across several lineage with a range of divergence times. They found that phylogenetic distance was not a good predictor of the number of SNPs recovered; uneven sequence coverage was found to have the most impact on missing data. They also found that as sample size increased, some loci that were initially found as singletons in small datasets were recovered in more individuals, thereby becoming phylogenetically informative. This suggests that sample size might be the biggest driver of the weak results, as potentially informative SNPs may only occur in one of the sampled individuals. The small size of the dataset, both in terms of numbers of samples and numbers of species (two sets of three species), also influenced the ability to use additional programs available for analyzing SNP data. For example, phylogenetically-informed approaches such as HyDe and PhyloNet are unlikely to be informative with so few samples and clusters of only three taxa. In some of the datasets examined by Eaton et al. (2017), sequencing fewer loci at a higher coverage (essentially the strategy I employed in the present study) resulted in large amounts of missing data for highly divergent lineages. Even though I did not have a large amount of missing data across samples, the high specificity of the double digest method may have been problematic for such a highly divergent group as *Dryopteris*.

3.5.4 Summary and future directions

Ferns are known to hybridize across vast evolutionary distances (Rothfels, Johnson, et al., 2015; Sessa, Vicent, Chambers, & y Galán, 2018), and so sequencing tools are needed that are informative across diverse and highly divergent groups. RAD methods may be more informative for studying polyploid or hybrid complexes that have formed more recently than the North American *Dryopteris* complex, but by adjusting the RAD methodology as discussed in Eaton et al. (2017), this type of reduced representation may indeed be a good

option for future investigations of deeply divergent fern species and their hybrids (assuming that adequate sample size can be achieved).

In addition to reduced representation methods, there are several next generation sequencing techniques that could potentially be utilized to explore complexes involving deeply divergent fern lineages. Ultraconserved elements (UCEs) have proven useful for investigating deep divergences in animal lineages, but are not a viable option for plants due to many ancient polyploidy events (Jiao et al., 2011), which can fracture and rearrange the genome, making UCEs difficult or impossible to isolate (Reneker et al., 2012). Target sequence capture methods (TSC) and exome sequencing currently seem to be the most promising methods for use in this field. TSC methods, which use probes or baits to preferentially amplify and sequence specific regions of the genome (often low or single copy nuclear genes) have been shown to be effective at resolving phylogenetic relationships across ferns (Wolf et al., 2018), and several methodologies are available for developing baits from transcriptome or genome sequences (Wolf et al., 2018; Yang & Smith, 2014; Zimmer & Wen, 2015). Exome capture kits have been designed for several polyploid crop plants (Sulonen et al., 2011), and could potentially be useful for fern population genetics as well. Both TSC and exome capture methods provide large amounts of data that should be informative across the evolutionary time scales needed to investigate the hybrid complexes that are so common in ferns, such as in North American *Dryopteris*.

References

Andrews, K., Good, J., Miller, M., Luikart, G., & Hohenlohe, P. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92.

Bainard, J., Henry, T., Bainard, L., & Newmaster, S. (2011). DNA content variation in monilophytes and lycophytes: Large genomes that are not endopolyploid. *Chromosome Research*, *19*(6), 763–775.

Beck, J., Windham, M., Yatskievych, G., & Pryer, K. (2010). A diploids-first approach to species delimitation and interpreting polyploid evolution in the fern genus *Astrolepis* (pteridaceae). *Systematic Botany*, *35*(2), 223–234.

Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ramdoss, Y., & Buckler, E. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140.

Danecek, P., Auton, A., Abecasis, G., Albers, C., Banks, E., DePristo, M., ... Group, . G. P. A. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.

Eaton, D., Spriggs, E., Park, B., & Donoghue, M. (2017). Misconceptions on missing data in RADseq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, *66*(3), 399–412.

Gastony, G., & Yatskievych, G. (1992). Maternal inheritance of the chloroplast and mitochondrial genomes in Cheilanthoid ferns. *American Journal of Botany*, *79*(6), 716.

Gompert, Z., Lucas, L., Buerkle, C., Forister, M., Fordyce, J., & Nice, C. (2014). Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology*, *23*(18), 4555–4573.

Grusz, A., Windham, M., & Pryer, K. (2009). Deciphering the origins of apomictic polyploids in the *Cheilanthes yavapensis* complex (Pteridaceae). *American Journal of*

Botany, *96*(9), 1636–1645.

Institute, B. (2019). *Picard toolkit (version 2.9.0). computer software.* Broad Institute, GitHub Repository.

Ishikawa, H., Watano, Y., Kano, K., Ito, M., & Kurita, S. (2002). Development of primer sets for PCR amplification of the *PgiC* gene in ferns. *Journal of Plant Research*, *115*(1117), 65–70.

Jiao, Y., Wickett, N., Ayyampalayam, S., Chanderbali, A., Landherr, L., Ralph, P., ... dePamphilis, C. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, *473*(7345), 97–100.

Jombart, T. (2008). Adegenet: An R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Jorgensen, S., & Barrington, D. (2017). Two beringian origins for the allotetraploid fern *Polystichum braunii* (Dryopteridaceae). *Systematic Botany*, *42*(1), 6–16.

Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., de Vries, J., Delaux, P.-M., ... et al., P. H. (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants*, *4*(7), 460–472.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Subgroup, . G. P. D. P. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... DePristo, M. (2010). The Genome Analysis Toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303.

Montgomery, J., & Wagner, W. (1993). *Dryopteris in flora of North America north of Mexico. Vol. 2.* Oxford University Press, New York, New York.

Peterson, B., Weber, J., Kay, E., Fisher, H., & Hoekstra, H. (2012). Double digest RADseq:

An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *Plos One*, *7*(5), e37135.

Pritchard, J., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Reneker, J., Lyons, E., Conant, G., Pires, J., Freeling, M., Shyu, C.-R., & Korkin, D. (2012). Long identical multispecies elements in plant and animal genomes. *Proceedings of the National Academy of Sciences*, *109*(19), E1183-91.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahe, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584.

Rothfels, C., Johnson, A., Hovenkamp, P., Swofford, D., Roskam, H., Fraser-Jenkins, C., ... Pryer, K. (2015). Natural hybridization between genera that diverged from each other approximately 60 million years ago. *The American Naturalist*, *185*(3), 433–442.

Rothfels, C., Li, F.-W., Sigel, E., Huiet, L., Larsson, A., Burge, D., ... et al., P. K. (2015). The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *American Journal of Botany*, *102*(7), 1089–1107.

Rothfels, C., Pryer, K., & Li, F.-W. (2017). Next-generation polyploid phylogenetics: Rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *The New Phytologist*, *213*(1), 413–429.

Rowe, H., Renaut, S., & Guggisberg, A. (2011). RAD in the realm of next-generation sequencing technologies. *Molecular Ecology*, *20*(17), 3499–3502.

Schuettpelz, E., Grusz, A., Windham, M., & Pryer, K. (2008). The utility of nuclear *GapCp* in resolving polyploid fern origins. *Systematic Botany*, *33*(4), 621–629.

Sessa, E., Banks, J., Barker, M., Der, J., Duffy, A., Graham, S., ... Wolf, P. (2014). Between two fern genomes. *GigaScience*, *3*(15).

Sessa, E., & Givnish, T. (2014). Leaf form and photosynthetic physiology of *Dryopteris* species distributed along light gradients in eastern North America. *Functional Ecology*, *28*(1), 108–123.

Sessa, E., Vicent, M., Chambers, S., & y Galán, J. G. (2018). Evolution and reciprocal

origins in mediterranean ferns: The *Asplenium obovatum* and *a. adiantum-nigrum* complexes. *Annals of the Missouri Botanical Garden*, *103*, 13–25.

Sessa, E., Zhang, L.-B., Väre, H., & Juslén, A. (2015). What we do (and don't) know about ferns: *Dryopteris* (Dryopteridaceae) as a case study. *Systematic Botany*, *40*(2), 387–399.

Sessa, E., Zimmer, E., & Givnish, T. (2012a). Phylogeny, divergence times, and historical biogeography of new world *Dryopteris* (Dryopteridaceae). *American Journal of Botany*, *99*(4), 730–750.

Sessa, E., Zimmer, E., & Givnish, T. (2012b). Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology*, *12*(104).

Sessa, E., Zimmer, E., & Givnish, T. (2012c). Reticulate evolution on a global scale: A nuclear phylogeny for new world *Dryopteris* (Dryopteridaceae). *Molecular Phylogenetics and Evolution*, *64*(3), 563–581.

Soltis, D., Visger, C., & Soltis, P. (2014). The polyploidy revolution then...and now: Stebbins revisited. *American Journal of Botany*, *101*(7), 1057–1078.

Sulonen, A.-M., Ellonen, P., Almusa, H., Lepistö, M., Eldfors, S., Hannula, S., ... Saarela, J. (2011). Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biology*, *12*(9), R94.

Team, R. D. C. (2016). *R: A language and environment for statistical computing (version 3.3.2), computer software.* R Foundation for Statistical Computing, Vienna, Austria.

Testo, W., Watkins, J., & Barrington, D. (2015). Dynamics of asymmetrical hybridization in North American wood ferns: Reconciling patterns of inheritance with gametophyte reproductive biology. *New Phytologist*, *206*(2), 785–795.

Vogel, J., Russell, S., Rumsey, F., Barrett, J., & Gibby, M. (1998). Evidence for maternal transmission of chloroplast DNA in the genus *Asplenium* (Aspleniaceae, Pteridophyta). *Botanica Acta*, *111*, 247–249.

Walker, S. (1955). Cytogenetic studies in the *Dryopteris spinulosa* complex i. *Watsonia*, *3*, 193–209.

Walker, S. (1959). Cytotaxonomic studies of some american species of *Dryopteris*. *American Fern Journal*, *49*(3), 104.

Walker, S. (1961). Cytogenetic studies in the *Dryopteris spinulosa* complex. II. *American Journal of Botany*, *48*(7), 607.

Walker, S. (1962). Further studies in the genus *Dryopteris*: The origin of *D. clintoniana*, *D. celsa*, and related taxa. *American Journal of Botany*, *49*(5), 497–503.

Walker, S. (1969). Identification of a diploid ancestral genome in the *Dryopteris spinulosa* complex. *British Fern Gazette*, *10*(2), 97–99.

Wolf, P., Robison, T., Johnson, M., Sundue, M., Testo, W., & Rothfels, C. (2018). Target sequence capture of nuclear-encoded genes for phylogenetic analysis in ferns. *Applications in Plant Sciences*, *6*(5), e01148.

Wood, T., Takebayashi, N., Barker, M., Mayrose, I., Greenspoon, P., & Rieseberg, L. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, *106*(33), 13875–13879.

Yang, Y., & Smith, S. (2014). Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: Improving accuracy and matrix occupancy for phylogenomics. *Molecular Biology and Evolution*, *31*(11), 3081–3092.

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina Paired-End ReAd MergeR. *Bioinformatics*, *30*(5), 614–620.

Zhao, P., Zhou, H.-J., Potter, D., Hu, Y.-H., Feng, X.-J., Dang, M., ... Woeste, K. (2018). Population genetics, phylogenomics and hybrid speciation of *Juglans* in china determined from whole chloroplast genomes, transcriptomes, and genotyping-by-sequencing (GBS). *Molecular Phylogenetics and Evolution*, *126*, 250–265.

Zimmer, E., & Wen, J. (2015). Using nuclear gene data for plant phylogenetics: Progress and Prospects II. Next-gen approaches. *Journal of Systematics and Evolution*, *53*(5), 371–379.

CHAPTER 4

Cryptic diversity in the model fern genus *Ceratopteris* (Pteridaceae)

Published in Molecular Phylogenetics and Evolution[1]

4.1 Abstract

Cryptic species are present throughout the tree of life. They are especially prevalent in ferns, because of processes such hybridization, polyploidy, and reticulate evolution. In addition, the simple morphology of ferns limits phenotypic variation and makes it difficult to detect cryptic species. The model fern genus *Ceratopteris* has long been suspected to harbor cryptic diversity, in particular within the highly polymorphic *C. thalictroides*. Yet no studies have included samples from throughout its pan-tropical range or utilized genomic sequencing, making it difficult to assess the full extent of cryptic variation within this genus. Here, I present the first multilocus phylogeny of the genus using reduced representation genomic sequencing (RADseq) and examine population structure, phylogenetic relationships, and ploidy level variation. I recover similar species relationships found in previous studies, find support for the cryptic species *C. gaudichaudii* as genetically distinct, and identify a novel genomic variation within two of the mostly broadly distributed species in the genus, *C. thalictroides* and *C. cornuta*. Finally, I detail the utility of the approach for working on cryptic, reticulate groups of ferns. Specifically, it does not require a reference genome, of which there are very few available for ferns. Next generation sequencing like RADseq is a cost-effective way to obtain the thousands of nuclear markers needed untangle the many species complexes present in ferns.

[1]Authors: Sylvia Kinosian, William Pearse, and Paul Wolf

4.2 Introduction

Describing the formation of species is essential to understanding the patterns and processes that create biodiversity. Detecting differences among species becomes increasingly challenging when taxa are morphologically similar, but genetically distinct and reproductively isolated (Bickford, Lohman, Sodhi, Ng, Meier, Winker, Ingram, & Das, 2007; Masuyama, 1992; Paris, Wagner, & Wagner, 1989). Such cryptic species have often been historically described as one larger species, or species complex, due to morphological similarities (Paris et al., 1989). However, advances in molecular methods have revealed that cryptic species can be monophyletic with disparate morphologies, ecological niches and functions (Amato, Kooistra, Ghiron, Mann, Proschold, & Montresor, 2007; Hebert, Penton, Burns, Janzen, & Hallwachs, 2004; Sattler, Bontadina, Hirzel, & Arlettaz, 2007; Southgate, Patel, & Barrington, 2019). In addition, cryptic species complexes can also be paraphyletic and separated by considerable evolutionary time, yet look similar and occupy comparable niches (Amor, Norman, Cameron, & Strugnell, 2014; Cunnington, Lawrie, & Pascoe, 2005).

Despite the recent shift in molecular approaches to investigate cryptic species, much is still unknown about their evolutionary history and ecosystem functions. Cryptic species complexes often have very broad distributions (Der, Thomson, Stratford, & Wolf, 2009; Knowlton, 1993; Nygren, 2014) and can occur in sympatry, allopatry, or parapatry (Bickford et al., 2007), making adequate sampling challenging. Studies have shown that cryptic species are prevalent across the Tree of Life (e.g. Bickford et al. (2007); del Carmen Molina, Divakar, Millanes, Sanchez, Ruth, Hawksworth, & Crespo (2011); Hebert et al. (2004); Nygren (2014)), and ferns in particular are known to harbor many lineages with cryptic variation (Adjie, Masuyama, Ishikawa, & Watano, 2007; Paris et al., 1989; Yatabe, Shinohara, Matsumoto, & Murakami, 2009). However, deciphering species boundaries in ferns can be quite difficult due to their ease of dispersal via spores, making gene flow possible across vast distances (Barrington, 1993; Tryon, 1970). Compared to seed plants, ferns have a high prevalence of polyploidy and reticulate evolution (Barrington, Haufler, & Werth, 1989; Otto & Whitton, 2000; Paris et al., 1989; Sigel, 2016). In addition, changes in ploidy

can contribute to cryptic variation in ferns by altering niche space or offspring viability (Masuyama, Yatabe, Murakami, & Watano, 2002; Otto, 2007; Southgate et al., 2019), yet have small effects on phenotype (Patel, Fawcett, & Gilman, 2019). One example is the widespread fern genus *Ceratopteris* Brong. (Pteridaceae), which is understood to have considerable cryptic variation, hybridization and polyploidy (Adjie et al., 2007; Lloyd, 1974; Masuyama & Watano, 2010), making it an ideal system to investigate the origins and ecology of cryptic species.

The fern genus *Ceratopteris* is a pan-tropical aquatic clade consisting of seven named species, three of which are cryptic and tetraploid (Lloyd, 1974; Masuyama & Watano, 2010; PPGI, 2016). *Ceratopteris* is perhaps best known for the model organism *C. richardii*, which has been used as such since the late twentieth century (Banks, 1994; Hickok, Warne, & Fribourg, 1995; Hickok, Warne, & Slocum, 1987). Sometimes called the "*Arabidopsis* of the fern world" (Sessa, Banks, Barker, Der, Duffy, Graham, Hasebe, Langdale, Li, Marchant, Pryer, Rothfels, Roux, Salmi, Sigel, Soltis, Soltis, Stevenson, & Wolf, 2014), *C. richardii* is an ideal model system because of its fast life cycle and ease of cultivation (Hickok et al., 1987); in addition, it can be transformed with recombinant DNA (Muthukumar, Joyce, Elless, & Stewart, 2013; Plackett, Huang, Sanders, & Langdale, 2014), has a reference ontogeny framework (Conway & Di Stilio, 2019), and is currently the only homosporous fern to have a published genome sequence (Marchant, Sessa, Wolf, Heo, Barbazuk, Soltis, & Soltis, 2019). In addition to *C. richardii*, other *Ceratopteris* species have been studied in a lab environment (Hickok, 1977; Hickok & Klekowski, 1974), and have potential for further research. The species boundaries within the genus are blurry, however, and there is evidence that all species in the group (both diploid and tetraploid) hybridize to some extent (Adjie et al., 2007; Hickok, 1977,7; Hickok & Klekowski, 1974; Lloyd, 1974). In such a well-utilized model genus, there is still a need to better understand species boundaries, evolutionary history, and occurrences of cryptic species. In particular, I need to understand the evolutionary dynamics of the *Ceratopteris* genus as a whole in order to best utilize this model system for future work.

The first and only comprehensive monograph for *Ceratopteris* was written by Lloyd in 1974, and employed a matrix of morphological traits to identify four species in the genus. However, *Ceratopteris* is a notorious group in terms of morphology: there are relatively few informative physical characters (Lloyd, 1974); in addition, habitat and developmental stage can have a large effect on plant phenotype (Masuyama, 1992). Most challenging for species delimitation is that rampant hybridization in the genus is known to further alter morphology and so make field identification difficult (Hickok & Klekowski, 1974; Lloyd, 1974; Masuyama & Watano, 2010). While Lloyd named four species, he noted that *C. thalictroides*, the most widespread species in the genus, is "highly polymorphic" (Lloyd, 1974) and is likely a cryptic species complex. In response, Masuyama and colleagues conducted a series of studies on Asian plants under the name of *C. thalictroides*, examining chloroplast DNA and cross-breeding (Masuyama et al., 2002), cytological characteristics (Masuyama & Watano, 2005), morphological traits (Masuyama, 1992, 2008), and nuclear DNA (Adjie et al., 2007). As a result, three cryptic species and two varieties were named from entities originally described as *C. thalictroides* (Masuyama & Watano, 2010). Molecular evidence suggests that all three cryptic species of *C. thalictroides* have independent hybrid allopolyploid origins and are paraphyletic (Adjie et al., 2007), the latter providing the most compelling evidence that they cannot be described under the same name. However, that study included only one nuclear marker; multilocus analyses are generally needed to substantiate paraphyly and polyploid origins (Eaton & Ree, 2013; Jorgensen & Barrington, 2017), and particularly in a case as challenging as *C. thalictroides*. In addition, while cryptic species in *Ceratopteris* have been investigated in the Old World, there is a similar problem of morphological variability in New World *C. thalictroides* (Lloyd, 1974; Masuyama & Watano, 2010), indicating a potential for more cryptic species to be discovered.

As a well-studied, pan-tropical group known to include cryptic species, *Ceratopteris*, and in particular *C. thalictroides*, is an ideal system in which to study the process of cryptic speciation. In this study, I produce the first multilocus genomic analysis of *Ceratopteris* using restriction-site associated DNA sequencing (RADseq). RADseq is a cost-effective

way to generate a large amount of genomic data, and many downstream analysis tools are currently available. I apply three approaches to identify cryptic species. First, I estimate population structure and hybridization; second, I reconstruct phylogenetic relationships among samples; finally, I investigate ploidy levels across individuals. This study is a step towards resolving species boundaries within this convoluted genus. This work will also provide a phylogenetic reference for future studies on *Ceratopteris*, as well as find areas within the genus in need of additional research.

4.3 Materials and methods

To best analyze the cryptic diversity within *Ceratopteris* I gathered samples from across its pan-tropical distribution. I utilized RADseq to generate a large genomic dataset, which was processed using the *ipyrad* pipeline (Eaton & Overcast, 2020) for downstream analysis. The three-part approach for assessing cryptic diversity takes advantage of the flexibility of RADseq data, and provides a unique window into the processes of speciation and diversification in a complex genus.

Specimen collection information can be found in Appendix A. All parameter values and code for data processing and downstream analyses can be found on GitHub (github.com/sylviakinosian/ceratopteris_RADseq). Demultiplexed, unfiltered reads can be accessed from the NCBI GenBank Short Read Archive (PRJNA606596).

4.3.1 Taxon sampling

Because of its broad distribution, I chose to use a combination of herbarium and silica-dried material from field collections to sample *Ceratopteris* (Fig. 4.1). For all herbarium specimens included, I assigned them to species according to the labels provided by the collector. For the samples collected in the field from Costa Rica, Taiwan, and Australia I utilized the keys provided by Lloyd (1974) and Masuyama & Watano (2010) to identify each to species or subspecies, where applicable.

I collected 90 samples, 56 from herbarium specimens and 34 from silica-dried tissue. These samples represented five of seven named species of *Ceratopteris*, and one sample of

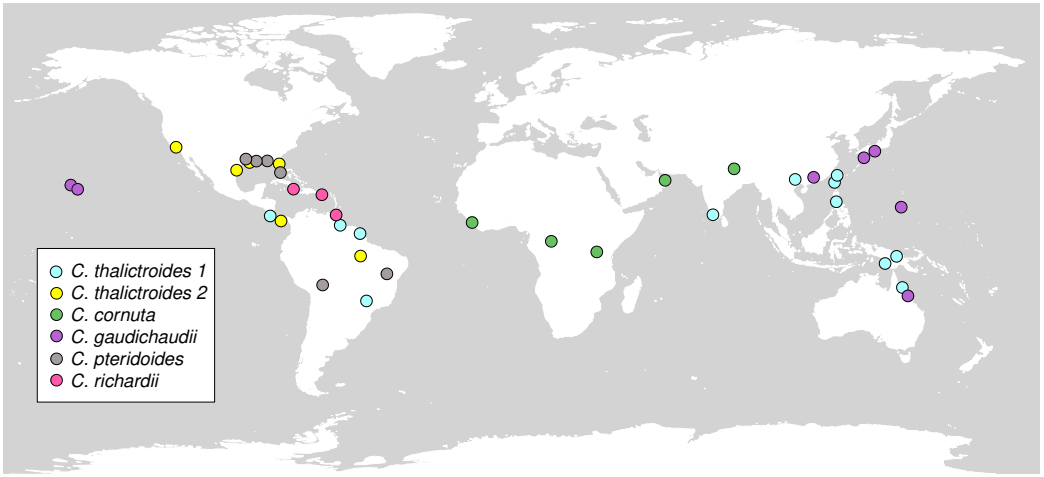


Fig. 4.1: Collection localities for samples of *Ceratopteris* included in the present study. These collections roughly indicate the range for each species. Not all accessions are included on this map because many herbarium specimens did not have GPS coordinates, or several samples were from nearly identical locations. Localities of *C. richardii* are those of sampled individuals, but were ultimately not included in the analyses.

Acrostichum aureum from Australia for use as an outgroup. Included in the silica-dried specimens was a sample of the common lab strain of *C. richardii* (*Hnn*). The three missing species of *Ceratopteris* were cryptic species of *C. thalictroides* (L.) Brongn. (Masuyama & Watano, 2010): *C. oblongiloba* Masuyama & Watano, known from Southeast Asia; *C. froesii* Brade, endemic to Brazil; and *C. shingii* Y. H. Yan & R. Zhang, endemic to Hainan Provence, China. These species have relatively narrow ranges and few herbarium species, especially *C. froesii*. I were unable to obtain any material from herbarium specimens labeled as either of these species, and none of the plants collected in the field keyed out to *C. oblongiloba* or *C. froesii*. *Ceratopteris shingii* was described as this manuscript was nearing completion, so no material could be included.

Herbarium specimens were chosen based on age (less than about 30 years old) and color (leaves still green). These specimens were collected from the Harvard University Herbaria (HUH), the Steere Herbarium at the New York Botanic Garden (NY), the University of California, Berkeley (UC), the United States National Herbarium at the Smithsonian (US), and the Pringle Herbarum at the University of Vermont (VT). Fresh tissue collections were

obtained from Taiwan, China, Costa Rica, and Australia. All field collections were stored on silica gel and vouchers were deposited at the Intermountain Herbarium (UTC) and James Cook University (JCT). See Appendix A for full details of specimen sources.

4.3.2 DNA extraction

Silica-dried plant tissue samples were supplied to the University of Wisconsin-Madison Biotechnology Center. DNA was extracted using the QIAGEN DNeasy mericon 96 QIAcube HT Kit, and quantified using the Quant-iT^{TM} PicoGreen® dsDNA kit (Life Technologies, Grand Island, NY). 17 specimens were extracted using a modified CTAB method (Doyle & Doyle, 1987) by SPK at Utah State University. These specimens were sent to the University of Wisconsin-Madison Biotechnology Center, analyzed for quality and then pooled with the rest of the samples.

4.3.3 Library construction and sequencing

Because species of *Ceratopteris* are known to hybridize (Hickok & Klekowski, 1974), but also form reproductively isolated cryptic species (Masuyama et al., 2002), I wanted to examine population structure of the genus across its distribution. The program STRUCTURE v. 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) is designed to determine admixture (hybridization) between populations of one or several closely related taxa. The program assumes that each individual's genome is a mosaic from K source populations and uses genotype assignments from SNPs to infer population structure and admixture. I utilized a set of 8478 SNPs, created in *pyrad* by selecting one SNP per loci. I ran STRUCTURE for K = 2 - 6 with 50 chains for each K. I then used CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) to process the STRUCTURE output and estimate the best K values (Evanno, Regnaut, & Goudet, 2005; Pritchard et al., 2000).

4.3.4 Species tree inference

To construct a phylogeny for the samples of *Ceratopteris*, I used the program TETRAD, which is included in the *ipyrad* analysis toolkit and is based on the software SVDQuartets

(Chifman & Kubatko, 2015). TETRAD utilizes the theory of phylogenetic invariants to infer quartet trees from a SNP alignment. Species relationships are estimated for all quartet combinations of individuals, then all quartet trees are joined into a species tree with the software wQMC (Avni, Cohen, & Snir, 2015). Quartet methods were designed to reduce computational time for tree-building with maximum likelihood methods (Ranwez & Gascuel, 2001), making them particularly useful for SNP datasets such as these.

I ran TETRAD via Python 2.7.16 for 50 samples of *Ceratopteris*, and one sample of *Acrostichum aureum* as the outgroup. The input consisted of 51 individuals and 26891 SNPs; I ran 100 bootstrap iterations for the final consensus tree. As a comparison, I also ran RAxML v. 8.2.11 (Stamatakis, 2014) with 1000 bootstrap iterations. I then plotted the final trees in R using the package phytools (Revell, 2012) and custom plotting functions.

4.3.5 Polyploidy analysis

Due to the presence of numerous cryptic species and morphological variation across the genus, I hypothesize there may be cryptic cytotype variation and/or ploidy type (allo vs. auto) within species, especially the polymorphic *C. thalictroides*. Since a majority the specimens were from herbarium accessions, I could not perform chromosome squashes. I utilized the R v. 3.5.2 (R Core Team 2018) package gbs2ploidy to infer ploidy variation across samples. This package was designed to detect ploidy levels in variable cytoptype populations (specifically quaking aspen, *Populus tremuloides*) using low-coverage (2X) genotyping-by-sequencing (GBS) or RADseq data; it was also tested on simulated data, and can detect diploids, triploids, or tetraploids (Gompert & Mock, 2017). This package estimates ploidy from allele ratios, which are calculated from genome-average heterozygosity and bi-allelic SNPs isolated from a variant calling format (VCF) file, which is included in the output from *ipyrad*.

To perform the analyses with gbs2ploidy, I first used the script vcf2hetAlleleDepth.py (github.com/carolrowe666/vcf2hetAlleleDepth) to convert the VCF file produced by *ipyrad* to the format needed for gbs2ploidy. I then ran gbs2ploidy, using the function estprops to estimate ploidy for each individual. I plotted the output from gbs2ploidy in R, with the

mean posterior probability for allelic ratios on the y axis and the 1:1, 2:1, and 3:1 ratios on the x axis (see Fig. 4.2); I also included error bars for the 95% equal tail probability intervals (ETPIs). I assigned ploidy to each individual using the highest posterior mean estimate for a certain allelic proportion. If the ETPIs overlapped, I considered it an ambiguous assignment. I required sequencing depth to be 6X or greater during the data processing, which suggests that the ploidy assignments are likely accurate.

4.3.6 Simulated RADseq analysis of the *Ceratopteris richardii* genome

Our sampling included relatively few individuals of *Ceratopteris richardii*, so I utilized the published genome from Marchant *et al.* (2019) to perform an *in silico* RADseq digest, creating a pseudo-individual to include in the analyses. I used the program ddRADseq-Tools (Mora-Márquez, García-Olivares, Emerson, & López de Heredia, 2017) to digest the genome with the same enzymes used in the *in vivo* digest (PstI and BfaI). Then I performed a simulated ddRADseq run with paired-end reads, demultiplexing, and trimming using ddRADseqTools. This simulated individual was added to the other samples, processed via the *ipyrad* pipeline, and analyzed via STRUCTURE.

4.4 Results

I retrieved an average of 2.58 x 10^6 raw reads per sample. From the 60 herbarium samples and 36 silica-dried specimens, I retrieved an average of 5954 and 15029 loci, respectively. On average, I recovered nearly three times the number of loci from silica-dried specimens compared to the herbarium specimens. The 17 herbarium samples had been previously extracted before sending to the University of Wisconsin; the remaining samples (of both herbarium and silica-dried tissue) were extracted from leaf tissue at UW facilities. The UW facilities obtained much higher-quality extracted DNA than I did in the own lab, which contributed to some of the disparity between herbarium and silica-dried DNA quality.

4.4.1 Population and genomic structure analysis

To determine the best K value across STRUCTURE I combined several metrics. I

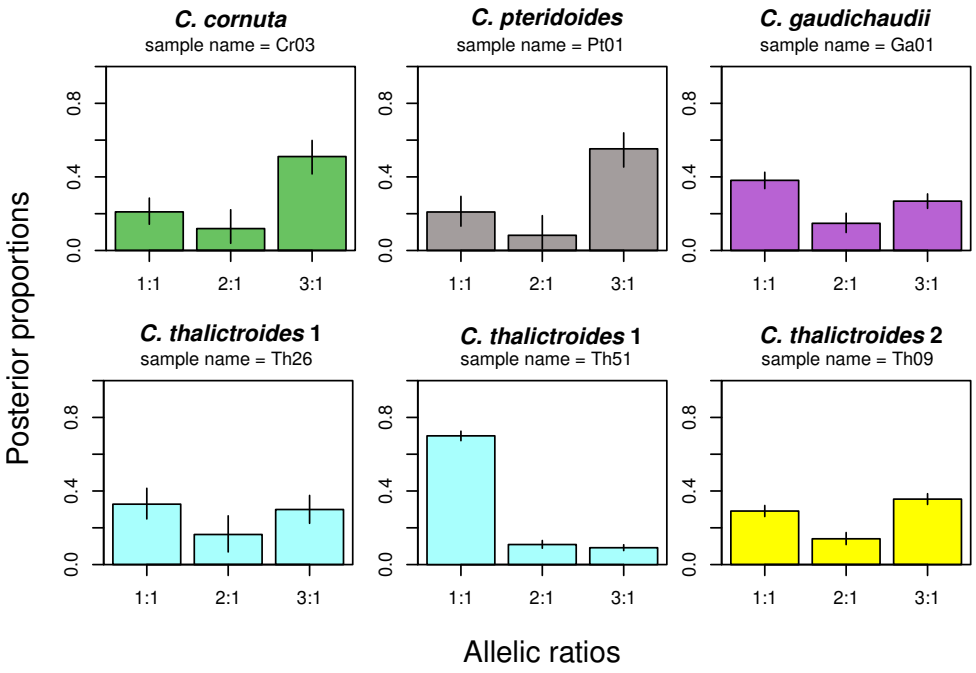


Fig. 4.2: Ploidy estimates for a subset of individuals. The X axis displays the estimated allele ratios (1:1, 2:1, 3:1), and the Y axis shows the Bayesian posterior proportion of allelic ratios. 95% equal-tail probability intervals (ETPIs) are shown on each bar; wherever these ETPIs do not overlap, there is a significant difference between the associated bars. The allelic ratio with the highest posterior estimate without an overlapping ETPI was used to assign ploidy to an individual. If the ETPIs overlapped, ploidy was assigned as ambiguous. *Ceratopteris cornuta* (green) and *C. pteridoides* (purple) are both known diploids ($n = 39$; (Hickok, 1977)). All specimens of these species had a 3:1 allele ratio, indicating potential homoploid hybridization, or genomic restructuring within diploids to yield such unbalanced allele ratios (Sigel, 2016). *C. gaudichaudii* and *C. thalictroides* are known tetraploids ($n = 77, 78$; (Masuyama & Watano, 2010)). *Ceratopteris thalictroides* 1 (grey) had two distinct patterns of allele ratios, a 1:1 (autopolyploid) group and a mixed to ambiguous group. The former is entirely from the Old World, whereas the latter group contains individuals from the Old and New Worlds. *Ceratopteris thalictroides* 2 (yellow) is only found in the New World, and all individuals had mixed 1:1 and 3:1 (allopolyploid) allele ratios. See Appendix A for all ploidy estimates.

used the best K method by Evanno *et al.* (2005), which determined a K value of 4. The best K estimate described in the STRUCTURE manual (Pritchard et al., 2000) determined a K value of 5. I also visually examined $K = 2$ - 6 and determined that $K = 5$ was the most biologically meaningful because it separated each named species, and also showed some intra-specific variability (Fig. 4.3). Increasing K past 5 did not add any meaningful population clusters (see Appendix A).

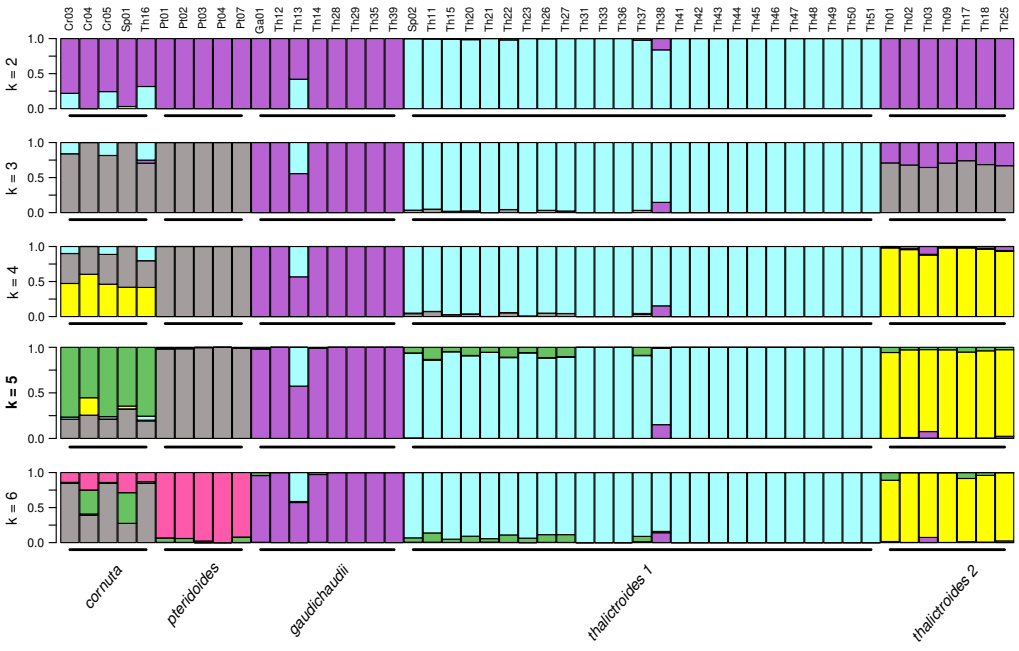


Fig. 4.3: STRUCTURE plots for $K = 2$ - 6. Each bar represents an individual, and each color is representative of a single genomic population. A single bar with multiple colored segments indicates ancestral admixture from K source populations. $K = 5$ was determined to be the most biologically informative. At $K = 5$, all named species of *Ceratopteris* fall into distinct populations. *Ceratopteris thalictroides* is split into three groups: Old World (blue), New World (yellow), and individuals grouping with *C. gaudichaudii* (purple). *Ceratopteris cornuta* is the only species with evidence of ongoing or recent admixture, most evident with *C. pteridoides*. At $K = 6$, the yellow bars may be indicative of *C. richardii*, but I not have a sample here to compare it to.

Named species of *Ceratopteris* mostly clustered together at $K = 5$ (Fig. 4.3). *Ceratopteris cornuta* and *C. pteridoides* appeared as distinct populations, although there was

some minor introgression from *C. pteridoides* into the *C. cornuta* population. *Ceratopteris thalictroides sensu latu* split in to three groups at $K = 5$. The first group consisted of several individuals identified as *C. thalictroides* (on herbarium sheets or in the field) that clustered with *C. gaudichaudii*, a cryptic species of *C. thalictroides* (Masuyama & Watano, 2010). The remaining individuals of *C. thalictroides* grouped into two populations, one consisting of mostly Old World individuals, and the other as entirely New World individuals. These two populations were strongly differentiated by STRUCTURE, being separated at all K values.

4.4.2 Species tree inference

The phylogeny resulting from the species quartet inference shows a similar pattern to the STRUCTURE output (Fig. 4.4). Individuals of *Ceratopteris pteridoides* form a monophyletic group, sister to the rest of the genus. *Ceratopteris cornuta* came out in two places in the phylogeny: sister to Old World *C. thalictroides*, and sister to a clade containing *C. gaudichaudii*, Old World and New World *C. thalictroides* clade. The phylogeny grouped all individuals placed in the *C. gaudichaudii* population by STRUCTURE in a monophyletic clade, sister to New World *C. thalictroides*.

The RAxML tree shows a very similar topology to the TETRAD tree. All clades present in the TETRAD tree are also present in the RAxML tree, with a slightly different topology: instead of New World *C. thalictroides* and *C. gaudichaudii* forming a monophyletic group, the former comes out as sister to all other species except *C. pteridoides*. Since these results are nearly identical, I present a graphical comparison of the two trees in Appendix A.

4.4.3 Ploidy assignment across species

Ceratopteris cornuta, *C. pteridoides*, and *C. richardii* are known as a diploid ($n = 39$) (Adjie et al., 2007; Hickok, 1977); *Ceratopteris thalictroides* and *C. gaudichaudii* are known as tetraploid, with some cytotypic variation ($n = 77, 78$) (Adjie et al., 2007; Masuyama & Watano, 2010). The ploidy analysis used allelic ratios to estimate the ploidy of samples. Diploids would have a 1:1 allele ratio in accordance with the two parental genomes

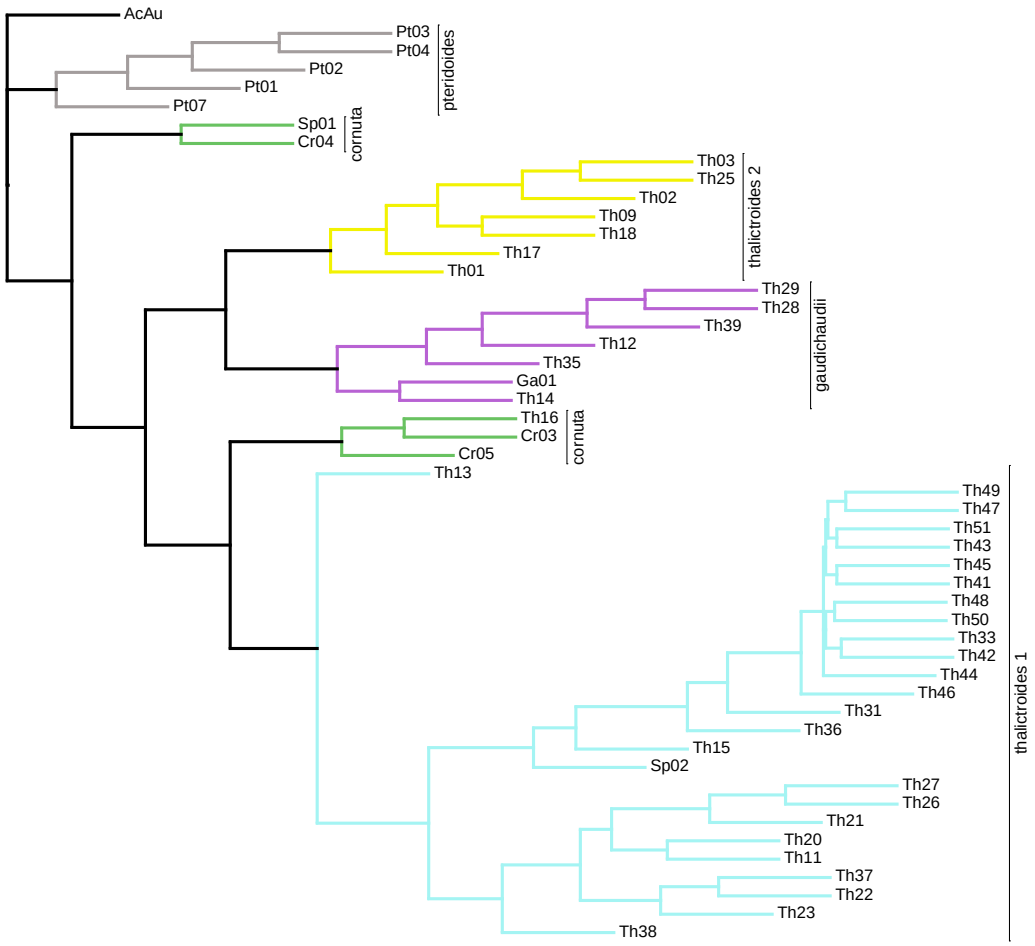


Fig. 4.4: Phylogeny of *Ceratopteris* generated via quartet methods. The tree is rooted with the sister genus to *Ceratopteris*, *Acrostichum*. *Ceratopteris pteridoides* is the most basal clade of the genus, and a clade of *C. cornuta* is the next to diverge. This clade might actually be a sampling of *C. richardii*; however, future sampling is needed to confirm this. Several individuals originally identified as *C. thalictroides* grouped with *C. gaudichaudii*, indicating that this cryptic species is indeed prolific. The other individuals of *Ceratopteris thalictroides sensu latu* breaks into two groups here: an Old World clade (1) and a New World clade (2). A clade of *C. cornuta* is sister to Old World *C. thalictroides* 1, which supports it as a potential progenitor of the latter species (Adjie et al., 2007).

present. Triploids (derived via hybridization) would have a 2:1 allele ratio, with two sets of chromosomes from a unreduced spore, and another from a haploid spore. Tetraploids could either have a 1:1 or 3:1 allele ratio: the former would be representative of an autopolyploid derived from a genome double event, or an allopolyploid with two sets of homeologous chromosomes; the latter would be representation of an allopolyploid that has undergone genomic restructuring (i.e. compensated aneuploidy, (Sigel, 2016)).

Three of four samples identified as *Ceratopteris cornuta* showed relatively equal allelic proportions of 1:1 and 3:1, indicating a possible tetraploid; the fourth sample was ambiguous. Some samples of the known diploid *C. pteridoides* showed a similar pattern, with one putative diploid and three tetraploids. However, both species are known as a diploid so this may be a mis-assignment of ploidy by the analysis, or detecting of some interesting genomic structure. *Ceratopteris gaudichaudii* had a high 1:1 allele ratio, suggesting that it is an allotetraploid. *Ceratopteris thalictroides* is also known to be a tetraploid (Lloyd, 1974; Masuyama & Watano, 2010), and individuals in the analysis were estimated to have 1:1, 2:1, and 3:1 allele ratios. Despite this variation, there was a distinct pattern between Old World and New World individuals: Old World *C. thalictroides* most commonly had a 1:1 allele ratio (autopolyploid), whereas New World *C. thalictroides* had a 3:1 allele ratio (allopolyploid).

Mis-assignment of ploidy could be due to coverage being at the lower end of the scale needed for true ploidy assignments, but it also could indicate hybridization leading to genomic complexity across species. Including chromosome counts and flow cytometry data in future analyses would help better understand the potential variation in polyploidy origins of *C. thalictroides*.

4.4.4 Analysis of simulated RADseq data

I attempted to create a pseudo-individual to include in the analysis using the *Ceratopteris richardii* genome assembly. I were able to obtain 9,047,554 simulated raw sequences utilizing an *in silico* RADseq digest. These reads, in combination with the rest of the *in vivo* data, were processed with the *ipyrad* pipeline. When constructing the final

assembly, *ipyrad* requires that all individuals are related relatively similar to one another, so that they share a large number of sequences. Very few sequences were retained for the pseudo-individual of *C. richardii*, indicating that the sequences obtained are too different from the *in vivo* samples. In order to retain enough SNPs to run STRUCTURE, I required that only 3 individuals needed to have data at a given locus for it to be included in the final data set. In the entirely *in vivo* dataset, this minimum number of individuals was 30; I designated such a high number to reduce missing data. In the dataset with the pseudo-individual of *C. richardii*, 80% of the retained loci had missing data.

The differences between the pseudo-individual derived from the *Ceratopteris richardii* and the other samples is mostly likely due to the fact that the current assembly (v. 1.1) only represents about one third of the total genome, and consists of mostly short reads. Long reads that can span the many repetitive elements in the genome comprise only 0.03% of the final assembly (Marchant et al., 2019). Because of these factors, the current genome assembly contains a small amount of short sequences that are difficult to accurately compare to the fragments generated by *in vivo* RADseq.

To explore the utility of the simulated data, I ran a population structure analysis (Pritchard et al., 2000) for this new dataset including the simulated *C. richardii* and one additional specimen of the same species. This dataset had 80% missing data, compared to 26% without *C. richardii*. Both of these individuals had mixed ancestry from several different species in the genus, which is not what I were expecting: *Ceratopteris richardii* has been assumed to be a diploid species, and although it hybridizes with other species in the genus, it was not thought to be of hybrid origin itself. The individual derived from the *C. richardii* genome aligned closely with *C. cornuta*; this is interesting because the two species ranges' overlap in western Africa, and they have very similar morphology. I can certainly speculate about potential hybrid origins or mis-identifications of specimens, but without higher quality data, I are hesitant to make any major claims. The population structure of all other species in the genus remained virtually identical to the analysis including only the highest-quality individuals (as seen in Fig. 4.3). The data simulation pipeline and the

population structure analysis including *C. richardii* can be found in chapter 4.

4.5 Discussion

To the best of the knowledge, this is the first application of next-generation sequencing techniques to the genus *Ceratopteris*. Using population genomic, phylogenetic, and ploidy inferences, I recover similar species relationships found in previous studies, find support for *C. gaudichaudii* as a genetically distinct cryptic species, identify a novel cryptic lineage from within *C. thalictroides sensu latu* in Central and South America, and show strong genetic variation across the broad range of *C. cornuta* (Fig. 4.1).

Not included in the results are specimens of the model fern *Ceratopteris richardii*, including the lab strain *Hnn*. All individuals samples I had of this species yielded very poor sequencing results. When included in the analyses, they significantly increased the amount of missing data. I felt it was appropriate to remove these samples because I did not want to try and infer anything about this important species with poor quality data. Future work will be aimed at including wild collections of *C. richardii* as well as the lab strain *Hnn* to determine the relationship of this species to the rest of the genus.

4.5.1 Species boundaries in *Ceratopteris*

Variation within *Ceratopteris cornuta*

Ceratopteris cornuta is found from western Africa to Northern Australia. The sterile fronds are consistently deltoid to lanceolate in shape, and pinnate to bipinnate in dissection (Lloyd, 1974). The population genomic analyses distinguished all specimens of *C. cornuta* as a distinct population (at K greater than 4, Fig. 4.3). All samples of *C. cornuta* grouped together at $K = 2$ - 6, along with one sample of *C. thalictroides* from Nepal (*Fraser-Jenkins 1564*) and an unidentified sample from Brazil (*L. Camargo de Abreu 21*).

Interestingly, the population genomic analysis and phylogenetic reconstruction show that individuals of *C. cornuta* from the western portion of its range (Sierra Leon, Brazil) are genetically different from individuals in the eastern portion of its range (Tanzania, Oman,

Nepal) (Fig. 4.3, 4.4). At $K = 6$, individuals from Sierra Leon (Cr04) and Brazil (Sp01) are separated slightly from the other three individuals of *C. cornuta*, and contain some genetic influence that is not attributed to another sampled species. In addition, this small clade is placed as sister to the clade containing *C. thalictroides sensu latu* and *C. gaudichaudii* (Fig. 4.4). It is important to note that there are no known collections of *C. cornuta* from the Americas. However, *Ceratopteris richardii* has been found in western Africa (Lloyd, 1974). I may have sampled African and South American individuals of *C. richardii*. The morphology of the two species is somewhat similar (Lloyd, 1974), and so these two specimens may have been mis-identified by their collectors. Similar patterns of Africa-South American gene flow have been seen in other broadly distributed pteridophytes. Recent gene flow between the two continents was observed in cosmopolitan *Pteridium* (Wolf, Rowe, Kinosian, Der, Lockhart, Shepherd, McLenachan, & Thomson, 2019), phylogenetic reconstruction has shown multiple migrations of grammitid ferns from South America to Africa (Sundue, Parris, Ranker, Smith, Fujimoto, Zamora-Crosby, Morden, Chiou, Chen, Rouhan, Hirai, & Prado, 2014), and several species (e.g. *Asplenium monanthes*, *Polypodium polypodioides*) have a disjunct distribution between the two continents (Kornas, 1993). Therefore, it is possible that there could be some trans-Atlantic gene flow in *Ceratopteris richardii*. Unfortunately, I do not have a known sample of *C. richardii* to confidently compare to this group. When I did include samples of *C. richardii* (from an herbarium specimen and the simulated digest of the *C. richardii* genome) in a STRUCTURE analysis, they were placed roughly in the same population as *C. cornuta* (Appendix A). As I have mentioned before, these data were poor quality and low coverage. While it is a window into potential relationships, increased sampling is needed to draw any meaningful conclusions.

Genomic conservation in *Ceratopteris pteridoides*

Endemic to the New World tropics, *Ceratopteris pteridoides* has perhaps the most unique morphology of the genus: sterile fronds are simple, palmately or pinnately lobed, with enlarged, air-filled stipes (petioles). They are most commonly found unrooted and floating on the surface of slow-moving water (Lloyd, 1974). *Ceratopteris pteridoides* is

also very strongly differentiated by the STRUCTURE analysis, with samples identified as *C. pteridoides* showing little to no introgression from other populations. Influence from *C. pteridoides* does appear in *C. cornuta*, *C. gaudichaudii*, and some samples of *C. thalictroides*. It is difficult to say whether this is true introgression into other species, or if it is an artifact of slightly lower sequencing coverage in these individuals. It is known, however, that *C. pteridoides* readily hybridizes with *C. richardii* where their ranges overlap in northern South America (Hickok & Klekowski, 1974). Including the latter in future work may help to reveal the extent to which natural hybridization is occurring between diploid *C. pteridoides* and *C. richardii* in the New World.

Cryptic species of *C. thalictroides*

There was one named cryptic species of *C. thalictroides* included in the study, *Ceratopteris gaudichaudii*, which has two varieties: *C. g. gaudichaudii* and *C. g. vulgaris* (Masuyama & Watano, 2010); only the latter was included in the analysis. *Ceratopteris g. vulgaris* can be found throughout the Pacific in Japan, the Philippines, Hawaii, Guam, Taiwan, Southeast Asia, and Northern Australia; *Ceratopteris g. gaudichaudii* is narrowly endemic to Guam (Masuyama & Watano, 2010). Individuals identified as *C. thalictroides* from Taiwan, Hawaii, Japan, China, and Australia were identified via the population genomic and phylogenetic analyses as being *C. gaudichaudii*, although it is impossible to tell from the limited sampling which subspecies they align with. The individual from Australia was examined and keyed out morphologically to *C. g. vulgaris*, using the key from Masuyama and Watano (2010). This same individual was growing a few meters away from an individual of Old World *C. thalictroides*, indicating that these two morphologically similar species are growing sympatrically but with limited to no gene flow.

All of the samples from Hawaii included in this study show alignment to *C. gaudichaudii* in the population structure analysis. Wagneri (1950) hypothesized that *Ceratopteris* was not native to Hawaii, and may have been introduced from Asia either as a food source or accidentally as a weed in taro patches. In addition, work by Hickok (1979) supported Hawaiian *Ceratopteris* as a distinct species from *C. thalictroides*. Japan was hypothesized to

be the source of Hawaiian *Ceratopteris* by Lloyd (1973). This is plausible since *Ceratopteris gaudichaudii* is relatively common in taro patches in Japan, and could have been brought to Hawaii via settlers or trade. Interestingly, one sample from Hawaii was shown to be a potential hybrid between *C. gaudichaudii* and *C. thalictroides*. The specimen (*L. M. Crago 2005-058* US) is highly dissected and has many proliferous buds, which could indicate a hybrid or perhaps just an older plant. This sample was also placed as sister to Old World *C. thalictroides* in the TETRAD and RAxML trees. Hybrid ancestry would explain grouping with *C. gaudichaudii* in one analysis, and *C. thalictroides* in another.

Three named cryptic species of *C. thalictroides* were not included in the analysis: *C. oblongiloba*, *C. froesii*, and *C. shingii*. I were not able to request type material, I had to rely on the herbarium sampling of their ranges to attempt to include *C. oblongiloba* and *C. froesii*. *Ceratopteris shingii* was described as this manuscript was nearing completion.

Ceratopteris oblongiloba is found throughout Cambodia, Indonesia (Sumatra and Java), the Philippines (Luzon), and Thailand (Malay) (Masuyama & Watano, 2010). The only specimen from this range is from the Philippines, does not key to *C. oblongiloba*, and aligns entirely with Old World *C. thalictroides*. This either indicates that I did not unknowingly sample *C. oblongiloba*, or it has a lesser degree of genomic distinction than the other cryptic species, *C. gaudichaudii*. *Ceratopteris froesii* is a Brazilian endemic characterized by very small fronds: about 4 cm or less for fertile leaves (Masuyama & Watano, 2010). I only had four accessions from Brazil, and none were small plants. Inclusion of this unique species in future studies may help further understand New World diversity of *Ceratopteris*. *Ceratopteris shingii* was recently described by Zhang *et al.* (2020), endemic to Hainan Province, China. It is the only known species in the genus to have creeping rhizomes, and is estimated to be sister to all other species of *Ceratopteris* (Zhang, Yu, Shao, Wang, & Yan, 2020).

Ceratopteris thalictroides, sensu latu

Individuals of *Ceratopteris thalictroides* that did not group with *C. gaudichaudii* break in to two groups: an Old World and a New World clade. These two clades are separated

at all values of K in the STRUCTURE analysis (Fig. 4.3) and are paraphyletic (Fig. 4.4), indicating strong genetic, if not morphological, separation.

The Old World *C. thalictroides* clade includes mostly individuals from Asia and Australia, but also a few from South America (Fig. 4.1). Increasing K in STRUCTURE did not break up this large group, suggesting that there is recent gene flow across the Pacific via long distance dispersal or human-mediated transport. Several fern and angiosperm genera share such a disjunct distribution between eastern Asia and North to Central America, either by vicariance or long distance dispersal (Kato & Iwatsuki, 1983; Les, Crawford, Kimball, Moody, & Landolt, 2003). One individual from Hawaii (*L. M. Crago 2005-058* US) shows shared ancestry with *C. gaudichaudii* and Old World *C. thalictroides*, indicating that trans-Pacific gene flow might be possible in *Ceratopteris*. Another individual from Japan shows a similar pattern, but with a much larger proportion of its genome comprised of the Old World *C. thalictroides* population. There may be continued gene flow not only within Old World *C. thalictroides*, but between *C. thalictroides* and *C. gaudichaudii*. The Hawaiian archipelago, as well as other Pacific islands, have been hypothesized to be a stepping-stone for trans-Pacific dispersal in angiosperm taxa (Harbaugh, Wagner, Allan, & Zimmer, 2009; Wright, Yong, Wichman, Dawson, & Gardner, 2001), a theory that could also hold in ferns, which are known to have very high dispersal capabilities (Barrington, 1993; Tryon, 1970).

New World *C. thalictroides* is strongly differentiated from Old World *C. thalictroides* in the population structure analysis, and shows virtually no gene flow between any other species at K 4 - 6. This strong distinction between Old and New World is present in the phylogeny as well: New World *C. thalictroides* is placed in a clade as sister to *C. gaudichaudii*, and this clade is in turn sister to Old World *C. thalictroides* (Fig. 4.4). *Ceratopteris gaudichaudii* has been hypothesized to be an allopolyploid hybrid (Adjie et al., 2007); perhaps the derivation of its sister species, New World *C. thalictroides*, could have been from a hybridization event with one of the parents of *C. gaudichaudii*, or genome doubling to form an autotetraploid, followed by long distance dispersal. The ploidy inference suggests that either form of polyploidy is possible, but slightly more individuals of New

World *C. thalictroides* appeared to be of allopolyploid origin, potentially with subsequent genomic rearrangement (Fig. 4.2).

Origins of cryptic species

All of the cryptic species in the *Ceratopteris thalictroides* complex are known to be tetraploid (n = 77, 78; (Masuyama & Watano, 2010)). Several authors have proposed that different cryptic species have been derived via differing mechanisms of polyploidy (i.e., allo- vs. autopolyploid). Adjie *et al.* (2007) identified Asian *C. thalictroides sensu latu* as paraphyletic and hypothesized that *C. thalictroides sensu strictu* and the cryptic species *C. gaudichaudii* and *C. oblongiloba* are allopolyploid hybrids between *C. cornuta* and unnamed, potentially extinct, diploid progenitors. McGrath *et al.* (1994) suggested that New World *C. thalictroides* could either be an autopolyploid, or an allopolyploid hybrid between *C. richardii* and another diploid species.

The population structure analysis did not reveal any cryptic species of *C. thalictroides* as distinct allotetraploids (i.e., containing roughly equal genomic proportions from two parents). Each cryptic species (Old World *C. thalictroides*, New World *C. thalictroides*, and *C. gaudichaudii*) were represented as separate populations, sharing little to no genetic material with any other population (Fig. 4.3). I were not able to include the Asian cryptic species *C. oblongiloba*, and so cannot say whether or not it is of recent hybrid origin.

Our ploidy estimation, however, indicates that the Old World clade of *C. thalictroides* and *C. gaudichaudii* may be a balanced allotetraploid, or autotetraploid. This former supports the theory proposed by Adjie *et al.* (2007), but the fact that there is no gene flow between Old World *C. thalictroides* and any other species supports the latter. The New World clade of *C. thalictroides* may be of allopolyploid origin. In the TETRAD phylogeny, it is sister to *C. gaudichaudii*; however, in the RAxML phylogeny it is sister to a clade containing Old World *C. thalictroides*, western *C. cornuta*, and *C. gaudichaudii* (see Appendix A). New World *C. thalictroides* was the only clade to come out in two different places in the two phylogenetic reconstructions. An allopolyploidy origin might explain this, with one method placing New World *C. thalictroides* closer to the maternal progenitor, and

the other placing it closer to the paternal progenitor.

4.5.2 Evolutionary drivers in *Ceratopteris*

Biogeography and reproductive boundaries

Weak reproductive boundaries are a theme across ferns, evidenced by numerous hybrids and reticulate species complexes (*e.g.*, (Barrington et al., 1989; Paris et al., 1989; Sessa, Zimmer, & Givnish, 2012; Sigel, 2016; Yatabe et al., 2009)), and intergeneric hybridization across 60 MY of evolution (Rothfels, Johnson, Hovenkamp, Swofford, Roskam, Fraser-Jenkins, Windham, & Pryer, 2015). Although ferns do have some pre- and post-zygotic barriers (Haufler, Pryer, Schuettpelz, Sessa, Farrar, Moran, Schneller, Watkins, & Windham, 2016), they are not as strong as in angiosperms (e.g., (de Nettancourt, 1997; Lafon-Placette & Kohler, 2016)). In addition, fern spores are very easily dispersed (Smith, 1972; Tryon, 1970), which limits isolation by distance.

In the case of *Ceratopteris*, its large range may be influenced by the Intertropical Convergence Zone (ITCZ) (see (Dettmann & Clifford, 1992; Lloyd, 1974; Schneider, Bischoff, & Haug, 2014)). The ITCZ is created by the rising portion of the Hadley Cell circulation in both hemispheres; it moves with the seasons, following the thermal equator (Schneider et al., 2014). It has been hypothesized as a possible mechanism of transport of Asian ferns to Hawaii (Geiger, Ranker, Ramp Neale, & Klimas, 2007), and could carry spores great distances from one area of the tropics to another, provided they make it high enough into the atmosphere, perhaps by monsoons or hurricanes. Continuous dispersal and/or movement of propagules is recognized as an important factor in widespread angiosperm hydrophytes (Les et al., 2003; Spalik, Banasiak, Feist, & Downie, 2014). The ITCZ could be a means of dispersal for *Ceratopteris*, which would help explain the continued gene flow across the large range of its species such as Old World *C. thalictroides*, or *C. cornuta*.

Even with a high level of dispersal, there is potential evidence for reproductive boundaries and sympatric speciation in *Ceratopteris*. Two specimens collected meters apart in Australia keyed to *C. thalictroides sensu strictu* and *C. gaudichaudii* var. *vulgaris*. In

addition, specimens of Old World *C. thalictroides* and New World *C. thalictroides* were obtained from similar regions in South America. Sympatric cryptic species are known from other genera of ferns (Patel et al., 2019; Yatabe et al., 2009), lichens (del Carmen Molina et al., 2011), angiosperms (Les, Peredo, King, Benoit, Tippery, Ball, & Shannon, 2015; Soltis, Soltis, Schemske, Hancock, Thompson, Husband, & Judd, 2007), as well as in animals (Amor et al., 2014; Hebert et al., 2004; Nygren, 2014). Different polyploid origins in *Ceratopteris* could potentially be the cause of sympatric speciation, with allo- and autotetraploids becoming isolated from one another. Another mechanism of differentiation could be minor ecologically separation. All *Ceratopteris* species occupy very similar habitats (Lloyd, 1974; Masuyama & Watano, 2010). However, in the case of the two Australian specimens collected at the same locality, the individual of Old World *C. thalictroides* was growing in full sun, while *C. gaudichaudii* was growing in deep shade. The former is potentially an autotetraploid (see Fig. 4.2), whereas the latter may be an allotetraploid of hybrid origin (Adjie et al., 2007). How reproductive boundaries have evolved between these co-occurring cryptic species area of future work.

Morphology and ecology

One of the most confusing aspects of *Ceratopteris* is the disparity between morphological and genetic distinction between species. The three diploid species are relatively morphologically distinct (Lloyd, 1974), especially in comparison to the cryptic species complex of *C. thalictroides sensu latu*. However, all the diploid species in the genus are known to hybridize with one another (Hickok, 1977; Hickok & Klekowski, 1974; Lloyd, 1974). In contrast, *C. thalictroides* and associated cryptic species are much more similar morphologically, yet different populations (some named as cryptic species) can be almost entirely reproductively isolated (Hickok, 1979; Masuyama & Watano, 2010; Masuyama et al., 2002). Throughout the genus, however, the body plan, life history, and niche requirements of all species are quite similar, indicating limited ecological divergence.

A potential explanation for morphological stasis in *Ceratopteris* relies on the influence of ecological pressures (Bickford et al., 2007). For a fern, *Ceratopteris* has an unusually

short life cycle of only about four months from spore to spore-bearing adult (Stein, 1971). It grows in temporary water sources such as pond edges, swamps, or tarot patches (Lloyd, 1974). Because of its ephemeral habitat, there are likely selective pressures for maintaining morphological stasis and rapid generation times. Fossil spore and leaf evidence also support morphological stasis in the Ceratopteridoideae dating back about 47 MY (Dettmann & Clifford, 1992; Rozefelds, Dettmann, Clifford, & Lewis, 2016). Fossil leaf impressions of the extinct genus *Tecaropteris* from Australia are very similar morphologically to extant *Ceratopteris* (Rozefelds et al., 2016). Australia has become increasingly dry and seasonal since the Eocene, when *Tecaropteris* lived (McKenna, Hanna, Banks, Sivachenko, Cibulskis, Kernytsky, Garimella, Altshuler, Gabriel, Daly, & DePristo, 2010; Rozefelds et al., 2016). Today, *Ceratopteris* is exceedingly difficult to find during the dry season in northeastern Australia, but plentiful during the wet season, even growing as a weed in garden ponds (Dr. Ashley Field, personal communication). Almost all leaves of *Ceratopteris* species are very thin, delicate, and herbaceous (Lloyd, 1974). When removed from the plant they wilt into an unrecognizable state within minutes (SPK, personal observation). These low-energy investment leaves can be grown quickly and without an excess of resources from the plant (Reich, 2014; Wright, Reich, Westoby, Ackerly, Baruch, Bongers, Cavender-Bares, Chapin, Cornelissen, Diemer, Flexas, Garnier, Groom, Gulias, Hikosaka, Lamont, Lee, Lee, Lusk, Midgley, Navas, Niinemets, Oleksyn, Osada, Poorter, Poot, Prior, Pyankov, Roumet, Thomas, Tjoelker, Veneklaas, & Villar, 2004), potentially as a result of the fast life cycle employed by *Ceratopteris*; any major shift in leaf development might inhibit *Ceratopteris* from proceeding through its life cycle before its water source dries up. Therefore, despite the genetic diversity in the genus, morphological diversity has changed very little due to niche constraints.

4.6 Conclusion

In summary, the findings suggest that there may be more cryptic species of *Ceratopteris thalictroides* yet to be discovered, especially in the New World. The potential for cryptic species in the Neotropics has been noted by several authors (Lloyd, 1974; Masuyama &

Watano, 2010), and this area of the world remains one of the least studied parts of the range of *Ceratopteris*. Considering that this is where the model species *C. richardii* is native, more work on all species of the genus in the New World is especially warranted. Researching *Ceratopteris* in its natural environment is critical to inform future lab studies on *C. richardii*, and the potential incorporation of other species for comparative studies. Most importantly, understanding speciation and cryptic species within *Ceratopteris* is important for its use as model system, as well as increasing the knowledge of evolutionary processes in ferns and across vascular land plants.

References

Adjie, B., Masuyama, S., Ishikawa, H., & Watano, Y. (2007). Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. *Journal of Plant Research*, *120*(1), 129–138.

Amato, A., Kooistra, W. H. C. F., Ghiron, J. H. L., Mann, D. G., Proschold, T., & Montresor, M. (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist*, *158*(2), 193–207.

Amor, M. D., Norman, M. D., Cameron, H. E., & Strugnell, J. M. (2014). Allopatric speciation within a cryptic species complex of Australasian octopuses. *PLoS One*, *9*(6), e98982.

Avni, E., Cohen, R., & Snir, S. (2015). Weighted quartets phylogenetics. *Systematic Biology*, *64*(2), 233–242.

Banks, J. A. (1994). Sex-determining genes in the homosporous fern *Ceratopteris*. *Development*, *120*(7), 1949–1958.

Barrington, D. S. (1993). Ecological and historical factors in fern biogeography. *Journal of Biogeography*, *20*(3), 275–279.

Barrington, D. S., Haufler, C. H., & Werth, C. R. (1989). Hybridization, reticulation, and species concepts in the ferns. *American Fern Journal*, *79*(2), 55–64.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends Ecology and Evolution*, *22*(3), 148–155.

Chifman, J., & Kubatko, L. (2015). Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of Theoretical Biology*, *374*, 35–47.

Conway, S. J., & Di Stilio, V. S. (2019). An ontogenetic framework for functional studies in the model fern *ceratopteris richardii*. *Developmental Biology*.

Cunnington, J. H., Lawrie, A. C., & Pascoe, I. G. (2005). Genetic variation within *podosphaera tridactyla* reveals a paraphyletic species complex with biological specialization towards specific *prunus* subgenera. *Mycology Resources*, *109*(Pt 3), 357–362.

de Nettancourt, D. (1997). Incompatibility in angiosperms. *Sexual Plant Reproduction*, *10*(4), 185–199.

del Carmen Molina, M., Divakar, P. K., Millanes, A. M., Sanchez, E., Ruth, D.-P., Hawksworth, D. L., & Crespo, A. (2011). *Parmelia sulcata* (Ascomycota: Parmeliaceae), a sympatric monophyletic species complex. *Lichenologist*, *43*(6), 585–601.

Der, J. P., Thomson, J. A., Stratford, J. K., & Wolf, P. G. (2009). Global chloroplast phylogeny and biogeography of bracken (*Pteridium*; Dennstaedtiaceae). *American Journal of Botany*, *96*(5), 1041–1049.

Dettmann, M. E., & Clifford, H. T. (1992). Phylogeny and biogeography of *Ruffordia*, *Mohria* and *Anemia* (Schizaeaceae) and *Ceratopteris* (Pteridaceae): evidence from in situ and dispersed spores. *Alcheringa*, *16*(4), 269–314.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*.

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*.

Eaton, D. A. R., & Ree, R. H. (2013). Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology*, *62*(5), 689–706.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620.

Geiger, J. M. O., Ranker, T. A., Ramp Neale, J. M., & Klimas, S. T. (2007). Molecular biogeography and origins of the Hawaiian fern flora. *Brittonia*, *59*(2), 142–158.

Gompert, Z., & Mock, K. E. (2017). Detection of individual ploidy levels with genotyping-by-sequencing (GBS) analysis. *Molecular Ecology Resources*, *17*(6), 1156–1167.

Harbaugh, D., Wagner, W. L., Allan, G. J., & Zimmer, E. A. (2009). The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: An example from the plant genus *Melicope* (Rutaceae). *Journal of Biogeography*, *36*(2)), 230.

Haufler, C. H., Pryer, K. M., Schuettpelz, E., Sessa, E. B., Farrar, D. R., Moran, R., Schneller, J. J., Watkins, J. E., & Windham, M. D. (2016). Sex and the single gametophyte: Revising the homosporous vascular plant life cycle in light of contemporary research. *Bioscience*, *66*(11), 928–937.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, *101*(41), 14812–14817.

Hickok, L. G. (1977). Cytological relationships between three diploid species of the fern genus *ceratopteris*. *Canadian Journal of Botany*, *55*(12), 1660–1667.

Hickok, L. G. (1979). A cytological study of intraspecific variation in *ceratopteris thalictroides*. *Canadian Journal of Botany*, *57*(16), 1694–1700.

Hickok, L. G., & Klekowski, E. J. (1974). Inchoate speciation in *Ceratopteris*: An analysis of the synthesized hybrid *C. richardii* x *C. pteridoides*. *Evolution*, *28*(3), 439–446.

Hickok, L. G., Warne, T. R., & Fribourg, R. S. (1995). The biology of the fern *ceratopteris* and its use as a model system. *International Journal of Plant Sciences*, *156*(3), 332–345.

Hickok, L. G., Warne, T. R., & Slocum, M. K. (1987). *Ceratopteris richardii*: Applications for experimental plant biology. *American Journal of Botany*, *74*(8), 1304–1316.

Jorgensen, S. A., & Barrington, D. S. (2017). Two beringian origins for the allotetraploid fern *polystichum braunii* (dryopteridaceae). *Systematic Botany*, *42*(1), 6–16.

Kato, M., & Iwatsuki, K. (1983). Phytogeographic relationships of pteridophytes between temperate North America and Japan. *Annals of the Missouri Botanical Garden.*, *70*(4), 724–733.

Knowlton, N. (1993). Sibling species in the sea. *Annual Review of Ecology, Evolution, and Systematics*, *24*(1), 189–216.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191.

Kornas, J. (1993). The significance of historical factors and ecological preference in the distribution of African pteridophytes. *Journal of Biogeography*, *20*(3), 281–286.

Lafon-Placette, C., & Kohler, C. (2016). Endosperm-based postzygotic hybridization barriers: Developmental mechanisms and evolutionary drivers. *Molecular Ecology*, *25*(11), 2620–2629.

Les, D. H., Crawford, D. J., Kimball, R. T., Moody, M. L., & Landolt, E. (2003). Biogeography of discontinuously distributed hydrophytes: A molecular appraisal of intercontinental disjunctions. *International Journal of Plant Sciences*, *164*(6), 917–932.

Les, D. H., Peredo, E. L., King, U. M., Benoit, L. K., Tippery, N. P., Ball, C. J., & Shannon, R. K. (2015). Through thick and thin: cryptic sympatric speciation in the submersed genus *najas* (hydrocharitaceae). *Molecular Phylogenetics and Evolution*, *82 Pt A*, 15–30.

Lloyd, R. M. (1974). Systematics of the genus *Ceratopteris* Brongn. (Parkeriaceae) II. Taxonomy. *Brittonia*, *26*(2), 139–160.

Marchant, D. B., Sessa, E. B., Wolf, P. G., Heo, K., Barbazuk, W. B., Soltis, P. S., & Soltis, D. E. (2019). The C-Fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. *Scientific Reports*, *9*(1), 18181.

Masuyama, S. (1992). Clinal variation of frond morphology and its adaptive implication in the fern *Ceratopteris thalictroides* in Japan. *Plant Species Biology*, *7*(2-3), 87–96.

Masuyama, S. (2008). Cryptic species in the fern *Ceratopteris thalictroides* (Parkeriaceae). III. Referential diagnostic characters of three cryptic species. *J. Plant Res.*, *12*(3), 279–286.

Masuyama, S., & Watano, Y. (2005). Cryptic species in the fern *Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae). II. Cytological characteristics of three cryptic species. *Acta Phytotaxonomica et Geobotanica*, *56*(3), 231–240.

Masuyama, S., & Watano, Y. (2010). Cryptic species in the fern *Ceratopteris thalictroides* (l.) Brongn. Parkeriaceae. IV. Taxonomic revision. *Acta Phytotaxonomica et Geobotanica*, *61*(2), 75–86.

Masuyama, S., Yatabe, Y., Murakami, N., & Watano, Y. (2002). Cryptic species in the fern *Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae). I. Molecular analyses and crossing tests. *Journal of Plant Resources*, *115*(1118), 87–97.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Resources*, *20*(9), 1297–1303.

Mora-Márquez, F., García-Olivares, V., Emerson, B. C., & López de Heredia, U. (2017). ddradseqtools: a software package for in silico simulation and testing of double-digest RADseq experiments. *Molecular Ecology Resources*, *17*(2), 230–246.

Muthukumar, B., Joyce, B. L., Elless, M. P., & Stewart, C. N., Jr (2013). Stable transformation of ferns using spores as targets: *Pteris vittata* and *Ceratopteris thalictroides*. *Plant Physiology*, *163*(2), 648–658.

Nygren, A. (2014). Cryptic polychaete diversity: a review. *Zoologica Scripta*, *43*(2), 172–183.

Otto, S. P. (2007). The evolutionary consequences of polyploidy. *Cell*, *131*(3), 452–462.

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*, 401–437.

Paris, C. A., Wagner, F. S., & Wagner, W. H. (1989). Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. *American Fern Journal*, *79*(2), 46–54.

Patel, N. R., Fawcett, S., & Gilman, A. V. (2019). *Phegopteris excelsior* (Thelypteridaceae): A new species of North American tetraploid beech fern. *Novon*, *27*(4), 211–218.

Plackett, A. R. G., Huang, L., Sanders, H. L., & Langdale, J. A. (2014). High-efficiency stable transformation of the model fern species *Ceratopteris richardii* via microparticle bombardment. *Plant Physiology*, *165*(1), 3–14.

PPGI (2016). A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution*, *54*(6), 563–603.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

R Core Team (2018). R: A language and environment for statistical computing.

Ranwez, V., & Gascuel, O. (2001). Quartet-based phylogenetic inference: improvements and limits. *Molecular Biology and Evolution*, *18*(6), 1103–1116.

Reich, P. B. (2014). The world-wide 'fast-slow' plant economics spectrum: a traits manifesto. *Journal of Ecology*, *102*(2), 275–301.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecology and Evolution*, *3*(2), 217–223.

Rothfels, C. J., Johnson, A. K., Hovenkamp, P. H., Swofford, D. L., Roskam, H. C., Fraser-Jenkins, C. R., Windham, M. D., & Pryer, K. M. (2015). Natural hybridization between genera that diverged from each other approximately 60 million years ago. *American Naturalist*, *185*(3), 433–442.

Rozefelds, A. C., Dettmann, M. E., Clifford, H. T., & Lewis, D. (2016). Macrofossil evidence of early sporophyte stages of a new genus of water fern *tecaropteris* (ceratopteridoideae: pteridaceae) from the paleogene redbank plains formation, southeast queensland, australia. *Alcheringa*, *40*(1), 1–11.

Sattler, T., Bontadina, F., Hirzel, A. H., & Arlettaz, R. (2007). Ecological niche modelling of two cryptic bat species calls for a reassessment of their conservation status. *Journal of Applied Ecology*, *44*(6), 1188–1199.

Schneider, T., Bischoff, T., & Haug, G. H. (2014). Migrations and dynamics of the intertropical convergence zone. *Nature*, *513*(7516), 45–53.

Sessa, E. B., Banks, J. A., Barker, M. S., Der, J. P., Duffy, A. M., Graham, S. W., Hasebe, M., Langdale, J., Li, F.-W., Marchant, D. B., Pryer, K. M., Rothfels, C. J., Roux, S. J., Salmi, M. L., Sigel, E. M., Soltis, D. E., Soltis, P. S., Stevenson, D. W., & Wolf, P. G. (2014). Between two fern genomes. *Gigascience*, *3*, 15.

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012). Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology*, *12*, 104.

Sigel, E. M. (2016). Genetic and genomic aspects of hybridization in ferns. *Journal of Systematics and Evolution*.

Smith, A. R. (1972). Comparison of fern and flowering plant distributions with some evolutionary interpretations for ferns. *Biotropica*, *4*(1), 4–9.

Soltis, D. E., Soltis, P. S., Schemske, D. W., Hancock, J. F., Thompson, J. N., Husband, B. C., & Judd, W. S. (2007). Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon*, *56*(1), 13–30.

Southgate, M. W., Patel, N. R., & Barrington, D. S. (2019). Ecological outcome of allopolyploidy in *adiantum* (pteridaceae): niche intermediacy and expansion into novel habitats. *Rhodora*, *121*(986), 108–135.

Spalik, K., Banasiak, Ł., Feist, M. A. E., & Downie, S. R. (2014). Recurrent short-distance dispersal explains wide distributions of hydrophytic umbellifers (Apiaceae tribe Oenantheae). *Journal of Biogeography*, *41*(8), 1559–1571.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Stein, D. B. (1971). Gibberellin-induced fertility in the fern *ceratopteris thalictroides* (l.) brongn. *Plant Physiologygy*, *48*(4), 416–418.

Sundue, M. A., Parris, B. S., Ranker, T. A., Smith, A. R., Fujimoto, E. L., Zamora-Crosby, D., Morden, C. W., Chiou, W.-L., Chen, C.-W., Rouhan, G., Hirai, R. Y., & Prado, J. (2014). Global phylogeny and biogeography of grammitid ferns (polypodiaceae).

Molecular Phylogenetics and Evolution, *81*, 195–206.

Tryon, R. (1970). Development and evolution of fern floras of oceanic islands. *Biotropica*, *2*(2), 76–84.

Wolf, P. G., Rowe, C. A., Kinosian, S. P., Der, J. P., Lockhart, P. J., Shepherd, L. D., McLenachan, P. A., & Thomson, J. A. (2019). Worldwide relationships in the fern genus *Pteridium* (bracken) based on nuclear genome markers. *American Journal of Botany*, *106*(10), 1365–1376.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., & Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827.

Wright, S. D., Yong, C. G., Wichman, S. R., Dawson, J. W., & Gardner, R. C. (2001). Stepping stones to hawaii: a trans-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS+ETS). *Journal of Biogeography*, *28*(6), 769–774.

Yatabe, Y., Shinohara, W., Matsumoto, S., & Murakami, N. (2009). Patterns of hybrid formation among cryptic species of bird-nest fern, *Asplenium nidus* complex (Aspleniaceae), in West Malesia. *Botanical Journal of the Linnean Society*, *160*(1), 42–63.

Zhang, R., Yu, J.-H., Shao, W., Wang, W.-Q., & Yan, Y.-H. (2020). *Ceratopteris shingii*, a new species of *Ceratopteris* with creeping rhizomes from Hainan, China. *Phytotaxa*, *449*(1), 23–30.

CHAPTER 5

There and back again: Reticulate evolution in *Ceratopteris*

Published in the American Fern Journal[1]

5.1 Abstract

Examples of reticulate evolution are known from throughout the tree of life, but are particularly common in ferns due to their unique reproductive biology and few prezygotic reproductive barriers. However, untangling the complex evolution history of these groups can be challenging. Often, several different types of data are needed to understand the full story of reticulate evolution; for example, chloroplast markers trace maternal inheritance, while nuclear markers complete the picture by suggesting paternal inheritance. Next generation sequencing can provide thousands of nuclear loci, which are informative for estimating reticulate evolutionary histories. The model fern genus *Ceratopteris* is known to have cryptic allotetraploid taxa, and hybridization is common between many species in the genus. To better understand the patterns of hybridization and reticulate evolution in the genus, I constructed a split network analysis using thousands of single nucleotide polymorphisms from samples collected throughout the pan-tropical range of *Ceratopteris*. The split network organizes taxa based on genomic similarity, revealing potential introgression between lineages. Combining this analysis with Patterson's D to measure gene flow, plus works by previous authors, I show extensive hybridization and reticulate evolution in Old World *Ceratopteris*, and also provide evidence for natural hybridization events involving the model species *C. richardii*.

[1]Authors: Sylvia Kinosian, William Pearse, and Paul Wolf

5.2 Introduction

Phylogenetic relationships are often depicted with neatly bifurcating trees, but such phylogenies do not always portray the full picture of evolution, especially within closely related groups (Otto & Whitton, 2000; Stebbins, 1950). When estimating the topology of an evolutionary tree, one should consider not just the steady accumulation of lineages through divergent evolution, but also the rejoining of branches via reticulate evolution. Reticulate species complexes are known from many lineages including animals (Hebert, Penton, Burns, Janzen, & Hallwachs, 2004; Nygren, 2014), fungi (del Carmen Molina, Divakar, Millanes, Sanchez, Ruth, Hawksworth, & Crespo, 2011), and angiosperms (Ma, Wen, Tian, Jamal, Chen, & Liu, 2018; Nauheimer, Cui, Clarke, Crayn, Bourke, & Nargar, 2019). Species complexes are particularly common in ferns, because spore-dispersed plants lack the prezygotic barriers created by intricate floral morphology of many angiosperms (Ranker & Sundue, 2015; Rothfels, Johnson, Hovenkamp, Swofford, Roskam, Fraser-Jenkins, Windham, & Pryer, 2015; Smith, 1972), leading to high rates of hybridization and polyploidy (Barrington, Haufler, & Werth, 1989; Sigel, 2016). This lack of prezygotic barriers in ferns allows hybridization to occur even across deep divergences (e.g., 60 Mya in $\times$*Cystocarpium* (Rothfels et al., 2015); 50 Mya in $\times$*Lindsaeosoria* (Lehtonen, 2018); 40 Mya in *Dryopteris* (Sessa, Zimmer, & Givnish, 2012b)). Hybridization is often followed by polyploidization, which can occur either via a triploid bridge (Gastony, 1986; Yamauchi, Hosokawa, Nagata, & Shimoda, 2004), or genome-doubling event (Glover, Redestig, & Dessimoz, 2016; Soltis & Soltis, 2009). Polyploidy also helps to mask genetic load (Haufler, Pryer, Schuettpelz, Sessa, Farrar, Moran, Schneller, Watkins, & Windham, 2016) and reduce inbreeding depression (Masuyama & Watano, 1990).

While prevalent in ferns, reticulate evolution can be difficult to discern. Historically, a combination of morphological, cytological, chloroplast (plastid), nuclear data, and experimental crossing tests was required to disentangle reticulate hybrid complexes (Barrington et al., 1989; Masuyama & Watano, 2010; Masuyama, Yatabe, Murakami, & Watano, 2002; Sessa, Zimmer, & Givnish, 2012a). Such work was time-consuming and laborious, because

detecting and parsing the complex signals of historic, reticulate evolution requires careful data curation (e.g., (Linder & Rieseberg, 2004; Masuyama, 2008; Masuyama & Watano, 2005; Masuyama et al., 2002)). The advent of next generation sequencing has created techniques that can generate up to tens of thousands of loci to test hypotheses about reticulate evolution. In particular, restriction site-associated DNA sequencing (RADseq) is a cost-effective way to generate these data (Blanco-Pastor, Bertrand, Liberal, Wei, Brummer, & Pfeil, 2019; Eaton & Ree, 2013; Ma et al., 2018; Moura, Shreves, Pilot, Andrews, Moore, Kishida, Moller, Natoli, Gaspari, McGowen, Chen, Gray, Gore, Culloch, Kiani, Willson, Bulushi, Collins, Baldwin, Willson, Minton, Ponnampalam, & Hoelzel, 2020). Such approaches are revolutionizing the study of reticulate evolution in other taxa, and also have the potential to do so for ferns (see chapters 3 and 4).

The fern genus *Ceratopteris* is best known for its utility as a model organism. The species *Ceratopteris richardii* Brong. was the first homosporous fern to have a partial whole genome sequence available (Marchant, Sessa, Wolf, Heo, Barbazuk, Soltis, & Soltis, 2019), and has been important to research topics such as sex determination (Banks, 1997), leaf development (Vasco, Smalls, Graham, Cooper, Wong, Stevenson, Moran, & Ambrose, 2016), and apogamy (Bui, Pandzic, Youngstrom, Wallace, Irish, Szovenyi, & Cheng, 2017). In addition to its contributions as a model organism and genomic resource, *Ceratopteris* provides an excellent system for the study of reticulate evolution. The genus consists of three diploid and five tetraploid species (see Table 5.1; (Lloyd, 1974; Masuyama & Watano, 2005; Zhang, Yu, Shao, Wang, & Yan, 2020)). Many of these are hypothesized to be of hybrid origin (Adjie, Masuyama, Ishikawa, & Watano, 2007; Lloyd, 1974), with potential backcrossing between some lineages (Hickok & Klekowski, 1974). Although there are studies of reticulate evolution in Old World *Ceratopteris* (Adjie et al., 2007), such relationships in other areas of its pan-tropical range have not been examined thoroughly.

Here, I present a network analysis of the species relationships within *Ceratopteris* using a multi-locus nuclear dataset. This information, along with estimations of gene flow, previously published gene trees and population genomic analyses, enable me to understand

reticulate evolution in this genus to a finer degree. I present evidence for extensive, ancient hybridization and reticulate evolution in both Old and New World *Ceratopteris*. In addition, the data suggest natural hybridization events involving the model species *C. richardii* (discussed in Hickok & Klekowski (1974); Nakazato, Jung, Housworth, Rieseberg, & Gastony (2007)). Although hybridization and reticulations were suspected in *Ceratopteris*, the network analysis shows how important these events have been throughout the evolutionary history of the genus.

5.3 Material and Methods

The sampling, DNA extraction and sequencing, and data processing (but not data analysis) are described in detail in Appendix A; however, I give an overview here. A detailed description of all steps and parameters of the data processing pipeline can be found on GitHub (github.com/sylviakinosian/ceratopteris_RADseq). Demultiplexed sequences can be accessed via the NCBI GenBank Sequence Read Archive (PRJNA606596).

5.3.1 Sampling

To obtain samples from throughout the pan-tropical range of *Ceratopteris*, I used a combination of herbarium and silica-dried fresh samples. I collected a total of 90 samples, 56 from herbarium specimens in the U.S. and 34 from the field in Taiwan, China, Costa Rica, and Australia. These samples represent five of the eight named species of *Ceratopteris* (see Table 5.1 for species descriptions), and one sample of *Acrostichum aureum* L. from Australia for use as an outgroup. Full specimen collections are given in Appendix A.

I identified specimens in the field by consulting the dichotomous keys and illustrations in Lloyd (1974) and Masuyama & Watano (2010). For herbarium specimens, I used the name given by the collector, or the most recent annotation. If a specimen was identified as a different species by the genomic analyses, I utilized the above keys to double check this identification with the digitized specimen or photographs taken by SPK during material collection.

Table 5.1: Named species of *Ceratopteris*. There are currently three diploid and five tetraploid species with published names; additional cryptic variation is suspected in *C. thalictroides* and *C. cornuta*. The authority, ploidy, spores per sporangium, range, potential hybridization, and relevant citations are included here.

Species name	Authority	Ploidy, spores	Geographic range	Evidence of hybridization	Notes	Citations
Ceratopteris richardii	Brong.	diploid ($n = 39$), 16 spores per sporangium	Central America, Western Africa	Yes, with *C. pteridoides* and *C. thalictroides* (PN8)	Model species; deltoid lamina	Hickok & Klekowski (1974); Lloyd (1974); Nakazato et al. (2007)
C. pteridoides	(Hook.) Hieron.	diploid ($n = 39$), 32 spores per sporangium	Southern North America (Gulf of Mexico coast), Central America, northern South American	Yes, with *C. richardii*	Plants often floating; inflated stipes; distinct genomic population	Hickok & Klekowski (1974); Lloyd (1974)
C. cornuta	(Pal. Beauv.) Le Prieur	diploid ($n = 39$), 32 spores per sporangium	Sub-Saharan Africa, Arabian Peninsula, India, SE Asia, N. Australia	Sub-Saharan Africa, Arabian Peninsula, India, SE Asia, N. Australia	Plants rooted or floating; highly gemmiferous	Adjie et al. (2007); Lloyd (1974)
C. thalictroides	(L.) Brong.	tetraploid ($n = 77$), 32 spores per sporangium	India, SE Asia, Australia, Japan, China, Taiwan, Central America	Yes, with *C. gaudichaudii* and *C. richardii* (PN8)	Highly polymorphic; Central American population may be separate species	Lloyd (1974); Masuyama & Watano (2010); Nakazato et al. (2007)
C. gaudichaudii var. gaudichaudii	Brong.	tetraploid ($n = 78$), 32 spores per sporangium	Endemic to Guam	unknown	Cryptic sp. of C. thalictroides; larger variety; highly gemmiferous	Masuyama & Watano (2010)
C. gaudichaudii var. vulgaris	Brong. (Masuyama & Watano)	tetraploid ($n = 78$), 32 spores per sporangium	China, Japan, Nepal, Hawaii, Guam, N. Australia	Yes, with *C. oblongiloba* and *C. thalictroides*	Cryptic sp. of *C. thalictroides*; smaller variety, more common	Masuyama & Watano (2005,1)
C. oblongiloba	Masuyama & Watano	tetraploid ($n = 78$), 32 spores per sporangium	Indonesia, Philippines, Thailand	Yes, with *C. gaudichaudii*	Cryptic sp. of *C. thalictroides*; highly dissected	Masuyama & Watano (2005,1)
C. froesii	Brade	Unknown, probably tetraploid	Endemic to Brazil	unknown	Very small plants, ¡3cm tall; considered dwarf *C. thalictroides* by Lloyd	Lloyd (1974); Masuyama & Watano (2010)
C. shingii	Y. H. Yan & R. Zhang	tetraploid ($n = 77$)	Hainan Province, China	unknown	Cryptic sp. of *C. thalictroides*; plants with creeping rhizomes; grows on volcanic rock	Zhang et al. (2020)

5.3.2 DNA extraction and sequencing

Most samples were supplied as silica-dried leaf tissue to the University of Wisconsin-Madison Biotechnology Center, where DNA was extracted using the QIAGEN DNeasy mericon 96 QIAcube HT Kit. The remainder (17 samples) were extracted from silica-dried leaf tissue using a modified CTAB method by SPK at Utah State University. Both the University of Wisconsin-Madison Biotechnology Center and SPK used variants of the CTAB DNA extraction method (Doyle & Doyle, 1987), as research suggests it is the best for extracting high-purity DNA from herbarium specimens (Sarkinen, Staats, Richardson, Cowan, & Bakker, 2012). All specimens were sent to the University of Wisconsin-Madison Biotechnology Center, analyzed for quality, and then pooled with the rest of the samples. Restriction site-associated DNA sequencing (RADseq) libraries were prepared by University of Wisconsin-Madison Biotechnology Center, following Elshire, Glaubitz, Sun, Poland, Kawamoto, Buckler, & Mitchell (2011) with minimal modification; sequencing was done on Illumina NovaSeq 6000 2x150 S2.

5.3.3 Data processing

RADseq yields millions of sequences that need to be filtered and processed before data analysis. Raw data were demultiplexed using stacks v. 2.4 process_radtags (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011). The resulting demultiplexed FASTQ files were paired, and low quality bases, adapters, and primers removed using ipyrad version 0.9.52 (Eaton & Overcast, 2020). I used default settings for the remainder of data processing, with the exception that I required a locus to have a sequencing depth of at least 6, and data from a minimum of 30 samples for inclusion in the final assembly to reduce missing data, due to the unbalanced nature of the sampling (discussed in Eaton, Spriggs, Park, & Donoghue (2017); Hou, Nowak, Mirré, Bjorå, Brochmann, & Popp (2016)).

For the purpose of this study I use the ipyrad pipeline's definition of a locus as a short sequence present across samples. Single nucleotide polymorphisms (SNPs) are selected from each locus and used in downstream analyses. The initial sampling contained 90 unique

specimens, plus six replicate samples taken randomly from the 90 unique samples. Samples that were low quality (yielding ¡1000 loci) were removed from downstream analysis to limit missing data. The final data set included 50 samples of *Ceratopteris* and one sample of *Acrostichum aureum*. No specimens identified by their collectors as the model species *C. richardii* made it through the data processing due to low quality and sequence coverage. However, I believe that some specimens identified as other taxa were misidentified, and potentially be *C. richardii*.

5.3.4 Estimation of interspecies gene flow

I calculated Patterson's D (also known as the ABBA-BABA or D-statistic) using the program Dsuite (Malinsky, Matschiner, & Svardal, 2021). This statistic measures divergence from a bifurcating evolutionary history (see Durand, Patterson, Reich, & Slatkin (2011), Green *et al.* 2010). Dsuite uses biallelic SNPs to estimate gene flow within species quartets, including three ingroup and one outgroup taxa; a D-statistic is reported for each quartet. I used the clades that were identified in chapter 4: *C. pteridoides*, *C. gaudichaudii*, *C. cornuta*, Old World *C. thalictroides*, and New World *C. thalictroides*, with *Acrostichum aureum* as the outgroup in all quartets.

5.3.5 Phylogenetic network analysis

Huson & Bryant (2006) define three types of phylogenetic networks: traditional phylogenetic trees, split networks, and reticulate neworks. Split networks are implicit representations of evolution, and can depict multiple phylogenetic hypotheses in one figure, as well as incongruencies between them. Reticulate networks, unlike split networks, are explicit evolutionary reconstructions and so their internal nodes represent ancestral individuals (Huson & Bryant, 2006). Split networks are a good way to explore data that may not be suitable for traditional phylogenetic approaches, such as a group known to have a reticulate evolutionary history.

To investigate the complex evolutionary history of *Ceratopteris*, I utilized the NeighborNet (Bryant & Moulton, 2002) split network algorithm in SplitsTree v. 4.16.1 ((Huson

& Bryant, 2006); Fig. 5.1). NeighborNet is similar to a neighbor joining (NJ) tree in that it pairs samples based on similarity, agglomerating taxa into larger and larger groups (Bryant & Moulton, 2002). Unlike NJ, however, NeighborNet does not produce a bifurcating tree: it creates a split network, where parallel lines indicate splits of taxa, and boxes created by these lines indicate conflicting signals (Bryant & Moulton, 2002). The handful of phylogenetic studies on Ceratopteris have yielded slightly different topologies (Adjie et al., 2007; Zhang et al., 2020), so I wanted to explore a method outside of a bifurcating tree to better understand the evolutionary history of the group. In addition, I conducted a Phi test (Bruen, Philippe, & Bryant, 2006) using SplitsTree. This test is used to detect evidence of recombination in a dataset, separating it from other processes such as recurrent mutation and incomplete lineage sorting.

Utilizing this novel split network analysis of *Ceratopteris*, in conjunction with other published phylogenies of the group (Adjie et al., 2007; Zhang et al., 2020), I drew a hypothesized reticulate network (Fig. 5.2). The rationale for each branch or clade is outlined below:

- *Acrostichum* is the sister genus to *Ceratopteris* ((PPG I, 2016)), so it was used as the outgroup for the reticulogram.

- Zhang et al. (2020) showed that *Ceratopteris shingii* Y. H. Yan and R. Zhang is the sister to the rest of the genus, so I placed it as such in the reticulogram. This placement is discussed in detail below.

- Adjie et al. (2007) and Zhang et al. (2020) recovered *C. pteridoides* (Hook.) Hieron. and *C. richardii* as sister species; Hickok & Klekowski (1974) demonstrated the natural hybridization of these two lineages. I placed these two species as sister to one another on the next diverging branch, and indicated hybridization between them with dotted lines.

- I inferred the position of New World *C. thalictroides* from chapter 4, placing it sister to the Old World species in the genus. I also indicated hybridization between *C.*

richardii and New World *C. thalictroides* as suggested by Nakazato et al. (2007), which I discuss in more detail below.

- Work by Adjie et al. (2007) informed the hybrid relationships of *C. cornuta* (Pal. Beauv.) Le Prieur, *C. gaudichaudii* Brong., *C. oblongiloba* Masuyama and Watano, and Old World *C. thalictroides* (L.) Brong. Species Y, a putative progenitor of *C. gaudichaudii* suggested by Adjie et al. (2007), was placed as the first diverging lineage of Old World *C. thalictroides*, with Species X (suggested by Adjie et al. (2007)) and *C. cornuta* forming a clade.

- There is evidence of hybrid backcrossing between *C. gaudichaudii* and *C. thalictroides* (Masuyama & Watano, 2005), as well as between *C. gaudichaudii* and *C. oblongiloba* (Masuyama & Watano, 2005). These reticulate events were added to the Old World *Ceratopteris* clade, detailed in the previous bullet point.

- The Brazilian endemic species *C. froesii* Brade may be closely related to New World *C. thalictroides* as they are both cryptic species of *C. thalictroides* sensu lato; however, no genetic information for this species is available, so it is not included in the reticulate network.

5.4 Results

Raw data consisted of 2.58 x 106 raw reads per sample, with slightly more reads from silica-dried tissue compared to herbarium samples. The final dataset of 51 individuals had 24% missing data. A Phi test found significant evidence for recombination in the dataset (p = 6.513 x 10-11). The split network analysis was performed with a dataset of 26,593 SNPs, and separated each clade (species) of *Ceratopteris* into its own genomic group, but also revealed historical gene flow between all groups (Fig. 5.1; See Appendix B for full specimen lables). Using the same set of SNPs, I calculated Patterson's D for all combinations of three species plus the outgroup, and found evidence of gene flow within four groupings of species (Table 5.2). These results are congruent with previous studies that explored species

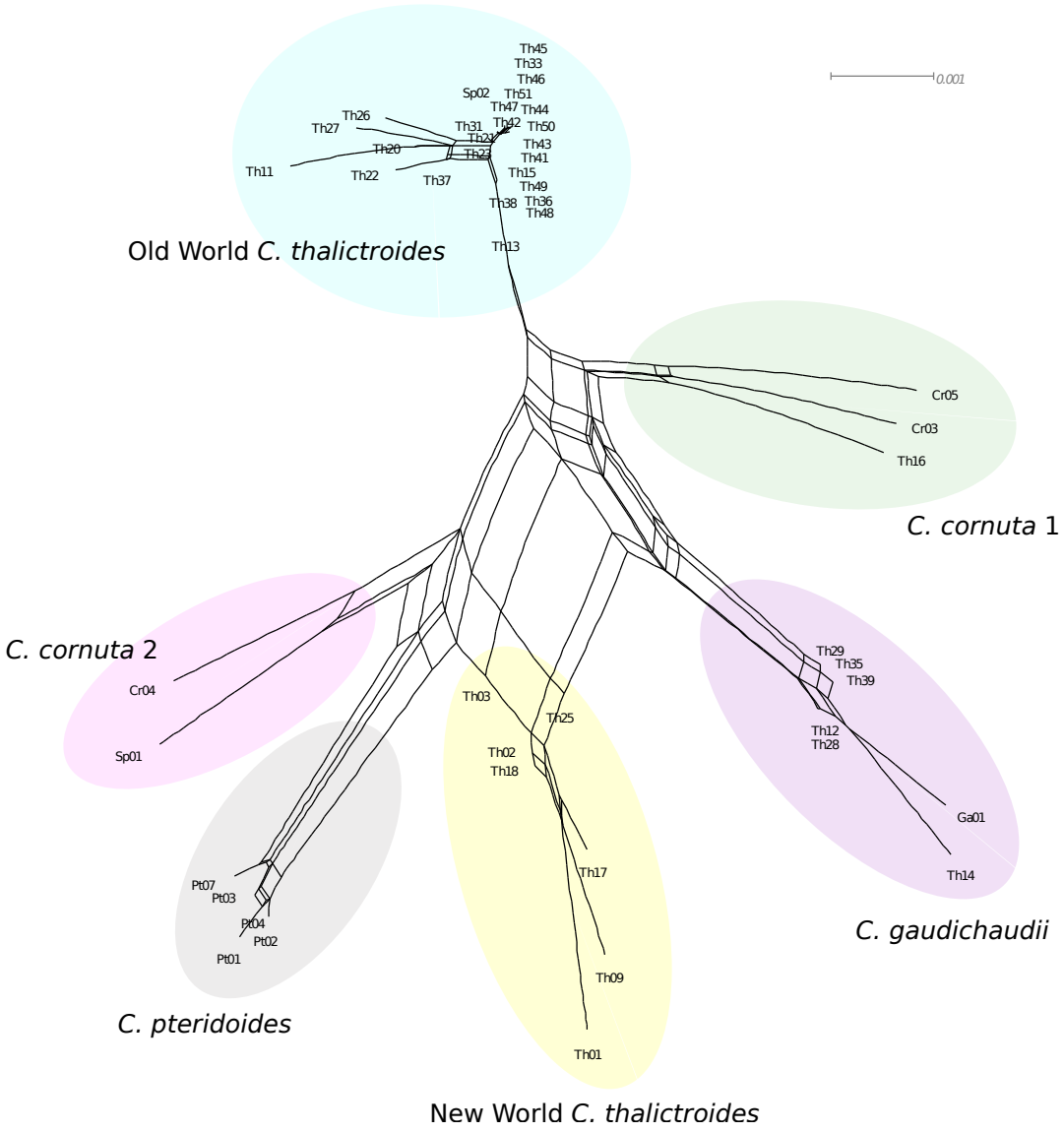


Fig. 5.1: Network analysis of *Ceratopteris* using SplitsTree4. Branch tips represent individuals, and each major group is labeled by species. Groups that share more links/sections of the network are hypothesized to have reticulate gene flow.

Table 5.2: Estimation of gene flow between species of *Ceratopteris*. Four groups have significant gene flow, emphasised in bold. The program Dsuite (Malinsky et al., 2021) was used to calculate the D-statistic from all possible quartets of species, including three ingroup taxa (P1-3, shown in this table) and one outgroup taxa (*Acrostichum aureum* for all groups, not shown) in each quartet. All species included in this study have gene flow with at least one other species as indicated by the D-statistics. This points towards rampant ancestral and/or modern gene flow throughout the genus. Specifically, *C. cornuta* has been hypothesized as a progenitor of *C. gaudichaudii* and Old World *C. thalictroides* (Adjie et al., 2007), which is substantiated by these estimates of gene flow. I suspect that with a larger sample size, and more even sampling across species, I would detect additional evidence of gene flow between taxa.

P1	P2	P3	D-statistic	p-value
C. gaudichaudii	*C. cornuta*	*C. pteridoides*	0.833002	0.00043111
Old World *C. thalictroides*	*C. cornuta*	*C. gaudichaudii*	0.420208	0.0773966
New World *C. thalictroides*	*C. cornuta*	*C. gaudichaudii*	0.17197	0.292506
Old World *C. thalictroides*	*C. cornuta*	*C. pteridoides*	0.868512	0.014955
New World *C. thalictroides*	*C. cornuta*	*C. pteridoides*	0.291677	0.241956
Old World *C. thalictroides*	*C. cornuta*	New World *C. thalictroides*	0.737024	0.0326966
Old World *C. thalictroides*	*C. gaudichaudii*	*C. pteridoides*	0.169435	0.346293
C. gaudichaudii	New World *C. thalictroides*	*C. pteridoides*	0.422691	0.0389534
New World *C. thalictroides*	*C. gaudichaudii*	Old World *C. thalictroides*	0.145423	0.245784
Old World *C. thalictroides*	New World *C. thalictroides*	*C. pteridoides*	0.401949	0.136029

relationships and hybridization within the genus (Adjie et al., 2007), but offer new insights on the prevalence of interspecies gene flow.

The split network retrieves the same two clades of *Ceratopteris thalictroides* as in chapter 4. Old World *C. thalictroides* is the most distinct species, as it is placed on a long branch in the split network, and has limited gene flow with other species as estimated by Patterson's D. In comparison, New World *C. thalictroides* also appears to be relatively distinct, but with potential shared ancestry with several other species. New World *C. thalictroides* was included in two of the four species groups with a significant Patterson's D, and potentially has shared ancestry with *C. cornuta*, *C. pteridoides*, Old World *C. thalictroides*, and *C. gaudichaudii* (Table 5.2).

The *Ceratopteris cornuta* 1 clade is relatively close to both Old World *C. thalictroides* and *C. gaudichaudii* in the split network; it also appears to have significant ancestral gene flow with other species (Table 5.2). Interestingly, the three individuals within the *C. cornuta* 1 seem to have slightly different evolutionary histories; these samples were from Tanzania, Oman, and Nepal, respectively. Since they are separated by such vast distances, slight genetic divergence is expected. There was also a second clade of *C. cornuta*, containing one individual of *C. cornuta*, and one unidentified specimen. Labeled as *C. cornuta* 2, it is very close to New World *C. thalictroides* and *C. pteridoides* in the split network. Both samples of the *C. cornuta* 2 are from within the known range of *C. richardii*, so may have been mis-identified by their collectors, or indicate undescribed variation with *C. cornuta*.

The outgroup, *Acrostichum aureum*, is mostly closely grouped with *C. pteridoides*. I chose to remove the outgroup from the final split network figure because it made the figure very difficult to read. *Acrostichum* was placed on an extremely long branch (10x longer than any other branch), and the samples of *Ceratopteris* were clumped so close together it was impossible to see the relationships between them.

Perhaps the most interesting finding is that there is a significant amount of ancestral, but not recent, gene flow among all species, as indicated by the large network in the center of Fig. 1. This is also supported by Patterson's D, which found evidence of gene flow between

four groups of species (Table 5.2), which are as follows: 1) *Ceratopteris gaudichaudii*, *C. cornuta*, and *C. pteridoides*; 2) Old World *C. thalictroides*, *C. cornuta*, and *C. pteridoides*; 3) New World *C. thalictroides*, *C. cornuta*, and Old World *C. thalictroides*; and 4) *C. gaudichaudii*, New World *C. thalictroides*, and *C. pteridoides*. All species included in this study have potential gene flow with at least one other species. Old World *C. thalictroides* and *Ceratopteris gaudichaudii* were present in two of these groups; New World *C. thalictroides* and *C. cornuta* appeared three times.

I have combined the findings from the split network analysis, as well as work by Adjie et al. (2007); Hickok & Klekowski (1974); Masuyama & Watano (2010); Zhang et al. (2020) to create an updated hypothesis of the reticulate relationships within *Ceratopteris* (Fig. 5.2).

5.5 Discussion

Here I provide evidence for reticulate evolution within the model fern genus *Ceratopteris*. I build on previous work exploring hybridization and species concepts in the group((Adjie et al., 2007; Lloyd, 1974; Masuyama & Watano, 2010), and provide the first hypothesis for the reticulate evolutionary history of the genus (Fig. 5.2). I find support for *Ceratopteris gaudichaudii* as a morphologically cryptic but genomically distinct species; separate New and Old World populations of *C. thalictroides*; plus eastern and western clades of *C. cornuta*, with the western clade (*C. cornuta* 2) potentially containing mis-identified specimens of *C. richardii*. I also find evidence of ancestral gene flow between species, as well as confirmation of more recent introgression (Fig. 5.1, Table 5.2; reviewed in Masuyama & Watano (2005); Nakazato et al. (2007)).

5.5.1 Hybridization in context for a model system

The neotropics are home to model species *Ceratopteris richardii*, as well as *C. pteridoides* and *C. thalictroides*. These species have received less attention than paleotropical species in recent molecular work, but there is some evidence of natural hybridization among New World taxa. Hickok & Klekowski (1974) showed that *C. richardii* hybridizes readily

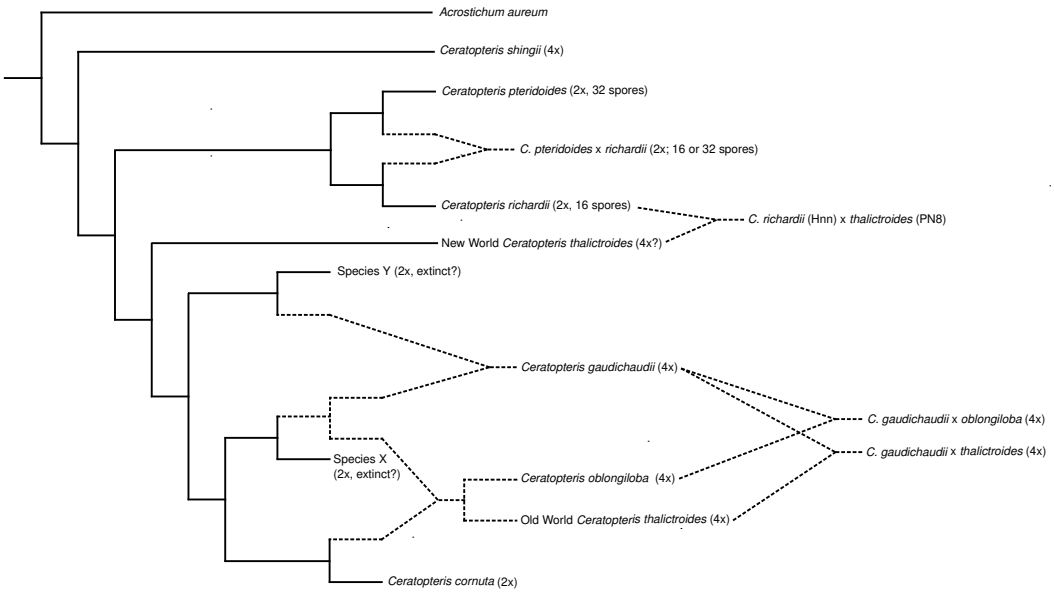


Fig. 5.2: Reticulogram showing hypothesized relationships in *Ceratopteris*. Solid lines show evolution by traditional diversification, dotted lines indicate hybridization. This diagram was created using phylogenies published by Adjie et al. (2007); Zhang et al. (2020), and work by Hickok & Klekowski (1974); Lloyd (1974). Species X and Y represent unknown or extinct diploid progenitor species, hypothesized by Adjie et al. (2007). Not included is *C. froesii*, a Brazilian endemic species, for which no genetic data exist.

with diploid *C. pteridoides* in lab conditions, with approximately 60% spore viability. Natural hybrids between these two species have also been documented, with an intermediate morphology that sometimes looks similar to *C. thalictroides* (Hickok & Klekowski, 1974). The split network results hint at the close relationship between these two taxa: the *C. cornuta 2* and *C. pteridoides* clades are placed close to one another and have recent, but not on-going gene flow (Fig. 1). Three of four species groups found to have a non-bifurcating evolutionary history included *C. cornuta* and *C. pteridoides* (Table 5.2). However, I included *C. cornuta* as just one group when estimating gene flow: when the groups were separated, the sample size for each was too small to calculate Patterson's D. Therefore, I cannot speak to which group (1 or 2) is involved in this gene flow. These results do, however, suggest that *C. richardii* and *C. pteridoides* might indeed be distinct species, yet lack reproductive boundaries to prevent interbreeding in natural conditions.

There is no evidence of recent admixture between *Ceratopteris richardii*, *C. pteridoides*,

and New World *C. thalictroides* in the split network analysis, or the population structure analysis in chapter 4. I did, however, find evidence of gene flow between these species via Patterson's D, which can detect ancient as well as recent admixture (Durand et al., 2011). This is supported by the fact that a few individuals of New World *C. thalictroides* (Th02, Th03, Th18, and Th25) are placed much closer to *C. pteridoides* than other individuals of New World *C. thalictroides* (Fig. 5.1). Perhaps New World *C. thalictroides* is the result of a past hybridization event between *C. richardii* (or *C. cornuta*) and *C. pteridoides*, where polyploidization of the hybrid lineage has led to limited gene flow with its progenitors (see Otto & Whitton (2000)). Further investigation of this lineage will be important to understand the origins of New World *Ceratopteris*, and which species are actively hybridizing.

There is evidence that the Hnn lab strain of *C. richardii* can hybridize with the lab strain PN8, but with some postzygotic barriers (Nakazato et al., 2007). Several authors have referred to PN8 as a collection of *C. richardii* with 32 spores per sporangium (Hickok, Warne, & Fribourg, 1995; Nakazato et al., 2007); however, Lloyd (1974) described *C. richardii* as having 16 spores per sporangium, and *C. thalictroides* as having 32 spores per sporangium (Table 5.2). The original specimen of PN8 (Nichols 1719, GH; legacy.tropicos.org/Specimen/1771652) is labeled as *C. thalictroides*. The fact that there are postzygotic reproductive barriers between this specimen and *C. richardii* supports this original identification by Nichols, suggesting that PN8 is not *C. richardii*. However, inclusion of this sample and Hnn in future molecular work is critical for identification. The discrepancy in identification of PN8 emphasizes the importance of updating the *Ceratopteris* taxonomy. While lab organisms can be indispensable for many studies, it is important to be aware of their phylogenetic placement, and to understand the complexities that exist in defining species boundaries. This is exemplified by work on the phylogeny and biogeography of *Arabidopsis* after the publication of its reference genome (Beck, Al-Shehbaz, O'Kane, & Schaal, 2007; Beck, Schmuths, & Schaal, 2008).

5.5.2 Polyploidy and reticulation in Old World *Ceratopteris*

Lloyd (1974) described *Ceratopteris thalictroides* as "highly polymorphic," and this

seems to be an accurate description in both morphology and genomic composition (Masuyama & Watano, 2010). Three cryptic species have been described from paleotropical plants (Masuyama & Watano, 2010), and there may be another cryptic species in the New World (see chapter 4), in addition to *C. froesii*.

The diploid species *Ceratopteris cornuta* has been hypothesized as an important progenitor of cryptic tetraploid taxa of Old World *Ceratopteris* (Adjie et al., 2007), but there is conflicting evidence as to the evolutionary history of these taxa (see chapter 4). Adjie et al. (2007) hypothesized *C. cornuta* to be the paternal progenitor, and an unknown, potentially extinct, diploid referred to as Species X, to be the material progenitor of Old World *C. thalictroides*. However, the population structure analysis in chapter 4 does not indicate that either population of *C. thalictroides* (Old World or New World) has recently hybridized with, or is the result of recent hybridization between *C. cornuta* with any other species. In fact, the opposite is potentially true: *Ceratopteris cornuta* is the only species that has a large amount of admixture within its population structure (chapter 4). In addition, the split network shows that Old World *C. thalictroides* is closely related to, but has not had recent gene flow with *C. cornuta* 1; this is evident by the long branch separating Old World *C. thalictroides* from the other species in the genus (Fig. 5.1). I did, however, find support that Old World *C. thalictroides* and *C. cornuta* may share a historic, reticulating history (Table 5.2; see Durand et al. (2011)). In addition, it is possible that progenitor Species X is extinct, or perhaps that I have just not sampled it. The data processing software (ipyrad) selects loci based on similarity throughout the dataset (Eaton & Overcast, 2020). If the progenitor(s) of Old World *C. thalictroides* were not included in this study, the loci only present from its progenitors and not shared with other species may have been left out.

A second tetraploid species thought to have arisen from a hybridization event between *C. cornuta* and Species X is *Ceratopteris oblongiloba* (Adjie et al. (2007); Fig. 5.2). However, in this case *C. cornuta* is the material progenitor, and the unknown diploid Species X is the paternal progenitor of *C. oblongiloba*. These two species may be the result of reciprocal crossing between two diploids forming two distinct polyploid species. Another instance of

reciprocal crossing between two diploids is known from the allotetraploid species *Polypodium hesperium* (Haufler, Windham, & Rabe, 1995). There are two allopatric populations of *P. hesperium*, identified by their plastid haplotypes; this may be a result of different genomes conferring advantages in separate habitats (Sigel, Windham, & Pryer, 2014). In the case of *C. oblongiloba* and Old World *C. thalictroides*, perhaps these reciprocal polyploids have had time to diversify both morphologically and ecologically since their hybrid origin.

The hybrid origin of *Ceratopteris gaudichaudii* does not seem to involve *C. cornuta* (Adjie et al., 2007), although they may have shared ancestry (Fig. 5.1; Table 5.2). Adjie et al. (2007) hypothesized this allotetraploid was derived from two potentially extinct diploids: Species X and Y (Fig. 5.2). In addition, backcrossing has been described between *C. gaudichaudii* and *C. oblongiloba* (Masuyama & Watano, 2005); and between *C. gaudichaudii* and Old World *C. thalictroides* (Masuyama & Watano, 2005); Fig. 5.2). Perhaps because *C. gaudichaudii* shares one progenitor (Species X) with *C. oblongiloba* and Old World *C. thalictroides*, their genomic similarity enables these hybrid species to reproduce with one another, although chromosome pairing is not always complete (Masuyama & Watano, 2005).

The oldest fossil from the Ceratopteridoideae sub-family is estimated to be about 42-56 My old (Bonde & Kumaran, 2002; Dettmann & Clifford, 1992; Rozefelds, Dettmann, Clifford, & Lewis, 2016), so there is potential for some deep divergences, perhaps with more recent hybridization creating the complex relationships I see today. A similar pattern of hybridization across deep time scales is seen in North American *Dryopteris* (Sessa et al., 2012a,1), which is also known to have an extinct diploid progenitor species, important in the hybrid origin of three taxa (Sessa et al., 2012a). Further investigation into divergence times, and progenitor Species X and Y of Old World *Ceratopteris* is important to understand the timeline of reticulate evolution in the genus.

5.5.3 Outlying taxa

The recently discovered *Ceratopteris shingii* was not included in this study, but Zhang et al. (2020) did not find any evidence of hybridization in this lineage. *Ceratopteris shingii*

is endemic to the island province of Hainan in southern China; molecular evidence shows that it is sister to all other species in the genus. The study by Zhang et al. (2020) used both Bayesian inference (MrBayes) and maximum likelihood (RAxML) to reconstruct a plastid phylogeny for the genus. Both trees showed high support for *C. shingii* as sister to the rest of the genus; however, the relationships between the remaining taxa had less support. Bayesian inference and maximum likelihood methods are often in accordance, but they can disagree in some cases, such as very short branch lengths (Alfaro, Zoller, & Lutzoni, 2003). A potential explanation is that *C. shingii* represents a distinct morphological and genetic lineage, exemplified by its unique creeping rhizomes (Zhang et al., 2020). The remaining species in *Ceratopteris* may be the results of a recent, rapid radiation, which would explain the discordance between the Bayesian inference and maximum likelihood trees.

The Brazilian endemic *Ceratopteris froesii* is the only species for which I were neither able to infer its phylogenetic position from other studies, nor find any available material. Masuyama & Watano (2010) described *C. froesii* as a cryptic species of *C. thalictroides*, characterized by very small plants (about 3 cm tall) known only from a few localities in Brazil. Including this unusual species in future works would help to better understand New World diversity of *Ceratopteris*, especially in the context of morphological plasticity. Masuyama (1992) found that different populations of Japanese *Ceratopteris* had distinct morphologies in the wild, but the phenotype of their progeny could be altered by manipulating greenhouse conditions. There is some evidence that *C. froesii* maintains its small size in cultivation, but it would be interesting to repeat the work done by Masuyama (1992) on South American *Ceratopteris*, to see if the New and Old World species have a similar level of morphological plasticity.

5.6 Conclusion

The present study builds on previous work on *Ceratopteris*, and investigates reticulate relationships that do not fit a traditional bifurcating tree. The split network I generated recovered similar relationships to previously studies, and shows the large amount of historical gene flow between all species of *Ceratopteris*: the web-like nature of the center of

the network indicates that there was past hybridization, leading to conflicting evolutionary histories of different loci (Fig. 5.1). Estimation of Patterson's D also confirms ancestral gene flow between almost all species in the genus (Table 5.2). These connections help explain the complicated relationships between many species of *Ceratopteris*, and why some species hybridize so readily today (e.g., Hickok & Klekowski (1974)). For future work, it will be important to include both strains of the model species *C. richardii*, as well as the newly described *C. shingii*. The taxonomic and evolutionary relevance of these species will hopefully help better understand the complex history of this genus.

References

Adjie, B., Masuyama, S., Ishikawa, H., & Watano, Y. (2007). Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. *Journal of Plant Research*, *120*(1), 129–138.

Alfaro, M. E., Zoller, S., & Lutzoni, F. (2003). Bayes or bootstrap? a simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution*, *20*(2), 255–266.

Banks, J. A. (1997). Sex determination in the fern *CCeratopteris*. *Trends in Plant Science*, *2*(5), 175–180.

Barrington, D. S., Haufler, C. H., & Werth, C. R. (1989). Hybridization, reticulation, and species concepts in the ferns. *American Fern Journal*.

Beck, J. B., Al-Shehbaz, I. A., O'Kane, S. L., Jr, & Schaal, B. A. (2007). Further insights into the phylogeny of *Arabidopsis* (Brassicaceae) from nuclear atmyb2 flanking sequence. *Molecular Phylogenetics and Evolution*, *42*(1), 122–130.

Beck, J. B., Schmuths, H., & Schaal, B. A. (2008). Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Molecular Ecology*, *17*(3), 902–915.

Blanco-Pastor, J. L., Bertrand, Y. J. K., Liberal, I. M., Wei, Y., Brummer, E. C., & Pfeil, B. E. (2019). Evolutionary networks from RAD seq loci point to hybrid origins of *Medicago carstiensis* and *Medicago cretacea*. *American Journal of Botany*, *17*, 81.

Bonde, S. D., & Kumaran, K. P. N. (2002). A permineralized species of mangrove fern *Acrostichum* L. from deccan intertrappean beds of india. *Review of Palaeobotany and Palynology*, *120*(3), 285–299.

Bruen, T. C., Philippe, H., & Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination. *Genetics*, *172*(4), 2665–2681.

Bryant, D., & Moulton, V. (2002). NeighborNet: An agglomerative method for the construction of planar phylogenetic networks. In *Algorithms in Bioinformatics*, (pp. 375–391).

Springer Berlin Heidelberg.

Bui, L. T., Pandzic, D., Youngstrom, C. E., Wallace, S., Irish, E. E., Szovenyi, P., & Cheng, C.-L. (2017). A fern AINTEGUMENTA gene mirrors BABY BOOM in promoting apogamy in *Ceratopteris richardii*. *Plant Journal*, *90*(1), 122–132.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *Genes, Genomes, Genetics*, *1*(3), 171–182.

del Carmen Molina, M., Divakar, P. K., Millanes, A. M., Sanchez, E., Ruth, D.-P., Hawksworth, D. L., & Crespo, A. (2011). *Parmelia sulcata* (Ascomycota: Parmeliaceae), a sympatric monophyletic species complex. *Lichenologist*, *43*(6), 585–601.

Dettmann, M. E., & Clifford, H. T. (1992). Phylogeny and biogeography of *Ruffordia, Mohria* and *Anemia* (Schizaeaceae) and *Ceratopteris* (pteridaceae): evidence from in situ and dispersed spores. *Alcheringa*, *16*(4), 269–314.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*.

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252.

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*(8), 2592–2594.

Eaton, D. A. R., & Ree, R. H. (2013). Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology*, *62*(5), 689–706.

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, *66*(3), 399–412.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, *6*(5), e19379.

Gastony, G. J. (1986). Electrophoretic evidence for the origin of fern species by unreduced spores. *American Journal of Botany*, *73*(11), 1563–1569.

Glover, N. M., Redestig, H., & Dessimoz, C. (2016). Homoeologs: What are they and how do we infer them? *Trends in Plant Science*, *21*(7), 609–621.

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prufer, K., Meyer, M., Burbano, H. A., Good, J. M., Schultz, R., Aximu-Petri, A., Butthof, A., Hober, B., Hoffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E. S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Z., Guvsic, I., Doronichev, V. B., Golovanova, L. V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R. W., Johnson, P. L. F., Eichler, E. E., Falush, D., Birney, E., Mullikin, J. C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., & Paabo, S. (2010). A draft sequence of the neandertal genome. *Science*, *328*(5979), 710–722.

Haufler, C. H., Pryer, K. M., Schuettpelz, E., Sessa, E. B., Farrar, D. R., Moran, R., Schneller, J. J., Watkins, J. E., & Windham, M. D. (2016). Sex and the single gametophyte: Revising the homosporous vascular plant life cycle in light of contemporary research. *Bioscience*, *66*(11), 928–937.

Haufler, C. H., Windham, M. D., & Rabe, E. W. (1995). Reticulate evolution in the *Polypodium vulgare* complex. *Systematic Botany*, *20*(2), 89–109.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, *101*(41), 14812–14817.

Hickok, L. G., & Klekowski, E. J. (1974). Inchoate speciation in *Ceratopteris*: An analysis of the synthesized hybrid *C. richardii* x *C. pteridoides*. *Evolution*, *28*(3), 439–446.

Hickok, L. G., Warne, T. R., & Fribourg, R. S. (1995). The biology of the fern *Ceratopteris* and its use as a model system. *International Journal of Plant Sciences*, *156*(3), 332–345.

Hou, Y., Nowak, M. D., Mirré, V., Bjorå, C. S., Brochmann, C., & Popp, M. (2016). RAD-seq data point to a northern origin of the arctic-alpine genus cassiope (ericaceae). *Molecular Phylogenetics and Evolution*, *95*, 152–160.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254–267.

Lehtonen, S. (2018). × *Lindsaeosoria flynnii* (Lindsaeaceae), another confirmed example of deep hybridization among the ferns. *American Fern Journal*, *108*(1), 7–18.

Linder, C. R., & Rieseberg, L. H. (2004). Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany*, *91*(10), 1700–1708.

Lloyd, R. M. (1974). Systematics of the genus *Ceratopteris* Brongn. (Parkeriaceae) II. Taxonomy. *Brittonia*, *26*(2), 139–160.

Ma, Z.-Y., Wen, J., Tian, J.-P., Jamal, A., Chen, L.-Q., & Liu, X.-Q. (2018). Testing reticulate evolution of four *Vitis* species from east asia using restriction-site associated DNA sequencing. *Journal of Systematics and Evolution*, *56*(4), 331–339.

Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite - fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, *21*(2), 584–595.

Marchant, D. B., Sessa, E. B., Wolf, P. G., Heo, K., Barbazuk, W. B., Soltis, P. S., & Soltis, D. E. (2019). The C-Fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. *Scientific Reports*, *9*(1), 18181.

Masuyama, S. (1992). Clinal variation of frond morphology and its adaptive implication in the fern *Ceratopteris thalictroides* in Japan. *Plant Species Biol.*, *7*(2-3), 87–96.

Masuyama, S. (2008). Cryptic species in the fern *Ceratopteris thalictroides* (Parkeriaceae).

III. referential diagnostic characters of three cryptic species. *Journal of Plant Research*, *121*(3), 279–286.

Masuyama, S., & Watano, Y. (1990). Trends for inbreeding in polyploid pteridophytes. *Plant Species Biology*, *5*(1), 13–17.

Masuyama, S., & Watano, Y. (2005). Cryptic species in the fern *Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae). II. Cytological characteristics of three cryptic species. *Acta phytotaxonomica et geobotanica*, *56*(3), 231–240.

Masuyama, S., & Watano, Y. (2010). Cryptic species in the fern *Ceratopteris thalictroides* (L.) Brongn. Parkeriaceae. IV. Taxonomic revision. *Acta Phytotaxonomica et Geobotanica*, *61*(2), 75–86.

Masuyama, S., Yatabe, Y., Murakami, N., & Watano, Y. (2002). Cryptic species in the fern *CCeratopteris thalictroides* (L.) Brongn. (Parkeriaceae). I. Molecular analyses and crossing tests. *Journal of Plant Research*, *115*(1118), 87–97.

Moura, A. E., Shreves, K., Pilot, M., Andrews, K. R., Moore, D. M., Kishida, T., Moller, L., Natoli, A., Gaspari, S., McGowen, M., Chen, I., Gray, H., Gore, M., Culloch, R. M., Kiani, M. S., Willson, M. S., Bulushi, A., Collins, T., Baldwin, R., Willson, A., Minton, G., Ponnampalam, L., & Hoelzel, A. R. (2020). Phylogenomics of the genus tursiops and closely related delphininae reveals extensive reticulation among lineages and provides inference about eco-evolutionary drivers. *Molecular Phylogenetics and Evolution*, *146*, 106756.

Nakazato, T., Jung, M.-K., Housworth, E. A., Rieseberg, L. H., & Gastony, G. J. (2007). A genomewide study of reproductive barriers between allopatric populations of a homosporous fern, *Ceratopteris richardii*. *Genetics*, *177*(2), 1141–1150.

Nauheimer, L., Cui, L., Clarke, C., Crayn, D. M., Bourke, G., & Nargar, K. (2019). Genome skimming provides well resolved plastid and nuclear phylogenies, showing patterns of deep reticulate evolution in the tropical carnivorous plant genus *Nepenthes* (Caryophyllales). *Australian Systematic Botany*, *32*(3), 243–254.

Nygren, A. (2014). Cryptic polychaete diversity: a review. *Zoologica Scripta*, *43*(2), 172–183.

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*, 401–437.

PPG I (2016). A community-derived classification for extant lycophytes and ferns: PPG I. *Journal of Sytematics Evolution*, *54*(6), 563–603.

Ranker, T. A., & Sundue, M. A. (2015). Why are there so few species of ferns? *Trends in Plant Science*, *20*(7), 402–403.

Rothfels, C. J., Johnson, A. K., Hovenkamp, P. H., Swofford, D. L., Roskam, H. C., Fraser-Jenkins, C. R., Windham, M. D., & Pryer, K. M. (2015). Natural hybridization between genera that diverged from each other approximately 60 million years ago. *American Naturalist*, *185*(3), 433–442.

Rozefelds, A. C., Dettmann, M. E., Clifford, H. T., & Lewis, D. (2016). Macrofossil evidence of early sporophyte stages of a new genus of water fern *Tecaropteris* (Ceratopteridoideae: Pteridaceae) from the Paleogene Redbank Plains Formation, southeast Queensland, Australia. *Alcheringa*, *40*(1), 1–11.

Sarkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to open the treasure chest? optimising DNA extraction from herbarium specimens. *PLoS One*, *7*(8), e43808.

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012a). Phylogeny, divergence times, and historical biogeography of new world *Dryopteris* (Dryopteridaceae). *American Journal of Botany*, *99*(4), 730–750.

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012b). Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology*, *12*, 104.

Sigel, E. M. (2016). Genetic and genomic aspects of hybridization in ferns. *Journal of Systematics and Evolution*.

Sigel, E. M., Windham, M. D., & Pryer, K. M. (2014). Evidence for reciprocal origins in *Polypodium hesperium* (Polypodiaceae): a fern model system for investigating how

multiple origins shape allopolyploid genomes. *American Journal of Botany*, *101*(9), 1476–1485.

Smith, A. R. (1972). Comparison of fern and flowering plant distributions with some evolutionary interpretations for ferns. *Biotropica*, *4*(1), 4–9.

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, *60*, 561–588.

Stebbins, G. L., Jr. (1950). *Varition and Evolution in Plants*. Columbia University Press.

Vasco, A., Smalls, T. L., Graham, S. W., Cooper, E. D., Wong, G. K.-S., Stevenson, D. W., Moran, R. C., & Ambrose, B. A. (2016). Challenging the paradigms of leaf evolution: Class III HD-Zips in ferns and lycophytes. *New Phytologist*, *212*(3), 745–758.

Yamauchi, A., Hosokawa, A., Nagata, H., & Shimoda, M. (2004). Triploid bridge and role of parthenogenesis in the evolution of autopolyploidy. *American Naturalist*, *164*(1), 101–112.

Zhang, R., Yu, J.-H., Shao, W., Wang, W.-Q., & Yan, Y.-H. (2020). *Ceratopteris shingii*, a new species of *Ceratopteris* with creeping rhizomes from Hainan, China. *Phytotaxa*, *449*(1), 23–30.

CHAPTER 6

The hybrid history of the C-fern (*Ceratopteris richardii*)

6.1 Abstract

Model organisms have an important role in the biological sciences. They serve as a representative of a taxon for research on any number of biological processes. Model organisms, and specifically model plants, are selected for study because of they are conducive to lab research: a short life cycle, can be cultivated easily, and have developmental pathways that are easy to observe and manipulate. However, sometimes the evolutionary history of model organisms is often one of the least understood aspects of their biology. It is critical to have a solid evolutionary context and taxonomic framework for perform any kind of comparative or developmental studies with a model species. The model fern *Ceratopteris richardii*, also known as the C-fern, is one such model that is lacking an evolutionary context. The C-fern has been suspected as a diploid hybrid, but this has never been investigated thoroughly. In addition, only the lab strain Hnn has been used for any phylogenetic studies, and no wild collections have been studied to understand this species across its full range. In this chapter, I place the model fern *C. richardii* in an evolutionary context for the first time using a combination of lab strains and wild-collected specimens. I demonstrate that *C. cichardii* may not be a distinct species, but rather a collection of hybrid plants with similar morphotypes. This research exemplifies the importance of understanding the taxonomy and evolutionary history of a model organism.

6.2 Introduction

A model organism is a species used to study various aspects of biology, often in a controlled setting (Leonelli & Ankeny, 2013). While model organisms are essential to many fields of modern science, there are disadvantages to the intensive study of just one representative of a taxon. Model species are often chosen because of convenience: many models are physically small, have fast life cycles, and canalized morphology, development, and genetics (Bolker, 1995). For example, the common fruit fly (*Drosophila melanogaster*) was chosen as a model organism because of its close proximity to humans and their food (Kohler, 1994). In addition, *Drosophila* is easy to work with in a laboratory setting; the same is true for the model angiosperm *Arabidopsis thaliana* (Somerville & Koornneef, 2002). A weakness of model organisms is that their natural and evolutionary history is often poorly understood (Alfred & Baldwin, 2015; Otero, Fernandez-Mazuecos, & Vargas, 2021): one example is the model nematode *Caenorhabditis elegans*, which has been studied extensively since the 1970s, but it was not until the 2010s that anything was known about its natural history or closely related species (Felix & Braendle, 2010; Frezal & Felix, 2015). Understanding the evolutionary context of a model species is important for any type of comparative study, research on adaptation, or trait evolution (Alfred & Baldwin, 2015).

The model fern genus *Ceratopteris* is one group where our understanding of its evolutionary history has lagged far behind research in other areas. *Ceratopteris* is a small, pantropical genus of about eight species that grows in or near seasonal water sources (Lloyd, 1974; Masuyama & Watano, 2010; Zhang, Yu, Shao, Wang, & Yan, 2020) (Fig. 6.1). The species *C. richardii*, also known as the "C-fern", was developed as a model organism in 1980s and 90s for research on plant development, mating systems, and physiology (Cooke, Hickok, & Sugai, 1995; Eberle, Nemacheck, Wen, Hasebe, & Banks, 1995; Hickok, Warne, & Slocum, 1987). More recently, a linkage map (Nakazato, Jung, Housworth, Rieseberg, & Gastony, 2006) and reference genome (Marchant et al., 2019) have been published for the lab strain Hnn of *C. richardii*. This is the only homosporous fern for which these resources are available, making it a valuable asset for fern genomics. Despite its importance as a

model species, the evolutionary history of *C. richardii* is largely unknown.

Ceratopteris was first described as *Acrostichum thalictroides* by Linneaus (Linnaeus, 1764). The genus name was later segregated from *Acrostichum*, and several more species described (Benedict, 1909; A. Brongniart, 1821; A. T. Brongniart, 1823; Hooker, 1825; Underwood, 1907). The first monograph of the genus was written by Lloyd (1974), including four species: *C. thalictroides*, *C. cornuta*, *C. pteridoides*, and *C. richardii*. In 2010, Masuyama and colleagues named three cryptic species of *C. thalictroides* from Asia and Oceania: *C. gaudichaudii*, *C. oblongiloba*, and *C. froesii* (from Brazil). Most recently, Zhang *et al.* (2020) described *C. shingii*, endemic to Hainan Province in China, and the only species in the genus with creeping rhizomes. In addition, there is potentially more cryptic variation within the species *C. cornuta* (see chapters 4 and 5), as well as many documented hybrid plants among several species in the genus (Adjie, Masuyama, Ishikawa, & Watano, 2007; Masuyama & Watano, 2005).

Missing from the evolutionary picture of *Ceratopteris*, however, is the model species *C. richardii*. This species is native to the Caribbean, as well as throughout tropical Africa (Lloyd, 1974). Several studies have placed the Hnn lab strain of *C. richardii* (collected in Cuba) as sister to the New World species *C. pteridoides* (Adjie et al., 2007; Zhang et al., 2020), but none have used wild or topotypic collections. This is perhaps because despite its use as a lab organism, there are very few herbarium collections and known localities of *C. richardii*, which makes studying it in its natural state very challenging. Incorporating wild collections of *C. richardii* into evolutionary studies is critical because in order to best utilize the findings derived from studies of a model organism, there needs to be a wild type (or natural state) with which to compare results. in the case of *C. richardii*, understanding its evolutionary history is even more important because it is suspected of being a hybrid species (Lloyd, 1974). If it is a recently derived hybrid species, this would mean that its genome structure is fundamentally different from a typical diploid species (e.g., Chaturvedi et al. (2020)), and might affect how its genomic resources are used.

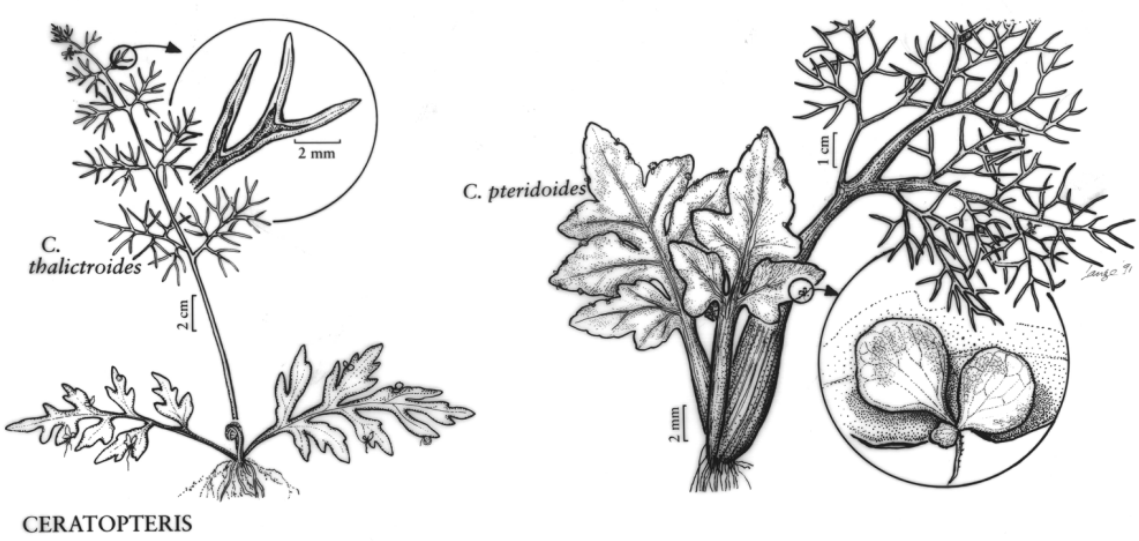


Fig. 6.1: *Ceratopteris thalictroides* and *C. pteridoides*, from Flora of North America: Volume 2 (Flora of North America Editorial Committee, 1993). *Ceratopteris thalictroides* is mainly terrestrial, cosmopolitan; *C. pteridoides* is mainly aquatic, New World.

There are several characteristics that suggest that *C. richardii* is a hybrid species, and may potentially have multiple origins given its broad range. In the original monographic description of *C. richardii*, Lloyd notes that it is potentially a hybrid species between *C. cornuta* and *C. thalictroides*, based on its intermediate morphology. In addition, *C. richardii* has 16 spores per sporangium, whereas all other species of *Ceratopteris* have 32 spore per sporangium. A lower spore number in ferns is often indicative of hybrid history or an apomictic reproductive status (Grusz, 2016). In lab conditions, apogamy is easy to induce in *C. richardii*, although it mainly reproduces sexually (Cordle, Irish, & Cheng, 2007; Hickok, Warne, & Fribourg, 1995). It is possible that their may be multiple origins of *C. richardii* via hybridization events across this large geographic area, a phenomenon that has been documented in other hybrid plant species (Harrison & Larson, 2016; Wolf et al., 2015).

In this study, I use restriction site-associated DNA sequencing (RADseq) to investigate the evolutionary history of the model fern *Ceratopteris richardii*. I assess whether the identification of *C. richardii* specimens is accurate, and how this affects hybridization

detection among all species in the genus. To accomplish this, I assess population structure and hybridization with five of eight species in the genus. In addition, I use a split network analysis to reconstruct evolutionary relationships in the genus, because previous work has shown that species relationships within *Ceratopteris* are highly reticulate (see chapter 5). The findings are corroborate previous work (see chapters 4 and 5), with the addition of placing *C. richardii* in evolutionary context for the first time. I show that the C-fern may not be a single species, but rather a hybrid morphotype with multiple origins. None of the specimens of *C. richardii* group together in the population structure or split network analysis, and two of these specimens show significant evidence of hybridization. The original collection of the Hnn strain is identical in genome composition to the Old World diploid *C. cornuta*. Finally, I evaluate the complex evolutionary history of *C. richardii* in the context of a model organism, and discuss the potential for alternative model species of *Ceratopteris*.

6.3 Methods

To assess the evolutionary history of the model fern *Ceratopteris richardii* I sampled a wide range of specimens, representing five of eight species of *Ceratopteris* from throughout the pantropical range of the genus. I constructed a genomic dataset with RADseq, which was processed using the ipyrad pipeline (Eaton & Overcast, 2020) for downstream analysis. I investigated population structure (Pritchard, Stephens, & Donnelly, 2000), hybridization among species and individual samples (Blischak, Chifman, Wolfe, & Kubatko, 2018), and utilized a split network approach to reconstruct species relationships (Huson & Bryant, 2006). All parameter values and code for data processing and analyses can be found on GitHub (github.com/sylviakinosian/ceratopterisRADseq). Demultiplexed, unfiltered reads can be accessed from the NCBI GenBank Short Read Archive (PRJNA606596).

6.3.1 Sample collection

I used a combination of silica-dried and herbarium samples to obtain a world-wide collection of samples of *Ceratopteris*. This dataset includes 108 samples of *Ceratopteris*, and one sample from its sister genus *Acrostichum*. Of the eight named species of *Ceratopteris*,

I were able to collect five: *C. thalictroides*, *C. cornuta*, *C. pteridoides*, *C. richardii*, and *C. gaudichaudii*. The species I were not able to include are: *C. froesii*, a Brasilian endemic; *C. oblongiloba*, endemic to the Philippines; and *C. shingii*, a recently described species endemic to Hainan Provence, China. Full specimen collection details are available in Appendicies A and C.

6.3.2 Sequencing

Most samples were supplied as silica-dried leaf tissue to the University of Wisconsin-Madison Biotechnology Center, where DNA was extracted using the QIAGEN DNeasy mericon 96 QIAcube HT Kit. The remainder (17 samples) were extracted from silica-dried leaf tissue using a modified CTAB method by SPK at Utah State University. Both the University of Wisconsin-Madison Biotechnology Center and SPK used variants of the CTAB DNA extraction method (Doyle & Doyle, 1987), as research suggests it is the best for extracting high-purity DNA from herbarium specimens (Sarkinen, Staats, Richardson, Cowan, & Bakker, 2012)). All specimens were sent to the University of Wisconsin-Madison Biotechnology Center, analyzed for quality, and then pooled with the rest of the samples. Restriction site-associated DNA sequencing (RADseq) libraries were prepared by University of Wisconsin-Madison Biotechnology Center, following Elshire et al. (2011) with minimal modification; sequencing was performed on Illumina NovaSeq 6000 2x150.

6.3.3 Data processing

Raw data were demultiplexed using stacks v. 2.4 process_radtags (J. Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; J. M. Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011). The resulting demultiplexed FASTQ files were paired, and low quality bases, adapters, and primers removed using ipyrad version 0.9.52 (Eaton & Overcast, 2020). I used default settings for the remainder of data processing, with the exception that I required a locus to have a sequencing depth of at least 6, and data from a minimum of 10 samples for inclusion in the final assembly to reduce missing data. For the purpose of this study I use the ipyrad pipeline's definition of a locus as a short sequence present across

samples. Single nucleotide polymorphisms (SNPs) are selected from each locus and used in downstream analyses. Samples that were low quality (yielding less than 1000 loci) were removed from downstream analyses.

Because the data are a mix of diploid and tetraploid species (Lloyd, 1974; Masuyama & Watano, 2010), I partitioned the data based on ploidy within the ipyrad pipeline. For diploid species, I set the maximum number of alleles per locus at 2, and 4 for tetraploids. All samples were run together for the demultiplexing and initial clustering steps. Then, each ploidy level was run separately through estimation of heterozygosity and error rate, and consensus base calling and filtering steps. All samples were merged for alignment and SNP calling. During these final steps, I set the maximum number of alleles per locus to 4 to attempt to account for the potential variability between samples.

6.3.4 Population structure

I used the population structure analysis program STRUCTURE v. 2.3.4 (Pritchard et al., 2000) to investigate relationships between species. The program assumes that each individual's genome is a mosaic from K ancestral populations and uses genotype assignments from SNPs to infer population structure and admixture. I utilized a set of 41,469 unlinked SNPs, created in ipyrad by selecting one SNP per loci. I ran STRUCTURE for $K = 2 – 6$ with 50 chains for each K. I then used CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015) to process the STRUCTURE output and estimate the best K values; I also visually inspected each K value to determine which was the most biologically meaningful (Evanno, Regnaut, & Goudet, 2005; Pritchard et al., 2000). I plotted STRUCTURE output in R v. 3.6.3 (R Core Team, 2018), using custom plotting functions. Using the STRUCTURE output, I adjusted naming of specimens according to which major group they clustered with.

6.3.5 Introgression tests

The program HyDe was developed to detect hybridization using phylogenetic invariants (Blischak et al., 2018). HyDe uses D-statistics to analyze hybridization at the population or

individual level for triplets of taxa (two progenitors and one hybrid). Admixture proportions are reported as γ-values to indicate contribution of progenitor taxa to a hybrid taxa. Equal values of γ (50:50) indicate a balanced hybrid, with equal contributions from both parents. If γ-values are highly unbalanced (i.e., close to 0.99 or 0.01), this could indicate more ancient hybridization, incomplete lineage sorting, or backcrossing between hybrids and progenitors (N. D. Wagner, He, & Horandl, 2020). HyDe reports p-values to test for significance of admixture.

I wanted to see how the naming of *C. richardii* affected the detection of hybridization. I did this by analyzing the data in two groups. Group 1 had samples of *C. richardii* labeled as they were identified by their collectors. All individuals in Group 2 were named according to their population assignment in STRUCTURE; all samples of *C. richardii* were grouped with other species, and did not cluster together (Fig. 6.2). For both groups, I first looked for evidence of hybridization among all samples using the run_hyde.py program. The output from run_hyde.py provided statistical support for which species showed evidence of hybridization, and which were potentially progenitors. I used this to inform the individual_hyde.py program, which detects individuals that have evidence of admixture. I followed the methodology of Wagner *et al.* (2020) and characterized recent hybrids as specimens with γ=0.4-0.6, and specimens with γ=0.1-0.4 and γ=0.6-0.9 ancient hybrids.

6.3.6 Split network analysis

Split networks are implicit representations of evolution, and can depict multiple phylogenetic hypotheses in one figure, as well as incongruencies between them. They are a good way to explore data that may not be suitable for traditional phylogenetic approaches, such as a group known to have a reticulate evolutionary history (Huson & Bryant, 2006). Previous phylogenetic studies on *Ceratopteris* have yielded slightly different topologies (Adjie et al., 2007; Zhang et al., 2020), so I wanted to explore a method outside of a bifurcating tree to better understand the evolutionary history of the group.

To investigate the complex evolutionary history of *Ceratopteris*, I utilized the NeighborNet (Bryant & Moulton, 2002) split network algorithm in SplitsTree v. 4.16.1 (Huson

& Bryant, 2006). NeighborNet is similar to a neighbor joining (NJ) tree in that it pairs samples based on similarity, agglomerating taxa into larger and larger groups (Bryant & Moulton, 2002). Unlike NJ, however, NeighborNet does not produce a bifurcating tree: it creates a split network, where parallel lines indicate splits of taxa, and boxes created by these lines indicate conflicting signals (hybridization) (Bryant & Moulton, 2002).

6.4 Results

Sequencing returned an average 3.9 x 10^6 across all samples, with an average of 13,980 loci in the final assembly with 70% missing data. I removed all samples that retained less than 1000 loci in the final assembly. After initial data analysis, I decided than samples with fewer than 1000 loci were being grouped together based on missing data, rather than any true relation. The final dataset contained 89 samples of *Ceratopteris*, and one sample of *Acrostichum*.

6.4.1 Population structure and hybridization

I determined the best K value for STRUCTURE to be 5, because it separated most named species into distinct populations (Fig. 6.2); in addition, K values higher than 5 did not add any biologically meaningful groups. The analysis recovered the same populations in chapter 4, but with the increased sampling I recovered more potential hybrid specimens. Most named species were grouped into distinct populations, with some evidence of hybridization and gene flow among them. Within the *C. cornuta* population, all specimens shared ancestry with *C. pteridoides*, and about half with New World *C. thalictroides*. Two individuals showed recently hybridization between *C. gaudichaudii* and New World *C. thalictroides* (Th81, Th110), and one between *C. gaudichaudii* and Old World *C. thalictroides* (Th13). About half of Old World *C. thalictroides* showed a small amount of shared ancestry with *C. cornuta*, *C. gaudichaudii*, and/or New World *C. thalictroides*. None of the individual of *C. richardii* were separated as a distinct species. They either clustered with *C. cornuta* (Ra10), *C. pteridoides* (Ra14), or Old World *C. thalictroides* (Ra03).

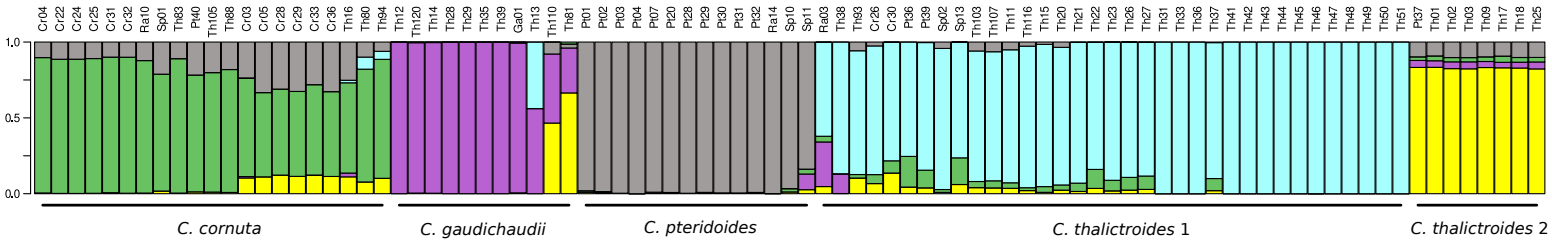


Fig. 6.2: Population structure of *Ceratopteris* for $K = 5$. At this K value, all species of *Ceratopteris* were grouped into distinct genomic populations, except for *C. richardii*. All specimens of *C. richardii* (Ra03, Ra10, Ra14, and Th88) clustered with different species: Ra10 (Hnn, Cuba) and Th88 (PN8, Nicaragua) clustered with *C. cornuta*; Ra14 (Louisiana, USA) clustered with *C. pteridoides*; and Ra03 (Trinidad) clustered with Old World *C. thalictroides*

6.4.2 Hybridization among *Ceratopteris* species

Group 1 – This group included all species of *C. richardii* as they were named by their collectors. To detect hybridization at the species level, HyDe used 3,157,886 sites to perform 60 tests, and found 20 combinations with significant evidence of hybridization (Table 6.1). I used these species groups to inform the individual-level analysis, where 402 combinations were tested, yielding 337 significant combinations among 48 specimens. Three specimens of *C. richardii* has significant evidence of hybridization with the following average γ values: Ra03 (Trinidad), 0.40; Ra14 (Louisiana, USA), 0.45; and Th88 (Nicaragua), 0.36. The former two had fairly even contributions from potential progenitor species suggesting recent hybrid ancestry. All species expect Old World *C. thalictroides* were included as potential progenitors for these two samples. The sample from Nicaragua (Th88) was the original collection of the lab strain PN8 had a lower γ value, showing a larger contribution of the progenitor *C. pteridoides*. A topotypic specimen for the lab strain Hnn (Ra10) did not show any evidence of hybridization.

Group 2 – This group used the output of STRUCTURE (Fig. 6.2) to name samples of *C. richardii*. To test for hybridization at the species level, HyDe used 3,157,886 sites to perform 30 tests, and found 10 combinations with significant evidence of hybridization (Table 6.2). I used these species groups to inform the individual-level analysis. At the individual level, HyDe tested 255 combinations and found 218 significant combinations among 47 specimens. The only sample of *C. richardii* to have significant evidence of hybridization was Ra03, a sample from Trinidad. This sample had an average γ value of 0.39.

6.4.3 Split network analysis

The split network analysis with SplitsTree4 yielded similar groupings of individuals as STRUCTURE (Fig. 6.2). The main difference was that individuals grouped into the *Ceratopteris cornuta* population in STRUCTURE were split into four groups in SplitsTree. All three samples labeled as *C. richardii* were placed in separate clades in the split network.

Table 6.1: HyDe statistics for species-level hybridization detection for Group 1. γ indicates the proportion of ancestry derived from P2; 1-γ indicated the proportion from P1. Combinations with *C. richardii* are shown in bold text.

P1	Hybrid	P2	Z-score	P-value	Gamma (γ)
pteridoides	**richardii**	**thalictroides2**	**18.370**	**0**	**0.155**
pteridoides	thalictroides1	thalictroides2	57.129	0	0.207
gaudichaudii	**richardii**	**thalictroides2**	**17.697**	**0**	**0.218**
pteridoides	cornuta	gaudichaudii	14.767	0	0.249
pteridoides	thalictroides2	gaudichaudii	18.251	0	0.262
gaudichaudii	thalictroides1	thalictroides2	82.136	0	0.311
cornuta	thalictroides1	thalictroides2	79.967	0	0.316
cornuta	thalictroides1	richardii	61.124	0	0.316
cornuta	**richardii**	**thalictroides2**	**25.168**	**0**	**0.324**
pteridoides	**richardii**	**gaudichaudii**	**37.497**	**0**	**0.341**
pteridoides	thalictroides1	gaudichaudii	67.083	0	0.359
richardii	thalictroides1	thalictroides2	94.223	0	0.452
cornuta	thalictroides1	gaudichaudii	78.384	0	0.516
cornuta	**richardii**	**gaudichaudii**	**32.845**	**0**	**0.609**
cornuta	thalictroides2	gaudichaudii	14.817	0	0.629
cornuta	thalictroides1	pteridoides	59.089	0	0.669
richardii	thalictroides1	gaudichaudii	68.692	0	0.679
cornuta	**richardii**	**pteridoides**	**25.511**	**0**	**0.759**
richardii	thalictroides1	pteridoides	42.257	0	0.809
cornuta	thalictroides2	pteridoides	5.999	9.9×10^{-10}	0.888

Table 6.2: HyDe statistics for species-level hybridization detection for Group 2. γ indicates the proportion of ancestry derived from P2; 1-γ indicated the proportion from P1.

P1	Hybrid	P2	Z-score	P-value	Gamma (γ)
pteridoides	thalictroides1	thalictroides2	61.556	0	0.232
gaudichaudii	thalictroides1	thalictroides2	80.798	0	0.311
cornuta	thalictroides1	thalictroides2	78.220	0	0.312
pteridoides	cornuta	gaudichaudii	15.247	0	0.323
pteridoides	thalictroides2	gaudichaudii	21.922	0	0.343
pteridoides	thalictroides1	gaudichaudii	72.033	0	0.401
cornuta	thalictroides1	gaudichaudii	77.658	0	0.512
cornuta	thalictroides2	gaudichaudii	15.978	0	0.607
cornuta	thalictroides1	pteridoides	64.435	0	0.620
cornuta	thalictroides2	pteridoides	9.552	0	0.799

This placement was very similar to that of the same individuals in the STRUCTURE analysis. The split network shows that there has been ancestral gene flow between all species of *Ceratopteris*, with modern hybridization within some specimens. *Ceratopteris cornuta* has several specimens that may have recent hybridization with samples of *C. pteridoides*. Between specimens of *C. pteridoides*, its seems that there might be some current diversification, but continued gene flow.

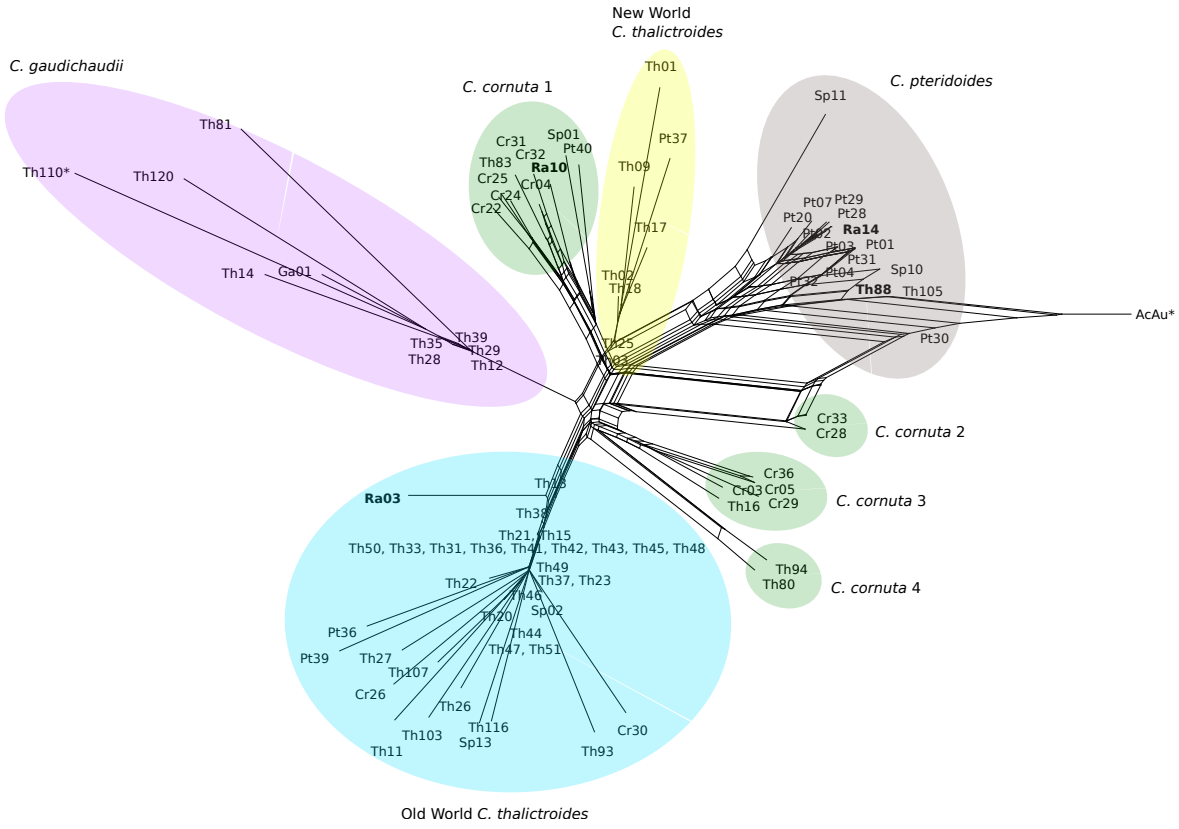


Fig. 6.3: Split network analysis of relationships in *Ceratopteris*. Clades are colored the same as STRUCTURE populations: green, *C. cornuta*; purple, *C. gaudichaudii*; grey, *C. pteridoides*, blue, Old World *C. thalictroides*; and yellow, New World *C. thalictroides*. The *C. cornuta* clades 1-4 are distinguished by geographical location, respectively: Americas and Western Africa; Arabian Peninsula; Tanzania, Arabian Peninsula, Nepal; and Australia. Samples in bold text are the three accessions of *C. richardii*: Ra10, original Hnn specimen from Cuba; Ra03 from Trinidad; Ra14 from Louisiana, USA; and Th88, the original specimen of PN8 from Nicaragua. The outgroup, *Acrostichum* (AcAu), and sample Th110 are noted with an asterisk, as the original branch lengths were trimmed for this figure.

6.5 Discussion

Model organisms are a fundamental part of research in evolutionary biology. Despite their use in many fields of biological research, their evolutionary history is often missing from the picture (Alfred & Baldwin, 2015; Otero et al., 2021). This study provides the first look at the evolutionary history and hybrid ancestry of the model fern species *Ceratopteris richardii*. I find that collections of *C. richardii* are often misidentified, and modern-day localities of this species are difficult to find. Three of the four specimens of *C. richardii* in this study showed evidence of recent hybrid ancestry, although the exact progenitors are hard to identify with certainty (Tables 6.1, 6.2). The original collection of the lab strains Hnn and PN8 were identified by the analyses as *C. cornuta*. None of the specimens that were initially identified as *C. richardii* grouped together (Fig. 6.2), or showed similar evolutionary histories (Fig. 6.3), bringing in to question the true nature of this model organism.

6.5.1 Hybrid history of *Ceratopteris richardii*

Ceratopteris richardii has been utilized as a model organisms for several decades (e.g., Hickok and Klekowski (1974); Hickok et al. (1995); Masuyama and Watano (1990)), but very little is known about its evolutionary or potential hybrid ancestry. It has been suggested that *C. richardii* is a hybrid species derived from diploid *C. cornuta* and tetraploid *C. thalictroides* (Lloyd, 1974). However, single marker analyses have placed it as sister to the diploid *C. pteridoides*, and no evidence of hybridization has been detected in the lab strain Hnn (Adjie et al., 2007; Zhang et al., 2020), including by the analyses. A topotypic specimen of the original collection of Hnn (*Killip 45469*, HUH) was placed in the *C. cornuta* population in both the population structure and split network analyses (Fig. 6.2, 6.3). No hybridization was detected for this sample via HyDe. Despite the hypothesis that *C. richardii* may be a hybrid species between *C. cornuta* and *C. thalictroides* (Lloyd, 1974), at least the Hnn specimen seems to be a misidentified accession of *C. cornuta*. Further, although there was evidence of hybridization in the other three specimens initially identified as *C. richardii*, no specimens of *C. richardii* has significant evidence of hybridization

between *C. cornuta* and Old World *C. thalictroides* (Table 6.1).

In addition to Hnn, the lab strain PN8 is also sometimes used as a model lineage of *C. richardii*. It was collected in Nicaragua and was originally identified as *C. thalictroides*, but later named as a variant of *C. richardii* with 32 spore per sporangium (Hickok et al., 1995; Nakazato et al., 2006); typical *C. richardii* has 16 spores (Lloyd, 1974)). The hybridization detection analysis found evidence for ancient hybrid ancestry from *C. pteridoides*; in addition, this specimen is grouped with *C. pteridoides* in the split network analysis. As suspected in chapter 5, PN8 may not be *C. richardii*. It is reported to have 32 spores per sporangium, where *C. richardii* has 16. In addition, there are significant postzyogtic reproductive barriers between PN8 and Hnn (Nakazato, Jung, Housworth, Rieseberg, & Gastony, 2007), suggesting they may not be closely related.

Ceratopteris has a highly reticulate evolutionary history (see chapter 5; Adjie et al. (2007)). This has made naming of species challenging within the genus, requiring multiple datasets to understand the relationships between species (Adjie et al., 2007; Masuyama, 2008; Masuyama & Watano, 2005, 2010; Masuyama, Yatabe, Murakami, & Watano, 2002). This is not uncommon in fern lineages, because of their low reproductive barriers and high incidence of hybridization, leading to highly reticulate species complexes (e.g., (Adjie et al., 2007; Schneider et al., 2013; Sessa, Zimmer, & Givnish, 2012; Sigel, 2016; Suissa et al., 2020)). This has made it very difficult to determine the total number of taxa in certain clades (Hickey, Taylor, & Luebke, 1989; Popovich, Farrar, & Gilman, 2020). In some cases, hybrid species within reticulate lineages have been grouped together, rather than given distinct names. For example, in a New World group of *Diplazium* ferns, there are several species pairs that yield hybrid species; all of the hybrid taxa look very similar, despite their independent origins. It was not until recently with the aid of careful morphological and molecular examinations that these distinct hybrid taxa were discovered (Testo, Sundue, & Vasco, 2017).

In the case of *Ceratopteris*, it is possible that hybrid plants derived from several species pairs may have similar morphologies, and therefore be grouped together under the name *C.*

richardii. Hybridization is known to be possible between species in the genus (in both laboratory and natural settings; Adjie et al. (2007); Hickok and Klekowski (1974); Masuyama and Watano (2005)), and the morphological identification of these plants is often very difficult (Hickok & Klekowski, 1974). In fact, the most reliable characteristics for identifying *Ceratopteris* species are not gross features such as leaf shape or habitat, but rather annulus (outer cells of the sporangium) cell number, number of spores per sporangium, as well as ploidy and cytotype (Lloyd, 1974; Masuyama & Watano, 2010).

Studying *C. richardii* outside of its role as a lab organism is also challenging because very few modern localities are known for this species. There are a few collections of *C. richardii* are known from Central America and the Caribbean Islands, but are rare compared to the other New World species in the genus. Additionally, a small number of collections are from tropical Africa, but these are also quite rare compared to other *Ceratopteris* species from these areas. The limited number of *Ceratopteris richardii* specimens could stem from several reasons. Firstly, *Ceratopteris* morphology is variable within species and with the age of the plant; identifying specimens can be difficult without the mature fertile and sterile fronds, and many herbarium specimens are misidentified. In addition, all species of *Ceratopteris* can be challenging to find in their native habitat: they grow in seasonally wet places, and so are not present on the landscape throughout the year. In the case of *C. richardii*, human development and climate change may also be altering or destroying the wetland habitats that it once grew in.

6.5.2 Biogeography of the C-fern

In Lloyd's 1974 monograph of the genus, it is proposed that *C. richardii* was transported to the New World during the trans-Atlantic slave trade, perhaps with taro crops (Lloyd, 1974). *Ceratopteris* is known to have been transported this way from Japan to Hawaii (W. H. Wagner Jr., 1950), so such a human-assisted migration is very possible. What is also possible, however, is that *C. cornuta* was transported in the same way, and therefore is misidentified in the New World because of this historically new disjunct distribution. I find support for this expanded distribution of *C. cornuta*: in addition to the Hnn

specimen from Cuba, I identified additional samples as *C. cornuta* from Brazil (*Camargo de Abreu 21*, NY), Costa Rica (*Crow 7436*, MO), El Salvador (*Carballo 899*, MO), and Nicaragua (*Nichols 1719*, HUH). These samples cluster with collections from several west African countries: Sierra Leone, Equatorial Guinea, Gabon, and Ghana, and Senegal (clade 1, Fig. 6.3).

The remaining samples of *C. cornuta* are from the Tanzania, the Arabian Peninsula, Nepal, and Australia. These form three clades in the split network analysis, and are not closely related to the American-west African clade of *C. cornuta* (clades 2-3, Fig. 6.3). However, all of *C. cornuta* samples cluster together in the population structure analysis (Fig. 6.2). This indicates that instead of several cryptic species within *C. cornuta*, there may be subspecies diversifying based on geographic separation.

In the context of the model C-fern, the model lineages Hnn and PN8 may not be accurate representations of their species. *Ceratopteris cornuta* is a very widespread and genetically diverse species. The lab collections could potentially have hybrid histories with other species (PN8), or be an unusual cytotype or an apotmictic lineage (Hnn).

6.6 Conclusion

The model fern *Ceratopteris richardii* is an important part of fern biology research (e.g., (Aragón-Raygoza, Vasco, Blilou, Herrera-Estrella, & Cruz-Ramírez, 2020; Bui et al., 2017; Conway & Di Stilio, 2019; Cordle et al., 2007; Ganger et al., 2019; Hickok et al., 1995; McGrath & Pichersky, 1997; Scott, Gastony, Jeremy W. Weatherford, & Nakazato, 2007; Warne & Hickok, 1987)). The evolutionary relationship of this model species within the rest of the genus *Ceratopteris* is necessary to best place such research in a broader context. This study provides the first look into the complex evolutionary history of *C. richardii*, and shows that this model species is potentially a lineage nested within the diploid *C. cornuta* (Fig. 6.3). I also find that specimens labeled as *C. richardii* are in fact hybrid lineages derived from different species pairs.

Looking towards the future of research with the C-fern, a reevaluation of *C. richardii* a model species is needed. *Ceratopteris richardii* was named before *C. cornuta* (A. T. Brong-

niart, 1823; Leprieur, 1830), and so has precedence if the two species are to be considered one lineage. Alternatively, the diploid *C. pteridoides* might be a more easily identifiable option for a model species. It is native to the Americas, is easily collected, and has a very distinct morphology making it easy to identify (Lloyd, 1974). In any case, incorporating these findings into new research with the C-fern is important to best utilize this model in an evolutionary context.

References

Adjie, B., Masuyama, S., Ishikawa, H., & Watano, Y. (2007). Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. *J. Plant Res.*, *120*(1), 129–138.

Alfred, J., & Baldwin, I. T. (2015). New opportunities at the wild frontier. *Elife*, *4*.

Aragón-Raygoza, A., Vasco, A., Blilou, I., Herrera-Estrella, L., & Cruz-Ramírez, A. (2020). *Development and cell cycle dynamics of the root apical meristem in the fern Ceratopteris richardii.*

Benedict, R. C. (1909). The genus *Ceratopteris*: A preliminary revision. *Bull. Torrey Bot. Club*, *36*(8), 463–476.

Blischak, P. D., Chifman, J., Wolfe, A. D., & Kubatko, L. S. (2018). HyDe: A python package for Genome-Scale hybridization detection. *Syst. Biol.*, *67*(5), 821–829.

Bolker, J. A. (1995). Model systems in developmental biology. *Bioessays*, *17*(5), 451–455.

Brongniart, A. (1821). Description d'un nouveau genre de fougere, nomme *Ceratopteris*. *Bull. Sci. Soc. Philom. Paris ser*, *3*(8), 184–187.

Brongniart, A. T. (1823). *Dictionnaire classique d'histoire naturelle* (Vol. 351).

Bryant, D., & Moulton, V. (2002). NeighborNet: An agglomerative method for the construction of planar phylogenetic networks. In *Algorithms in bioinformatics* (pp. 375–391). Springer Berlin Heidelberg.

Bui, L. T., Pandzic, D., Youngstrom, C. E., Wallace, S., Irish, E. E., Szövényi, P., & Cheng, C.-L. (2017). A fern AINTEGUMENTA gene mirrors BABY BOOM in promoting apogamy in *Ceratopteris richardii*. *Plant Journal*, *90*(1), 122–132.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Mol. Ecol.*, *22*(11), 3124–3140.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3*, *1*(3), 171–182.

Chaturvedi, S., Lucas, L. K., Buerkle, C. A., Fordyce, J. A., Forister, M. L., Nice, C. C.,

& Gompert, Z. (2020). Recent hybrids recapitulate ancient hybrid outcomes. *Nature Communications*, *11*(1), 2179.

Conway, S. J., & Di Stilio, V. S. (2019). An ontogenetic framework for functional studies in the model fern *Ceratopteris richardii*. *Developmental Biology*.

Cooke, T. J., Hickok, L. G., & Sugai, M. (1995). The fern *Ceratopteris richardii* as a lower plant model system for studying the genetic regulation of plant photomorphogenesis. *Int. J. Plant Sci.*, *156*(3), 367–373.

Cordle, A. R., Irish, E. E., & Cheng, C. (2007). Apogamy induction in *Ceratopteris richardii*. *Int. J. Plant Sci.*, *168*(4), 361–369.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical bulletin*.

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*(8), 2592–2594.

Eberle, J., Nemacheck, J., Wen, C.-K., Hasebe, M., & Banks, J. A. (1995). *Ceratopteris*: A model system for studying sex-determining mechanisms in plants. *Int. J. Plant Sci.*, *156*(3), 359–366.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, *6*(5), e19379.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.*, *14*(8), 2611–2620.

Felix, M.-A., & Braendle, C. (2010). The natural history of *Caenorhabditis elegans*. *Current Biology*, *20*(22), R965–9.

Flora of North America Editorial Committee. (1993). *Flora of North America: Volume 2: Pteridophytes and Gymnosperms*. Oxford University Press.

Frezal, L., & Felix, M.-A. (2015). *C. elegans* outside the petri dish. *Elife*, *4*.

Ganger, M. T., Hiles, R., Hallowell, H., Cooper, L., McAllister, N., Youngdahl, D., ...

Ewing, S. J. (2019). A soil bacterium alters sex determination and rhizoid development in gametophytes of the fern *Ceratopteris richardii*. *AoB Plants*, *11*(2), lz012.

Grusz, A. L. (2016). A current perspective on apomixis in ferns. *Jnl of Sytematics Evolution*, *54*(6), 656–665.

Harrison, R. G., & Larson, E. L. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Mol. Ecol.*, *25*(11), 2454–2466.

Hickey, R. J., Taylor, W. C., & Luebke, N. T. (1989). The species concept in pteridophyta with special reference to *Isoëtes*. *American Fern Journal*, 78–89.

Hickok, L. G., & Klekowski, E. J. (1974). Inchoate speciation in *Ceratopteris*: An analysis of the synthesized hybrid *C. richardii* x *C. pteridoides*. *Evolution*, *28*(3), 439–446.

Hickok, L. G., Warne, T. R., & Fribourg, R. S. (1995). The biology of the fern *Ceratopteris* and its use as a model system. *Int. J. Plant Sci.*, *156*(3), 332–345.

Hickok, L. G., Warne, T. R., & Slocum, M. K. (1987). *Ceratopteris richardii*: Applications for experimental plant biology. *Am. J. Bot.*, *74*(8), 1304–1316.

Hooker, W. J. (1825). *Exotic flora*.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254–267.

Kohler, R. E. (1994). *Lords of the fly: Drosophila genetics and the experimental life*. University of Chicago Press.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.*, *15*(5), 1179–1191.

Leonelli, S., & Ankeny, R. A. (2013). What makes a model organism? *Endeavour*, *37*(4), 209–212.

Leprieur, F. M. R. (1830). Note sur le *Pteris cornuta* de Palisot-Beauvois espece du genre *Ceratopteris*. *Annales des Sciences Naturelles*, *19*(103), 4.

Linnaeus, C. v. (1764). *Species plantarum: Volume 2*. Imprensis Laurentii Salvii.

Lloyd, R. M. (1974). Systematics of the genus *Ceratopteris* Brongn. (Parkeriaceae) II. taxonomy. *Brittonia*, *26*(2), 139–160.

Marchant, D. B., Sessa, E. B., Wolf, P. G., Heo, K., Barbazuk, W. B., Soltis, P. S., & Soltis, D. E. (2019). The C-Fern (*Ceratopteris richardii*) genome: Insights into plant genome evolution with the first partial homosporous fern genome assembly. *Sci. Rep.*, *9*(1), 18181.

Masuyama, S. (2008). Cryptic species in the fern *Ceratopteris thalictroides* (Parkeriaceae). III. Referential diagnostic characters of three cryptic species. *J. Plant Res.*, *121*(3), 279–286.

Masuyama, S., & Watano, Y. (1990). Trends for inbreeding in polyploid pteridophytes. *Plant Species Biol.*, *5*(1), 13–17.

Masuyama, S., & Watano, Y. (2005). Cryptic species in the fern *Ceratopteris thalictroides* (l.) brongn. (Parkeriaceae). II. Cytological characteristics of three cryptic species. *Acta phytotaxonomica et geobotanica*, *56*(3), 231–240.

Masuyama, S., & Watano, Y. (2010). Cryptic species in the fern *Ceratopteris thalictroides* (l.) brongn. Parkeriaceae. IV. Taxonomic revision. *Acta. Phytotax. Geobot.*, *61*(2), 75–86.

Masuyama, S., Yatabe, Y., Murakami, N., & Watano, Y. (2002). Cryptic species in the fern *Ceratopteris thalictroides* (l.) brongn. (Parkeriaceae). I. Molecular analyses and crossing tests. *J. Plant Res.*, *115*(1118), 87–97.

McGrath, J. M., & Pichersky, E. (1997). Methylation of somatic and sperm DNA in the homosporous fern *Ceratopteris richardii*. *Plant Molecular Biology*, *35*(6), 1023–1027.

Nakazato, T., Jung, M.-K., Housworth, E. A., Rieseberg, L. H., & Gastony, G. J. (2006). Genetic Map-Based analysis of genome structure in the homosporous fern *Ceratopteris richardii*. *Genetics*, *173*(3), 1585–1597.

Nakazato, T., Jung, M.-K., Housworth, E. A., Rieseberg, L. H., & Gastony, G. J. (2007). A genomewide study of reproductive barriers between allopatric populations of a homosporous fern, *Ceratopteris richardii*. *Genetics*, *177*(2), 1141–1150.

Otero, A., Fernandez-Mazuecos, M., & Vargas, P. (2021). Evolution in the model genus *Antirrhinum* based on phylogenomics of topotypic material. *Front. Plant Sci.*, *12*, 35.

Popovich, S. J., Farrar, D. R., & Gilman, A. V. (2020). *Botrychium furculatum* (Ophioglossaceae), a new moonwort species from the Rocky Mountains of North America. *American Fern Journal*, *110*(4), 165–182.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria.

Sarkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to open the treasure chest? optimising DNA extraction from herbarium specimens. *PLoS One*, *7*(8), e43808.

Schneider, H., Navarro-Gomez, A., Russell, S. J., Ansell, S., Grundmann, M., & Vogel, J. (2013). Exploring the utility of three nuclear regions to reconstruct reticulate evolution in the fern genus *Asplenium*. *J. Syst. Evol.*, *51*(2), 142–153.

Scott, R. J., Gastony, G. J., Jeremy W. Weatherford, & Nakazato, T. (2007). Characterization of four members of the Alpha-Tubulin gene family in *Ceratopteris richardii*. *American Fern Journal*, *97*(2), 47–65.

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012). Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). *BMC Evol. Biol.*, *12*, 104.

Sigel, E. M. (2016). Genetic and genomic aspects of hybridization in ferns. *J. Syst. Evol.*.

Somerville, C., & Koornneef, M. (2002). A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews Genetics*, *3*(11), 883–889.

Suissa, J. S., Kinosian, S. P., Schafran, P. W., Bolin, J. F., Carl Taylor, W., & Zimmer, E. A. (2020). *Revealing the evolutionary history of a reticulate polyploid complex in the genus Isoëtes.*

Testo, W. L., Sundue, M., & Vasco, A. (2017). *Diplazium* hybrids involving *D. plantaginifolium* and *D. ternatum* from Mexico and Central America. *Brittonia*, *69*(3), 295–306.

Underwood, L. M. (1907). The names of some of our native ferns. *Torreya*, *7*(10), 193–198.

Wagner, N. D., He, L., & Horandl, E. (2020). Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Front. Plant Sci.*, *11*, 1077.

Wagner, W. H., Jr. (1950). Ferns naturalized in hawaii. *Occasional Papers of Bernice P. Bishop Museum*, *20*(8), 95–121.

Warne, T. R., & Hickok, L. G. (1987). Single gene mutants tolerant to NaCl in the fern *Ceratopteris*: Characterization and genetic analysis. *Plant Sci.*, *52*(1), 49–55.

Wolf, P. G., Rowe, C. A., Der, J. P., Schilling, M. P., Visger, C. J., & Thomson, J. A. (2015). Origins and diversity of a cosmopolitan fern genus on an island archipelago. *AoB Plants*, *7*.

Zhang, R., Yu, J.-H., Shao, W., Wang, W.-Q., & Yan, Y.-H. (2020). *Ceratopteris shingii*, a new species of *Ceratopteris* with creeping rhizomes from Hainan, China. *Phytotaxa*, *449*(1), 23–30.

CHAPTER 7

Conclusion

A bifurcating model of evolution does not fit all clades, and pteridophytes are no exception (e.g. Barrington, Haufler, and Werth (1989)). This dissertation builds upon a large body of work on reticulate evolution in pteridophytes (e.g. Adjie, Masuyama, Ishikawa, and Watano (2007); Haufler, Windham, and Rabe (1995); Rothfels (2021); Schneider et al. (2013); Sigel, Windham, and Pryer (2014)), and shows that even well-known groups can have a more complex evolutionary history when examined closely. In particular, this work provides the first genus-wide look at reticulate evolution in the model fern genus *Ceratopteris*, bringing to light the importance of having a solid understanding of the evolutionary relationships in such a well-studied group.

7.1 Evolutionary history in the study of model organisms

The naming of species, or taxonomy, is the structure upon which any comparative study is based. Modern phylogenetics seeks to build such a taxonomic structure using evolutionary relationships derived from DNA. However, before one can investigate genetic relationships, each species needs to be described in the context of morphology, environment, reproductive status, and other distinguishing characteristics. The *International Code of Nomenclature for algae, fungi, and plants* provides a list of rules for describing species, that ultimately links a name to a physical or "type" specimen. Type specimens provide a reference for identification of a plant species, for example, and help differentiate from close relatives or hybrid individuals. This is particularly important for model organisms, since sometimes a single specimen or locality can be the basis for a model species (Otero, Fernández-Mazuecos, & Vargas, 2021).

In the case of *Ceratopteris richardii*, the lab strain Hnn is from a single collection from Cuba (*Killip 44595*, GH), from which spores were collected for propagation (Scott &

Hickok, 1987). Other collections have been used for lab stains (e.g., PN8 from Nicaragua; *Nichols 1719*, GH), and sometimes hybridized (Nakazato, Jung, Housworth, Rieseberg, & Gastony, 2006, 2007; Scott & Hickok, 1987), further complicating the origins of these model plants. None of these specimens, however, are type material for *C. richardii* (*Richards s.n.*, P; Lloyd (1974)); in addition, some of them do not have suitable material for morphological identification (e.g., missing sterile or fertile fronds). This makes it difficult to confirm with morphology the identity of these specimens, particularly because it has been hypothesized that *C. richardii* may be a hybrid species derived from *C. cornuta* and *C. thalictroides* (Lloyd, 1974).

Re-collection from the topotypic localities for lab strains of *C. richardii* would be greatly beneficial to see if the plants still exist in these places, and also aid in more accurate identification (Otero et al., 2021). The findings from this dissertation suggest that the model strains Hnn and PN8 might actually belong to the species *C. cornuta*, or the widespread species *C. cornuta* may have more genomic variability than previously believed. Either way, model specimens must be rooted in collections that can be identified and compared to a type species. A reevaluation of *Ceratopteris* taxonomy and potential alternative model species is crucial for future work. Without a taxonomic foundation, samples cannot be accurately identified, and comparitve studies cannot accurately reconstruct the process of evolution.

7.2 Advantages and limitations of RADseq data in ferns

The Chapters 2-5 in this dissertation are some of the first works using restriction site-associated DNA sequencing (RADseq) in ferns (but also see Dauphin, Mossion, Wirth, Sandoz, and Grant (2017); Suissa et al. (2020); Wolf et al. (2019)). This technique has proven to be useful in generating genome-wide data for dozens of individuals with fairly low cost (Davey & Blaxter, 2010), and also can be used for tissue collected fresh or from herbarium samples (Särkinen, Staats, Richardson, Cowan, & Bakker, 2012). This has made collecting samples for a pantropical genus such as *Ceratopteris* much more feasible, since a mix of silica-dried and herbarium samples were used.

An important note is that this mix of tissue types did have an influence on the amount

of reads sequenced and loci retained for downstream analyses. Silica-dried specimens had fewer raw reads than herbarium samples, but the former retained more reads and loci in the final assembly than the latter (Table 7.1). DNA from herbarium specimens can be more fragmented than fresh collections, as time and preservation processes can degrade genetic material (Särkinen et al., 2012). However, work on missing data in RADseq studies has shown that if there are a few high-quality samples per clade, lower quality samples with higher amounts of missing data can still be useful; in other words, if there is balanced sampling at the tree tips with high- and low-quality individuals, accuracy can still be good along the tree edges and internal nodes (Eaton, Spriggs, Park, & Donoghue, 2017). In Chapters 2-4, some clades did not have any samples from fresh tissue collections; without a comparison, however, it is difficult to tell if this altered the results significantly. In the future, work with RADseq data would probably benefit from having a few silica-dried or fresh tissue collections per clade. This might help with the accuracy of phylogenetic or population genomic analyses, if other samples are taken from herbarium collections.

RADseq data are well-suited for population genomic analyses like STRUCTURE (Pritchard, Stephens, & Donnelly, 2000), that use single nucleotide polymorphisms for inferences about evolutionary relationships. Since RADseq data contain mostly short reads (100-400 base pairs), they are less informative for phylogeny building or allele phasing; however, they can be a good place to start building a hypothesis for reticulate relationships and hybrid progenitors. Methodologies like target sequence capture are more appropriate for building trees with polyploid and hybrid taxa, and investigating genomic contributions of progenitor species (Rothfels, 2021).

Table 7.1: RADseq raw data information for silica-dried and herbarium samples used in Chapters 2-5. Silica-dried samples had fewer raw reads, but retained more reads and loci after data processing. Herbarium samples had more raw reads, but retained fewer reads and loci in the final assembly. See Appendix A for full collection information.

Source	Raw reads	Consensus reads	Loci	Standard deviation (loci)
Silica-dried	2,981,907	50,967	28,165	3,144
Herbarium	4,123,759	61,176	11,143	7,122

7.3 Reticulate evolution and pteridophytes

Reticulate evolution is common in spore-dispersed vascular plants for several reasons. Compared to seed plants, ferns have a high prevalence of polyploidy and hybridization (Barrington et al., 1989; Otto & Whitton, 2000; Paris, Wagner, & Wagner, 1989; Sigel, 2016). This may be a result of the simple reproductive biology of ferns, and the lack of prezygotic barriers (i.e., flowers) compared to angiosperms (Rothfels et al., 2015; Smith, 1972; Sundue, Testo, & Ranker, 2015). Hybrid plants are often sterile, but fertility can be restored through genome doubling (polyploidzation) (Otto & Whitton, 2000; Rothfels, 2021; Suissa et al., 2020). In addition, fern spores can travel via wind current due to their small size, making gene flow possible across vast distances (Barrington, 1993; Tryon, 1970); this can aid to confusion about evolutionary histories if sampling does not include physically distant but genetically close populations.

While we may understand some of the prezygotic aspects of reticulate evolution in ferns, postzygotic barriers remain uncertain. Nakazato et al. (2007) compared reproductive barriers in two allopatric populations of *C. richardii*[1] and found that genetic barriers evolve slowly over time. The F2 generation was the most negatively affected, with extreme spore inviability (Nakazato et al., 2007). Work in angiosperms has shown that Dobzhanky-Muller incompatibilies (negative gene interactions) evolve via adaptation and often most severely affect the F2 generation (Dobzhansky, 1931, 1937; Fishman & Willis, 2001; Orr & Turelli, 2001). From this case study in *Ceratopteris*, fern reproductive isolation may be influenced by Dobzhanky-Muller incompatibilies, but other modes of isolation such as chromosomal inversions and translations should be investigated as well. As the genomic resources for ferns increase, this type of work on the structural nature of fern genetic incompatibilites will be more feasible, helping to begin a new chapter of work on reticulate evolution in pteridophytes.

[1]These allopatric populations (Hnn and PN8) may not, in fact, both be *C. richardii*. They have strong postzygotic barriers and their morphological identifications have varied over time; see Chapters 4 and 5

References

Adjie, B., Masuyama, S., Ishikawa, H., & Watano, Y. (2007). Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. *Journal of Plant Research*, *120*(1), 129–138.

Barrington, D. S. (1993). Ecological and historical factors in fern biogeography. *Journal of Biogeography*, *20*(3), 275–279.

Barrington, D. S., Haufler, C. H., & Werth, C. R. (1989). Hybridization, reticulation, and species concepts in the ferns. *American Fern Journal*, *79*(2), 55–64.

Dauphin, B., Mossion, V., Wirth, S., Sandoz, F., & Grant, J. (2017). Recent postglacial recolonization from in situ survival and intense gene flow within natural populations of the alpine fern *Botrychium lunaria* (Ophioglossaceae) revealed by ddRAD sequencing. *Monsieur Benjamin DAUPHIN*, 149.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Brief. Funct. Genomics*, *9*(5-6), 416–423.

Dobzhansky, T. (1931). The decrease of Crossing-Over observed in translocations, and its probable explanation. *American Naturalist*, *65*(698), 214–232.

Dobzhansky, T. (1937). Genetic nature of species differences. *American Naturalist*, *71*(735), 404–420.

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology.*, *66*(3), 399–412.

Fishman, L., & Willis, J. H. (2001). Evidence for Dobzhansky-Muller incompatibilites contributing to the sterility of hybrids between *Mimulus guttatus* and *M. nasutus*. *Evolution*, *55*(10), 1932–1942.

Haufler, C. H., Windham, M. D., & Rabe, E. W. (1995). Reticulate evolution in the *Polypodium vulgare* complex. *Systematic Botany*, *20*(2), 89–109.

Lloyd, R. M. (1974). Systematics of the genus *Ceratopteris* Brongn. (Parkeriaceae) II. Taxonomy. *Brittonia*, *26*(2), 139–160.

Nakazato, T., Jung, M.-K., Housworth, E. A., Rieseberg, L. H., & Gastony, G. J. (2006). Genetic Map-Based analysis of genome structure in the homosporous fern *Ceratopteris richardii*. *Genetics*, *173*(3), 1585–1597.

Nakazato, T., Jung, M.-K., Housworth, E. A., Rieseberg, L. H., & Gastony, G. J. (2007). A genomewide study of reproductive barriers between allopatric populations of a homosporous fern, *Ceratopteris richardii*. *Genetics*, *177*(2), 1141–1150.

Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution*, *55*(6), 1085–1094.

Otero, A., Fernández-Mazuecos, M., & Vargas, P. (2021). Evolution in the model genus *Antirrhinum* based on phylogenomics of topotypic material. *Frontiers in Plant Sciences*, *12*, 35.

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*, 401–437.

Paris, C. A., Wagner, F. S., & Wagner, W. H. (1989). Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. *American Fern Journal*, *79*(2), 46–54.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Rothfels, C. J. (2021). Polyploid phylogenetics. *New Phytologist*.

Rothfels, C. J., Johnson, A. K., Hovenkamp, P. H., Swofford, D. L., Roskam, H. C., Fraser-Jenkins, C. R., ... Pryer, K. M. (2015). Natural hybridization between genera that diverged from each other approximately 60 million years ago. *American Naturalist*, *185*(3), 433–442.

Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to open the treasure chest? optimising DNA extraction from herbarium specimens. *PLoS One*, *7*(8), e43808.

Schneider, H., Navarro-Gomez, A., Russell, S. J., Ansell, S., Grundmann, M., & Vogel, J. (2013). Exploring the utility of three nuclear regions to reconstruct reticulate

evolution in the fern genus asplenium. *Journal of Systematics and Evolution*, *51*(2), 142–153.

Scott, R. J., & Hickok, L. G. (1987). Genetic analysis of antheridiogen sensitivity in *Ceratopteris richardii*. *American Journal of Botany*, *74*(12), 1872–1877.

Sigel, E. M. (2016). Genetic and genomic aspects of hybridization in ferns. *Journal of Systematics and Evolution*.

Sigel, E. M., Windham, M. D., & Pryer, K. M. (2014). Evidence for reciprocal origins in *Polypodium hesperium* (Polypodiaceae): a fern model system for investigating how multiple origins shape allopolyploid genomes. *American Journal of Botany.*, *101*(9), 1476–1485.

Smith, A. R. (1972). Comparison of fern and flowering plant distributions with some evolutionary interpretations for ferns. *Biotropica*, *4*(1), 4–9.

Suissa, J. S., Kinosian, S. P., Schafran, P. W., Bolin, J. F., Carl Taylor, W., & Zimmer, E. A. (2020). *Revealing the evolutionary history of a reticulate polyploid complex in the genus Isoëtes.*

Sundue, M. A., Testo, W. L., & Ranker, T. A. (2015). Morphological innovation, ecological opportunity, and the radiation of a major vascular epiphyte lineage. *Evolution*, *69*(9), 2482–2495.

Tryon, R. (1970). Development and evolution of fern floras of oceanic islands. *Biotropica*, *2*(2), 76–84.

Wolf, P. G., Rowe, C. A., Kinosian, S. P., Der, J. P., Lockhart, P. J., Shepherd, L. D., ... Thomson, J. A. (2019). Worldwide relationships in the fern genus *Pteridium* (bracken) based on nuclear genome markers. *American Journal of Botany.*, *106*(10), 1365–1376.

APPENDICES

APPENDIX A

Supplementary tables and figures for chapter 4

Table A.1: Collection information for samples used in chapter 4. Sample ID (used in figures) species name, and collection information included. All collections by S. P. Kinosian are from silica-dried leaf tissue, the remaining samples are from herbarium specimens. *Continued on following pages.*

Sample name	Species	Collector	Number	Year	Location
AcAu	*Acrostichum aureum*	S. P. Kinosian	62	2018	Australia
Cr01	*Ceratopteris cornuta*	Fay	1116	1985	Sierra Leone
Cr02	*Ceratopteris cornuta*	Fay	1097	1984	Sierra Leone
Cr03	*Ceratopteris cornuta*	R. B Faden	96	1996	Tanzania
Cr04	*Ceratopteris cornuta*	A. D. A. Fay	1116	1985	Sierra Leone
Cr05	*Ceratopteris cornuta*	C. J. Rothfels	4298	2012	Oman
Cr06	*Ceratopteris cornuta*	D. S. Erdman	U48.165	1948	Saudi Arabia
Cr07	*Ceratopteris cornuta*	A. Peter	40479	1926	
Ga01	*Ceratopteris gaudichaudii*	L. Raulerson	15656	1987	Guam
Ga02	*Ceratopteris gaudichaudii*	R. C. Ching	7787	1928	China
Pt01	*Ceratopteris pteridoides*	John D. Bruza	3121	1985	USA
Pt02	*Ceratopteris pteridoides*	F. M. Givens	2229		USA
Pt03	*Ceratopteris pteridoides*	Mary W. Diddle	8289	1954	USA
Pt04	*Ceratopteris pteridoides*	F. B. Matos	2324	2013	USA
Pt05	*Ceratopteris pteridoides*	Robert R. Haynes	7640		Venezuela
Pt06	*Ceratopteris pteridoides*	F. Delascio	11181	1981	Venezuela
Pt07	*Ceratopteris pteridoides*	J. C. Solomon	7651	1982	Bolivia
Pt08	*Ceratopteris pteridoides*	D. S. Conant	948	1974	Brazil
Pt09	*Ceratopteris pteridoides*	D. P. Foudriat	13	1973	USA
Pt10	*Ceratopteris pteridoides*	Mejia	26264		Dominican Republic
Pt11	*Ceratopteris pteridoides*	Crow	7436	1989	Costa Rica
Ra01	*Ceratopteris richardii*	A. D. A. Fay	1097		Sierra Leone
Ra02	*Ceratopteris richardii*	P. Acevedo-Rodriguez	15497	2013	Puerto Rico
Ra03	*Ceratopteris richardii*	J. T. Mickel	9509	1984	Trinidad
Ra04	*Ceratopteris richardii*	S. P. Kinosian		2018	Hnn

Sample name	Species	Collector	Number	Year	Location
Ra05	*Ceratopteris richardii*	J. J. de Granville	10049	2002	Guyana
Sp01	*Ceratopteris* sp.	L. Camargo de Abreu	21	1974	Brazil
Sp02	*Ceratopteris* sp.	Takeuchi	5669	1990	New Guinea
Th01	*Ceratopteris thalictroides*	Aleta Jo Petrik-Ott	1002	1975	USA
Th02	*Ceratopteris thalictroides*	Aleta Jo Petrik-Ott	1003	1975	USA
Th03	*Ceratopteris thalictroides*	D. S. Correll	35295	1967	USA
Th05	*Ceratopteris thalictroides*	Maria F. Gonzales	603	2011	Colombia
Th07	*Ceratopteris thalictroides*	D. S. Correll	38315		USA
Th08	*Ceratopteris thalictroides*	Gerardo Herrera	1797	1988	Costa Rica
Th09	*Ceratopteris thalictroides*	Helen Kennedy	2377	1973	Panama
Th10	*Ceratopteris thalictroides*	Pete Wallace	s/n	1987	USA
Th11	*Ceratopteris thalictroides*	J. J. de Granville	15555	1987	Guyana
Th12	*Ceratopteris thalictroides*	K. Lynch	s/n	2004	Hawaii
Th13	*Ceratopteris thalictroides*	L M. Crago	2005-058	2005	Hawaii
Th14	*Ceratopteris thalictroides*	W. L. Wagner	1990	1990	Hawaii
Th15	*Ceratopteris thalictroides*	J. F. Barcelona	2505	2004	Philippines
Th16	*Ceratopteris thalictroides*	C. R. Fraser-Jenkins	1564	1997	Nepal
Th17	*Ceratopteris thalictroides*	D. S. Correll	38316	1970	USA
Th18	*Ceratopteris thalictroides*	D. S. Correll	35295	1967	USA
Th19	*Ceratopteris thalictroides*	L. H. Vogel	s/n	1970	USA
Th20	*Ceratopteris thalictroides*	J. J. de Granville	15555	2002	Guyana
Th21	*Ceratopteris thalictroides*	W. L. Testo	529	2014	Costa Rica
Th22	*Ceratopteris thalictroides*	Wen-Yu Wang	380	2009	Taiwan
Th23	*Ceratopteris thalictroides*	Pi-Fong Lu	9278	2005	Taiwan
Th24	*Ceratopteris thalictroides*	S. R. Hill	17743	1987	Costa Rica
Th25	*Ceratopteris thalictroides*	S. R. Hill	13054	1983	Brazil

Sample name	Species	Collector	Number	Year	Location
Th26	*Ceratopteris thalictroides*	S. P. Kinosian	3	2017	Costa Rica
Th27	*Ceratopteris thalictroides*	S. P. Kinosian	4	2017	Costa Rica
Th28	*Ceratopteris thalictroides*	Takahashi and Sawada	28366	1999	Japan
Th29	*Ceratopteris thalictroides*	Takahashi and Miyaha	26962	1998	Japan
Th30	*Ceratopteris thalictroides*	Hoover	5444	1990	Thailand
Th31	*Ceratopteris thalictroides*	Takeuchi and Towati	15253	2001	New Guinea
Th32	*Ceratopteris thalictroides*	Gui-Sheng Li	s/n	2018	China
Th33	*Ceratopteris thalictroides*	T. E. Almeida	4473	2016	Australia
Th34	*Ceratopteris thalictroides*	S. P. Kinosian	60a	2018	Australia
Th35	*Ceratopteris thalictroides*	S. P. Kinosian	37a	2018	Taiwan
Th36	*Ceratopteris thalictroides*	S. P. Kinosian	71a	2018	Australia
Th37	*Ceratopteris thalictroides*	C. R. Fraser-Jenkins	FN136	2007	India
Th38	*Ceratopteris thalictroides*	Taku Miyazaki	611074		Japan
Th39	*Ceratopteris thalictroides*	S. P. Kinosian	67	2018	Australia
Th40	*Ceratopteris thalictroides*	C. R. Fraser-Jenkins	30782	2004	Nepal
Th41	*Ceratopteris thalictroides*	S. P. Kinosian	63	2018	Australia
Th42	*Ceratopteris thalictroides*	S. P. Kinosian	64	2018	Australia
Th43	*Ceratopteris thalictroides*	S. P. Kinosian	65	2018	Australia
Th44	*Ceratopteris thalictroides*	S. P. Kinosian	66	2018	Australia
Th45	*Ceratopteris thalictroides*	S. P. Kinosian	68	2018	Australia
Th46	*Ceratopteris thalictroides*	S. P. Kinosian	69	2018	Australia
Th47	*Ceratopteris thalictroides*	S. P. Kinosian	72	2018	Australia
Th48	*Ceratopteris thalictroides*	S. P. Kinosian	73	2018	Australia
Th49	*Ceratopteris thalictroides*	S. P. Kinosian	74	2018	Australia
Th50	*Ceratopteris thalictroides*	S. P. Kinosian	75	2018	Australia
Th51	*Ceratopteris thalictroides*	S. P. Kinosian	76	2018	Australia

Sample name	Species	Collector	Number	Year	Location
Th52	*Ceratopteris thalictroides*	S. P. Kinosian	60b	2018	Australia
Th53	*Ceratopteris thalictroides*	S. P. Kinosian	60c	2018	Australia
Th54	*Ceratopteris thalictroides*	S. P. Kinosian	60d	2018	Australia
Th55	*Ceratopteris thalictroides*	S. P. Kinosian	60e	2018	Australia
Th56	*Ceratopteris thalictroides*	S. P. Kinosian	73b	2018	Australia
Th57	*Ceratopteris thalictroides*	S. P. Kinosian	73c	2018	Australia
Th58	*Ceratopteris thalictroides*	S. P. Kinosian	73d	2018	Australia
Th59	*Ceratopteris thalictroides*	S. P. Kinosian	73e	2018	Australia
Th60	*Ceratopteris thalictroides*	S. P. Kinosian	73f	2018	Australia
Th61	*Ceratopteris thalictroides*	S. P. Kinosian	71b	2018	Australia
Th62	*Ceratopteris thalictroides*	S. P. Kinosian	76b	2018	Australia
Th63	*Ceratopteris thalictroides*	S. P. Kinosian	76c	2018	Australia
Th64	*Ceratopteris thalictroides*	S. P. Kinosian	76d	2018	Australia

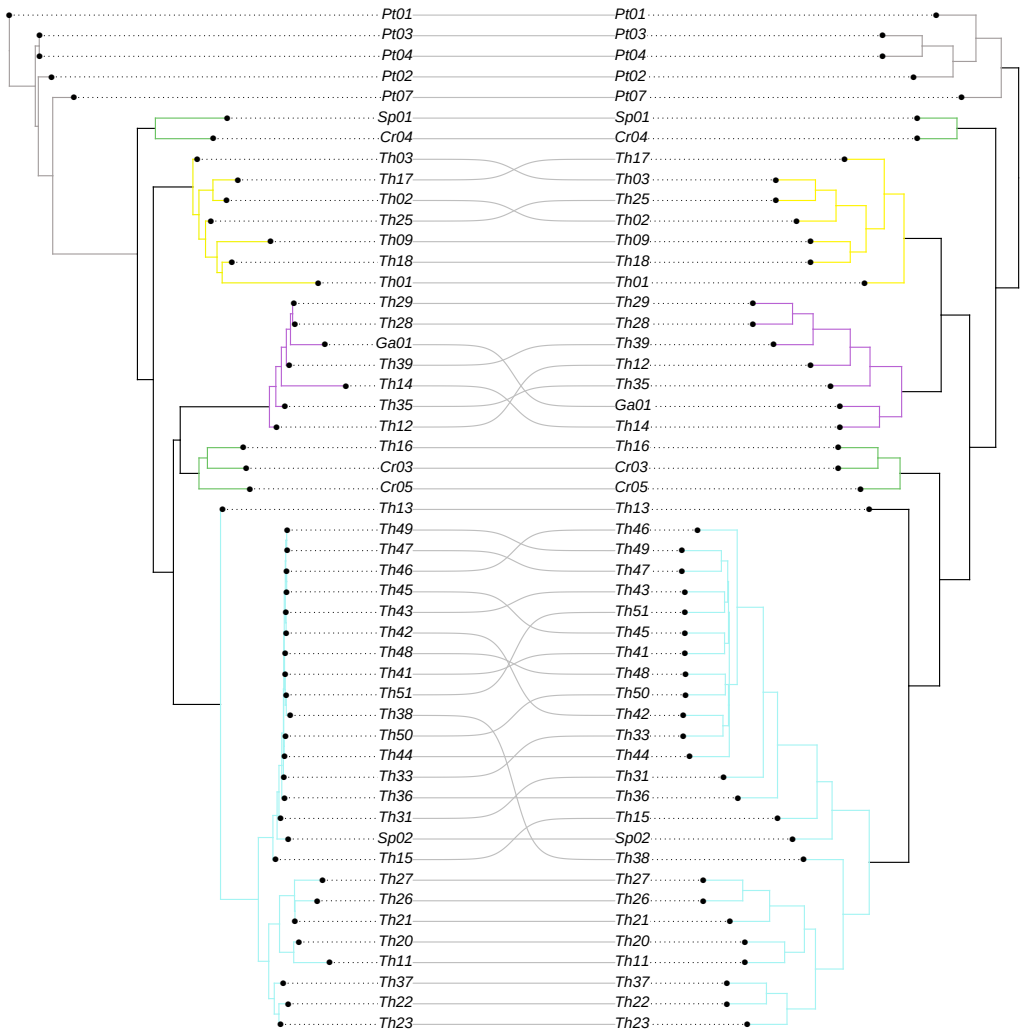


Fig. A.1: Comparison of phylogenetic trees inferred using RAxML (right) and TETRAD (left). Both trees are rooted with *Acrostichum aureum*; this tip was removed to conserve space. The two trees have very similar topologies: all species except *C. pteridoides* and *C. cornuta* are monophyletic in both trees. In the RAxML tree, *C. pteridoides* is a grade, whereas in the TETRAD tree it forms a clade. *Ceratopteris cornuta* is paraphyletic in both trees, but the same individuals forming each respective group. Another difference is that in the TETRAD phylogeny, New World *C. thalictroides* and *C. gaudichaudii* forming a monophyletic group; the former comes out as sister to all other species except *C. pteridoides* in the RAxML tree. There is certainly reticulate evolution and hybridization occurring in *Ceratopteris*, so for one allotetraploid clade to be in two different places is not unexpected.

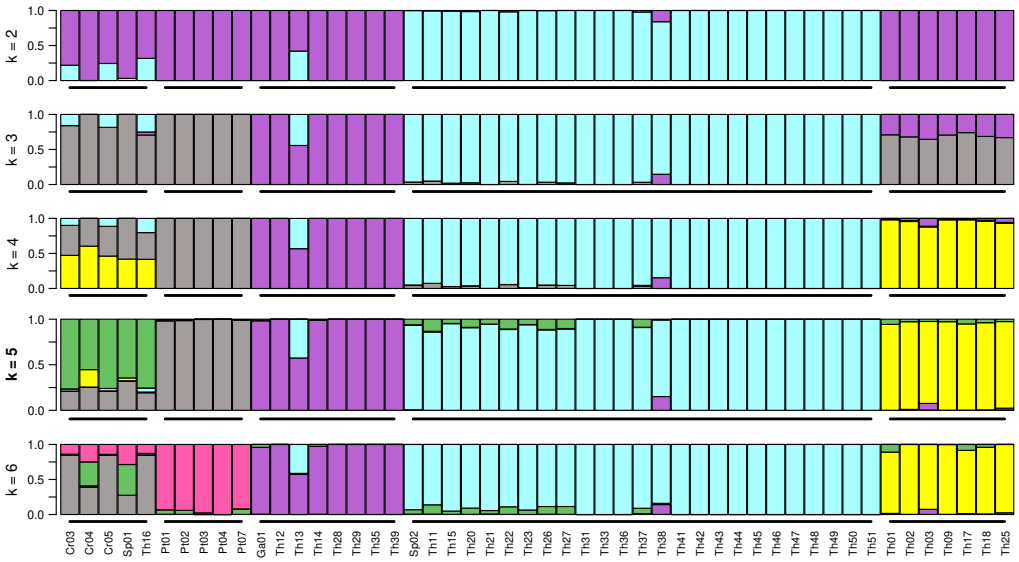


Fig. A.2: STRUCTURE plots for K = 2 - 6. Each bar represents an individual, and each color is representative of a single genomic population. A single bar with multiple colored segments indicates ancestral admixture from K source populations. K = 5 was determined to be the most biologically informative. At K = 5, all named species of *Ceratopteris* fall into distinct populations. *Ceratopteris thalictroides* is split into three groups: Old World (blue), New World (yellow), and individuals grouping with *C. gaudichaudii* (purple). *Ceratopteris cornuta* is the only species with evidence of ongoing or recent admixture, most evident with *C. pteridoides*. At K = 6, the yellow bars may be indicative of *C. richardii*, but do not have a sample here to compare it to.

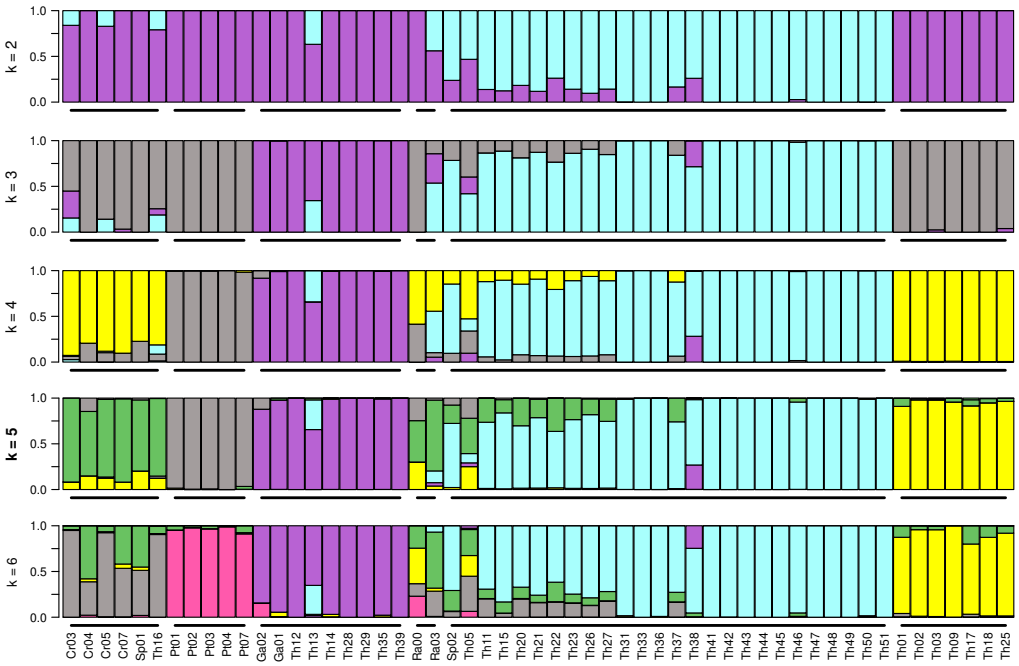


Fig. A.3: STRUCTURE plots for K = 2 - 6, including *C. richardii*. Individuals of *Ceratopteris richardii* are labeled as: Ra00 (simulated individual from the genome assembly), and Ra03 (herbarium specimen Trinidad; J. T. Mickel 9509). Note that both individuals have variable population assignment across all values of K; it is difficult to understand whether this is due to hybrid origin or poor data quality without a high-quality specimen of the same species for reference. There is more variability in population assignment across species; this is perhaps due to the fact that this dataset contained 80% missing data, compared to 26% missing data in the main dataset.

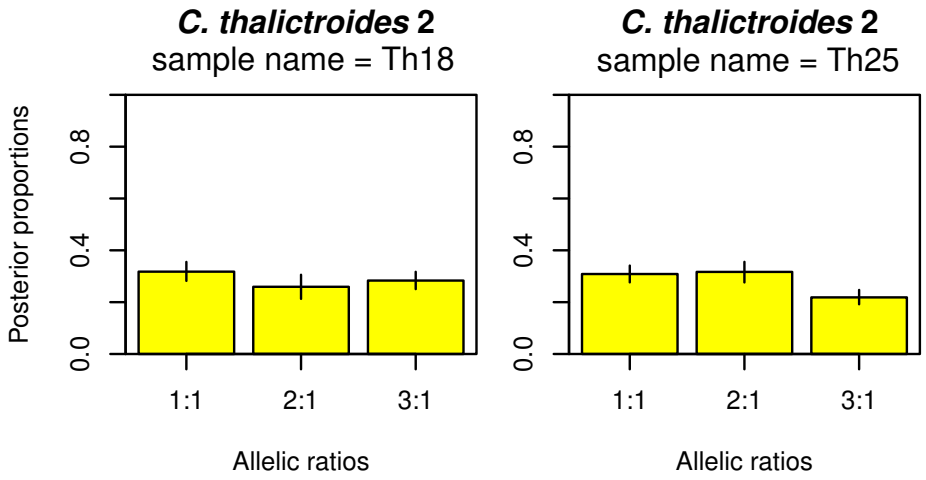
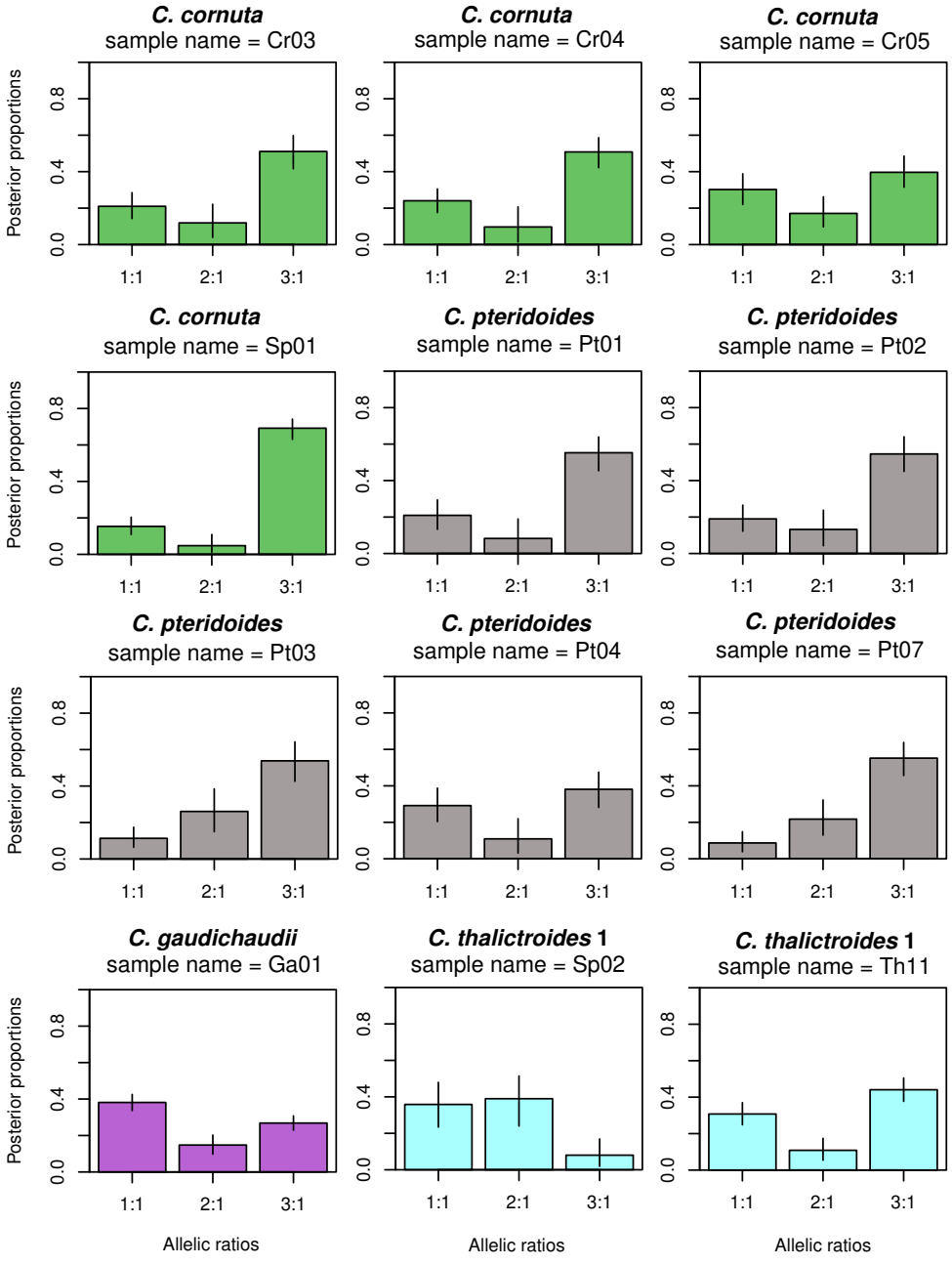
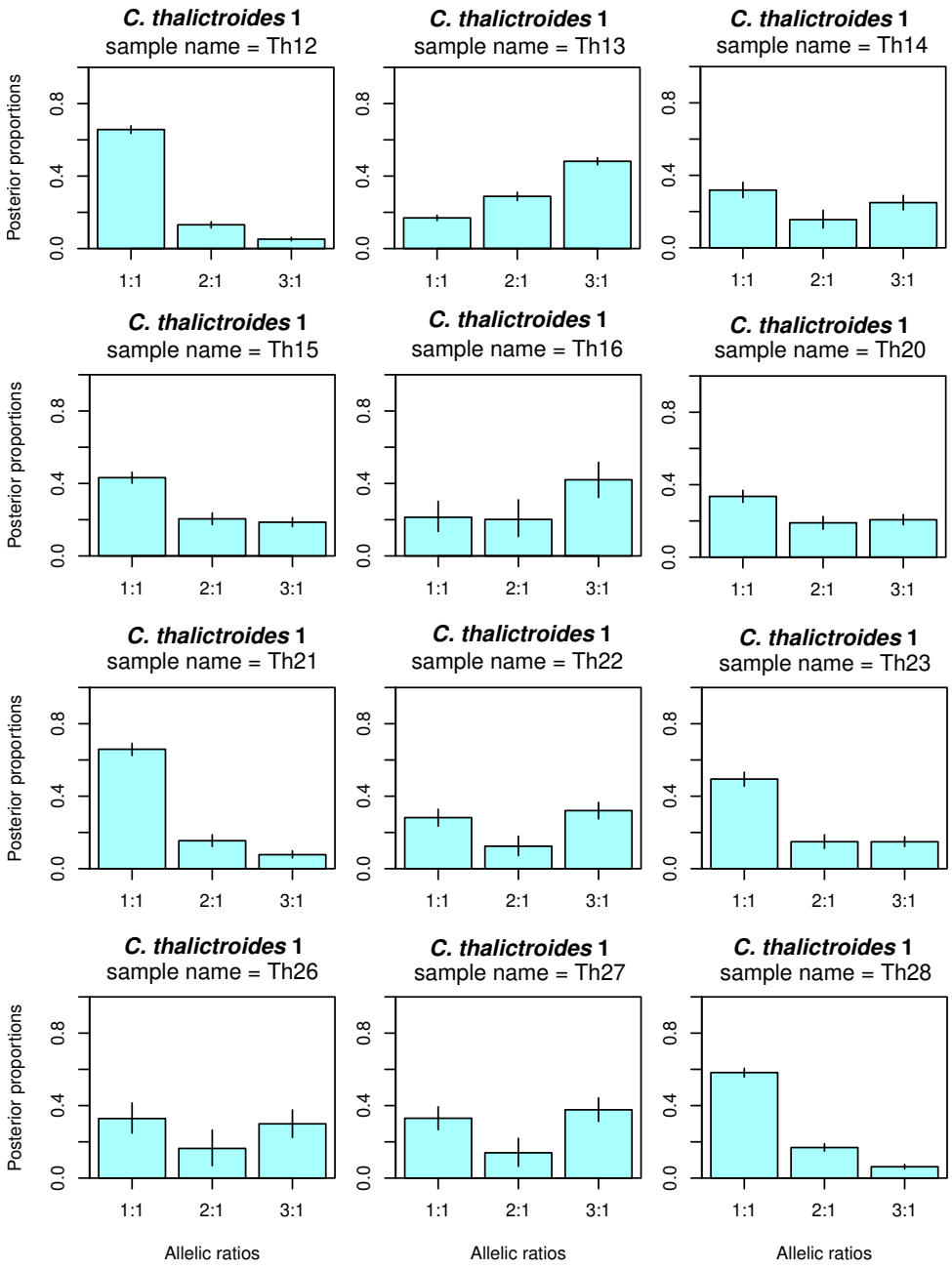
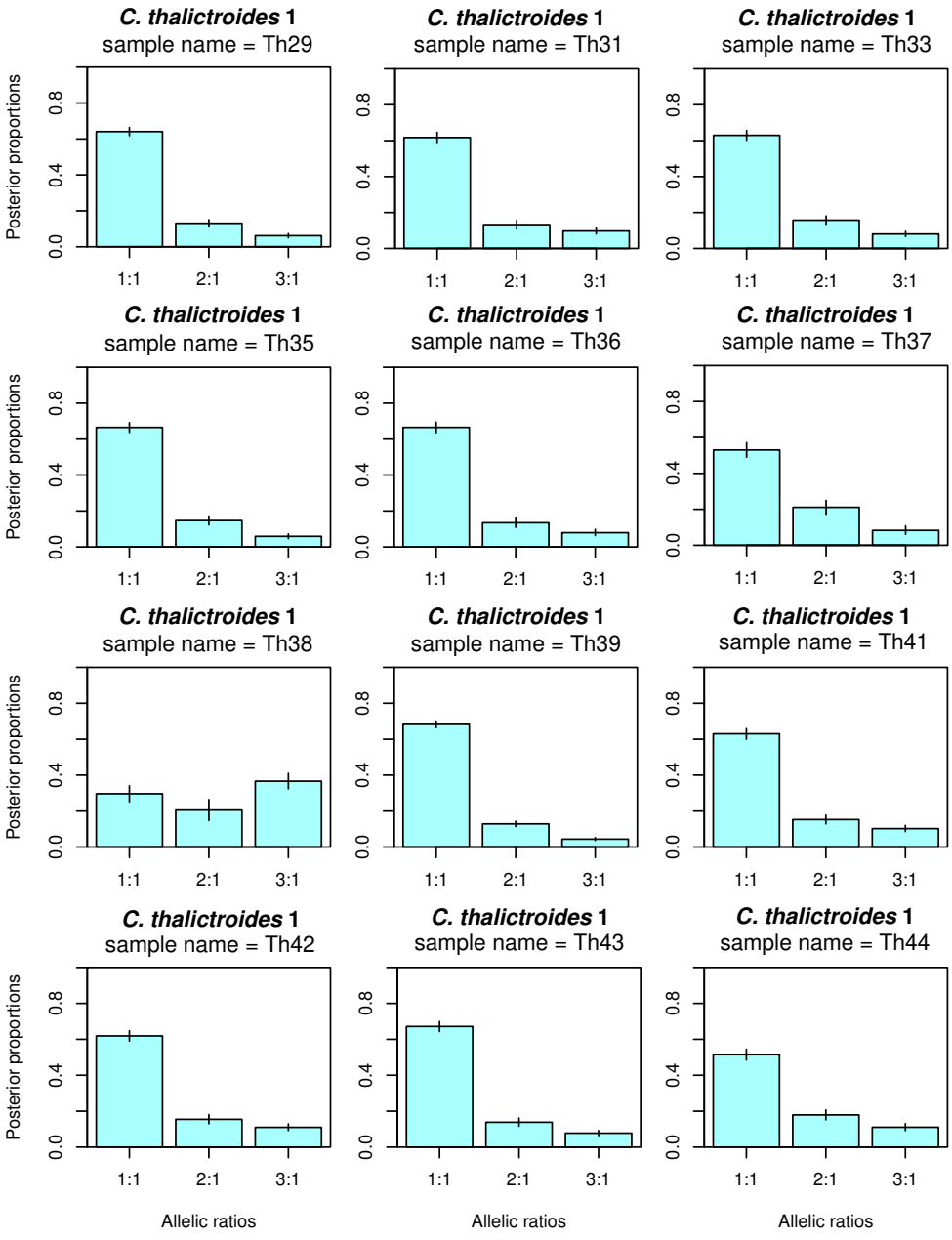
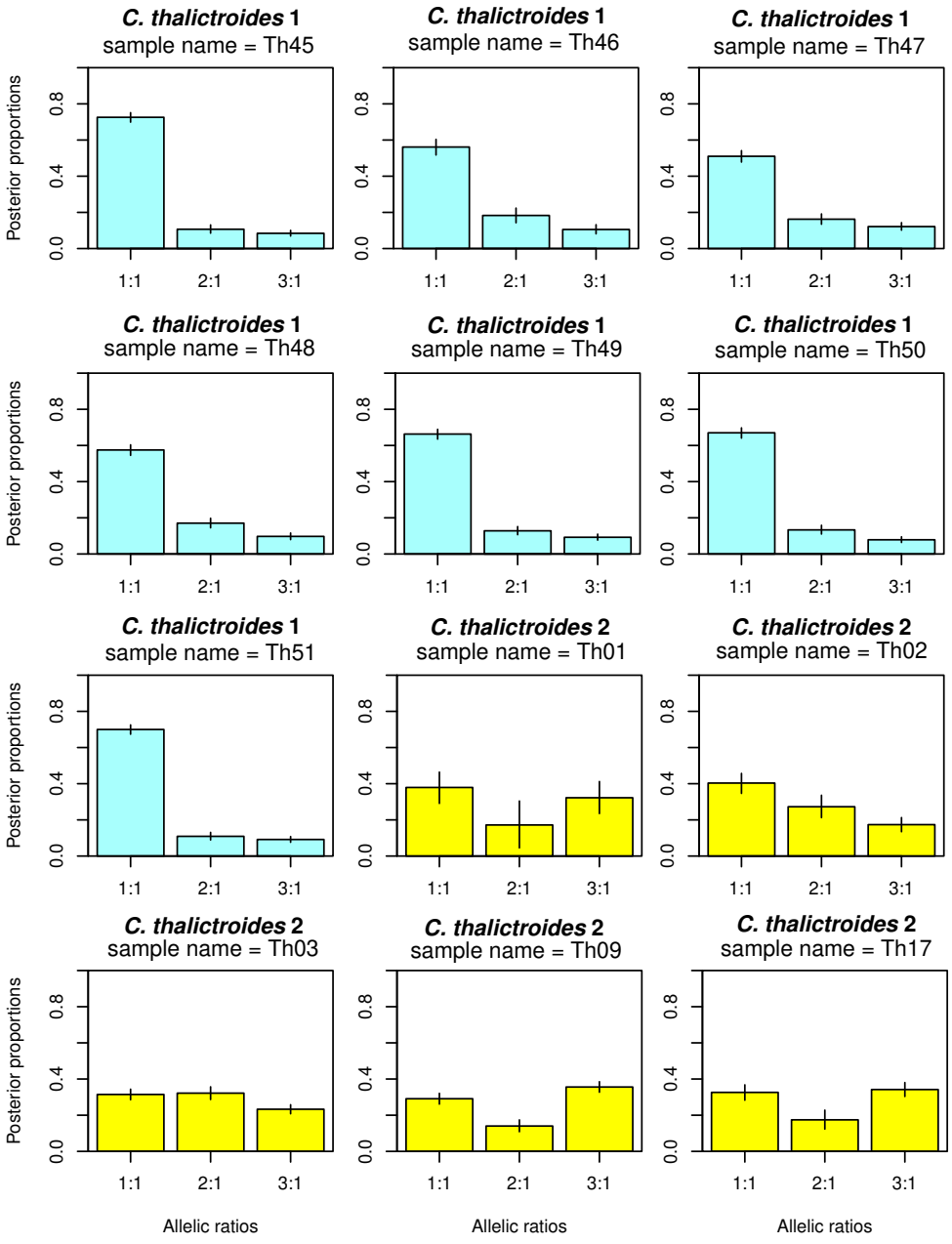


Fig. A.4: Ploidy inference for all individuals included in the final dataset (page 5/5), estimated via allelic ratios. Diploids would have a 1:1 allele ratio in accordance with the two parental genomes present. Triploid hybrids would have a 2:1 allele ratio, with two sets of chromosomes from a unreduced spore, and another from a haploid spore. Tetraploids could either have a 1:1 or 3:1 allele ratio. A 1:1 allelic ratio could indicate an autopolyploid derived from a genome double event, or an allopolyploid with two sets of chromosomes (one from each parent). A 3:1 allelic ratio would be representation of an allopolyploid that has undergone genomic restructuring. All individuals removed from the initial sampling had ambiguous ploidy assignments. Samples names correspond to rows in Table A.1


C. cornuta
sample name = Cr03
C. cornuta
sample name = Cr04
C. cornuta
sample name = Cr05
C. cornuta
sample name = Sp01
C. pteridoides
sample name = Pt01
C. pteridoides
sample name = Pt02
C. pteridoides
sample name = Pt03
C. pteridoides
sample name = Pt04
C. pteridoides
sample name = Pt07
C. gaudichaudii
sample name = Ga01
C. thalictroides 1
sample name = Sp02
C. thalictroides 1
sample name = Th11
Posterior proportions
0.0
0.4
0.8
1:1
2:1
3:1
Allelic ratios


C. thalictroides 1
sample name = Th12
C. thalictroides 1
sample name = Th13
C. thalictroides 1
sample name = Th14
C. thalictroides 1
sample name = Th15
C. thalictroides 1
sample name = Th16
C. thalictroides 1
sample name = Th20
C. thalictroides 1
sample name = Th21
C. thalictroides 1
sample name = Th22
C. thalictroides 1
sample name = Th23
C. thalictroides 1
sample name = Th26
C. thalictroides 1
sample name = Th27
C. thalictroides 1
sample name = Th28
Posterior proportions
0.0
0.4
0.8
1:1
2:1
3:1
Allelic ratios


C. thalictroides 1
sample name = Th29
C. thalictroides 1
sample name = Th31
C. thalictroides 1
sample name = Th33
C. thalictroides 1
sample name = Th35
C. thalictroides 1
sample name = Th36
C. thalictroides 1
sample name = Th37
C. thalictroides 1
sample name = Th38
C. thalictroides 1
sample name = Th39
C. thalictroides 1
sample name = Th41
C. thalictroides 1
sample name = Th42
C. thalictroides 1
sample name = Th43
C. thalictroides 1
sample name = Th44
Posterior proportions
0.0
0.4
0.8
1:1
2:1
3:1
Allelic ratios


C. thalictroides 1
sample name = Th45
C. thalictroides 1
sample name = Th46
C. thalictroides 1
sample name = Th47
C. thalictroides 1
sample name = Th48
C. thalictroides 1
sample name = Th49
C. thalictroides 1
sample name = Th50
C. thalictroides 1
sample name = Th51
C. thalictroides 2
sample name = Th01
C. thalictroides 2
sample name = Th02
C. thalictroides 2
sample name = Th03
C. thalictroides 2
sample name = Th09
C. thalictroides 2
sample name = Th17
Posterior proportions
0.0
0.4
0.8
1:1
2:1
3:1
Allelic ratios

APPENDIX B

Supplementary figures for chapter 5

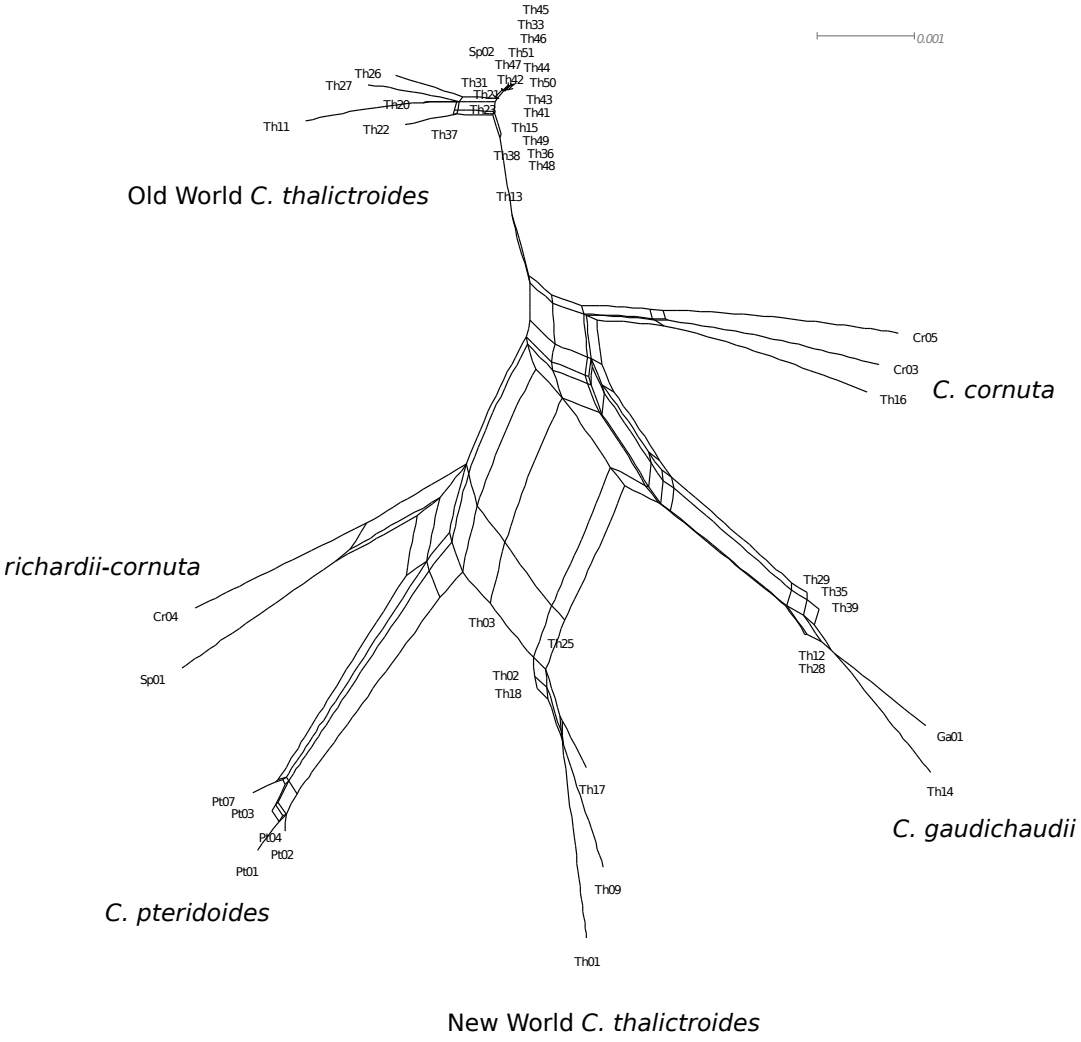


Fig. B.1: Network analysis of *Ceratopteris* using SplitsTree4. Branch tips represent individuals, and each major group is labeled by species. Each tip is labeled by the associated specimen; see Kinosian, Pearse, and Wolf (2020) for collection information. Groups that share more links/sections of the network are hypothesized to have reticulate gene flow.

APPENDIX C

Supplementary tables for chapter 6

Table C.1: Collection information for samples used in chapter 5. Sample ID (used in figures) species name, and collection information included. All collections by S. P. Kinosian are from silica-dried leaf tissue, the remaining samples are from herbarium specimens. *Continued on following pages.*

Sample name	Species	Collector	Number	Year	Location
Cr22	*Ceratopteris cornuta*	Carvalho	4396	1990	Equatorial Guinea
Cr24	*Ceratopteris cornuta*	Bigger, M.	2520	1971	Ghana
Cr25	*Ceratopteris cornuta*	C. C. H. Jongkind	2189	1971	Ghana
Cr26	*Ceratopteris cornuta*	A. Rasolohery	319	2001	Madagascar
Cr28	*Ceratopteris cornuta*	Maconochie, J.R.	3173	1982	Oman
Cr29	*Ceratopteris cornuta*	Miller, A.G.	7044	1985	Oman
Cr30	*Ceratopteris cornuta*	Brass, L.J.	22042	1953	Papua New Guinea
Cr31	*Ceratopteris cornuta*	C. Vanden Berghen	6318	1984	Senegal
Cr32	*Ceratopteris cornuta*	L. Fay	1109	1985	Sierra Leon
Cr33	*Ceratopteris cornuta*	Miller, A.G.	M.8236A	1989	Yemen
Cr34	*Ceratopteris cornuta*	Kilian, N.	NK6312	1999	Yemen
Cr36	*Ceratopteris cornuta*	J. R. I. Wood	2717	1979	Yemen
Pt20	*Ceratopteris pteridoides*	T. M. Pedersen	14800	1986	Argentina
Pt23	*Ceratopteris pteridoides*	R. Callejas	4996	1987	Colombia
Pt28	*Ceratopteris pteridoides*	D. A. White	543	1987	Louisiana
Pt29	*Ceratopteris pteridoides*	N. M. Gilmore	2675	1986	Louisiana
Pt30	*Ceratopteris pteridoides*	F. M. Givens	1740	1980	Louisiana
Pt31	*Ceratopteris pteridoides*	F. M. Givens	1740A	1982	Louisiana
Pt32	*Ceratopteris pteridoides*	R. Novelo	3430	1998	Mexico
Pt34	*Ceratopteris pteridoides*	H. Kennedy	3171	1973	Panama
Pt36	*Ceratopteris pteridoides*	M. Van Etten	953	1984	Trinidad
Pt37	*Ceratopteris pteridoides*	R. Abbott	25282	2008	Florida
Pt39	*Ceratopteris pteridoides*	R. Abbott	23948	2007	Florida
Pt40	*Ceratopteris pteridoides*	G. E. Crow	7436	1989	Costa Rica
Ra10	*Ceratopteris richardii*	E. P. Killip	45469	1956	Cuba

Sample name	Species	Collector	Number	Year	Location
Ra12	*Ceratopteris richardii*	G. R. Proctor	32940	1972	Jamaca
Ra13	*Ceratopteris richardii*	G. R. Proctor	38368	1979	Jamaca
Ra14	*Ceratopteris richardii*	Glen N. Montz	8339	1996	Louisiana
Ra15	*Ceratopteris richardii*	G. R. Proctor	33450	1973	Florida
Ra16	*Ceratopteris richardii*	S. P. Kinosian	Hnn, fresh collection	2019	Cuba
Sp10	*Ceratopteris* sp.	J. Saunders	507	1980	Honduras
Sp11	*Ceratopteris* sp.	F. Mereles	6837	1997	Paraguay
Sp13	*Ceratopteris* sp.	C. Congdon	245	1979	Thailand
Sp16	*Ceratopteris* sp.	F. Freidmann	4818	1983	Seychelles
Th103	*Ceratopteris thalictroides*	A. Bodenghien	3141	1987	Dem. Repub. Congo
Th104	*Ceratopteris thalictroides*	J. Monterrosa	1428	2007	El Salvador
Th105	*Ceratopteris thalictroides*	R. A. Carballo	899	2003	El Salvador
Th107	*Ceratopteris thalictroides*	E. A. Christenson	1869	1994	Guyana
Th109	*Ceratopteris thalictroides*	Wu Su-Gong	WS-2382	2011	Laos
Th110	*Ceratopteris thalictroides*	G. Thebaud	138	2003	Laos
Th111	*Ceratopteris thalictroides*	C. R. Birkinshaw	155	1995	Madagascar
Th116	*Ceratopteris thalictroides*	Frodin, D.G.	7007	1982	Papua New Guinea
Th120	*Ceratopteris thalictroides*	R. Knapp	2850	2012	Taiwan
Th80	*Ceratopteris thalictroides*	C. B. Hellquist	15100	1981	Australia
Th81	*Ceratopteris thalictroides*	K.C. Cheng	CL820	2007	Cambodia
Th83	*Ceratopteris thalictroides*	T. Stevart	4511	2012	Gabon
Th84	*Ceratopteris thalictroides*	Murata, G.	s.n.	1963	Japan
Th87	*Ceratopteris thalictroides*	Flora of East Nepal	9760572	1997	Nepal

Sample name	Species	Collector	Number	Year	Location
Th88	*Ceratopteris thalictroides*	C. E. Nichols	1719	1969	Nicaragua
Th89	*Ceratopteris thalictroides*	Wiakabu, J.	70372	1977	Papua New Guinea
Th90	*Ceratopteris thalictroides*	Croft, J.R.	61920	1974	Papua New Guinea
Th93	*Ceratopteris thalictroides*	Niyomdham, C.	1659	1988	Thailand
Th94	*Ceratopteris thalictroides*	S. W. L. Jacobs	8287	1997	Australia
Th95	*Ceratopteris thalictroides*	T. B. Croat	52304A	1981	Australia
Th96	*Ceratopteris thalictroides*	G. Murata	261	1979	Borneo
Th98	*Ceratopteris thalictroides*	D. F. Austin	7193	1979	Brasil

APPENDIX D

Permission-to-use letters

Jacob S. Suissa
Department of Organismic & Evolutionary Biology
Harvard University

20 April 2021

Subject: Permission to use paper for dissertation chapter

To whom it may concern,

I am a co-author on the paper “Mothers of Pteridology”, published in volume 110, issue 1 of the American Fern Journal. I give permission for Sylvia Kinosian to use this paper as part of her doctoral dissertation.

Sincerely,

Jacob S. Suissa

Dr. Weston Testo
Carl Skottsbergs Gata 22 A
Department of Biology
University of Gothenburg
Gothenburg, Sweden

20 April 2021

Subject: Permission to use paper for dissertation chapter

To whom it may concern,

I am a co-author on the paper “Using RAD Data to Confirm Parentage of Polyploids in a Reticulate Complex of Ferns”, published in volume 109, issue 3 of the American Fern Journal. I give permission for Sylvia Kinosian to use this paper as part of her doctoral dissertation.

Sincerely,

Dr. Sally Chambers
Marie Selby botanical Gardens
Botany Department
1534 Mound St.
Sarasota, FL 34236

29 April 2021

Subject: Permission to use paper for dissertation chapter

To whom it may concern,

I am a co-author on the paper “Using RAD Data to Confirm Parentage of Polyploids in a Reticulate Complex of Ferns”, published in volume 109, issue 3 of the American Fern Journal. I give permission for Sylvia Kinosian to use this paper as part of her doctoral dissertation.

Sincerely,

UF UNIVERSITY of FLORIDA
College of Liberal Arts and Sciences
Department of Biology

220 Bartram Hall
PO Box 118525
Gainesville, FL 32611

April 28, 2021

Subject: Permission to use paper for dissertation chapter

To whom it may concern,

I am a co-author on the paper "Using RAD Data to Confirm Parentage of Polyploids in a Reticulate Complex of Ferns", published in volume 109, issue 3 of the American Fern Journal. I give permission for Sylvia Kinosian to use this paper as part of her doctoral dissertation.

Sincerely,

Dr. Emily B. Sessa
Associate Professor
Department of Biology
University of Florida
Phone: +1 352 213 0334

CURRICULUM VITAE

Sylvia Page Kinosian

Utah State University Biology Department & Ecology Center

5305 Old Main Hill Logan Utah 84322 USA

518.708.5827 – sylvia.kinosian@aggiemail.usu.edu

Education

Ph.D., Utah State University (USU) 2021

Co-advisors: Paul G. Wolf and William D. Pearse

Bachelor of Science, University of Vermont (UVM) 2015

Major in Forestry & minor in Plant Biology, *summa cum laude*

Honors & Awards

National Science Foundation Graduate Research Fellowship ($138,000) April 2016

Award for Excellence in Plant Biology, UVM Plant Biology Department ($250) May 2015

W. R. Adams Award for Outstanding Academic Achievement in Forestry, UVM Rubenstein School of Environment and Natural Resources (RSENR) May 2015

Holcomb Natural Resource Prize, UVM RSENR May 2015

Dale Bergdahl Scholarship Award, UVM RSENR ($1,000) May 2014

Dean's Book Award for Outstanding Juniors, UVM RSENR May 2013

Dean's List, UVM RSENR Fall 2011 - Spring 2015

Research & Travel Grants

American Society of Plant Taxonomists Travel (ASPT) Grant ($335) July 2019

ASPT Graduate Student Research Award ($1,200) April 2019

Graduate Research Award, USU Ecology Center ($3,700) March 2019

Asian Symposium of Ferns and Lycophytes

2018 Student Travel Grant ($480) October 2018

ASPT Travel Grant ($335) July 2018

Joseph E. Greaves Endowed Scholarship, USU Biology Department ($4,550) April 2017

Organization for Tropical Studies, Barbara Joe Hoshizaki
Memorial Scholarship ($500) January 2017

PUBLICATIONS

1. J. S. Suissa, **S. P. Kinosian**, P. W. Schafran, J. Bolin, W. C. Taylor, and E. A. Zimmer. Revealing the evolutionary history of a reticulate polyploid complex in the genus *Isoetes*. Molecular Phylogenetics and Evolution. *Major Revisions*. BioRxiv: 10.1101/2020.11.04.363374

2. T. P. Smith, A. S. Gallinat, **S. P. Kinosian**, M. Stemkovski, H. J. T. Unwin, O. J. Watson, C. Whittaker, L. Cattarino, I. Dorigatti, M. Tristem, and W. D. Pearse. 2021. Temperature and populationdensity influence SARS-CoV-2 transmission in the absence of non-pharmaceutical interventions. Proceedings of the National Academy of Sciences. 118(25):e2019284118. DOI: 10.1073/pnas.2019284118

3. **S. P. Kinosian**, W. D. Pearse, and P. G. Wolf. 2020. There and back again: Reticulate evolution in *Ceratopteris*. American Fern Journal. 110(4):193-210. *Invited for special issue on biogeography and reticulate evolution in ferns.*

4. **S. P. Kinosian**, W. D. Pearse, and P. G. Wolf. 2020. Cryptic diversity in the model fern genus *Ceratopteris* (Pteridaceae). Molecular Phylogenetics and Evolution. 152:106938. DOI: 10.1016/j.ympev.2020.106938.

5. **S. P. Kinosian** and J. S. Suissa. 2020. Mothers of pteridology. American Fern Journal. 110(1):3-19. DOI: 10.1640/0002-8444-110.1.3.

6. P. G. Wolf, C. A. Rowe, **S. P. Kinosian**, Z. Gompert, J. Der, T. McLenachen, P. Lockhart, L. Shepherd, and J. Thomson. 2019. Relationships among worldwide groups in the fern genus *Pteridium* (bracken) based on nuclear genome markers. American Journal of Botany. 106(10):1365-1376. DOI:10.1002/ajb2.1365

7. **S. P. Kinosian**, W. L. Testo, S. M. Chambers, and E. B. Sessa. 2019. Using RAD data to confirm parentage of polyploids in a reticulate complex of ferns. American Fern Journal, 109(3):267-282. DOI: 10.1640/0002-8444-109.3.267 *Invited for special issue on fern genomics.*

Additional Publications, Not Peer-Reviewed

1. W. D. Pearse, A. S. Gallinat. C. Chelak. **S. P. Kinosian**, E. Simpson, J. Bravo, A. McManis, M. Hagadorn, and S. B. Hudson. 2019. MADcomm - Make A Database of Communities. DOI: 10.5281/zenodo.3361820.

2. W. D. Pearse, M. J. Farrell, K. Hafen, M. Hagadorn, S. B. Hudson, **S. P. Kinosian**, R. McCleary, A. Rego, and K. Welgarz. 2019. MADtraits - Make A Database of Traits. DOI: 10.5281/zenodo.3361815.

3. J. S. Suissa and **S. P. Kinosian**. 2018. Botany 2018: Exploring Minnesota's Driftless Area, with drifting pteridologists. Fiddlehead Forum. 45(4):73-77.

4. M. Barkworth, P. G. Wolf, **S. P. Kinosian**, C. Dyreson, W. Pearse, B. Brandt, and N. Cobb. 2017. The Value of Agricultural Voucher Specimens. Proceedings of TDWG 1: e19932.

Teaching Experience

Writing Mentor. September - October 2020

USU NSF GRF Workshop Series.

Helped students develop and refine NSF Graduate Research Fellowship proposals.

Instructor. January - May 2020

USU Biology Department, Plant Systematics and Diversity.

Designed and taught a mixed undergraduate/graduate level course covering land plant diversity and evolution; emphasis on Utah and Great Basin plants.

Teaching Assistant. September - December 2019

USU Biology Department, Evolutionary Genetics.

Helped students design final research papers, graded assignments.

Writing Mentor. September - October 2019

USU NSF GRF Workshop Series.

Helped students develop and refine NSF Graduate Research Fellowship proposals.

Teaching Assistant. January - May 2018

USU Biology Department, Biology II Laboratory.

Implemented weekly labs and held office hours for three lab sections.

Co-Instructor/TA. August - December 2017

USU Biology Department, Plant Systematics and Diversity.

Prepared and gave course lectures, designed and taught weekly labs, wrote exams.

Professional Experience

Field Technician. May - August 2016

USU Quinney College of Natural Resources.

Assisted with field measurements for a study on aspen regeneration.

Laboratory Technician. July - December 2015

UVM Plant Biology Department.

Performed lab work for various projects as well as identification of herbarium specimens.

Intern. January - May 2014

Vermont Urban and Community Forestry. Burlington, VT.

Raised awareness of invasive insects and planned a UVM Arbor Day celebration.

Intern. June - August 2013

The Land Stewardship Program. Burlington, VT.

Conducted property surveys to inform conservation and management decisions.

Peer Tutor. Fall 2012 - Spring 2014

UVM Learning Cooperative.

Tutored students in a variety of subjects including Botany, Astronomy, and Anthropology.

PRESENTATIONS

S. P. Kinosian, W. D. Pearse, and P. G. Wolf. Cryptic diversity in the model fern genus *Ceratopteris* (Pteridaceae). Virtual presentation, ESA 2020.

J. S. Suissa, **S. P. Kinosian**, Peter W. Schafran, Jay Bolin, W. Carl Taylor, and Elizabeth A. Zimmer. Revealing the origins of a polyploid complex of *Isoetes* in northwestern North America. Virtual presentation, Botany 2020. July 31 2020.

S. P. Kinosian and P. G. Wolf. Species boundaries, biogeography, and polyploidy in genus *Ceratopteris* (Pteridaceae). Oral presentation, Botany 2019. Tucson, AZ, USA. July 30 2019.

S. P. Kinosian. Population structure analysis of the pan-tropical fern genus *Ceratopteris* (Pteridaceae). Oral presentation, Asian Symposium of Ferns and Lycophytes. Taipei, Taiwan. October 18 2018.

S. P. Kinosian, Z. Gompert, P. G. Wolf, J. Der, C. Rowe, T. McLenachen, P. Lockhart, L. Shepherd, and J. Thomson. Population admixture in the cosmopolitan fern genus *Pteridium* (Dennstaeditaceae). Oral presentation, Botany 2018. Rochester, MN, USA. July 24 2018.

S. P. Kinosian, W. D. Pearse, and M. E. Barkworth. Spindle: SPecimen INformation Data capture and Label crEation. Contributed poster, Botany 2018. Rochester, MN, USA. July 23 2018.

S. P. Kinosian, Z. Gompert, P. G. Wolf, J. Der, C. Rowe, T. McLenachen, P. Lockhart, L. Shepherd, and J. Thomson. Population admixture in the cosmopolitan fern genus *Pteridium* (Dennstaeditaceae). Contributed poster, Evolution Conference. Portland, OR, USA. June 24 2017.

S. P. Kinosian. Admixture analysis of *Pteridium* (Dennstaeditaceae) in a global context. Presentation, Tropical Ferns and Lycophytes Course. Organization for Tropical Studies. San Vito, Costa Rica. January 12 2017.

S. P. Kinosian, N. Patel, and D. S. Barrington. Insights into the Hawaiian *Polystichum* (Dryopteridaceae) species. Contributed poster, Next Generation Pteridology. Smithsonian

Institution, Washington D.C., USA June 1 2015.

S. P. Kinosian., N. Patel, and D. S. Barrington. Insights into the Hawaiian *Polystichum* (Dryopteridaceae) species. Contributed poster, University of Vermont Student Research Conference. Burlington, VT, USA. April 23 2015.

Professional and Service Affiliations

Peer review for: American Fern Journal, Molecular Phylogenetics and Evolution, Phytotaxa, and Taxon.

Graduate student representative. August - October 2019

Plant Physiologist Search Committee, USU Biology Department.

Member. American Society of Plant Taxonomists. April 2018 - present

Member. American Fern Society. January 2018 - present

President & Travel Officer. USU Cycling Team. August 2017 - 2019

Organizing Committee. USU Biology Programming Club. January 2017 - December 2019

Vice President. UVM Forestry Club. January - May 2015

Field Courses & Trainings

AIARE I. January 2018

American Institute for Avalanche Research and Education. Logan, UT.

Learned how to make informed decisions and stay safe in avalanche terrain.

Wilderness First Aid. May 2017

Desert Mountain Medicine. Logan, UT

Trained in wilderness safety and emergency medical protocols.

Tropical Ferns and Lycophytes. January 2017

Organization for Tropical Studies. Costa Rica.

Course emphasizing field identification and systematics of tropical pteridophytes.

Tropical Plant Systematics. January 2015

UVM Plant Biology Department. Costa Rica.

Explored tropical land plant diversity at four field sites in Costa Rica.

SKILLS

Languages: English (primary), Spanish (conversational)

Programming Languages: R (proficient), BASH (proficient), LaTeX(proficient), Angular 6 (HTML & Typescript) (familiar), Perl (familiar), Python (familiar), C / C++ (familiar)