Int. J. Plant Sci. 182(5):325–343. 2021.

1058-5893/2021/18205-0001$15.00 DOI: 10.1086/713442

FROM GENOMES TO POPULATIONS: A META-ANALYSIS AND REVIEW OF FERN POPULATION GENETICS

Jessie A. Pelosi[1,*] and Emily B. Sessa*

*Department of Biology, University of Florida, Gainesville, Florida 32611, USA

Editor: Jeremie Fant

Premise of research. Numbering around 10,000 species, ferns are the second-largest group of vascular plants. Relatively little work, however, has focused on the population genetic processes driving fern evolution. Early work by pteridologists suggested that ferns should be highly inbred (via gametophytic selfing) and therefore should have exceptionally low genetic diversity, with most genetic variation derived from polyploidy.

Methodology. We conducted a meta-analysis, starting with literature searches using Google Scholar and Web of Science, to identify studies on fern population genetics. From these papers, we recorded population genetic metrics including expected heterozygosity under Hardy-Weinberg equilibrium (H_e), percent of polymorphic loci (%P), the inbreeding coefficient (F), Wright's fixation index (F_{ST}), the molecular marker(s) used, and the geographic origin of the samples. We compared these metrics among ferns with various mating systems and growth habits.

Pivotal results. We compiled a data set of 156 fern taxa from 87 publications. We found that both mating system and growth habit have a significant impact on the genetic diversity (%P) and population structuring of fern populations (F, F_{ST}). We found that temperate regions are the most thoroughly sampled and that most studies have used allozymes, with a shift in marker usage in the 2000s and 2010s.

Conclusions. Contrary to early hypotheses, natural fern populations are not restricted to gametophytic selfing and instead regularly outcross. As we anticipate rapid growth in this field in the near future, we present a review of the major findings in fern population genetics to date. We also suggest possible future directions such as expanding geographic and taxonomic sampling and highlighting methods that will capitalize on the rapidly expanding genomic resources for ferns.

Keywords: ferns, genome, population, population genetics, population genomics, pteridophyte.

Online enhancements: appendixes.

Introduction

Ferns are an ancient lineage of plants that likely originated in the Devonian or earlier (430 to 354 mya; Testo and Sundue 2016; Schneider et al. 2004, respectively). Today, they are the second-largest group of vascular plants, comprising around 10,000 species (PPG I 2016). Ferns play important roles in tropical and temperate ecosystems, where they can act as keystone species (George and Bazzaz 1999*a*, 1999*b*; Coomes et al. 2005; Gaxiola et al. 2008; Walker et al. 2010), alter ecosystem nutrient cycling (Russell et al. 1998; Vitousek et al. 2009), and control successional processes (Walker 1994; Walker et al. 2010). Human affairs are also heavily impacted by these plants; invasive ferns like *Lygodium* spp. (Pemberton and Ferriter 1998) and *Salvinia molesta* (Koutika and Rainey 2015) are economically and ecologically harmful in the southeastern United States, where they have disrupted natural ecosystem functioning and cost governments millions of dollars (e.g., Koop 2009). Other ferns, such as *Azolla*, have been extensively incorporated into agricultural practices (Lumpkin and Plucknett 1980; Pabby et al. 2004). Furthermore, the phylogenetic position of ferns as the sister lineage to the wildly diverse and speciose seed plants affords ferns a unique and crucial comparative role in understanding the evolution of land plants.

Despite their significance, ferns have historically received less scientific attention compared with other plant lineages (e.g., Sessa et al. 2014). Most of the work on this group has focused on the systematic description of extant diversity and understanding fern reproduction and ecology, while large-scale genetic studies (particularly at the population level), which are critical for understanding microevolutionary and diversification processes, have been limited in number. The early cytological work of pteridologists revealed that ferns tend to have high chromosome numbers (e.g., Manton 1950) and, subsequently, large genome sizes (Clark et al. 2016), which likely discouraged future studies in fern genetics because of the inherent difficulty of working with

[1] Author for correspondence; email: jessiepelosi@ufl.edu.

Manuscript received August 2020; revised manuscript received December 2020; electronically published April 21, 2021.

excessively large and complicated genomes. Although angiosperm and other seed plant genomes have been available for two decades (Arabidopsis Genome Initiative 2000), only recently, with the sequencing of several species (Li et al. 2018; Marchant et al. 2019), have genomic resources become available for ferns.

The major population genetic processes (i.e., genetic drift, nonrandom mating, gene flow, mutation, and—particularly relevant in light of climate change—adaptation and selection) guide the evolutionary trajectories of populations in different directions and with varying magnitudes (table 1). Many life-history traits of ferns, such as their unique life cycles with independent, often ephemeral gametophytes and relatively long-lived sporophytes, make it difficult to study one or more of these processes, such as mutation and selection, in these organisms. However, other traits, such as their ability to produce massive quantities of easily wind-dispersed spores (Wolf et al. 2001), make ferns ideal candidates for studying other processes, such as gene flow.

Much of the early work on fern population genetics focused on determining the prevalence of different mating systems across fern lineages, which is linked to their life cycle. All land plants undergo an alternation of generations, transitioning between diploid sporophyte and haploid gametophyte life stages. Through meiosis, the diploid sporophyte produces haploid spores, which undergo mitosis to become haploid gametophytes that in turn generate egg and sperm cells, also by mitosis. The union of egg and sperm at fertilization forms the diploid zygote, which will develop into a new sporophyte. In bryophytes (mosses, liverworts, and hornworts), the gametophyte is the dominant life stage, with a nutritionally and ecologically dependent sporophyte. The reverse is true in seed plants (gymnosperms and angiosperms), which have highly reduced gametophytes composed of only a few cells and dominant sporophytes. Unlike these other lineages, lycophytes and ferns have nutritionally independent sporophyte and gametophyte life phases (Haufler et al. 2016). Additionally, while seed plants are heterosporous, with meiosis producing two different spore types, the majority of ferns (99%; PPG I 2016) are homosporous, producing a single type of spore and therefore potentially bisexual gametophytes. For these homosporous plants, three mating systems are possible (reviewed by Haufler 2002; Ranker and Geiger 2008; Haufler et al. 2016). The three mating systems in ferns are, as follows:

1. Gametophytic selfing (intragametophytic selfing sensu Klekowski 1970, 1972): a single bisexual gametophyte produces both egg and sperm, which combine to form a completely homozygous sporophyte. Among the vascular plant lineages, this system is possible only in homosporous ferns and lycophytes.
2. Sporophytic selfing (intergametophytic selfing): the union of egg and sperm from two separate gametophytes derived from a single sporophyte (analogous to selfing in seed plants).
3. Sporophytic outcrossing (intergametophytic outcrossing): the egg is fertilized by sperm produced by a gametophyte derived from a separate sporophyte (analogous to outcrossing in seed plants).

The ability of ferns to reproduce along this continuum creates a hierarchy of possible genetic variation in daughter sporophytes. Continued gametophytic selfing would lead to the formation of populations of completely homozygous individuals that harbor low levels of genetic diversity and in which mutation is the only source of new genetic variation (Klekowski 1972). These populations would be highly structured and strongly differentiated from each other. In contrast, individuals derived from sporophytic selfing develop from sperm and eggs that, although they are from the same sporophyte parent, are produced in different meiotic events and are therefore not entirely homozygous, although variation would be relatively low compared with products of sporophytic outcrossing (Klekowski 1972). Populations that undergo only or primarily outcrossing are likely to have greater genetic diversity and heterozygosity than those using higher proportions of the other mating systems.

Klekowski and colleagues postulated that gametophytic selfing was the predominant mating system in ferns and would be followed by intense selective pressure (Klekowski and Baker 1966). They hypothesized that ferns could adapt to buffer against this pressure by retaining complete chromosome sets following whole-genome duplication events. These genomes would act as an additional storage mechanism for genetic variation (i.e., via fixed heterozygosity). This hypothesis simultaneously explained the high chromosome numbers and apparent prevalence of polyploidy across ferns (Klekowski and Baker 1966; Klekowski 1972). If ferns were truly predominantly selfing, the majority of fern sporophyte populations should be genetically depauperate unless they retained chromosomes following allopolyploidy

Table 1

Effects of Various Microevolutionary Forces in Isolation and When Interacting with Gene Flow

Microevolutionary force	Effect in isolation	Interactive effect with gene flow
Gene flow	Homogenizes populations, decreases variation between populations	NA
Genetic drift	Stochastic changes in allele frequencies, increase in among-population differentiation	Acts to prevent population divergence; cohesive
Mutation	Ultimate source of genetic variation, increases genetic diversity	If the rate of gene flow is greater than the rate of mutation, gene flow becomes the primary source of genetic variation
Negative/purifying selection	Reduces the frequency of deleterious alleles in a population	Prevents local adaptation by introducing possibly deleterious alleles
Positive selection	Increases the frequency of beneficial alleles in a population	Spreads beneficial alleles among populations
Nonrandom mating	Produces highly structured populations	Prevents inbreeding, particularly decreases in fitness; decreases population structuring

Note. Gene flow can act to enhance or counteract the influence of other forces. The table is modified in part from Ellstrand (2014). NA = not applicable.

events as a method to store genetic variation. Low (<10%; observed by Hickok [1978]) or absent recombination between chromosomes from different progenitors (homoeologous chromosomes) could be used to maintain heterozygosity in homosporous pteridophytes even through instances of extreme inbreeding. However, although Sessa et al. (2016) found that up to 70% of fern species studied (145 of 207, including both diploids and polyploids) are capable of gametophytic selfing in the laboratory, the limited data from ferns in the field suggest that most natural populations are predominantly outcrossing (reviewed in Haufler et al. 2016; Sessa et al. 2016). This leaves the pteridologist with a conundrum: Are ferns primarily inbreeding or outcrossing in natural conditions, and what are the consequences of inbreeding?

Investigating fern population genetics has been and will continue to be critical for understanding the evolutionary mechanisms driving the population dynamics of these plants. As genomic tools for ferns become more accessible, it is inevitable that authors will revisit classic questions about fern population structure and dynamics and will also address new questions that were previously unanswerable. The time, therefore, seems ripe to review the major existing findings on fern population genetics. Here we report on a data set assembled from previous publications on fern population biology, which we use to review and discuss a broad set of topics in the field of fern population biology. We explore the population genetic consequences of mating systems and life-history traits and survey the molecular approaches used in existing studies, as well as the geographic and taxonomic biases that exist in this body of work. Although significant progress has been made in fern population genetics since the 1980s, much has still been left unexplored. To facilitate the further advancement of this field, we make several suggestions regarding (1) understudied systems, both geographically and taxonomically, and (2) updated molecular approaches. In the age of widely available and increasingly inexpensive high-throughput sequencing technologies, we underscore the necessity for pteridologists to move studies of ferns from population genetics to population genomics.

Material and Methods

To review our current understanding of the population biology and microevolutionary forces of ferns, we began by revisiting Ranker and Geiger (2008), the most recent review of fern population genetics, and tracked publications that either were cited in this book chapter or cited papers listed in this chapter. To capture additional studies that did not cite papers from or were not cited in Ranker and Geiger (2008), we searched for "fern population genetics" and "fern gene flow" using Google Scholar and Web of Science. We did not include publications that focused on interspecific gene flow, as these studies are above the population level. Although some studies focused primarily on systematic and taxonomic revisions, most (e.g., Haufler 1985) included information about variation below the species level and were included in the data set we assembled.

From each publication, we recorded as much of the following data as were available: the taxon or taxa investigated, mean expected heterozygosity under Hardy-Weinberg equilibrium (H_e), mean percent of polymorphic loci (%P), mean inbreeding coefficient (F), mean fixation index (F_{ST}), molecular marker(s) used, and geographic origin of the samples. For classification purposes, we followed PPG I (2016). Growth habit was coded as terrestrial, epiphytic, epipetric, or aquatic on the basis of the literature and images posted on http://www.inaturalist.org and http://www.gbif.org. Ploidal level was determined by comparing chromosome counts for each taxon with the base chromosome number of the genus to determine whether each species was diploid, tetraploid, and so on (apps. 1, 2, both available online). We did not include ploidy for taxa that had multiple possible cytotypes, unless the specific cytotype was given in the publication from which the other data came. We also recorded the mating system for each taxon if this information was included either in the publication itself or in the references listed in table S1 of Sessa et al. (2016).

Depending on the data available in each publication, we also made one or more of the following calculations as necessary to compare equivalent metrics. Because multiple methods for indirectly measuring gene flow exist and the values generated by these different methods are not necessarily directly comparable, instead of recording these values from publications, we instead used reported F_{ST} values to calculate a consistent metric of gene flow as Nm, using a modification of Wright's (1951) equation as reported by Meirmans et al. (2018):

$$Nm = \frac{1}{2kF_{ST}} - \frac{1}{2k},$$

where Nm is the number of migrants (a proxy for gene flow), k is the ploidal level, and mutation is negligible. We used only pairwise measurements of F_{ST} since other metrics (such as G_{ST}, a measure commonly used for loci with more than two alleles) tend to overestimate differentiation and underestimate Nm (Slatkin and Barton 1989). If a publication did not report an average value for F or F_{ST} but provided individual population-level values instead, we averaged those population-level values to obtain a single mean F or F_{ST} value for that taxon and publication. If F values were not reported at all, we calculated an approximation using

$$F = 1 - \left(\frac{H_o}{H_e}\right),$$

where H_o is the observed heterozygosity and H_e is the expected heterozygosity under Hardy-Weinberg equilibrium, if these values were provided. If multiple studies were conducted on a single taxon, we calculated average values first within a publication and then across publications.

We compared H_e, %P, F, F_{ST}, and Nm among groups of ferns with three different mating systems (selfing, mixed, and outcrossing) and groups with different growth habits (epipetric, epiphytic, and terrestrial). We excluded taxa with either an apomictic mating system or an aquatic habit because of the relatively few studies that included representatives of these types of ferns ($n \leq 2$ for both categories). For the same reason, our analyses focused exclusively on homosporous ferns because of a paucity of studies on heterosporous taxa ($n = 2$). We used Kruskal-Wallis tests by rank to determine whether groups were significantly different from one another for each metric of interest, followed by pairwise comparisons using Wilcoxon rank sum tests with corrections for multiple testing using the Holm method (Holm

1979). All statistical analyses were performed in R version 3.6.1 (R Core Team 2019; script available at http://github.com/jessiepelosi/fernpopgen).

Results

We identified 87 published studies on fern population genetics using the procedure described above, and from these articles we assembled a data set that included 156 taxa (fig. 1; app. 1). Ranker and Geiger (2008) reported on 26 studies of population genetics; we built on that data set, adding an additional 61 studies. The studies are distributed unevenly across the fern phylogeny, with numerous lineages having received no study, while others have had numerous taxa included in population genetics research (fig. 1). Various molecular markers were employed across the 87 publications, but allozymes were the most frequently encountered (fig. 2). As with the phylogenetic coverage of the data, the geographic origins of samples were also uneven, with most studied taxa occurring in North America, Europe, or Asia (fig. 3). Across the data set, we were able to record or calculate our population genetic metrics of interest for the following numbers of taxa: mean H_e for 79 taxa, mean %P for 93 taxa, mean F for 57 taxa, mean F_{ST} for 63 taxa, and Nm for 52 taxa. Ploidy level was available for 135 taxa, growth habit for all 156 taxa, and mating system for 79 taxa. Regardless of mating system or growth habit, the average values of these population genetic statistics are $H_e = 0.19$ (range, 0–0.74; $n = 79$), %P = 38.4 (0–100; $n = 93$), $F = 0.172$ (-1 to 1; $n = 53$), $F_{ST} = 0.274$ (0–0.952; $n = 59$), and $Nm = 2.07$ (0–11.65; $n = 48$). These values are broadly comparable to those from seed plants; for example, average H_e and %P for seed plants are 0.13 and 34.2, respectively, as reported by Hamrick and Godt (1990; table 2).

For our comparisons of genetic diversity among mating systems, there were 42 taxa with data for both mating system and H_e (4 selfing, 10 mixed, 28 outcrossing), 49 with values of %P (4 selfing, 11 mixed, 34 outcrossing), 44 with values of F_{ST} (6 selfing, 11 mixed, 27 outcrossing), 44 with values of F (3 selfing, 10 mixed, 31 outcrossing), and 38 for which we were able to calculate Nm (5 selfing, 10 mixed, 23 outcrossing). Levels of heterozygosity between the three mating types (outcrossing taxa mean $H_e = 0.16$ [0.02–0.32; $n = 28$]; mixed mating mean $H_e = 0.22$ [0.01–0.72; $n = 10$]; selfing taxa mean $H_e = 0.38$ [0.06–0.73; $n = 4$]) were not significantly different from each other (Kruskal-Wallis test by ranks, $\chi^2 = 2.63$, df = 2, $P > 0.05$; fig. 4*A*). In contrast, the mean percents of polymorphic loci were significantly different (Kruskal-Wallis test by ranks, $\chi^2 = 12.339$, df = 2, $P < 0.01$); outcrossing taxa (mean %P = 50.8 [6.7–94.1; $n = 33$]) had significantly higher %P compared with mixed mating taxa (mean %P = 24.7 [0–81; $n = 12$]; pairwise Wilcoxon rank sums test, $P < 0.001$) and selfing taxa (mean %P = 22.5 [12.5–38.1; $n = 4$]; $P < 0.05$). No other pairwise comparisons were significant for this metric (fig. 4*B*). There were significant differences in mean F between all three mating systems (Kruskal-Wallis test by ranks, $\chi^2 = 8.72$, df = 2, $P < 0.05$), with outcrossing taxa having F values close to 0 (mean $F = 0.050$ [-0.34 to 0.35; $n = 31$), mixed mating taxa with values intermediate between those of outcrossing and selfing taxa (mean $F = 0.385$ (-0.167 to 0.961; $n = 10$), and selfing taxa with values close to 1 (mean $F = 0.953$ [0.947–0.962; $n = 3$]). We also found significant differences in mean F_{ST} (Kruskal-Wallis test by ranks, $\chi^2 = 17.37$, df = 2, $P < 0.01$). In particular, outcrossing taxa (mean $F_{ST} = 0.123$ [0–0.38; $n = 27$]) were significantly different from mixed (mean $F_{ST} = 0.487$ [0.057–0.779; $n = 11$]; pairwise Wilcoxon rank sums test, $P < 0.001$) and selfing taxa (mean $F_{ST} = 0.419$ [0.085–0.783; $n = 6$]; pairwise Wilcoxon rank sums test, $P < 0.05$; fig. 4*C*). We also found significant differences in levels of gene flow (Kruskal-Wallis test by ranks, $\chi^2 = 14.965$, df = 2, $P < 0.001$), particularly between outcrossing (mean $Nm = 3.60$ [0.28–11.65; $n = 23$]) and mixed mating taxa (mean $Nm = 0.55$ [0.04–4.14; $n = 10$]; pairwise Wilcoxon rank sums test, $P < 0.001$), but not between these and selfing taxa (mean $Nm = 1.15$ [0.01–2.69; $n = 5$]; pairwise Wilcoxon rank sums test, $P > 0.05$; fig. 4*D*).

For our comparisons of genetic diversity and gene flow among groups with different growth habits, there were 78 taxa that had growth habit information and mean H_e values (11 epipetric, 7 epiphytic, 60 terrestrial), 90 taxa with mean %P values (8 epipetric, 9 epiphytic, 73 terrestrial), 56 taxa with mean F values (9 epipetric, 11 epiphytic, 36 terrestrial), 60 taxa with mean F_{ST} values (8 epipetric, 11 epiphytic, 41 terrestrial), and 49 taxa for which we were able to calculate Nm (6 epipetric, 9 epiphytic, 34 terrestrial). Mean values of H_e were not significantly different among epipetric (mean $H_e = 0.23$ [0–0.72; $n = 11$]), epiphytic (mean $H_e = 0.19$ [0.14–0.27; $n = 7$]), and terrestrial (mean $H_e = 0.18$ [0–0.74; $n = 60$]) habits (Kruskal-Wallis test by ranks, $\chi^2 = 0.64$, df = 2, $P > 0.05$; fig. 5*A*). The average percent of polymorphic loci was significantly different among groups (Kruskal-Wallis test by ranks, $\chi^2 = 11.12$, df = 2, $P < 0.01$); epiphytic taxa (mean %P = 62.3 [33.3–80.0; $n = 10$]) were significantly different from the epipetric taxa (mean %P = 27.0 [0–65.0; $n = 8$]; pairwise Wilcoxon rank sums test, $P < 0.01$) and terrestrial taxa (mean %P = 36.9 [0–100; $n = 73$]; pairwise Wilcoxon rank sums test, $P < 0.05$; fig. 5*B*). We did not find a significant difference in F among the three growth habits (Kruskal-Wallis test by ranks, $\chi^2 = 4.299$, df = 2, $P > 0.05$), epiphytic (mean $F = 0.030$ [-0.013 to 0.115; $n = 11$]), epipetric (mean $F = 0.240$ [-1 to 0.961; $n = 9$]), and terrestrial (mean $F = 0.200$ [-0.619 to 1; $n = 37$]). There was a significant difference in mean F_{ST} among growth habits (Kruskal-Wallis test by ranks, $\chi^2 = 9.715$, df = 2, $P < 0.01$), with epiphytic (mean $F_{ST} = 0.130$ [0.021–0.770; $n = 11$]) and epipetric taxa being significantly different (mean $F_{ST} = 0.45$ [0.094–0.779; $n = 8$]; pairwise Wilcoxon rank sums test, $P < 0.01$), but not for other pairwise comparisons. There were significant differences in values of Nm among growth habits (Kruskal-Wallis test by ranks, $\chi^2 = 10.138$, df = 2, $P < 0.01$), but only the pairwise comparison of epiphytic (mean $Nm = 4.7$ [0.07–11.65; $n = 9$]) and epipetric taxa (mean $Nm = 0.38$ [0.03–1.20; $n = 7$]; pairwise Wilcoxon rank sums test, $P < 0.05$) was significant (fig. 5*D*).

Discussion

Below we discuss our data and results, with relevance to four topics: (1) mating systems and population structure, (2) life-history traits, (3) taxonomic and geographic biases, and (4) molecular methods. We focus in particular on several historical hypotheses and paradigms that have long been of interest in fern evolutionary studies, including classic work on fern

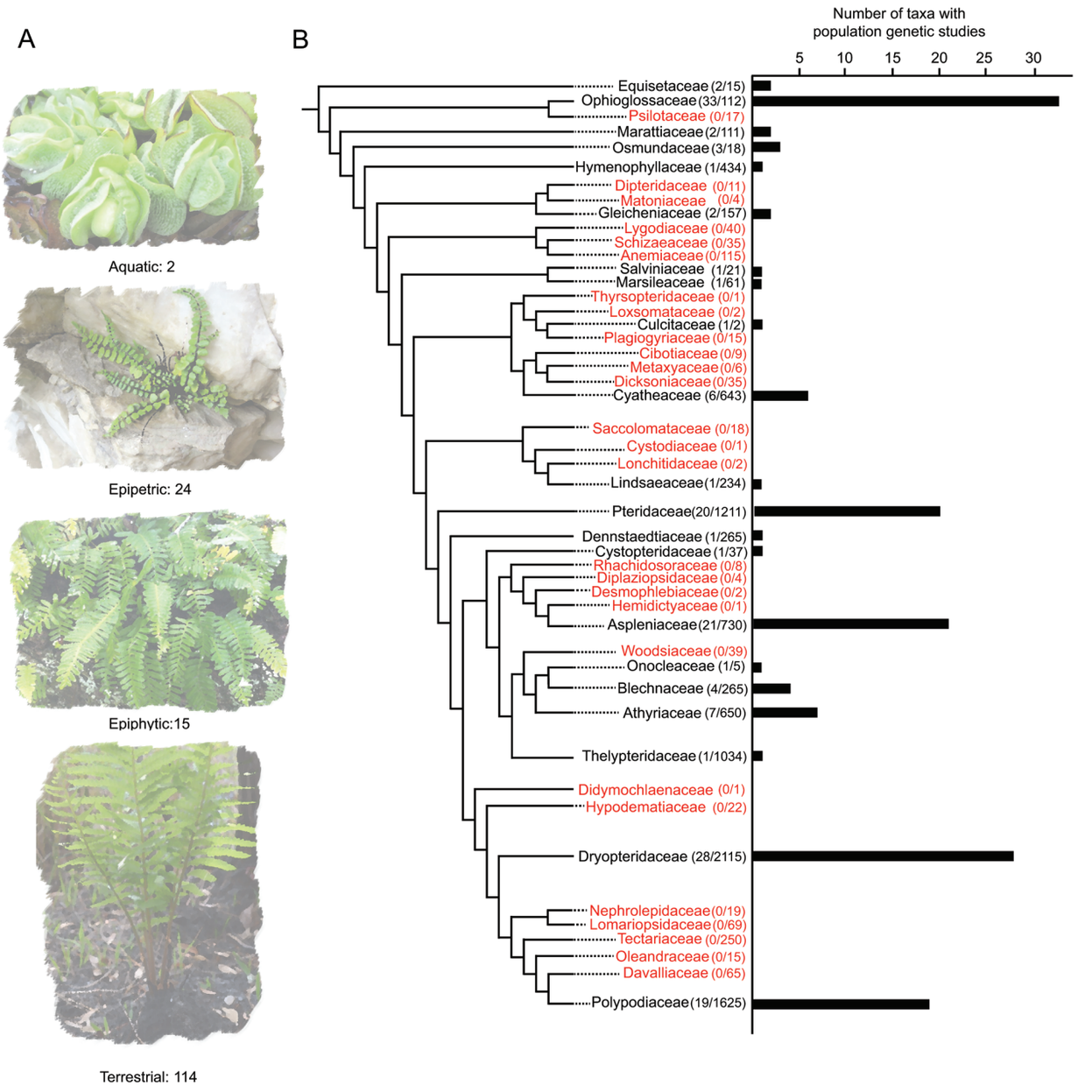

Fig. 1 *A*, Numbers of taxa for which population genetic studies have been published for four different growth habits, aquatic, epipetric, epiphytic, and terrestrial, with representative species pictured. From top: *Salvinia molesta*, *Asplenium trichomanes* subsp. *quadrivalens*, *Pleopeltis michauxiana*, and *Osmundastrum cinnamomeum*. The top three images are by E. B. Sessa, and the bottom image is by J. A. Pelosi. *B*, For each of the 48 fern families, the numbers of taxa for which population genetic studies have been published. The numbers in parentheses are the total number of taxa with population genetic studies/the total number of recognized species in the family, according to PPG I (2016). Families in red have not been the subjects of any population genetic studies. The phylogeny is modified from PPG I (2016).

mating systems, and implications for the large genomes of homosporous ferns.

Mating Systems and Population Structuring

If true, Klekowski's hypothesis that ferns are naturally predominantly inbreeding has profound consequences for the genetic variation and structuring of fern populations. This hypothesis and its underlying assumptions have three major implications for fern populations; these populations should (1) harbor low levels of genetic diversity as a result of inbreeding, (2) be deficient in heterozygous individuals, and (3) be highly genetically structured. We used our compiled data set to evaluate the evidence for these hypothetical consequences of selfing by addressing the following questions: (1) Are most ferns naturally inbreeding? (2) How does the

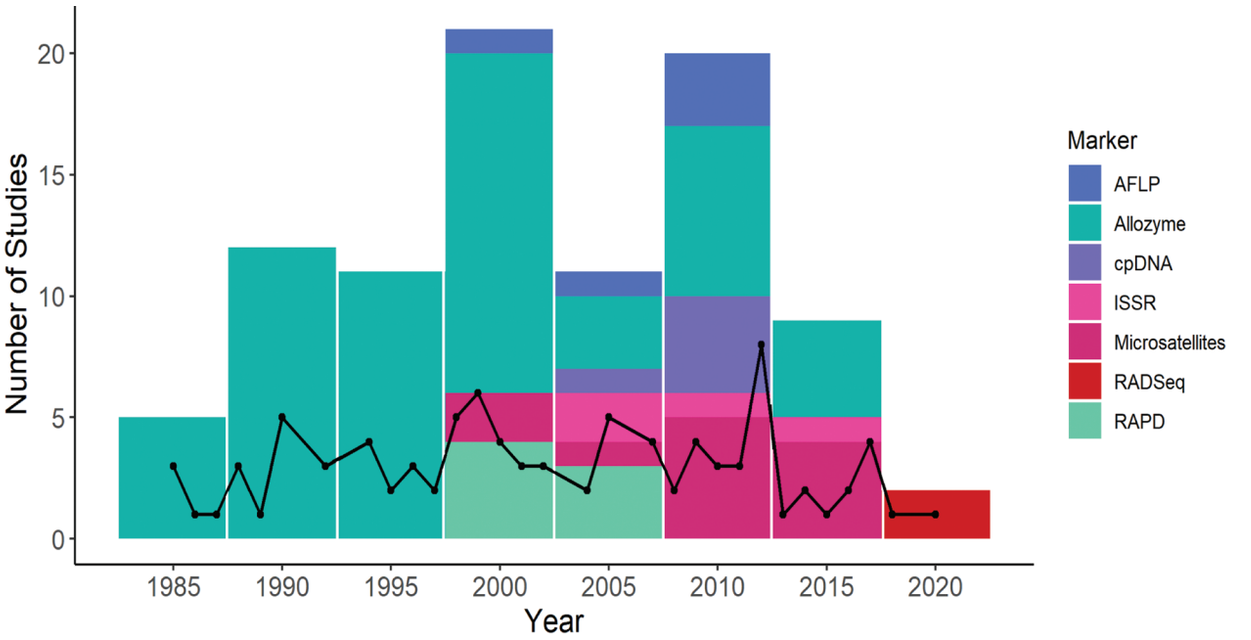

Fig. 2 Distribution of studies over time, with 5-yr divisions colored by the molecular marker used (allozymes, amplified fragment length polymorphisms [AFLPs], random amplified polymorphic DNA [RAPD], microsatellites, intersimple sequence repeats [ISSRs], chloroplast DNA sequencing [cpDNA], and restriction site–associated DNA sequencing [RADSeq]). The line traces the number of studies published by year.

mating system impact genetic diversity? (3) How do mating systems affect population structure? If ferns are predominantly inbreeding, we would expect to find evidence from natural populations for low levels of genetic diversity, few heterozygous individuals, and most genetic variation partitioned among (rather than within) populations.

Are most ferns naturally inbreeding? In contrast to Klekowski's assertion, ferns span the continuum of mating systems,

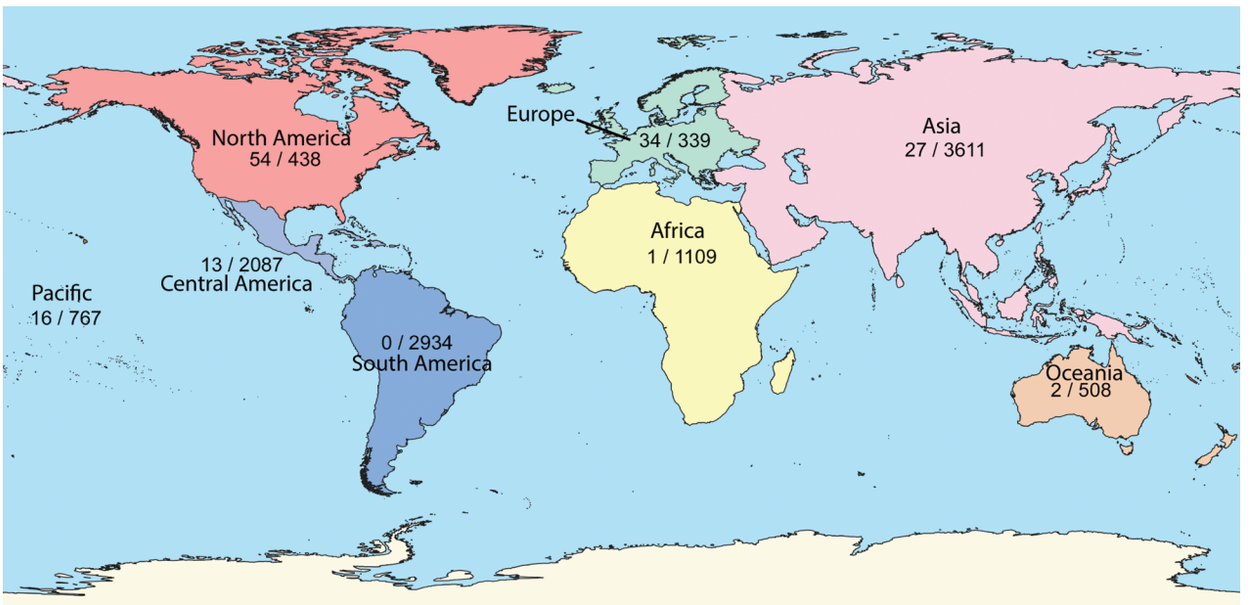

Fig. 3 Number of unique taxa studied thus far in each global region (first number) followed by the total number of taxa in each region (species richness estimates from Suissa et al. 2021), including Africa (yellow), Asia (pink), Australia and New Zealand (orange), Central America and the Caribbean (purple), Europe (green), North America (red), the Pacific Islands (black), and South America (blue).

Table 2

Percent Polymorphic Loci (%P) and Expected Heterozygosity (H_e) for Five Families of Ferns, All Ferns with Available Data, and Seed Plants

	Mating system(s)	Mean %P	Mean H_e
Aspleniaceae	Equal proportions of all three systems	18.5 (0–81; $n = 7$)	.16 (0–.49; $n = 9$)
Dryopteridaceae	Mixed or outcrossing	31.2 (0–81.3; $n = 15$)	.17 (.03–.57; $n = 8$)
Ophioglossaceae	Predominantly inbreeding	29.7 (0–66.7; $n = 25$)	.15 (0–.34; $n = 21$)
Polypodiaceae	Predominantly outcrossing	58.7 (33.3–80, $n = 12$)	.19 (.14–.27; $n = 10$)
Pteridaceae	Mixed or outcrossing	48.9 (19–70; $n = 14$)	.26 (.03–.72; $n = 10$)
All ferns	Selfing, mixed, outcrossing	38.4 (0–100; $n = 93$)	.19 (0–.74; $n = 79$)
Seed plants	Selfing, mixed, outcrossing	34.2	.13

Note. The ranges of each metric are provided in parentheses, followed by the sample size, *n*. The predominant mating system of each group is also indicated. Seed plant values are from Hamrick and Godt (1990).

from extremely inbreeding (gametophytic selfing) in *Botrypus* (Soltis and Soltis 1986) to highly outcrossing in *Polystichum* (Soltis and Soltis 1987*b*). By estimating the rate of gene flow with the number of migrants per generation (*Nm*) and the inbreeding coefficient (F_{IS}), Soltis and Soltis (1987*b*) demonstrated that, despite having the capacity for gametophytic selfing, *Polystichum munitum* regularly outcrosses (with up to 24 migrants per generation) over great distances. Other species within Dryopteridaceae, such as *Polystichum imbricans* (Soltis et al. 1988) and *Polystichum otomasui* (Maki and Asada 1998), are also primarily outcrossing, although others, like *Dryopteris expansa* (Soltis and Soltis 1987*a*), *Dryopteris remota* (Schneller et al. 1998), and *Dryopteris fragrans* (Bouchard et al. 2017), exhibit mixed mating systems (a combination of selfing and outcrossing). In Aspleniaceae, some species are highly inbreeding (Suter et al. 2000; Bucharová and Münzbergová 2012; De Groot et al. 2012; Fernando et al. 2015), whereas others regularly outcross (Hunt et al. 2009; Bucharová and Münzbergová 2012). Nearly all taxa in Ophioglossaceae regularly undergo gametophytic selfing (Soltis and Soltis 1986; Watano and Sahashi 1992; Hauk and Haufler 1999; Camacho and Liston 2001; Chung et al. 2010, 2013). Conversely, members of Polypodiaceae are predominantly outcrossing (Ranker 1992*b*, 1994; Hooper and Haufler 1997), while Pteridaceae contains a mix of inbreeding (Ranker 1992*a*) and outcrossing species (Gastony and Gottlieb 1985; Haufler 1985; Ranker 1992*a*; Pryor et al. 2001; Liu et al. 2007; Kang et al. 2008). Many authors have additionally calculated estimates of the rate of gametophytic selfing (e.g., Holsinger 1987; Soltis et al. 1988; Maki and Asada 1998) and found that most ferns have low frequencies of gametophytic selfing. While we had insufficient data to divide selfing taxa further into gametophytic and sporophytic selfing types, our findings nevertheless agree with these previous conclusions that selfing and therefore inbreeding are not the norm among ferns. While all homosporous ferns may theoretically be capable of extreme inbreeding in the form of gametophytic selfing (Sessa et al. 2016), natural fern populations clearly span the gamut of possible mating behaviors (table 2; apps. 1, 2). In accordance with the findings in Sessa et al. (2016) that polyploids have a greater proportion of gametophytic selfing than diploids, we also found that there are significantly different proportions of diploids and polyploids within each mating system, with diploids having less selfing (Pearson's χ^2 test, $\chi^2 = 9.036$, df $= 1$, $P < 0.05$). How this affects the population structure and diversity of polyploidy populations remains unanswered.

Despite ferns' capacity to produce bisexual gametophytes, only a handful of taxa appear to regularly undergo gametophytic selfing in natural conditions (table 2; apps. 1, 2). Various factors influencing the timing of gametangia (archegonia and antheridia) development in gametophytes, including temperature, light, and soil conditions, with the notable exception of genetic regulation (Schneller 2008), have been proposed. Döpp (1950) was the first to demonstrate the existence in ferns of pheromones called antheridiogens, which induce the production of antheridia in fern gametophytes. Typically, the first spore to germinate and reach the meristic stage (more than one cell layer) will become an archegoniate gametophyte and produce antheridiogens. Neighboring gametophytes that have yet to reach this developmental stage will respond to this chemical cue and become male, producing only antheridia. By suppressing the production of archegonia in surrounding plants, this process prevents gametophytic selfing and promotes mating between gametophytes from either the same sporophyte (sporophytic selfing) or other sporophyte(s) (sporophytic outcrossing). Although most studies of the antheridiogen system are laboratory based, several studies (Cousens 1979; Schneller 1988; Hamilton and Lloyd 1991; Quintanilla et al. 2005) have corroborated these findings in natural populations. Reviews on the antheridiogen system are given by Schneller (2008) and Hornych et al. (2021).

The numerous studies cited above make it clear that natural fern populations do not strictly or even primarily undergo gametophytic selfing in the field. Rather, ferns capitalize on all three potential mating systems. While some families may be primarily selfing (e.g., Ophioglossaceae), the majority tend to be primarily outcrossing (e.g., Polypodiaceae) and/or have mixed mating systems (Ranker and Geiger 2008).

How does mating system impact genetic diversity? Under Klekowski's hypothesis that natural fern populations are predominantly inbreeding, most homosporous fern populations should exhibit low levels of genetic diversity. Genetic diversity is typically evaluated by calculating %P and/or H_e. The results from our compiled data set suggest that ferns have levels of genetic diversity comparable to those of seed plants (table 2; see also Hamrick and Godt 1990). Levels of heterozygosity did not differ significantly among mating systems and were similar among outcrossing, mixed mating, and selfing taxa (fig. 4*A*). Interestingly, the mean expected heterozygosity for outcrossing species was less than that for the selfing species (fig. 4*A*), contrary to theoretical predictions, although this may be an artifact of low sample sizes. Increased heterozygosity in selfing taxa may

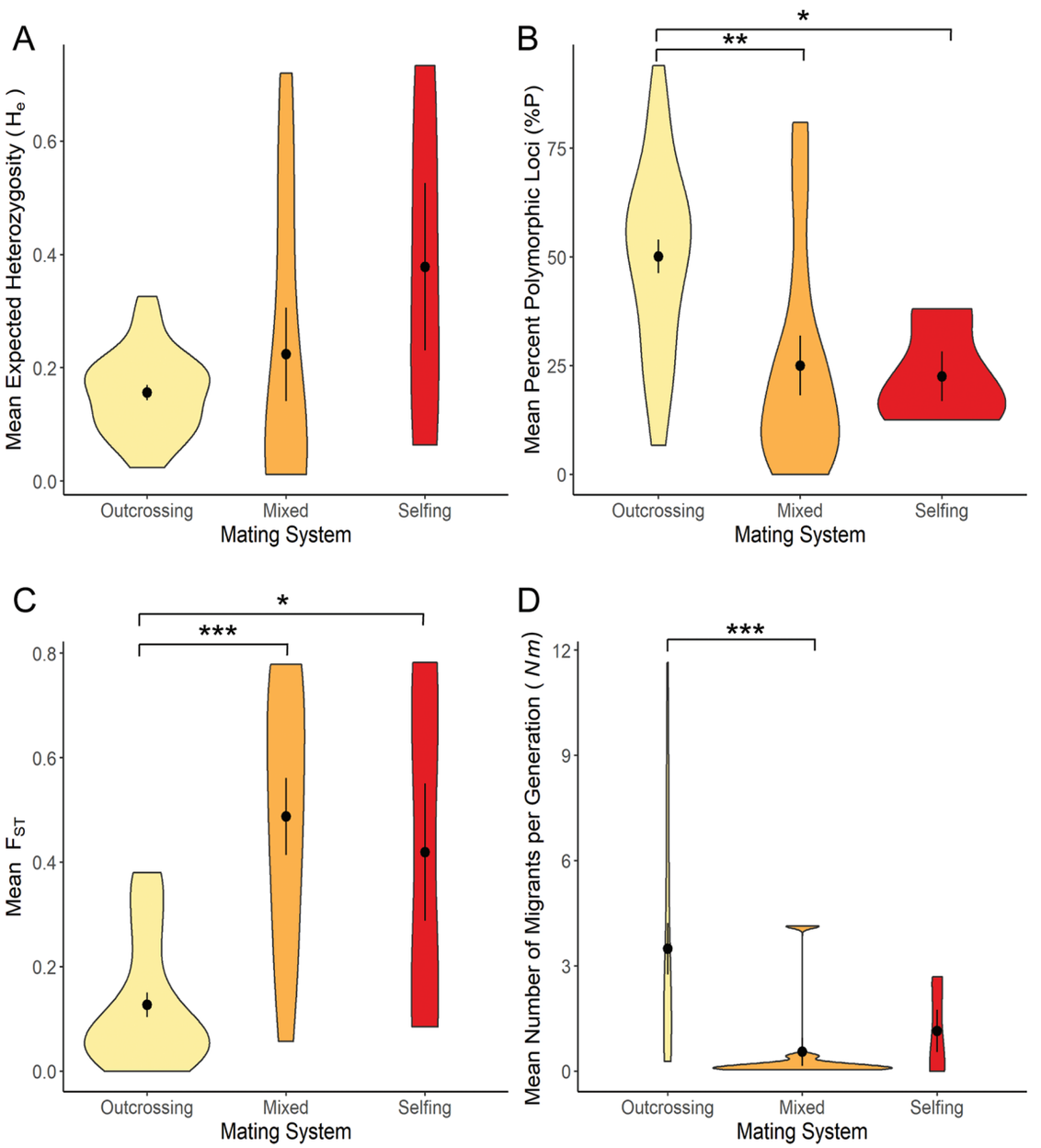

Fig. 4 Violin plots of estimates of population genetic metrics across the three mating systems of ferns: outcrossing (yellow), mixed (orange), and gametophytic and sporophytic selfing (red). Plots show the density and distribution of observations for each metric, with the width proportional to the density of observations. The inner circle indicates the mean, and the vertical lines indicate ±SE. *A*, Mean expected heterozygosity under Hardy-Weinberg equilibrium (H_e) from 42 total taxa (4 selfing, 10 mixed, 28 outcrossing). *B*, Mean percent of polymorphic loci (%P) from 49 total taxa (4 selfing, 11 mixed, 34 outcrossing). *C*, Mean fixation index (F_{ST}) from 44 total taxa (6 selfing, 11 mixed, 27 outcrossing). *D*, Mean number of migrants per generation (*Nm*) from 38 total taxa (5 selfing, 10 mixed, 23 outcrossing). Significance: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

also result from fixed heterozygosity in polyploids, which are more likely to be tolerant of selfing because of their extra genomes (Lande and Schemske 1985; Hedrick 1987; Soltis and Soltis 2000; Barringer 2007). This likely has a small effect on our data set, however, as there are only two polyploid selfing taxa included (*Asplenium adulterinum* and *Ophioglossum vulgatum*) and they differ substantially from one another in H_e: *O. vulgatum* has remarkably low levels of heterozygosity ($H_e = 0.06$; Chung et al. 2013), while *A. adulterinum* has levels of heterozygosity ($H_e = 0.49$; Bucharová and Münzbergová 2012) two and a half times greater than the average for all ferns ($H_e = 0.19$). While we cannot draw broad conclusions from a

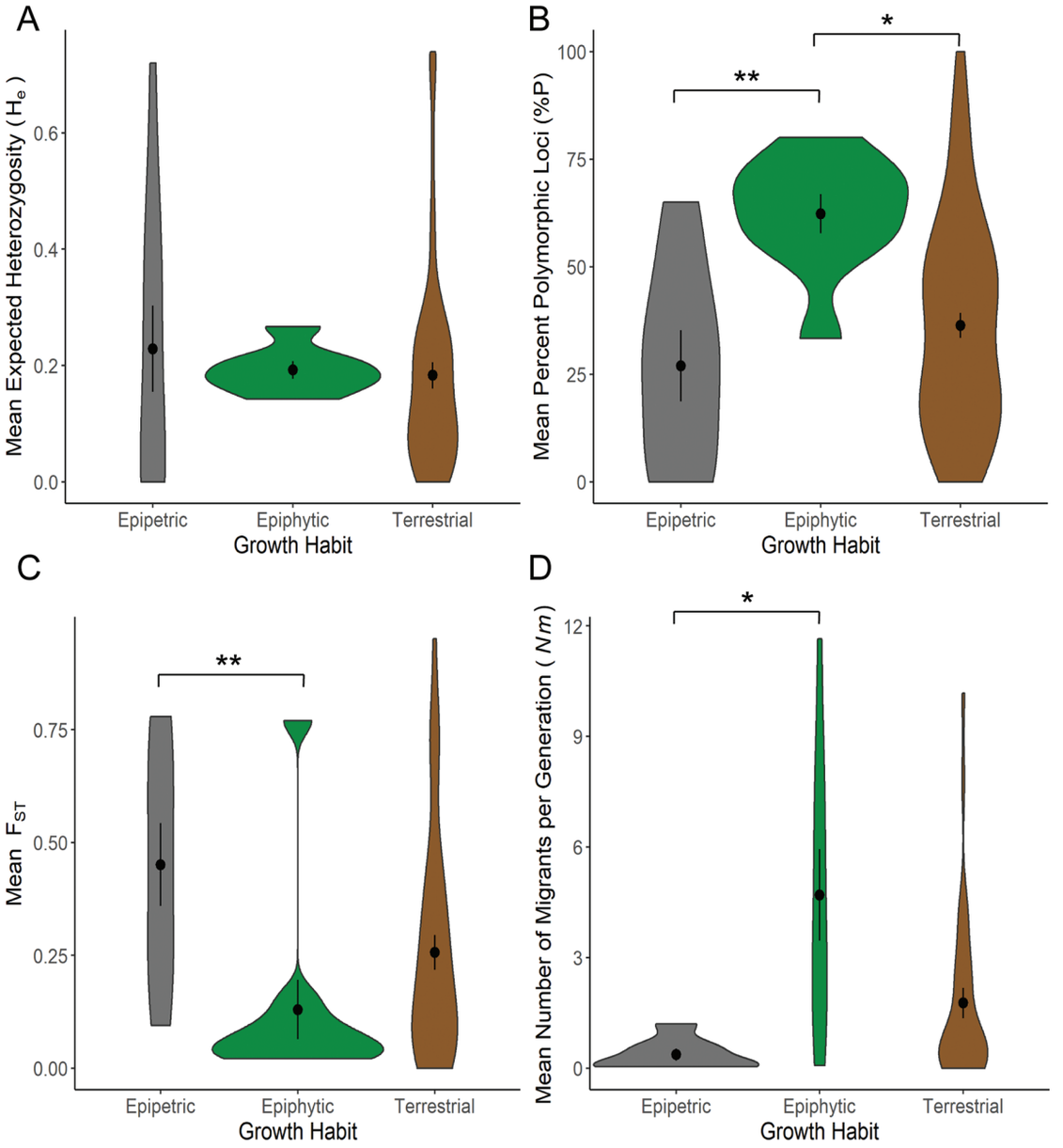

Fig. 5 Estimates of population genetic metrics across three fern growth habits: epipetric (gray), epiphytic (green), and terrestrial (brown). Plots show the density and distribution of observations for each metric, with the width proportional to the density of observations. The inner circle indicates the mean, and the vertical lines indicate ±SE. *A*, Mean expected heterozygosity under Hardy-Weinberg equilibrium (H_e) from 78 total taxa (11 epipetric, 7 epiphytic, 60 terrestrial). *B*, Mean percent of polymorphic loci (%P) from 90 total taxa (8 epipetric, 9 epiphytic, 73 terrestrial). *C*, Mean fixation index (F_{ST}) from 60 total taxa (8 epipetric, 11 epiphytic, 41 terrestrial). *D*, Mean number of migrants per generation (Nm) from 49 total taxa (6 epipetric, 9 epiphytic, 34 terrestrial). Significance: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

two-taxon sample, these data suggest that genetic diversity in polyploid ferns may be strongly influenced by individual, lineage, or population-specific effects.

Many authors have also used the inbreeding coefficient, F, to assess levels of heterozygosity within populations (Wright 1931, 1965). It is calculated as

$$F = 1 - \left(\frac{H_o}{H_e}\right),$$

where H_e is the expected heterozygosity under Hardy-Weinberg equilibrium and H_o is the observed heterozygosity of a population. It is a measure of the reduction of heterozygosity relative to if that population were in Hardy-Weinberg equilibrium (Hartl and Clark 1989). Values of F close to 0 are indicative of outcrossing, whereas values approaching 1 are suggestive of inbreeding. Using our compiled data, we found that ferns, on average, have F values close to 0 (average $F = 0.172$), suggesting that these plants are predominantly randomly mating. These

results reinforce the findings of Soltis and Soltis (1990), who compiled F values from nine taxa and found an average F of 0.156, ranging from −0.075 in *Polystichum dudleyi* (outcrossing) to 0.962 in *Botrypus virginianus* (selfing). This smaller set of values and the larger data set presented here emphasize that outcrossing appears to be the norm in natural fern populations.

In contrast to expected heterozygosity, the percent of polymorphic loci differed significantly among the different mating systems, with outcrossing taxa having higher levels of polymorphism than either mixed mating or selfing taxa (which did not differ significantly from one another; fig. 4*B*). These findings follow theoretical expectations, with selfing taxa having lower values of %P than outcrossing taxa because individuals produced from a single gametophyte or sporophyte are more likely to be homozygous at multiple loci than individuals produced from outcrossing events (Haufler 2002). Interestingly, taxa with mixed mating systems had the lowest %P but with a spread of values more similar to that of outcrossing than to that of selfing taxa (fig. 4*B*).

We further investigated how patterns of genetic diversity compare across the five families of ferns that have (relatively) extensive data available (fig. 1). We assigned each of these families a dominant mating system type based on the frequency each mating system is found in the taxa belonging to that family in the data set (table 2). We found that in general and as expected, families with primarily outcrossing taxa have higher levels of genetic diversity than families that are mostly composed of taxa with mixed and selfing mating systems (table 2). The average values of %P and H_e across all ferns in this data set are similar to those in seed plants (Hamrick and Godt 1990), which are heterosporous and can use either outcrossing or selfing, which is analogous to sporophytic selfing in ferns but not to the extreme gametophytic selfing that homosporous ferns are capable of. The similarity in values among all fern and seed plants thus suggests that gametophytic selfing is not the predominant reproductive mode of ferns. Furthermore, the family Aspleniaceae, which includes taxa that employ all three mating systems, had relatively low levels of genetic diversity even compared with the predominantly selfing Ophioglossaceae, suggesting that, as in the case of polyploidy, there are likely additional and perhaps lineage-specific factors other than mating system at play in shaping the genetic diversity of fern populations (table 2).

How do mating systems affect population structure? The amount of genetic differentiation that can be explained by population structure can be measured by F_{ST}, a widely used metric. It is calculated as

$$F_{ST} = \frac{\sigma_a^2}{\sigma^2},$$

where σ_a^2 is the variance of allele frequencies in a subpopulation and σ^2 is the total allele frequency variance in the population (Hartl and Clark 1989). Values of F_{ST} range from 0 to 1, with 0 indicating that populations are interbreeding freely and 1 suggesting complete differentiation (i.e., subpopulations are effectively not interbreeding, and there is no gene flow between them; fig. 6*C*). Theory suggests that species that outcross should have F_{ST} values close to 0 and partition most of their genetic variation within populations (i.e., fig. 6*A*), whereas inbreeders would have F_{ST} values significantly greater than 0, with most of their genetic variation partitioned among or between populations (fig. 6*B*, 6*C*).

We found, as expected, that outcrossing ferns have populations that are less structured than those of mixed mating and selfing taxa (fig. 4C). The lack of structure in these populations is likely related to the extraordinarily high levels of gene flow exhibited by ferns (fig. 4*D*), which range from values near 0 migrants per generation in *O. vulgatum* (selfing; Chung et al. 2013) to 11.6 migrants per generation in *Pleopeltis astrolepis* (outcrossing; Hooper and Haufler 1997) and 24 migrants per generation in *P. munitum* (outcrossing; Soltis and Soltis 1987*b*). Theoretically, one migrant per generation is sufficient to prevent the divergence between populations that would naturally result from stochastic changes in allele frequency due to genetic drift combined with inbreeding (where intrasubpopulation inbreeding leads to fixation and differentiation from other subpopulations; Spieth 1974; Hartl and Clark 1989). It therefore follows that values of $Nm > 1$ reflect levels of gene flow sufficiently high to maintain homogeneous populations, whereas values of $Nm < 1$ are indicative of restricted gene flow, which should result in heterogeneous populations. The high values of Nm seen in ferns (fig. 4*D*) and the consequently large amounts of genetic information potentially passing between populations in each generation should prevent differentiation among these populations, resulting in low values of F_{ST} (fig. 4*C*).

Mating system clearly has a significant impact on the genetic structuring of fern populations. Taxa that are primarily outcrossing have significantly less structured populations than do

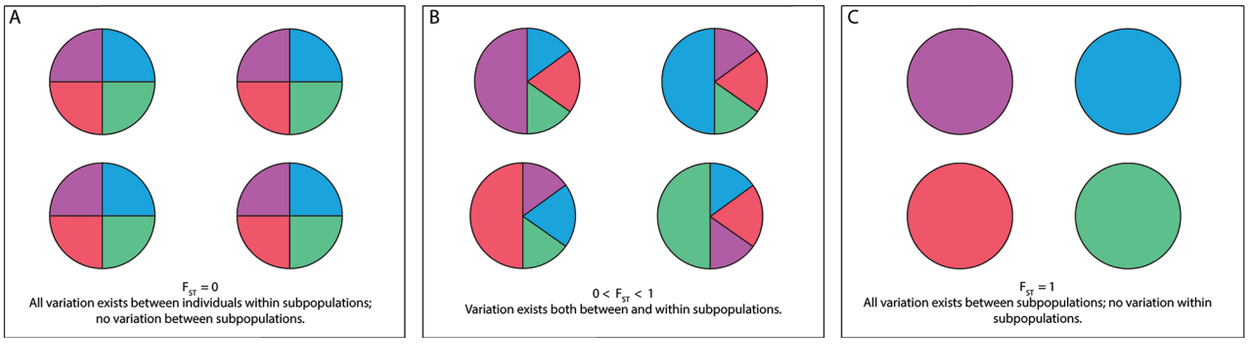

Fig. 6 Graphical representation of inter- versus intrapopulation structure as it relates to the fixation index (F_{ST}) value.

mixed and selfing taxa. There were, however, very few studies focused on primarily selfing taxa, and the terminology we have used here to describe mixed mating taxa may be too broad to describe these ferns. We believe that it is, therefore, necessary for further studies to focus on mixed mating taxa, as in Wang et al. (2012), which could further elucidate the specific impact of mating systems on populations of organisms, even while variation in the mating system may be present in a taxon. The interplay between habitat differences (e.g., soil types, precipitation) and mating system and how this interaction may impact genetic divergence and variation are understudied and deserve further attention. The data and results compiled here reiterate the findings of early publications that outcrossing taxa partition most of their variation within populations, while those that are predominantly selfing partition most variation among populations (e.g., Brown 1979; Soltis and Soltis 1990). The distribution of alleles within and among populations may have significant implications for the ability of populations to adapt, highlighting the interplay between mating systems, gene flow, and selection.

Life-History Traits

Synthesis. Despite Soltis and Soltis's (1990, p. 171) encouragement that "life-history characteristics should be examined for both levels and patterns of genetic variation," relatively few studies have assessed how life-history traits affect population genetic dynamics in ferns. It is not difficult to imagine how a change in mating strategy may influence the genetic diversity of a population, but life span, growth habit, and age at first reproduction may seem more subtle. Consider, however, comparing the genetic variation of long-lived trees and short-lived herbaceous annuals or of long- and short-distance dispersers—the likelihood that these different life-history strategies will significantly impact population genetic dynamics is high and has been demonstrated in both plants and animals (e.g., Hamrick and Godt 1996; Nybom 2004; Ellegren and Galtier 2016). Unsurprisingly, data connecting life-history traits and genetic diversity are scarce for ferns. We use the available data to first address questions regarding how different growth habits may influence fern population dynamics. We then explore other interesting aspects of fern life histories (for which data are not currently available) that would be fruitful avenues for future research.

Although ferns are substantially less species rich than the seed plants (ca. 10,000 species compared with ca. 300,000–400,000; PPG I 2016; Soltis et al. 2017), they fill many of the same ecological and life-history niches as members of the latter group. This is particularly true for growth habit: ferns may be terrestrial, epiphytic (growing on other plants, especially trees), hemiepiphytic (both epiphytic and terrestrial during their life span), epipetric (growing on rock faces; also known as lithophytes), vining, or fully aquatic. The growth habit of the sporophyte is relevant for population genetics because it likely affects spore dispersal and may therefore influence the way that genetic variation is partitioned. It is thus an important component of fern population genetics that has been largely overlooked. Many species (e.g., *Adiantum incisum*) and genera (e.g., *Asplenium*) exhibit multiple growth habits and provide a valuable opportunity to examine how a particular habit influences the genetic landscape of a population and species. For example, do epiphytes exhibit greater connectivity between populations than terrestrial species? Do epipetric species have higher levels of population structure due to habitat fragmentation? Is the dispersal distance of aquatic ferns greater than that of terrestrial species? To the best of our knowledge, these questions have not yet been addressed and are open areas of research.

To begin to address some of these questions, we assessed how different life-history traits might affect population genetic dynamics in ferns using our assembled data set. We found that genetic diversity (%P) was greatest in epiphytic ferns and that there were lower levels of population differentiation and increased gene flow in epiphytic taxa compared with epipetric taxa (fig. 5). Dassler and Farrar (2001) provide a compelling case that increased outcrossing and long-distance dispersal (LDD) in some epiphytic ferns may be linked to the production of gemmae (vegetative propagules from gametophytes). Furthermore, Sessa et al. (2016) reported lower selfing rates in epiphytic taxa compared with terrestrial taxa. While these results suggest that epiphytic ferns have elevated levels of genetic diversity and connectivity between populations, the limited size and phylogenetic breadth of the data set limit the generalizability of these conclusions. It is important to note that all the epiphytes in this data set are from the Polypodiaceae and are all outcrossing; these data points may not represent other epiphytic lineages (e.g., Hymenophyllaceae, Aspleniaceae) or epiphytes with different mating systems. Differences between epiphytic and other fern taxa presented here require further investigation.

Another key life-history characteristic of ferns is their nutritionally independent gametophyte phase, which has long been viewed as a fragile handicap in the fern life cycle. Only relatively recently has this notion been reevaluated; fern gametophytes have now been shown to be physiologically resilient (Watkins et al. 2007*b*), and they can be long-lived (Watkins et al. 2007*a*) and can even persist without a sporophyte (Farrar 1967), a phenomenon termed the "separation of generations" (Pinson et al. 2017). As Watkins et al. (2007*a*, p. 701) stated, "Every sporophyte population owes its beginnings to the gametophyte generation," and yet little is understood about the underlying genetic dynamics of gametophyte populations. While the minute size of the gametophyte may make it difficult to study in many taxa, several authors (Farrar 1990; Watkins et al. 2007*a*; Chambers and Emery 2016; Pinson and Schuettpelz 2016) have analyzed naturally occurring gametophyte populations, revealing that some gametophytes can persist indefinitely without producing a sporophyte. Others (e.g., Ebihara et al. 2013) have shown that distributions of gametophytes may be larger than those of their sporophyte counterparts. Further, Nitta et al. (2016) demonstrated that the community structure of the gametophyte and sporophyte life stages can be significantly different and that gametophyte and sporophyte communities can vary in their phylogenetic composition along elevational gradients. Of the studies examined here, only 3.4% (3 of 87 studies) investigated the genetics of gametophyte populations; quantitative comparison of the two generations has, therefore, yet to be explored thoroughly (but see Rumsey et al. 1999 for a qualitative assessment of differences in sporophyte and gametophyte genetic variation in *Vandenboschia speciosa*). Despite its relevance to population establishment, the population genetics of gametophytes represents a substantial deficiency in our knowledge of ferns and deserves further inquiry.

Future directions. Although ferns can take on several growth habits, we suggest that future work focus in particular on the population dynamics of epiphytic ferns. Epiphytes make up around one-third of extant fern diversity (Schuettpelz and Pryer 2009) and play important roles in canopy nutrient cycling and microclimate maintenance (Nadkarni et al. 2004; Cardelús et al. 2006; Watkins et al. 2007*c*; Cardelús 2010; Cardelús and Mack 2010). Despite their prevalence throughout the fern phylogeny, all epiphytic taxa thus far considered in population genetic studies have been from the Polypodiaceae. To explore the role of epiphytism in the population genetic dynamics of ferns, we must expand the taxonomic sampling of epiphytes outside this single family. We suggest investing efforts into the Davalliaceae, Oleandraceae, and Aspleniaceae, families in which epiphytes make up at least 25% of the taxa (Watkins and Cardelús 2012). In addition, Lomariopsidaceae represents a unique opportunity to study the evolution of epiphytism, as many members of this family are hemiepiphytic, a potential intermediate step in the evolution of the epiphytic growth habit (Watkins and Cardelús 2012). We also recommend that future work focus on the Hymenophyllaceae, which both lacks an extensive characterization of population genetics (fig. 1) and represents a lineage with a potentially quite early origin of epiphytism. Schuettpelz and Pryer (2009) posited that epiphytism in the Polypodiales originated primarily in the Cenozoic (i.e., within the past 65 Myr). The Hymenophyllaceae includes many epiphytic taxa, and it diverged from its sister clade, which includes the majority of leptosporangiate ferns, approximately 350–275 mya (Schuettpelz and Pryer 2009; Testo and Sundue 2016). Epiphytism may thus have evolved much earlier in this lineage than in the Polypodiales, and it would be interesting to know how population structure and gene flow may vary between these independent origins of the epiphytic habit.

Another fern lineage that represents both an understudied clade and a unique growth habit is the order Salviniales, the water ferns. Members of this clade are strikingly different from other ferns both in their fully aquatic life cycle and because they are heterosporous and produce two separate spore types, like seed plants. While they make up less than 1% of extant fern diversity, species of *Azolla* and *Salvinia* are influential in human affairs, for either their positive (as in the case of *Azolla*) or negative (*Salvinia*) economic value. Their substantially smaller genomes (0.25–1.76 Gb; Li et al. 2018) and lower chromosome counts, which accompanied the shift to heterospory, make these ferns ideal candidates for future population genomic studies. They are also some of the only ferns with readily available genomic resources (Li et al. 2018), and so these two genera provide excellent opportunities to answer questions about how the evolution of heterospory may affect population dynamics.

While growth habit appears to play a role in driving genetic diversity and structuring fern populations, other life-history characteristics may also be important factors in shaping population dynamics. In seed-free homosporous plants, a single spore is capable of founding a new population (Klekowski 1970, 1972; Haufler et al. 2016; Sessa et al. 2016). After spore germination, the gametophyte is theoretically able to undergo gametophytic selfing to produce a new sporophyte. This process exposes deleterious recessive alleles, effectively purging the new population of lethal alleles (genetic load; Klekowski 1969, 1979). Frequently colonizing ferns are often but not always capable of reproducing through gametophytic selfing and therefore harbor lower levels of genetic load than outcrossing lineages (Ranker and Geiger 2008). This mode of colonization may seem impractical; the subsequent population would be initially homozygous and genetically depauperate, with occasional mutations and meiotic recombinants the only means of gradually increasing genetic diversity unless other conspecific spores arrive to permit outcrossing. However, LDD is common in ferns, and LDD and single-spore colonizations have played an important role in shaping biogeographic patterns and biodiversity (reviewed by Moran 2008). In one of the few studies to document in detail the colonization history of an island system by ferns, Ranker et al. (1994) estimated that there had been up to 17 independent colonizations of the Hawaiian archipelago by *Asplenium adiantum-nigrum* from mainland populations in North America, Europe, western Asia, and northern Africa. This result demonstrates not only that ferns are exceptional long-distance dispersers capable of reaching the most remote oceanic island chains but also that repeated colonization of the same island is entirely possible, and a population initially established by a single spore may in fact not have long to wait before the means of achieving genetic diversity arrives in the form of additional propagules. Fern biogeography is a rich area of active research and is often conducted at or above the species level; data on the population-level dynamics of ferns that either occur on islands or are otherwise thought to be the result of limited numbers of colonization events are, therefore, scarce (Tryon 1970, 1986) and would be fascinating subjects for future research.

Taxonomic and Geographic Biases

Synthesis. Although there are ca. 10,000 extant species of ferns (PPG I 2016), only 156 taxa (or approximately 1% of the extant diversity) from 21 of the 48 fern families have been the focus of population genetic studies (fig. 1). Of these 156 taxa, 118 (77%) belong to five families, with Ophioglossaceae receiving the most attention (33 taxa), followed by Dryopteridaceae (28), Aspleniaceae (21), Pteridaceae (20), and Polypodiaceae (19). Other families either lack studies entirely or have had relatively few (one to seven) taxa treated (fig. 1). Some of these groups have received special attention because of their fascinating biology and conservation status. For example, the gametophytes of *Botrychium* spp. (Ophioglossaceae) are entirely subterranean (Bierhorst 1958), a feature that likely has profound impacts on the genetic structure and mating systems of these plants (Soltis and Soltis 1986). Other taxa have been of interest because they are threatened or endangered (e.g., *Marsilea strigosa* Vitalis et al. 2002, *Mankyua chejuense* Chung et al. 2010, *Asplenium scolopendrium* var. *americanum* Fernando et al. 2015) or are otherwise rare (e.g., *Adenophorus periens* Ranker 1994). While these are fascinating groups of organisms, a broader, more representative sampling is needed to understand the factors influencing genetic diversity in ferns.

To determine whether there were any geographic biases in the literature, we also evaluated studies on the basis of the geographic region of the focal taxa (fig. 3). Nearly all taxa examined (87%) occur in North America, Europe, or eastern Asia, with very few publications focusing on areas outside the Northern Hemisphere. These data reveal that the tropics have been

largely neglected, with the exception of a handful of studies in Central America, where species richness is much greater than in the temperate regions (Moran 2008). Additionally, with the exception of one publication (Vitalis et al. 2002), Africa and South America, which harbor around 40% of global fern diversity, have been entirely neglected.

Future directions. There are clear deficiencies in the taxonomic sampling of fern population genetic studies, with the vast majority focusing on only five of the 48 recognized families (PPG I 2016; fig. 1). We recommend that future work focus on orders and families that have been neglected, particularly those with interesting phylogenetic positions. These include the numerous nonpolypod leptosporangiate fern families, some of whose members have intriguing life-history characteristics (e.g., members of the Gleicheniales and Schizaeales, the climbing and vining ferns, and the Cyatheaceae, the tree ferns). Within the eupolypod ferns, the families making up Eupolypods II (Aspleniineae sensu PPG I) have been somewhat more evenly studied than those in Eupolypods I (Polypodiineae sensu PPG I); in the latter clade, only two of the nine families have been the focus of population genetics studies, compared with six of the 11 families in Eupolypods II. In addition, there are several species-rich families with disproportionately fewer studies than other similarly sized groups, including Athyriaceae, Blechnaceae, Cyatheaceae, Dennstaedtiaceae, Hymenophyllaceae, Lindsaeaceae, and Thelypteridaceae. A variety of growth habits, including epipetric, epiphytic, terrestrial, and aquatic lifestyles, are included in these families.

The boundary separating population genetics and taxonomy is permeable, and population genomics may be important in informing taxonomy (e.g., Kinosian et al. 2020). Population-level information may be critical for resolving cryptic taxa or relationships of subspecies, varieties, and cytotypes that remain largely unsettled. For instance, there are unresolved relationships within the *Asplenium trichomanes* complex (e.g., Vogel et al. 1999) that offer an excellent opportunity to use population genetics and cytogenetics to inform the taxonomic status of the cytotypes, subspecies, and varieties in this species complex.

In addition to taxonomic sampling deficiencies, there are also substantial deficits in our geographic sampling for fern population genetics, as noted above (fig. 3). Many of the families mentioned above are cosmopolitan and therefore represent opportunities to ask questions about groups that are underrepresented in the literature on multiple fronts. The global biogeographic distributions of various ferns also offer opportunities to investigate additional questions about population structure in taxa with unique ranges. For example, the phylogenetic affinity of many South American, African, and/or Australian fern species (Moran and Smith 2001; Morero et al. 2019) provides a framework to compare the population genetics of closely related species separated by thousands of miles. Furthermore, given the propensity of ferns to frequently disperse to and among oceanic islands, the Pacific is an exceptional theater to test the widely used models of population divergence. Testing the assumptions and accuracy of these models will enhance our understanding of the evolution of fern populations. The accumulation and maintenance of genetic diversity following bottlenecks and habitat fragmentation events, such as those noted for African *Dryopteris* (Sessa et al. 2017), are of great interest for population and conservation geneticists. Depending on the questions posed, we implore investigators to consider regions of the globe traditionally overlooked in the literature, particularly in the tropics of Africa, Asia, and South America.

Molecular Methods

The approaches used to determine an individual's genotype (and therefore our ability to determine allelic variation for population genetic analyses) have changed drastically over the past five decades. The vast majority of fern population genetic studies have used allozyme electrophoresis to determine variation among and within populations (fig. 2). With this technique, only nonsynonymous mutations in specific protein-coding regions can be assessed. Much of what we know about fern population genetics is derived from this method, in part because allozymes were the dominant system used for population genetic studies in the 1980s and 1990s, when the bulk of the research on fern population genetics was conducted (fig. 2). This method is still used today (e.g., Williams et al. 2016; Stensvold and Farrar 2017) despite its constraints, and several measures of genetic variation discussed above, including %P, Wright's *F*-statistics (F and F_{ST}), and measures of heterozygosity (e.g., H_e), can easily be calculated using allozymes.

In the mid-1980s and early 1990s, DNA fingerprinting techniques, including restriction fragment length polymorphisms (RFLPs), amplified fragment length polymorphisms, and microsatellites, gained traction. In contrast with allozyme analyses, these processes directly compare DNA sequences, but large fern genomes can make the application of DNA fingerprinting difficult, as homology is harder to assess (especially in RFLPs). Although the relative popularity of allozymes in population genetics had largely diminished by the early 1990s (Schlötterer 2004), higher-resolution DNA-based markers such as microsatellites were not applied in ferns until the early 2000s (Pryor et al. 2001). Microsatellites have gradually displaced allozymes and have emerged as the marker of choice in plant (including fern) population genetics (fig. 2). Unlike other DNA fingerprinting techniques, microsatellite analyses are highly informative, as these markers are codominant and the methods are reproducible (Vieira et al. 2016). Despite their utility, however, multilocus genotyping of the hundreds of samples required for population genetic studies using microsatellites can be time-consuming and labor-intensive.

The advent of high-throughput DNA sequencing has greatly increased our knowledge of land plant evolution, extending our sequencing capabilities from single genes (Chase et al. 1993) to thousands (One Thousand Plant Transcriptomes Initiative 2019). Individual genes may contain hundreds of characters for comparison and have been used extensively in fern systematics. Both phylogenetic and population genetic studies in plants have chiefly relied on sequencing genes from the chloroplast genome, which are uniparentally inherited organelles that may act as a single locus (Lynch 2007; but see Gonçalves et al. 2019 for a discussion of the varying evolutionary histories of plastid loci), therefore capturing only part of the evolutionary history. The nuclear genome poses a more promising, albeit complex, landscape with which to tackle evolutionary processes. The decreasing cost of high-throughput sequencing has largely advanced us from the field of population genetics to that of population genomics. Transcriptomic and genomic data can be used to

analyze thousands of loci at a relatively low cost and allow for powerful new types of analyses. Unlike earlier methods, markers developed from this method do not necessarily rely on large stretches of DNA but rather focus on single nucleotides. Substitutions at a single base in the genome that occur at a frequency of greater than 1% in a population are classified as single-nucleotide polymorphisms (SNPs), and analyses of these markers have revolutionized the field of population genetics. Yet, to the best of our knowledge, only two publications (Rowe et al. 2018; Wang et al. 2020) have used this methodology to investigate fern population genetics using reduced representation sequencing (discussed below), although this technique has been used in various systematic contexts (e.g., Massatti et al. 2016; Fitz-Gibbon et al. 2017; Kinosian et al. 2019). The lack of genomic data due to the large genomes of ferns has largely impeded the advancement of this field. However, with several new candidate taxa for whole-genome sequencing recently identified, many families are likely to gain genomic resources in the near future (Kuo and Li 2019). With this work anticipated, we foresee great strides in fern population genetics and genomics and make the following recommendations to facilitate this process.

Future directions. We emphasize the need to transition from allozymes to large-scale genomic data. The advent of high-throughput sequencing technologies has drastically reduced the expense of acquiring highly informative genomic data. Fern genomes are exceedingly large, with a mean 1C size of 12 Gb (range, 0.25–148 Gb; Sessa and Der 2016; Li et al. 2018), and they contain large numbers of highly repetitive sequences (Wolf et al. 2015; Li et al. 2018; Marchant et al. 2019), making assembly from short-read data (e.g., 454 and Illumina) difficult. Advances in long-read sequencing technologies such as Oxford Nanopore and PacBio have improved our ability to assemble homosporous fern genomes, as long-read technology can contend better with repeat content (Marchant et al. 2019). In many cases, a hybrid approach using both long- and short-read data is required and enhances the resulting assembly quality (Zimin et al. 2017). Sequencing haploid gametophytes in place of diploid (or polyploid) sporophytes or sequencing completely homozygous sporophytes produced from gametophytic selfing is another avenue available for ferns that may reduce problems during genome assembly by reducing or eliminating concerns about heterozygosity in particular. Resequencing efforts will greatly enhance our ability to detect fine-scale population genetic variation and microevolutionary processes. Although the initial reference should have relatively deep coverage, resequencing coverage can be as low as 5–10 × coverage (Ellegren 2014) in homozygous or haploid individuals, although confident SNP calling in highly heterozygous genomes requires higher coverage, at least 13–15 × (Meynert et al. 2013; Song et al. 2016).

While whole-genome sequencing and resequencing can be expensive as well as computation and time intensive, reduced representation approaches greatly decrease the cost and amount of data for assembly while theoretically retaining homologous regions for comparison. Transcriptome sequencing, restriction site–associated DNA sequencing (RAD-seq), and sequence capture methods are all viable alternatives to whole-genome sequencing for population genetic studies and can provide hundreds to thousands of loci for downstream analyses at a relatively low cost (da Fonseca et al. 2016). For questions focused on selection or when the presence of selection is not an issue, transcriptomic data, which include primarily expressed exonic and regulatory regions, should be satisfactory. RAD-seq may be preferable when addressing questions that rely on (nearly) neutral variation in the genome, that is, markers not under selective pressure (Holderegger et al. 2006). RAD-seq, however, may not be suitable for taxa with especially large genomes, as the fragments that are sequenced may not be homologous (but see Weisrock et al. 2018 for work with RAD-seq in large salamander genomes). Other RAD-related methods that incorporate sequence capture, such as RADcap (Ali et al. 2016; Hoffberg et al. 2016), may be more appropriate for ferns.

As an alternative to transcriptome sequencing and RAD approaches, sequence capture techniques that target specific loci have proven to be useful for studies of both loci under selection (George et al. 2011; Aagaard et al. 2013; Christmas et al. 2016) and those that require neutral markers (e.g., Christmas et al. 2017; de La Harpe et al. 2019). Sequence capture methods have also been successful in recovering high-quality genomic data from herbarium collections (Hart et al. 2016). These "genomic treasure troves" (Staats et al. 2013) are invaluable resources for analyzing genetic diversity over time. Given the recent release of effective probe sets designed for reconstructing fern phylogenies (Wolf et al. 2018; Breinholt et al. 2021) and the development of baits for population genetic studies (e.g., Christmas et al. 2017), targeted sequence capture will almost certainly be increasingly used in a variety of fern evolutionary studies, and we encourage its continued development for population genetics in particular.

Using different molecular markers (e.g., allozymes, microsatellites, SNPs) may impact calculations for metrics of genetic variation (e.g., Russell et al. 1997; Sun et al. 2004; Lemopoulos et al. 2019; but see Coates et al. [2009], who found no differences between microsatellites and SNPs in humans). Microsatellites are not under strong purifying selection and evolve at faster rates than do allozymes and (potentially) SNPs. The extent to which these methods cover the genome also differ, with allozymes restricted to coding regions, while microsatellites and SNPs can be found throughout the genome. As reviewed by Ellegren (2014) and Welles and Dlugosch (2018), genome-wide SNP analyses have been instrumental in inferring demographic changes (e.g., Barker et al. 2017) and admixture/hybridization (e.g., Ma et al. 2019), loci under selection (e.g., Christmas et al. 2016) or associated with certain phenotypes (e.g., Prince et al. 2017), and the prevalence of mutations (e.g., Exposito-Alonso et al. 2018). The identification of SNPs in functional loci (e.g., promoters, genes), such as through Gene Ontology enrichment tests, has become popular in population genomics. Genome-wide markers such as SNPs provide valuable information about how different parts of the genome evolve, such as which parts might be the product of introgression and/or adaptive evolution (e.g., Edelman et al. 2019). Using a mix of genome-wide markers subject to either neutral or adaptive evolution, researchers are able to construct a more complete picture of the processes shaping population and genomic diversity.

Conclusions

The field of fern biology has seen remarkable advances in the past few decades. Most recently, the publication of the first fern genomes has greatly enhanced our knowledge of land plant

evolution both across major plant lineages and within ferns. While the field of fern population genetics has developed relatively slowly, data compiled from publications over the past 35 years reveal great strides in this area. Contrary to initial hypotheses in the 1960s and 1970s, natural fern populations are not restricted to gametophytic selfing and instead regularly outcross. We found that mating system has a significant impact on the genetic diversity and population structuring of fern populations, with predominantly outcrossing taxa having the greatest genetic diversity and least structure. We found that the percent of polymorphic loci, a measure of genetic diversity, and the fixation index are both significantly different among ferns with different mating systems. We also investigated the potential role of growth habit in shaping the genetic diversity and microevolutionary patterns of ferns and found that epiphytic taxa may have higher rates of gene flow and genetic diversity than other growth habits. Our compiled data set revealed a clear taxonomic bias in population genetic studies toward five of the 48 fern families and a geographic tendency toward the study of temperate taxa found in North America, Europe, and eastern Asia, and we strongly recommend that future research focus on addressing these gaps and biases. We have suggested several fern lineages that should pose fruitful avenues for future work and methods that have been employed successfully in other plant and animal lineages. Fern population genetics sits today on a solid foundation of theory and data, but much is still left unexplored, and we are excited to see the launching of the era of fern population genomics.

Acknowledgments

We are grateful to Cody Howard, Lindsey Riibe, Weston Testo, and three anonymous reviewers for comments and suggestions on an earlier version of the article.

Literature Cited

Aagaard JE, RD George, L Fishman, MJ Maccoss, WJ Swanson 2013 Selection on plant male function genes identifies candidates for reproductive isolation of yellow monkeyflowers. PLoS Genet 9: e1003965.

Ali OA, SM O'Rourke, SJ Amish, MH Meek, G Luikart, C Jeffres, MR Miller 2016 RAD capture (Rapture): flexible and efficient sequence-based genotyping. Genetics 202:389–400.

Arabidopsis Genome Initiative 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815.

Barker BS, K Andonian, SM Swope, DG Luster, KM Dlugosch 2017 Population genomic analyses reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (*Centaurea solstitialis*). Mol Ecol 26:1131–1147.

Barringer BC 2007 Polyploidy and self-fertilization in flowering plants. Am J Bot 94:1527–1533.

Bierhorst DW 1958 Observations on the gametophytes of *Botrychium virginianum* and *B. dissectum*. Am J Bot 45:1–9.

Bouchard JR, DD Fernando, SW Bailey, J Weber-Townsend, DJ Leopold 2017 Contrasting patterns of genetic variation in central and peripheral populations of *Dryopteris fragrans* (fragrant wood fern) and implications for colonization dynamics and conservation. Int J Plant Sci 178:607–617.

Breinholt JW, SB Carey, GP Tiley, EC Davis, L Endara, SF McDaniel, LG Neves, et al 2021 A target enrichment probe set for resolving the flagellate land plant tree of life. Appl Plant Sci 9:e11406.

Brown AHD 1979 Enzyme polymorphism in plant populations. Theor Popul Biol 15:1–42.

Bucharová A, Z Münzbergová 2012 Gene flow among populations of two rare co-occurring fern species differing in ploidy level. PLoS ONE 7:e45855.

Camacho FJ, A Liston 2001 Population structure and genetic diversity of *Botrychium pumicola* (Ophioglossaceae) based on inter-simple sequence repeats (ISSR). Am J Bot 88:1065–1070.

Cardelús CL 2010 Litter decomposition within the canopy and forest floor of three tree species in a tropical lowland rain forest, Costa Rica. Biotropica 42:300–308.

Cardelús CL, RK Colwell, JE Watkins 2006 Vascular epiphyte distribution patterns: explaining the mid-elevation richness peak. J Ecol 94:144–156.

Cardelús CL, MC Mack 2010 The nutrient status of epiphytes and their host trees along an elevational gradient in Costa Rica. Plant Ecol 207:25–37.

Chambers SM, NC Emery 2016 Population differentiation and countergradient variation throughout the geographic range in the fern gametophyte *Vittaria appalachiana*. Am J Bot 103:86–98.

Chase MW, DE Soltis, RG Olmstead, D Morgan, DH Les, BD Mishler, MR Duvall, et al 1993 Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. Ann Mo Bot Gard 80:528.

Christmas MJ, E Biffin, MF Breed, AJ Lowe 2016 Finding needles in a genomic haystack: targeted capture identifies clear signatures of selection in a nonmodel plant species. Mol Ecol 25:4216–4233.

——— 2017 Targeted capture to assess neutral genomic variation in the narrow-leaf hopbush across a continental biodiversity refugium. Sci Rep 7:41367.

Chung MY, J López-Pujol, JM Chung, MO Moon, MG Chung 2013 Genetic diversity in the homosporous fern *Ophioglossum vulgatum* (Ophioglossaceae) from South Korea: inference of mating system and population history. J Hered 104:263–272.

Chung MY, JD Nason, B-Y Sun, M-O Moon, JM Chung, C-W Park, MG Chung 2010 Extremely low levels of genetic variation in the critically endangered monotypic fern genus *Mankyua chejuense* (Ophioglossaceae) from Korea: implications for conservation. Biochem Syst Ecol 38:888–896.

Clark J, O Hidalgo, J Pellicer, H Liu, J Marquardt, Y Robert, M Christenhusz, et al 2016 Genome evolution of ferns: evidence for relative stasis of genome size across the fern phylogeny. New Phytol 210:1072–1082.

Coates BS, DV Sumerford, NJ Miller, KS Kim, TW Sappington, BD Siegfried, LC Lewis 2009 Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. J Hered 100:556–564.

Coomes DA, RB Allen, WA Bentley, LE Burrows, CD Canham, L Fagan, DM Forsyth, et al 2005 The hare, the tortoise and the crocodile: the ecology of angiosperm dominance, conifer persistence and fern filtering. J Ecol 93:918–935.

Cousens MI 1979 Gametophyte ontogeny, sex expression, and genetic load as measures of population divergence in *Blechnum spicant*. Am J Bot 66:116–132.

da Fonseca RR, A Albrechtsen, GE Themudo, J Ramos-Madrigal, JA Sibbesen, L Maretty, ML Zepeda-Mendoza, PF Campos, R Heller, RJ Pereira 2016 Next-generation biology: sequencing and data analysis approaches for non-model organisms. Mar Genomics 30:3–13.

Dassler CL, DR Farrar 2001 Significance of gametophyte form in long-distance colonization by tropical, epiphytic ferns. Brittonia 53:352–369.

De Groot GA, HJ During, SW Ansell, H Schneider, P Bremer, ERJ Wubs, JW Maas, H Korpelainen, RHJ Erkens 2012 Diverse spore rains and limited local exchange shape fern genetic diversity in a recently created habitat colonized by long-distance dispersal. Ann Bot 109:965–978.

de La Harpe M, J Hess, O Loiseau, N Salamin, C Lexer, M Paris 2019 A dedicated target capture approach reveals variable genetic markers across micro- and macro-evolutionary time scales in palms. Mol Ecol Resour 19:221–234.

Döpp W 1950 Eine die antheridienbildung bei farnen fordernde substanz in den Prothallien von *Pteridium aquilinum* (L.) Kuhn. Ber Dtsch Bot Ges 63:139.

Ebihara A, A Yamaoka, N Mizukami, A Sakoda, JH Nitta, R Imaichi 2013 A survey of the fern gametophyte flora of Japan: frequent independent occurrences of noncordiform gametophytes. Am J Bot 100:735–743.

Edelman NB, PB Frandsen, M Miyagi, B Clavijo, J Davey, RB Dikow, G García-Accinelli, et al 2019 Genomic architecture and introgression shape a butterfly radiation. Science 366:594–599.

Ellegren H 2014 Genome sequencing and population genomics in non-model organisms. Trends Ecol Evol 29:51–63.

Ellegren H, N Galtier 2016 Determinants of genetic diversity. Nat Rev Genet 17:422–433.

Ellstrand NC 2014 Is gene flow the most important evolutionary force in plants? Am J Bot 101:737–753.

Exposito-Alonso M, C Becker, VJ Schuenemann, E Reiter, C Setzer, R Slovak, B Brachi, et al 2018 The rate and potential relevance of new mutations in a colonizing plant lineage. PLoS Genet 14:e1007155.

Farrar DR 1967 Gametophytes of four tropical fern genera reproducing independently of their sporophytes in the southern Appalachians. Science 155:1266–1267.

——— 1990 Species and evolution in asexually reproducing independent fern gametophytes. Syst Bot 15:98–111.

Fernando DD, JJ Discenza, JR Bouchard, DJ Leopold 2015 Genetic analysis of the threatened American hart's-tongue fern (*Asplenium scolopendrium* var. *americanum* [Fernald] Kartesz and Gandhi): insights into its mating system and implications for conservation. Biochem Syst Ecol 62:25–35.

Fitz-Gibbon S, AL Hipp, KK Pham, PS Manos, VL Sork 2017 Phylogenomic inferences from reference-mapped and de novo assembled short-read sequence data using RADseq sequencing of California white oaks (*Quercus* section *Quercus*). Genome 60:743–755.

Gastony GJ, LD Gottlieb 1985 Genetic variation in the homosporous fern *Pellaea andromedifolia*. Am J Bot 72:257.

Gaxiola A, LE Burrows, DA Coomes 2008 Tree fern trunks facilitate seedling regeneration in a productive lowland temperate rain forest. Oecologia 155:325–335.

George LO, FA Bazzaz 1999*a* The fern understory as an ecological filter: emergence and establishment of canopy-tree seedlings. Ecology 80:833–845.

——— 1999*b* The fern understory as an ecological filter: growth and survival of canopy-tree seedlings. Ecology 80:846–856.

George RD, G McVicker, R Diederich, SB Ng, AP MacKenzie, WJ Swanson, J Shendure, JH Thomas 2011 Trans genomic capture and sequencing of primate exomes reveals new targets of positive selection. Genome Res 21:1686–1694.

Gonçalves DJP, BB Simpson, EM Ortiz, GH Shimizu, RK Jansen 2019 Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. Mol Phylogenet Evol 138:219–232.

Hamilton RG, RM Lloyd 1991 Antheridiogen in the wild: the development of fern gametophyte communities. Funct Ecol 5:804.

Hamrick JL, MJW Godt 1990 Allozyme diversity in plant species. Pages 43–63 *in* AHD Brown, MT Clegg, AL Kahler, BS Weir, eds. Plant population genetics, breeding, and genetic resources. Sinauer, Sunderland, MA.

——— 1996 Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc B 351:1291–1298.

Hart ML, LL Forrest, JA Nicholls, CA Kidner 2016 Retrieval of hundreds of nuclear loci from herbarium specimens. Taxon 65:1081–1092.

Hartl DL, AG Clark 1989 Principles of population genetics. 2nd ed. Sinauer, Sunderland, MA.

Haufler CH 1985 Enzyme variability and modes of evolution in *Bommeria* (Pteridaceae). Syst Bot 10:92.

——— 2002 Homospory 2002: an odyssey of progress in pteridophyte genetics and evolutionary biology. Bioscience 52:1081.

Haufler CH, KM Pryer, E Schuettpelz, EB Sessa, DR Farrar, R Moran, JJ Schneller, JE Watkins, MD Windham 2016 Sex and the single gametophyte: revising the homosporous vascular plant life cycle in light of contemporary research. Bioscience 66:928–937.

Hauk WD, CH Haufler 1999 Isozyme variability among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). Am J Bot 86:614–633.

Hedrick PW 1987 Genetic load and the mating system in homosporous ferns. Evolution 41:1282–1289.

Hickok LG 1978 Homoeologous chromosome pairing and restricted segregation in the fern *Ceratopteris*. Am J Bot 65:516.

Hoffberg SL, TJ Kieran, JM Catchen, A Devault, BC Faircloth, R Mauricio, TC Glenn 2016 RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. Mol Ecol Resour 16:1264–1278.

Holderegger R, U Kamm, F Gugerli 2006 Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landsc Ecol 21:797–807.

Holm S 1979 A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70.

Holsinger KE 1987 Gametophytic self-fertilization in homosporous plants: development, evaluation, and application of a statistical method for evaluating its importance. Am J Bot 74:1173.

Hooper E, C Haufler 1997 Genetic diversity and breeding system in a group of Neotropical epiphytic ferns (*Pleopeltis*; Polypodiaceae). Am J Bot 84:1664.

Hornych O, WL Testo, EB Sessa, JE Watkins, CE Campany, J Pittermann, L Ekrt 2021 Insights into the evolutionary history and widespread occurrence of antheridiogen systems in ferns. New Phytol 229:607–619.

Hunt HV, SW Ansell, SJ Russell, H Schneider, JC Vogel 2009 Genetic diversity and phylogeography in two diploid ferns, *Asplenium fontanum* subsp. *fontanum* and *A. petrarchae* subsp. *bivalens*, in the western Mediterranean. Mol Ecol 18:4940–4954.

Kang M, H Huang, M Jiang, AJ Lowe 2008 Understanding population structure and historical demography in a conservation context: population genetics of an endangered fern. Divers Distrib 14:799–807.

Kinosian SP, WD Pearse, PG Wolf 2020 Cryptic diversity in the model fern genus *Ceratopteris* (Pteridaceae). Mol Phylogenet Evol 152:106938.

Kinosian SP, WL Testo, SM Chambers, EB Sessa 2019 Using RAD data to confirm parentage of polyploids in a reticulate complex of ferns. Am Fern J 109:267.

Klekowski EJ 1969 Reproductive biology of the Pteridophyta. II. Theoretical considerations. Bot J Linn Soc 62:347–359.

——— 1970 Populational and genetic studies of a homosporous fern—*Osmunda regalis*. Am J Bot 57:1122–1138.

——— 1972 Genetical features of ferns as contrasted to seed plants. Ann Mo Bot Gard 59:138.

——— 1979 The genetics and reproductive biology of ferns. Pages 133–170 *in* AF Dyer, ed. The experimental biology of ferns. Academic Press, London.

Klekowski EJ, HG Baker 1966 Evolutionary significance of polyploidy in the Pteridophyta. Science 153:305–307.

Koop AL 2009 *Lygodium microphyllum* (Old World climbing fern), *Lygodium japonicum* (Japanese climbing fern), and *Lygodium flexuosum*: weed risk assessment. A Jackson, ed. USDA, Raleigh, NC.

Koutika L-S, HJ Rainey 2015 A review of the invasive, biological and beneficial characteristics of aquatic species *Eichhornia crassipes* and *Salvinia molesta*. Appl Ecol Environ Res 13:263–275.

Kuo L-Y, F-W Li 2019 A roadmap for fern genome sequencing. Am Fern J 109:212.

Lande R, DW Schemske 1985 The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24–40.

Lemopoulos A, JM Prokkola, S Uusi-Heikkilä, A Vasemägi, A Huusko, P Hyvärinen, M-L Koljonen, J Koskiniemi, A Vainikka 2019 Comparing RADseq and microsatellites for estimating genetic diversity and relatedness: implications for brown trout conservation. Ecol Evol 9:2106–2120.

Li F-W, P Brouwer, L Carretero-Paulet, S Cheng, J de Vries, P-M Delaux, A Eily, et al 2018 Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants 4:460–472.

Liu X-Q, RW Gituru, LQ Chen 2007 Genetic variation in the endangered fern *Adiantum reniforme* var. *sinense*. Ann Bot Fenn 44:25–32.

Lumpkin TA, DL Plucknett 1980 *Azolla*: botany, physiology, and use as a green manure. Econ Bot 34:111–153.

Lynch M 2007 The origins of genome architecture. Sinauer, Sunderland, MA.

Ma Y, J Wang, Q Hu, J Li, Y Sun, L Zhang, RJ Abbott, J Liu, K Mao 2019 Ancient introgression drives adaptation to cooler and drier mountain habitats in a cypress species complex. Comm Biol 2:213.

Maki M, Y Asada 1998 High genetic variability revealed by allozymic loci in the narrow endemic fern *Polystichum otomasui* (Dryopteridaceae). Heredity 80:604–610.

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

Marchant DB, EB Sessa, PG Wolf, K Heo, WB Barbazuk, PS Soltis, DE Soltis 2019 The C-fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. Sci Rep 9:18181.

Massatti R, AA Reznicek, LL Knowles 2016 Utilizing RADseq data for phylogenetic analysis of challenging taxonomic groups: a case study in *Carex* sect. *Racemosae*. Am J Bot 103:337–347.

Meirmans PG, S Liu, PH van Tienderen 2018 The analysis of polyploid genetic data. J Hered 109:283–296.

Meynert AM, LS Bicknell, ME Hurles, AP Jackson, MS Taylor 2013 Quantifying single nucleotide variant detection sensitivity in exome sequencing. BMC Bioinformatics 14:195.

Moran RC 2008 Diversity, biogeography, and floristics. Pages 367–394 *in* TA Ranker, CH Haufler, eds. Biology and evolution of ferns and lycophytes. Cambridge University Press, Cambridge.

Moran RC, AR Smith 2001 Phytogeographic relationships between Neotropical and African-Madagascan pteridophytes. Brittonia 53: 304–351.

Morero RE, R Deanna, GE Barboza, DS Barrington 2019 Historical biogeography of the fern genus *Polystichum* (Dryopteridaceae) in Austral South America. Mol Phylogenet Evol 137:168–189.

Nadkarni NM, D Schaefer, TJ Matelson, R Solano 2004 Biomass and nutrient pools of canopy and terrestrial components in a primary and a secondary montane cloud forest, Costa Rica. For Ecol Manag 198:223–236.

Nitta JH, J-Y Meyer, R Taputuarai, CC Davis 2016 Life cycle matters: DNA barcoding reveals contrasting community structure between fern sporophytes and gametophytes. Ecol Monogr 87:278–296.

Nybom H 2004 Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13: 1143–1155.

One Thousand Plant Transcriptomes Initiative 2019 One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574:679–685.

Pabby A, R Prasanna, PK Singh 2004 Biological significance of *Azolla* and its utilization in agriculture. Proc Indian Natl Sci Acad B70:299–333.

Pemberton RW, AP Ferriter 1998 Old World climbing fern (*Lygodium microphyllum*), a dangerous invasive weed in Florida. Am Fern J 88:165.

Pinson JB, SM Chambers, JH Nitta, L-Y Kuo, EB Sessa 2017 The separation of generations: biology and biogeography of long-lived sporophyteless fern gametophytes. Int J Plant Sci 178:1–18.

Pinson JB, E Schuettpelz 2016 Unraveling the origin of the Appalachian gametophyte, *Vittaria appalachiana*. Am J Bot 103:668–676.

PPG I (Pteridophyte Phylogeny Group I) 2016 A community-derived classification for extant lycophytes and ferns. J Syst Evol 54:563–603.

Prince DJ, SM O'Rourke, TQ Thompson, OA Ali, HS Lyman, IK Saglam, TJ Hotaling, AP Spidle, MR Miller 2017 The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. Sci Adv 3:e1603198.

Pryor KV, JE Young, FJ Rumsey, KJ Edwards, MW Bruford, HJ Rogers 2001 Diversity, genetic structure and evidence of outcrossing in British populations of the rock fern *Adiantum capillus-veneris* using microsatellites. Mol Ecol 10:1881–1894.

Quintanilla LG, E Pangua, J Amigo, S Pajarón 2005 Comparative study of the sympatric ferns *Culcita macrocarpa* and *Woodwardia radicans*: sexual phenotype. Flora 200:187–194.

Ranker TA 1992*a* Genetic diversity, mating systems, and interpopulation gene flow in Neotropical *Hemionitis palmata* L. (Adiantaceae). Heredity 69:175–183.

——— 1992*b* Genetic diversity of endemic Hawaiian epiphytic ferns: implications for conservation. Selbyana 13:131–137.

——— 1994 Evolution of high genetic variability in the rare Hawaiian fern *Adenophorus periens* and implications for conservation management. Biol Conserv 70:19–24.

Ranker TA, SK Floyd, MD Windham, PG Trapp 1994 Historical biogeography of *Asplenium adiantum-nigrum* (Aspleniaceae) in North America and implications for speciation theory in homosporous pteridophytes. Am J Bot 81:776–781.

Ranker TA, JMO Geiger 2008 Population genetics. Pages 107–133 *in* TA Ranker, CH Haufler, eds. Biology and evolution of ferns and lycophytes. Cambridge University Press, Cambridge.

R Core Team 2019 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Rowe CA, DP Hauber, PG Wolf 2018 Genomic variation of introduced *Salvinia minima* in southeastern United States. Aquat Bot 151:38–42.

Rumsey FJ, JC Vogel, SJ Russell, JA Barrett, M Gibby 1999 Population structure and conservation biology of the endangered fern *Trichomanes speciosum* Willd. (Hymenophyllaceae) at its northern distributional limit. Biol J Linn Soc 66:333–344.

Russell AE, JW Raich, PM Vitousek 1998 The ecology of the climbing fern *Dicranopteris linearis* on windward Mauna Loa, Hawaii. J Ecol 86:765–779.

Russell JR, JD Fuller, M Macaulay, BG Hatz, A Jahoor, W Powell, R Waugh 1997 Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor Appl Genet 95:714–722.

Schlötterer C 2004 The evolution of molecular markers—just a matter of fashion? Nat Rev Genet 5:63–69.

Schneider H, E Schuettpelz, KM Pryer, R Cranfill, S Magallón, R Lupia 2004 Ferns diversified in the shadow of angiosperms. Nature 428:553–557.

Schneller JJ 1988 Spore bank, dark germination and gender determination in *Athyrium* and *Dryopteris*: results and implications for population biology of Pteridophyta. Bot Helv 98:77–86.

——— 2008 Antheridiogens. Pages 134–158 *in* TA Ranker, CH Haufler, eds. Biology of ferns and lycophytes. Cambridge University Press, Cambridge.

Schneller JJ, R Holderegger, F Gugerli, K Eichenberger, E Lutz 1998 Patterns of genetic variation detected by RAPDs suggest a single origin with subsequent mutations and long-distance dispersal in the apomictic fern *Dryopteris remota* (Dryopteridaceae). Am J Bot 85:1038.

Schuettpelz E, KM Pryer 2009 Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. Proc Natl Acad Sci USA 106:11200–11205.

Sessa EB, JA Banks, MS Barker, JP Der, AM Duffy, SW Graham, M Hasebe, et al 2014 Between two fern genomes. Gigascience 3:15.

Sessa EB, J Der 2016 Evolutionary genomics of ferns and lycophytes. Pages 215–254 *in* SA Rensing, ed. Genomes and evolution of charophytes, bryophytes, and ferns. Vol 78. Advances in Botanical Research. Elsevier, San Diego, CA.

Sessa EB, A Juslén, H Väre, SM Chambers 2017 Into Africa: molecular phylogenetics and historical biogeography of sub-Saharan African woodferns (*Dryopteris*). Am J Bot 104:477–486.

Sessa EB, WL Testo, JE Watkins 2016 On the widespread capacity for, and functional significance of, extreme inbreeding in ferns. New Phytol 211:1108–1119.

Slatkin M, NH Barton 1989 A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349–1368.

Soltis DE, PS Soltis 1986 Electrophoretic evidence for inbreeding in the fern *Botrychium virginianum* (Ophioglossaceae). Am J Bot 73: 588.

——— 1987*a* Breeding system of the fern *Dryopteris expansa*: evidence for mixed mating. Am J Bot 74:504–509.

——— 1987*b* Population structure and estimates of gene flow in the homosporous fern *Polystichum munitum*. Evolution 41:620–629.

——— 1990 Genetic variation within and among populations of ferns. Am Fern J 80:161.

——— 2000 The role of genetic and genomic attributes in the success of polyploids. Proc Natl Acad Sci USA 97:7051–7057.

Soltis DE, PS Soltis, P Endress, M Chase, S Manchester, W Judd, L Majure, E Mavrodiev 2017 Phylogeny and evolution of the angiosperms. University of Chicago Press, Chicago.

Soltis PS, DE Soltis, KE Holsinger 1988 Estimates of intragametophytic selfing and interpopulational gene flow in homosporous ferns. Am J Bot 75:1765.

Song K, L Li, G Zhang 2016 Coverage recommendation for genotyping analysis of highly heterologous species using next-generation sequencing technology. Sci Rep 6:35736.

Spieth PT 1974 Gene flow and genetic differentiation. Genetics 78: 961–965.

Staats M, RHJ Erkens, B van de Vossenberg, JJ Wieringa, K Kraaijeveld, B Stielow, J Geml, JE Richardson, FT Bakker 2013 Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. PLoS ONE 8:e69189.

Stensvold MC, DR Farrar 2017 Genetic diversity in the worldwide *Botrychium lunaria* (Ophioglossaceae) complex, with new species and new combinations. Brittonia 69:148–175.

Suissa JS, MA Sundue, WL Testo 2021 Mountains, climate and niche heterogeneity explain global patterns of fern diversity. J Biogeogr (forthcoming). https://doi.org/10.1111/jbi.14076.

Sun G, O Díaz, B Salomon, R Bothmer 2004 Microsatellite variation and its comparison with allozyme and RAPD variation in *Elymus fibrosus* (Schrenk) Tzvel. (Poaceae). Hereditas 129:275–282.

Suter M, JJ Schneller, JC Vogel 2000 Investigations into the genetic variation, population structure, and breeding systems of the fern *Asplenium trichomanes* subsp. *quadrivalens*. Int J Plant Sci 161: 233–244.

Testo W, M Sundue 2016 A 4000-species dataset provides new insight into the evolution of ferns. Mol Phylogenet Evol 105:200–211.

Tryon R 1970 Development and evolution of fern floras of oceanic islands. Biotropica 2:76–84.

——— 1986 The biogeography of species, with special reference to ferns. Bot Rev 52:117–156.

Vieira MLC, L Santini, AL Diniz, C de F Munhoz 2016 Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol 39:312–328.

Vitalis R, M Riba, B Colas, P Grillas, I Olivieri 2002 Multilocus genetic structure at contrasted spatial scales of the endangered water fern *Marsilea strigosa* Willd. (Marsileaceae, Pteridophyta). Am J Bot 89:1142–1155.

Vitousek P, GP Asner, OA Chadwick, S Hotchkiss 2009 Landscape-level variation in forest structure and biogeochemistry across a substrate age gradient in Hawaii. Ecology 90:3074–3086.

Vogel JC, FJ Rumsey, SJ Russell, CJ Cox, JS Holmes, W Bujnoch, C Stark, JA Barrett, M Gibby 1999 Genetic structure, reproductive biology and ecology of isolated populations of *Asplenium csikii* (Aspleniaceae, Pteridophyta). Heredity 83:604–612.

Walker LR 1994 Effects of fern thickets on woodland development on landslides in Puerto Rico. J Veg Sci 5:525–532.

Walker LR, FH Landau, E Velázquez, AB Shiels, AD Sparrow 2010 Early successional woody plants facilitate and ferns inhibit forest development on Puerto Rican landslides. J Ecol 98:625–635.

Wang J, S Dong, L Yang, AJ Harris, H Schneider, M Kang 2020 Allopolyploid speciation accompanied by gene flow in a tree fern. Mol Biol Evol 37:2487–2502.

Wang T, Y Su, Y Li 2012 Population genetic variation in the tree fern *Alsophila spinulosa* (Cyatheaceae): effects of reproductive strategy. PLoS ONE 7:e41780.

Watano Y, N Sahashi 1992 Predominant inbreeding and its genetic consequences in a homosporous fern genus, *Sceptridium* (Ophioglossaceae). Syst Bot 17:486.

Watkins JE, CL Cardelús 2012 Ferns in an angiosperm world: Cretaceous radiation into the epiphytic niche and diversification on the forest floor. Int J Plant Sci 173:695–710.

Watkins JE, M Mack, SS Mulkey 2007*a* Gametophyte ecology and demography of epiphytic and terrestrial tropical ferns. Am J Bot 94:701–708.

Watkins JE, M Mack, TR Sinclair, SS Mulkey 2007*b* Ecological and evolutionary consequences of desiccation tolerance in tropical fern gametophytes. New Phytol 176:708–717.

Watkins JE, PW Rundel, CL Cardelús 2007*c* The influence of life form on carbon and nitrogen relationships in tropical rainforest ferns. Oecologia 153:225–232.

Weisrock DW, PM Hime, SO Nunziata, KS Jones, MO Murphy, S Hotaling, JD Kratovil 2018 Surmounting the large-genome "problem" for genomic data generation in salamanders. Pages 115–142 *in* PA Hohenlohe, OP Rajora, eds. Population genomics. Springer, Cham, Switzerland.

Welles SR, KM Dlugosch 2018 Population genomics of colonization and invasion. Pages 655–683 *in* PA Hohenlohe, OP Rajora, eds. Population genomics. Springer, Cham, Switzerland.

Williams EW, DR Farrar, D Henson 2016 Cryptic speciation in allotetraploids: lessons from the *Botrychium matricariifolium* complex. Am J Bot 103:740–753.

Wolf PG, TA Robison, MG Johnson, MA Sundue, WL Testo, CJ Rothfels 2018 Target sequence capture of nuclear-encoded genes for phylogenetic analysis in ferns. Appl Plant Sci 6:e01148.

Wolf PG, H Schneider, TA Ranker 2001 Geographic distributions of homosporous ferns: does dispersal obscure evidence of vicariance? J Biogeogr 28:263–270.

Wolf PG, EB Sessa, DB Marchant, F-W Li, CJ Rothfels, EM Sigel, MA Gitzendanner, et al 2015 An exploration into fern genome space. Genome Biol Evol 7:2533–2544.

Wright S 1931 Evolution in Mendelian populations. Genetics 16:97–159.

——— 1951 The genetical structure of populations. Ann Eugen 15:323–354.

——— 1965 The interpretation of population structure by *F*-statistics with special regard to systems of mating. Evolution 19:395–420.

Zimin AV, D Puiu, M-C Luo, T Zhu, S Koren, G Marçais, JA Yorke, J Dvořák, SL Salzberg 2017 Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. Genome Res 27:787–792.