UNIVERSITÉ DE NEUCHÂTEL

Evolution of moonwort ferns (*Botrychium*, Ophioglossaceae) on local to global scales

Thèse présentée à la Faculté des sciences

Institut de biologie

Laboratoire de génétique évolutive

Université de Neuchâtel, Suisse

Pour l'obtention du grade de DOCTEUR ÈS SCIENCES

Par

Benjamin Dauphin

Présenté aux membres du jury de thèse:

P.D. Dr Grant Jason, directeur de thèse et président du jury

Prof. Daniel Croll, rapporteur

Prof. Donald Farrar, rapporteur

Prof. Felix Kessler, rapporteur

Dr Michael Kessler, examinateur

Prof. Carl Rothfels, examinateur

Soutenue le 17 octobre 2017

UNIVERSITÉ DE NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Benjamin DAUPHIN

Titre:

"Evolution of moonwort ferns (*Botrychium*, Ophioglossaceae) on local to global scales"

sur le rapport des membres du jury composé comme suit:

- MER Jason Grant, directeur de thèse, Université de Neuchâtel
- Prof. Daniel Croll, Université de Neuchâtel
- Prof. Donald R. Farrar, Iowa State University, USA
- Prof. Felix Kessler, Université de Neuchâtel
- Dr Michael Kessler, Universität Zürich
- Prof. Carl Rothfels, University of California, Berkeley, USA

Neuchâtel, le 9 novembre 2017

Le Doyen, Prof. R. Bshary

«Fais de ta vie un rêve, et d'un rêve, une réalité»

Antoine de Saint-Exupéry (1900–1944)

Acknowledgments

This PhD was an intense and marvelous life experience for me. I am grateful to my spouse Lisa who has provided me emotional support and full confidence during the completion of this dissertation. A very special gratitude goes out to my parents Marie-Christine and Daniel who encouraged and supported me during my university studies, especially at a time when I had other interests than science.

I thank my friends for your unconditional support despite the strange world in which I work.

For three years, I was lucky to work in such a scientific environment and lead a project that really meant a lot to me. Several people and institutions have contributed greatly in the success of this project and I would like to express my sincerely gratitude in the simplest words.

I warmly thank you Jason for giving me confidence to carry out this project and introducing me to the fern world. Also, thanks a lot for your patience. I know, sometimes it was demanding…

Thank you Don for everything!

Carl, Berkeley is a captivating city, and I am lucky to have studied there on the fascinating topic of polyploidy. Many thanks for your hospitality in California and your precious collaboration, especially your tricks and secrets about the PacBio seq and the use of PURC.

I am also grateful to members of the Evolutionary Genetics lab at the University of Neuchâtel, the Genetic Diversity Center from ETH Zürich, and the Rothfels lab at the University of California at Berkeley, with a special mention to Olivier Bachmann, Forrest Freund, Stella Huynh, Ingrid Jordon-Thaden, Silvia Kobel, Aria Minder, Vinciane Mossion, Frédéric Sandoz, Julien Vieu, Guillaume Wos, and Niklaus Zemp. It was fantastic to work with each of them, especially on NGS data!

Without funding, we can't go very far in academic research. Therefore, I thank the Swiss National Science Foundation, the University of Neuchâtel, and the University of California at Berkeley for supporting this thesis.

Thanks to everyone for your positive spirit!

Abstract

Plants have long fascinated biologists by their ability to form cryptic species, hybrids between distinct species, large and stable genome sizes, and a wide variety of mating systems among taxa that have driven to the current plant biodiversity. More than 450 million year ago, vascular plants emerged and colonized the land. Rapidly, they dispersed to explore new habitats and diversified in suitable environments to form over 300,000 extant species. A better understanding of past evolutionary forces that led to the current plant morphologies, ecologies, and genetic diversity is critical for predicting their evolution, especially with the global changes ongoing. In this context, the main goals of this thesis were to investigate the phylogenetic relationships among the early divergent *Botrychium* taxa and how allopolyploidy and the alternation of mating systems have led to the speciation of these species.

In a multidisciplinary framework, we combined phylogenetics with population genetics and flow cytometry to provide relevant cellular and molecular data for exploring key biological mechanisms taking place at local or global scale. As the backbone of our project, we reconstructed the plastid phylogeny of the genus *Botrychium* based on a worldwide sampling to investigate relationships among diploid taxa and the maternal origins of allopolyploids. As a second step, we estimated the genome sizes of *Botrychium* diploids and polyploids to study fluctuations of DNA amounts after allopolyploidization and between ploidy levels. Then, we applied PacBio sequencing to infer a nuclear phylogeny for revealing the evolutionary history for both, the maternal and the paternal lineages of allopolyploids, and drawing the timescale of their bi-parental inheritance based on divergence time estimates. Focusing at a local and regional scale, we genotyped *B. lunaria* populations in Alps using co-dominant allozymes and ddRADseq data to identify the predominant mode of reproduction as well as its demographic history during the last glacial maximum.

We found unexpected results and major discoveries for the understanding of the biology of early vascular plants. For the species diversity itself, we identified thirteen possibly new taxa and characterized the exceptional haplotype diversity occurring in the Lunaria complex. Also, we presented evidence for multiple origins of several polyploid taxa and highlighted incontestable cases of inter-continental dispersal from North America to Europe and Asia. With flow cytometry, we reported a new hexaploid in *Botrychium* and described different genome sizes between diploid species of the two major clades Lanceolatum and Lunaria. Besides, our results supported the genome size stability after allopolyploidization, therefore rejecting the scenario of genome downsizing widespread accepted for angiosperms. Probably the most striking outcome of our phylogenetic investigations is the recurrent allopolyploidy in that genus and the strong bias of parental donors in the formation of allopolyploid taxa. Furthermore, our divergence time estimates revealed the recent and rapid

speciation via allopolyploidy in the last two million years, which constitutes a first case of radiation in the old lineage of Ophioglossaceae. For our population genetics study, we found unprecedented genetic diversity within *B. lunaria* populations with a large number of heterozygotes that supports the outcrossing mating system. Thus, we presented the capabilities of dispersion and diversification of these *Botrychium* species to better understand the ancestral vascular plant mating system. Being a key element of this speciation model, we found a genetic signature indicative of a refuge for *B. lunaria* in the central Alps during the last glacial maximum, which has hosted individuals having a high allele richness that was secondarily dispersed after deglaciation with the maintenance of outcrossing in alpine grasslands.

Undeniably, the genus *Botrychium* offers a unique opportunity to address the role of allopolyploidy and the importance of alternation of mating systems in plant speciation. This work is intended to be the starting point for further studies in evolutionary biology that ultimately will provide a better understanding of the life style of these enigmatic fern species.

Key words: *Botrychium*; ddRAD sequencing; divergence time dating; ferns; flow cytometry; genome size; low-copy makers; mating system; molecular phylogeny; Ophioglossaceae; PacBio; polyploidy; polyploid network; population genetics; PURC; reticulate evolution.

Résumé

Les plantes ont depuis longtemps fasciné les biologistes par leur capacité à former des espèces cryptiques, des hybrides entre des espèces distinctes, de grandes tailles de génomes stables, et une large variété de systèmes reproducteurs parmi les taxons qui ont conduit à l'actuelle biodiversité des plantes. Il y a plus de 450 millions d'années, les plantes vasculaires ont émergé et colonisé la terre. Rapidement, elles se sont dispersées pour explorer de nouveaux habitats et se sont diversifiées dans des environnements favorables pour former plus de 300 000 espèces vivantes. Une meilleure compréhension des forces évolutives passées qui ont mené aux morphologies, aux écologies, et aux diversités génétiques actuelles des plantes est essentielle pour prédire leur évolution, particulièrement dans le contexte des changements globaux. Ainsi, les objectifs principaux de cette thèse consistaient à étudier les relations phylogénétiques parmi les taxons de *Botrychium* et comment l'allopolyploïdie et l'alternance des systèmes reproducteurs ont mené à la spéciation de ces espèces.

Dans un cadre multidisciplinaire, nous avons combiné la phylogénétique avec la génétique des populations et la cytométrie en flux pour produire des données cellulaires et moléculaires pertinentes pour l'exploration des mécanismes biologiques clefs prenant place à des échelles locales ou globales. Etant les fondements de notre projet, nous avons reconstruit la phylogénie chloroplastique du genre *Botrychium* basée sur un échantillonnage mondiale pour examiner les relations de parentés parmi les taxons diploïdes et les origines maternelles des allopolyploïdes. Dans une seconde étape, nous avons estimé les tailles de génome des *Botrychium* diploïdes et polyploïdes pour comparer les variations du contenu en ADN après l'allopolyploïdisation et entre les niveaux de ploïdie. Ensuite, nous avons utilisé le séquençage PacBio pour inférer une phylogénie moléculaire dans le but de révéler l'histoire évolutive des deux, les lignées maternelles et paternelles des allopolyploïdes, et d'établir l'échelle temporelle de leur héritage bi-parentale basée sur les estimations des temps de divergence. Travaillant à une échelle régionale et locale, nous avons génotypé des populations de *B. lunaria* des alpes en utilisant des variants enzymatiques co-dominants et des données de séquençage ddRAD pour caractériser le mode prédominant de reproduction ainsi que leur histoire démographique durant le dernier maximum glaciaire.

Nous avons trouvé des résultats inattendus et des découvertes majeures pour la compréhension de la biologie des plantes vasculaires précoces. Pour la diversité spécifique elle-même, nous avons identifié treize nouveaux taxons potentiels et caractérisé la diversité exceptionnelle d'haplotypes survenant dans le complexe d'espèce Lunaria. De plus, nous avons présenté des évidences pour l'origine multiple de plusieurs taxons allopolyploïdes et révélé des cas incontestables de dipersion intercontinentale d'amérique du nord vers l'europe et l'asie. Avec la cytométrie en flux, nous avons rapporté un nouveau taxa hexploïde chez *Botrychium* et avons décrit différentes tailles de génome

entre espèces diploïdes des deux clades majeurs, Lanceolatum et Lunaria. Par ailleurs, nos résultats ont supporté le scénario de stabilité des tailles de génome après l'allopolyploidization, rejetant ainsi l'hypothèse de la réduction des tailles de génome largement acceptée chez les angiospermes. Le résultat le plus frappant dans nos investigations phylogénétiques est probablement la récurrence de l'allopolyploïdie dans ce genre et les contributions asymétriques des donneurs parentaux dans la formation des taxons allopolyploïdes. De plus, nos estimations du temps de divergence ont révélé la spéciation récente et rapide par allopolyploïdie au cours des deux derniers millions d'années, ce qui constitue un premier cas de radiation dans l'ancienne lignée des Ophioglossaceae. Pour notre étude sur la génétique des populations, nous avons trouvé une diversité génétique sans précédent dans les populations de *B. lunaria* avec un grand nombre d'hétérozygotes qui soutient la prédominance pour la fécondation croisée. Ainsi, nous avons présenté les capacités de dispersion et de diversification des espèces de *Botrychium* afin de mieux comprendre le système d'accouplement des plantes vasculaires ancestrales. En tant qu'élément clef de ce modèle de spéciation, nous avons trouvé une signature génétique indicative d'un refuge pour *B. lunaria* dans les alpes centrales lors du dernier maximum glaciaire, lequel a abrité des individus ayant une richesse allélique élevée, qui se sont ensuite dispersés durant la déglaciation tout en maintenant la fécondation croisée.

Indéniablement, le genre *Botrychium* offre une occasion unique d'aborder le rôle de l'allopolyploïdie et l'importance de l'alternance des systèmes d'accouplement dans la spéciation des plantes. Ce travail est destiné à être le point de départ pour d'autres études en biologie évolutive, lesquelles permettront peut-être de mieux comprendre le cycle de vie de ces espèces de fougères énigmatiques.

Mots clefs: *Botrychium*; séquençage ddRAD; datation des temps de divergence; fougères; cytométrie en flux; taille de génome; marqueurs en faible copie; système de reproduction; phylogénie moléculaire; Ophioglossaceae; PacBio; polyploïdie; réseau polyploïde; génétique des populations; PURC; évolution réticulée.

Table of contents

General introduction

The genetic diversity hidden within cryptic species

Morphological diversity does not systematically reflect genetic diversity or species richness within a group of organisms and therefore challenges the classification system. Testified by deep phylogenetic divergences, similar morphological plants may be unrelated even though botanists may consider them as a single taxon since no clear taxonomic characters were found to differentiate the species. Advances in Next-Generation Sequencing (NGS) technology and relatively inexpensive DNA sequencing have allowed the characterization of cryptic species containing genetically well-differentiated groups or populations (Bickford et al. 2007). The term "cryptic species" is defined as "two or more distinct species that are erroneously classified (and hidden) under one species name" (Bickford et al. 2007). This phenomenon is widespread to all organisms since it has been reported in animals (Pfenninger and Schwenk 2007), fungi (Crespo and Thorsten 2010), protists (Amato et al. 2007), and plants (Bickford et al. 2007). The discovery of cryptic species has several important implications for the understanding of biology of these organisms because ecological conditions and the belonging to a phylogenetic group would determine the frequency of such undetected diversity. In this framework, systematic studies are a powerful approach to reveal hidden species richness in any organism group and track key morphological characteristics for better taxonomic recognition. Similarly in ferns, three criteria are used to define a "cryptic species": "1) poor morphological differentiation, 2) reproductive isolation, and 3) misinterpretation of taxa as members of a single broader species" (Paris et al. 1989). Based on these considerations, the genus *Botrychium* is a classical example of cryptic species in ferns with its simple organization and subtle morphological traits that encompass a large species group containing unsuspected haplotype diversity (Dauphin et al. 2014).

Evolution of genome sizes among the earliest vascular plants

Plant genomes are known to be highly labile through evolutionary scales (Leitch et al. 2005). More than 450 million year ago, primitive genomes arose from a common ancestor and pursued different evolutionary trajectories with dramatic genome size reductions or expansions, or both, occurring multiple times, which has led to the current genome diversity. Among the early divergent vascular plants, species have a large range of DNA amount, with the smallest genome size of 0.086 pg found in lycophytes (*Selaginella*) and the among the biggest ones characterized in monilophytes (also named pteridophytes) with 1C = 72.68 pg for *Psilotum nudum* (Leitch and Leitch 2013). Fern genomes are particularly large and heterogeneous, especially in whisk ferns and ophioglossoids that have among the largest genomes in vascular plants. Interestingly, enormous genomes of those two

well-distinct lineages, sister to the Marattiidae-Polypodiidae clade (Rothfels et al. 2015), have probably evolved through two different cytological mechanisms. As representative of ophioglossoid ferns, *Ophioglossum* was characterized by a large number of small chromosomes (1.5–4.5 µm; Abraham et al. 1962), up to 2n = 1440 (Khandelwal 1990), which constitutes the highest chromosome number discovered in the plant kingdom. In contrast, the slightly bigger genomes of whisk ferns were reported with smaller karyotypes (2n = 208 for *Psilotum nudum*) but larger chromosomes (4.5–18 µm) (Brownsey and Lovis 1987) that are indicative of different cytological mechanisms in fern genomes evolution. Hence, how is such an amount of DNA maintained in a stable karyotype with as many chromosomes? What are the biological mechanisms that allow the genomes evolution?

Polyploidization as the fuel for speciation

Polyploid organisms occur in all major groups of animals, fungi, protists, and plants, and Whole-Genome Duplication (WGD) is recognized as a major process in eukaryote evolution. In fact, WGD events are at the origin of vertebrates (Dehal and Boore 2005) and all flowering plants have undergone one to several polyploid events in their evolutionary history (Jiao et al. 2011). Interestingly, independent WGD events have occurred simultaneously in most plant lineages during the last massive extinction, suggesting that polyploids may have evolutionary advantages to overcome crises (Fawcett et al. 2009). However, on a broader phylogenetic timescale, polyploids have shown higher extinction rates than diploids and recent studies have presented strong evidence in favor of the evolutionary dead end hypothesis of polyploids (Mayrose et al. 2011; Arrigo and Barker 2012).

Polyploid organisms result from one of the two main biological processes that are accompanied by WGD, either via autopolyploidy or allopolyploidy (Stebbins 1947). Autopolyploids are considered to have arisen from intraspecific hybridization by the doubling of similar homologous chromosome sets (AAAA), whereas allopolyploids have involved interspecific hybridization associated with the doubling of homeologous chromosome sets (AABB). This categorization is not intended to be limiting and it is now widely accepted that polyploidization occurred by the merging of more or less divergent original genomes along a continuum. In complement to this traditional view, a more complex model of polyploid formation has been proposed with multiple origins generating genetically variable polyploid lineages from distinct parental genotypes (Soltis and Soltis 1999). Illustrating the evolutionary implications on a short timescale, the allopolyploid *Tragopogon miscellus* has been independently formed at least 20 times during the past 70 years (Soltis and Soltis 1999), which provided the main source of the genetic diversity found among those allopolyploid lineages and perhaps the genetic basis supporting the success of allopolyploids (Soltis and Soltis 2000). Importantly, allopolyploidy is a drastic speciation process that allows the maintenance of high genetic variation via fixed

heterozygosity (Soltis and Soltis 2009) and the generation of evolutionary novelties through sub- and neo-functionalization arising from gene redundancy (Rastogi and Liberles 2005).

From molecular phylogeny to population genetics

Molecular phylogenetics is a powerful tool to investigate species relationships and build the most comprehensive view of the tree of life. It is the backbone for describing morphological and genetic diversity among related species, and testing many hypotheses in biology based on robust statistical frameworks. However, this approach is inapplicable to study the biological processes themselves taking place in natural populations, which initiate the micro-evolutionary changes that are span along the leaves of the tree of life. How can we fully understand and interpret the phylogenetic relationships of organisms without basis indications on their life cycles and the impact of environment or demographic history on mating systems?

Filling in this gap, population genetics is a promising approach to consider the role of the environment in reproductive isolation and differentiation among populations. Importantly, genetic diversity and gene flow can be investigated by genotyping data to identify the predominant mating system maintained in natural populations and compare the different strategies of life cycles among closely related species. Thus, the use of co-dominant nuclear markers allows the characterization of the heterozygote fraction and bi-parental inheritance of diploids and polyploids (Stensvold and Farrar 2017). Combined, molecular phylogenetics informs us about the demographic history within and among taxa that is a crucial element for interpreting of genetic variation at the population level.

The genus *Botrychium* as a model in evolutionary biology

The genus *Botrychium* (Ophioglossaceae) has unique biological features to address the prevalence of cryptic taxa in ferns and how allopolyploidy and the alternation of mating systems have led to speciation. It has been known for a long time that *Botrychium* has cryptic diploid and polyploid taxa (Paris et al. 1989). The genus still challenges botanists in species identification based on subtle morphological characters such as trophophore shape and pinnae number, orientation, and margin. Also, *Botrychium* has a broad circumtemperate to circumboreal distribution, with several species having experienced long-range dispersal between Asia, Europe, and North America. Thus, geographic distribution and range size varies considerably among species, although our knowledge of their distribution is limited in Europe and Asia, which greatly complicates species delineation and the characterization of hybrid zones forming polyploids. To date, 35 species have been described and at least three centers of species diversity characterized; the Alps in Europe (Dauphin et al. 2014), and both the Rocky Mountains and the Great Lakes region in North America (Hauk et al. 2012).

In addition to species richness, *Botrychium* genomes contain a large number of chromosomes (2n = 90; Wagner and Lord 1956), multiple ploidy levels (di, tetra, hexa, and probably octoploid; Wagner

and Wagner 1993), and large genome sizes (2C-value ranging from 22.05 pg to 53.68 pg; Williams and Waller 2012). At broad evolutionary scales, *Botrychium* species have likely experienced multiple rounds of polyploidy associated with diploidization—restoration of secondary diploid-like behavior of polyploid genomes—that led to the evolution of paleopolyploids, resulting in the current diploid taxa. More recently, neopolyploids—young polyploids without diploidized genomes—have arisen from the interspecific hybridizations between diploid progenitors that currently form half of the species diversity of *Botrychium*. The rampant allopolyploidy is postulated to play a central role in the contemporary speciation of these species as well as in the older paleopolyploidization events. Those hybrids were hypothesized from morphological characters (Figure 1) (Wagner and Wagner 1981, 1983, 1986), karyological data (Wagner and Lord, 1956; Wagner and Wagner, 1993), allozyme-based investigations (Farrar 2011), and plastid molecular phylogenies (Hauk et al., 2012; Williams and Waller, 2012; Dauphin et al., 2014).

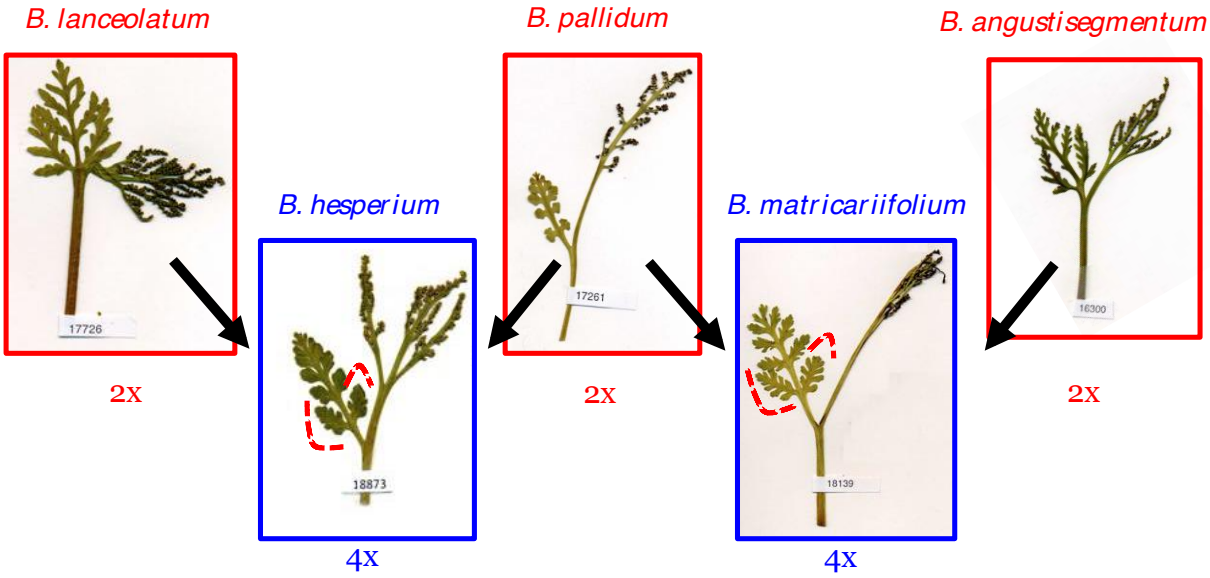


Figure 1 An example of morphological characters intermediate between known putative diploid progenitors. Diploids are in red and allotetraploids are in blue. The "lanceolate" trait of *B*. *lanceolatum* and *B*. *angustisegmentum*, and the pinna span of *B*. *pallidum* given to allopolyploids are symbolized in dashed red lines.

Mating systems of *Botrychium* are of great interest for understanding the biology of early divergent vascular plants and how primitive plants have colonized and explored new terrestrial habitats. Species have the capability of vegetative reproduction via gemmae produced in leaf axils of their underground stems (*B*. *campestre* and *B*. *pallidum*, Farrar and Johnson-Groh 1990; *B*. *pumicola*, Camacho 1996), and they theoretically support three sexual systems of reproduction; gametophytic selfing, sporophytic selfing, and sporophytic outcrossing (Sessa et al. 2016). Although gametophytic selfing appears predominant in *Botrychium* (Hauk and Haufler 1999), a recent study has shown evidence of a mixed mating system in several populations of *B*. *campestre* var. *lineare* sampled at high elevations in Rocky Mountains (Farrar and Gilman 2017), which has sparked our curiosity about the Swiss Alps *B*. *lunaria* populations that occur in similar habitats along Alpine grassland.

Outline of the chapters

This PhD thesis aims to better understand the evolution of early divergent vascular plants using the genus *Botrychium* as a model. It is organized in five chapters and summarized in a last section of conclusion.

Chapter 1 represents the starting point for assessing phylogenetic relationships and genetic diversity among taxa in the genus *Botrychium*. We tested a series of hypotheses for the recognition of new taxa and reconstructed a chloroplast (maternal) phylogeny for the genus. We characterized the exceptional haplotype diversity occurring in the Lunaria complex and found evidence for multiple origins of several polyploid taxa. In total, our study investigated 47 taxa of which thirteen are discussed for possible taxonomic recognition.

In chapter 2, we applied cytological methods to determine the ploidy level and genome size of representatives of the *Botrychium* s.l. clade, including *Botrychium* s.s., *Botrypus*, and *Sceptridium*. We found a new hexaploid cytotype for *Botrychium boreale*, representing only the second report of hexaploidy in the genus, and identified different genome sizes between diploid species of the two major clades in *Botrychium* (Lanceolatum versus Lunaria). Our results support the hypothesis of genome size stability after allo-polyploidization, therefore rejecting the scenario of genome downsizing, which is generally accepted for angiosperms.

Chapter 3 complements Chapter 1's chloroplast (maternal) phylogeny, in that we investigate the nuclear (paternal) phylogeny of nearly all species of *Botrychium*, allowing us to identify the parental lineages of polyploid taxa. Using a new sequencing approach, we inferred phylogenetic relationships among taxa and estimated the divergence time between lineages. We also tested hypotheses of the independent origins of allopolyploid lineages and identified their parental progenitors. Interestingly, we found non-random parental donors in the formation of allopolyploid taxa with strong bias in favor of several key diploid species, which suggests biological mechanisms that determine preferential parental couples in hybrid populations.

Chapter 4 was a major breakthrough of my PhD. Based on a sampling of Swiss Alps populations of *Botrychium lunaria*, we conducted a population genetics study using allozyme data to characterize their predominant mode of reproduction and put in perspective the alternation of outcrossing and inbreeding mating system among the taxa of the Lunaria complex. We presented the Swiss Alps populations as the second report of the mixed mating system in *Botrychium* and discussed the

important implications in evolution of their closely related homozygous taxa belonging to the species complex. We propose that *Botrychium* may be a relevant evolutionary model for a better understanding of dispersion and diversification of early divergent vascular plants.

As a continuation of chapter 4, chapter 5 addresses the role of the last glacial maximum on population genetic structure within the Val d'Hérens valley in the Swiss Alps. We used the new genotyping approach ddRAD (double-digest restriction-associated DNA) sequencing to characterize genetic diversity, gene flow among populations, and breeding system within populations. We found strong evidence for the maintenance of the mixed mating system and a recent colonization of four populations from a central Alps glacial refuge.

Finally, we highlight the main results and conclusions of my PhD dissertation and discuss the perspectives for future work on the evolutionary biology of ferns.

References

Abraham A, Ninan CA, Mathew PM (1962) Studies on the cytology and phylogeny of the pteridophytes VII. Observations on one hundred species of South Indian ferns. J Indian Bot Soc 41: 339–421.

Amato A, Kooistra WH, Ghiron JH, Mann DG, Pröschold T, Montresor M (2007) Reproductive isolation among sympatric cryptic species in marine diatoms. Protist 158: 193–207.

Arrigo N, Barker MS (2012) Rarely successful polyploids and their legacy in plant genomes. Curr Opin Plant Biol 15: 140–146.

Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. Trends Ecol Evol 22: 148–155.

Brownsey PJ, Lovis JD (1987) Chromosome numbers of the New Zealand species of *Psilotum* and *Tmesipteris*, and the phylogenetic relationships of the Psilotales. New Zealand J Bot 25: 439–454.

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

Crespo A, Lumbsch HT (2010) Cryptic species in lichen-forming fungi. IMA fungus 1: 167–170.

Dauphin B, Vieu J, Grant JR (2014) Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* s.s., Ophioglossaceae). Am J Bot 101: 128–140.

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biology 3: e314.

Farrar D (2011) Moonwort (*Botrychium*) Systematics. http://www.public.iastate.edu/~herbarium/botrychium

Farrar DR, Gilman AV (2017) Relationships in the *Botrychium campestre* (Ophioglossaceae) complex. Brittonia 1–11.

Farrar DR, Johnson-Groh CL (1990) Subterranean sporophytic gemmae in moonwort ferns, *Botrychium* subgenus *Botrychium*. Am J Bot 77: 1168–1175.

Fawcett J, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. PNAS 106: 5737–5742.

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

Hauk WD, Kennedy L, Hawke HM (2012) A phylogenetic investigation of *Botrychium* s.s. (Ophioglossaceae): evidence from three plastid DNA sequence datasets. Syst Bot 37: 320–330.

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473: 97–100.

Khandelwal S (1990) Chromosome evolution in the genus *Ophioglossum* L. Bot J Linn Soc 102: 205–217.

Leitch IJ, Leitch AR (2013) Genome size diversity and evolution in land plants. In Plant Genome Diversity Volume 2 (pp. 307–322). Springer Vienna.

Leitch IJ, Soltis DE, Soltis PS, Bennett MD (2005) Evolution of DNA amounts across land plants (Embryophyta). Ann Bot 95: 207–217.

Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, Otto SP (2011) Recently formed polyploid plants diversify at lower rates. Science 333: 1257–1257.

Paris CA, Wagner FS, Wagner WH (1989) Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. Am Fern J 79: 46–54.

Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol Biol 7: 121.

Rastogi S, Liberles DA (2005) Subfunctionalization of duplicated genes as a transition state to neofunctionalization. BMC Evol Biol 5: 28.

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

Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. Trends Ecol Evol 14: 348–352.

Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. PNAS 97: 7051–7057.

Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. Ann Review Plant Biol 60: 561–588.

Stebbins GL (1947) Types of polyploids: their classification and significance. Adv Genet 1: 403–429.

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

Wagner FS 1993. Chromosomes of North American grapeferns and moonworts (Ophioglossaceae: *Botrychium*). Contributions from the University of Michigan Herbarium 19: 83–92.

Wagner WH, Lord LP (1956) The morphological and cytological distinctness of *Botrychium minganense* and *B. lunaria* in Michigan. Bulletin of the Torrey Botanical Club 83: 261–280.

Wagner WH, Wagner FS (1981) New species of moonworts, *Botrychium* subg. *Botrychium* (Ophioglossaceae), from North America. Am Fern J 71: 20–30.

Wagner WH, Wagner FS (1983) Two moonworts of the Rocky Mountains; *Botrychium hesperium* and a new species formerly confused with it. Am Fern J 73: 53–62.

Wagner WH, Wagner FS (1986). Three new species of moonworts (*Botrychium* subg. *Botrychium*) endemic in western North America. Am Fern J 76: 33–47.

Williams EW, Waller DM (2012) Phylogenetic placement of species within the genus *Botrychium* s.s. (Ophioglossaceae) on the basis of plastid sequences, amplified fragment length polymorphisms, and flow cytometry. Int J Plant Sci 173: 516–531.

Chapter 1

A Worldwide Molecular Phylogeny provides New Insight on Cryptic Diversity within the Moonworts (*Botrychium* S.S., Ophioglossaceae)

Benjamin Dauphin[1,3], Donald R. Farrar[2], Alessio Maccagni[1], Jason Grant[1]

[1]Université de Neuchâtel, Laboratoire de botanique évolutive, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

[2]Iowa State University, Department of Ecology, Evolution, and Organismal Biology, 2200 Osborn Dr., Room 251 Bessey Hall, Ames, Iowa 50011–4009, U.S.A.

This chapter is in press in *Systematic Botany* 42.

Abstract

The moonwort genus, *Botrychium* s. s., includes diploid and polyploid taxa that occur primarily in the northern hemisphere. Their evolutionary history, morphologically cryptic taxa and deep divergence of the family in the phylogeny of ferns has long fascinated pteridologists. Previous molecular studies did not include a complete taxonomic sampling of the taxa in the genus, nor multiple specimens from throughout the known geographical range of each taxon. Therefore, to investigate evolutionary relationships of the major clades of *Botrychium* s. s., we increased both taxonomic representativeness (multiple accessions per taxa), as well as phylogenetic resolution by including additional new chloroplast markers. To confirm identification and provide evidence from both maternal and paternal parentage of allopolyploids we also included specimens that have been characterized by allozyme profiles determined by electrophoretic analysis of 20 nuclear enzyme loci for each taxon. We analyzed four chloroplast regions (*matK* intron, $trnH^{GUG}$–*psbA* and $trnL^{UAA}$-$trnF^{GAA}$ intergenic spacers, and *rpL16* intron region) of 365 specimens from Asia, Europe, North America, Oceania, and South America, sampling the geographical range of 34 of 35 accepted *Botrychium* s. s. taxa and thirteen putatively new taxa. We conducted a phylogenetic analysis of maternal lineages based on 2,385 aligned nucleotides using maximum likelihood and Bayesian inference to explore genetic diversity and phylogenetic relationships among taxa. We found strong support for the monophyly of three major clades: Lanceolatum, Lunaria, and Simplex-Campestre, and resolved 15 subclades. Our results suggest multiple origins for at least four polyploid taxa (*B. boreale*, *B. michiganense*, *B. yaaxudakeit*, and *B. watertonense*). The Simplex-Campestre clade had the largest number of species, despite having a similar total number of haplotypes as the Lunaria clade (62 and 59, respectively), which has the broadest worldwide distribution. In total, our new molecular phylogeny comprises 47 taxa, of which thirteen are discussed for possible taxonomic recognition.

Keywords: *Botrychium*; fern; molecular phylogeny; moonwort; Ophioglossaceae; polyploid.

Introduction

Botrychium Sw. s. s. is a genus of the fern family Ophioglossaceae with 35 accepted taxa. Ten genera are currently accepted in the family: *Botrychium* s. s. (*B*.), *Botrypus* Michx., *Cheiroglossa* C. Presl, *Helminthostachys* Kaulf., *Japanobotrychium* Masam., *Mankyua* B. Y. Sun, M. H. Kim & C. H. Kim, *Ophioderma* (Blume) Endl., *Ophioglossum* L. s. s., *Rhizoglossum* C. Presl, and *Sceptridium* Lyon (Clausen 1938; Kato 1987; Hauk 1995; Sun et al. 2001; Hauk et al. 2003; Sun et al. 2009; Meza-Torres and Ferrucci 2013; Shinohara et al. 2013; Schuettpelz et al. 2016). *Botrypus* and *Sceptridium* are segregates of *Botrychium* s. l., and *Cheiroglossa*, *Ophioderma*, and *Rhizoglossum* are segregates of *Ophioglossum* s. l. (Hauk et al. 2003). The divergence time of the Ophioglossaceae from other extant ferns is still uncertain at ca. 160 ± 80 MY (Pryer et al. 2004; Rothfels et al. 2015), but two macrofossils attest to an ancient lineage divergence between genera (57 mya, Rothwell and Stockey 1989; 23 mya, Bozukov et al. 2010).

Botrychium occurs mainly in perennial herbaceous meadows of mountainous regions, tundra, taiga, grasslands, and open forests, and secondarily in similar maintained habitat of older roadsides and utility corridors. The genus has a broad circumtemperate to circumboreal distribution in North America including Mexico (Farrar et al. *in press*; Wagner and Mickel 1991), Europe, North Africa (Atlas Mountains) (Haroni et al. 2009), and Asia, with several notable range disjunctions to the southern hemisphere in Oceania (Australia, Tasmania, and New Zealand) and temperate South America (Patagonia and the Falkland Islands). At least three centers of species diversity occur; the Alps in Europe (Dauphin et al. 2014), and both the Rocky Mountains and the Great Lakes region in North America (Hauk et al. 2012). Nevertheless, the geographic distribution and range size varies considerably among species. Highlighting the dispersal ability of fern spores, some species display long-range disjunctions, for example the extreme southern hemisphere disjunctions of the indistinguishable *B*. *spathulatum* W. H. Wagner and *B*. *dusenii* (Christ) Alston from North America to Patagonia and Falkland Islands (Meza-Torres et al. 2017), and *B*. *neolunaria* Stensvold & Farrar from Alaska to Australia, Tasmania, and New Zealand (Stensvold 2008; Stensvold and Farrar 2016), which are probably facilitated by way of trans-polar migratory birds (Ballard and Sytsma 2000; Schönswetter et al. 2008; Piñeiro et al. 2012; Lewis et al. 2014). Also, several species have an inter-continental distribution (e.g. *B*. *lanceolatum* (Gmel.) Ångstr., *B*. *lunaria* (L.) Sw., *B*. *matricariifolium* (Döll) A. Braun, *B*. *simplex* E. Hitchc., and *B*. *tenebrosum* A. A. Eaton) while others have much more restricted ranges (e.g. *B*. *gallicomontanum* Farrar & Johnson-Groh, *B*. *montanum* W. H. Wagner, and *B*. *pseudopinnatum* W. H. Wagner) (Wagner and Wagner 1990; Farrar and Johnson-Groh 1991).

Botrychium s. s. is morphologically simple and each year an underground bud produces a single leaf that is dichotomously divided into a trophophore (blade) and a sporophore (fertile segment) (Fig. 1). It has a symbiotic relationship with various species of arbuscular mycorrhizal fungi (AMF) (Glomeromycota) that colonize its roots, and some moonwort taxa tend to occur in various associated

soil types and floristic communities (Sandoz 2016). Subtle morphological characters, such as trophophore shape and pinnae number, orientation, and margin, often make it difficult to distinguish taxa and misidentification is frequent. These morphological challenges have led botanists to characterize moonworts as a morphologically cryptic group (Paris et al. 1989), in which [two or more distinct species may be erroneously classified or hidden under a single species name] (Bickford et al. 2007). *Botrychium minganense* Victorin illustrates these ambiguities, being first described by Victorin (1927) as a distinct species, then relegated to varietal rank (*B. lunaria* var. *minganense*) by Clausen (1938). Subsequent investigations of morphological characters combined with karyological analyses based on 14 differentiating morphological traits and different ploidy levels led Wagner and Lord (1956) to recognize *B. minganense* and *B. lunaria* as separate and distinct species.

Farrar and coworkers were able to circumvent morphological limitations by using applied genetic analyses of 20 enzyme loci to characterize individual taxa and test relationships within the genus (see Farrar 2011 for overview). These studies led to the discovery of additional cryptic species (Farrar and Johnson-Groh 1991; Stensvold et al. 2002; Stensvold and Farrar 2016), and, because enzyme electrophoresis visualizes codominant alleles, it allowed the detection of both parental taxa of allopolyploid species (Zika and Farrar 2009; Gilman et al. 2015; Williams et al. 2016; Meza-Torres et al. 2017). Based on genetic identity (GI) (Nei 1978) between species, DNA sequences (Dauphin et al. 2014), karyological data to determine ploidy level (Wagner 1993; Dauphin et al. 2016), and ecological and relevant morphological characteristics (Farrar 2011), all these authors used the biological species concept to define species and varieties, and we follow their circumscription of taxa.

Initial identification of taxa by genetic analysis of allozymes has also aided morphological identification enabling recognition of certain diagnostic morphological characters (Farrar and Popovich 2012; Williams et al. 2016). We further evaluated several recently published taxa characterized by genetic data (e.g. *B. lunaria* var. *melzeri* Stensvold & Farrar, *B. michiganense* W. H. Wagner ex A. V. Gilman, Farrar & Zika, *B. neolunaria*, and *B. nordicum* Stensvold & Farrar) (Gilman et al. 2015; Stensvold and Farrar 2016). This study also evaluates eight unpublished taxa (*B. alaskense* var. *salchaketense* J. R. Grant, var. nov. ined., *B. boreale* × *B. lunaria 4*, *B. farrarii* Legler & Popovich, sp. nov. ined., *B. furculatum* Popovich & Farrar, sp. nov. ined., *B. lanceolatum* "red" × *B. pallidum 1, B. lanceolatum* "red" × *B. pallidum 2, B. lanceolatum* "green", and *B. minganense* × *B. echo*), and presents evidence for five more cryptic taxa (*B. lunaria 2*, *B. lunaria 3*, *B. lunaria 4*, *B. lunaria 5*, and *B. simplex 2* (Table 1). The unpublished new taxa we name here are not intended to be validated in this paper as they are in preparation or in press elsewhere.

Previous molecular phylogenies described relationships in *Botrychium* s. s. based on more limited geographical representativeness and taxonomic sampling (Hauk et al. 2012; Williams and Waller 2012; Dauphin et al. 2014). The present study expands geographic and taxonomic sampling and uses new markers to support many of the hypotheses generated by earlier studies, but also introduces new

insights into 1) the cladistic relationships among known and proposed taxa, 2) maternal lineages and multiple origins of allopolyploid taxa, 3) congruence of chloroplast and nuclear data in supporting taxa and their relationships, and 4) geographic distributions of clades and clade members, and potential for further analysis of geographic migrations and origins of *Botrychium* taxa.

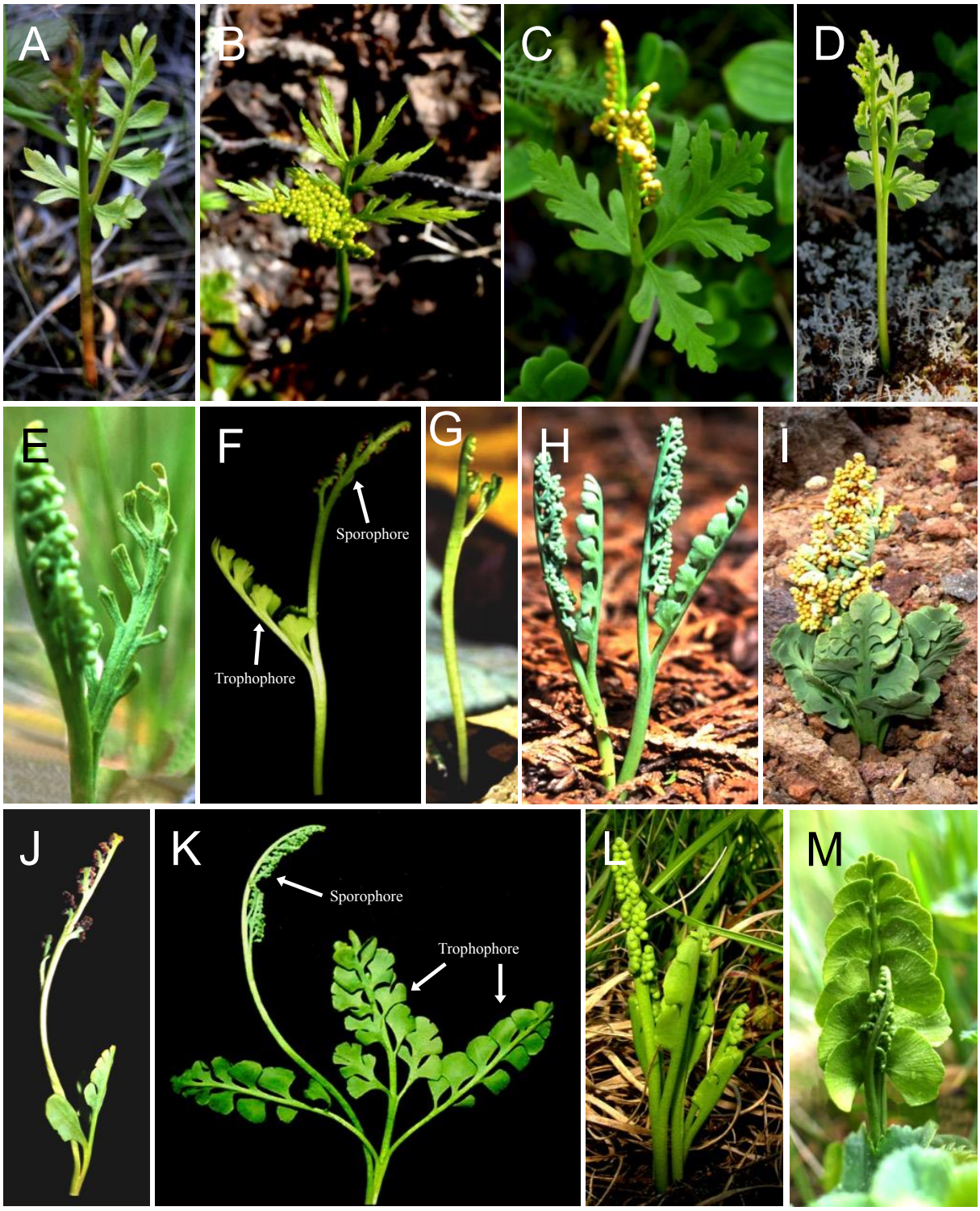


Fig. 1. A–M. Morphological diversity in *Botrychium*. A. *B. alaskense* var. *alaskense* (photo by B. Dauphin, U.S.A). Trophophore and sporophore are indicated with arrows on the photos F and K. B. *B. angustisegmentum* (photo by B. Dauphin, Sweden). C. *B. lanceolatum* "green" (photo by J. Grant, U.S.A.). D. *B. boreale* (photo by B. Dauphin, Sweden). E. *B. campestre* var. *lineare* (photo by D. Barton, U.S.A.). F. *B. minganense* (photo by D. Farrar, U.S.A.). G. *B. mormo* (photo by D. Farrar, U.S.A.). H. *B. pallidum* (photo by S. Mortensen, U.S.A.). I. *B. pumicola* (photo by D. Farrar, U.S.A.). J. *B. simplex* var. *simplex* (photo by D. Farrar, U.S.A.). K. *B. simplex* var. *compositum* (photo by D. Farrar, U.S.A.). L. *B. tenebrosum* (photo by D. Farrar, U.S.A.). M. *B. lunaria* var. *lunaria* (photo by A. Maccagni, Switzerland).

Materials and Methods

Sampling

Our sampling covers the geographical range of each clade and subclade of *Botrychium* s. s. as identified by Hauk et al. (2003; 2012) and Dauphin et al. (2014), 34 of the 35 accepted taxa (only *B. tolucaense* W. H. Wagner & Mickel from Mexico is missing), as well as several accessions from potential new species and varieties. We included a specimen from the type locality (Topotype) for most species. At least five species occur in Asia, fourteen in Europe, eighteen in North America, and one each in North Africa, Oceania, and temperate South America (Table 1). In addition, thirteen putatively new taxa are included: *B. alaskense* var. *salchaketense* var. nov. ined., *B. boreale* × *B. lunaria 4*, *B. farrarii* sp. nov. ined., *B. furculatum* sp. nov. ined., *B. lanceolatum* "green", *B. lanceolatum* "red" × *B. pallidum 1*, *B. lanceolatum* "red" × *B. pallidum 2*, *B. lunaria 2*, *B. lunaria 3*, *B. lunaria 4*, *B. lunaria 5*, *B. minganense* × *B. echo*, and *B. simplex 2*.

Taxonomic considerations

As determined through enzyme electrophoresis, the taxonomic rank of all published (and many unpublished) *Botrychium* taxa has been evaluated by Farrar and coauthors (cited herein) employing GI (Nei 1978). In doing so, these studies have followed a general guideline of GI values below 70% as characterizing taxon distinction at the species level and identities above 90% as characterizing populations within a species that do not warrant taxon distinction. Intermediate identities are considered evidence of intermediate rank, e.g., varieties (see Stensvold and Farrar 2016 for review of these criteria). When applied to published *Botrychium* species, these criteria have been generally supportive of the published ranks, with exception of the taxa discussed below. First, *B. lanceolatum* was confirmed to be actually composed of three morphologically and genetically distinct species, herein referred to as *B. lanceolatum* "red" (GI = 77%), *B. lanceolatum* "green" (GI = 74%), and *B. angustisegmentum* (Pease & A. H. Moore) Fernald (GI = 79%) (Stensvold and Farrar 2016). We accept the provisional name *B. lanceolatum* "green" (Stensvold and Farrar 2016) applied to the taxon segregated from *B. lanceolatum* "red" (resolution of whether "red" or "green" morphotypes represent published *B. lanceolatum* (Gmel.) Angstr. is under study). Second, *B. acuminatum* W. H. Wagner was not supported as distinct from *B. matricariifolium* (GI = 99%) (Williams et al. 2015). Third, *B. lineare* W. H. Wagner was reduced to varietal status under *B. campestre* W. H. Wagner & Farrar (GI = 83%) (Farrar and Gilman, 2017), and finally, preliminary analysis of both allozyme and chloroplast data suggest that the species pairs *B. mormo* W. H. Wagner and *B. montanum*, and *B. dusenii* and *B. spathulatum* have genetic similarity much higher than expected for distinct species and are proposed herein for reevaluation as to proper rank. Some newly recognized clades in this study that may warrant taxonomic recognition, but have yet to be evaluated for GI, are listed as potentially new taxa warranting further study (Table 1).

The majority of specimens from North America and many from Europe were provided by the herbarium (ISC) at Iowa State University (Ames, Iowa) with their identification confirmed by both morphological characters and genetic profiles based on allozyme analyses (for electrophoresis procedures see Stensvold and Farrar 2016). Most specimens from Alaska and Europe were collected by BD, AM, and JG, while other material, largely from Asia, Oceania, and Sweden, was received dried in silica gel from colleagues (see acknowledgements). Specimen voucher data and GenBank accession numbers are provided in Appendix 1. Three representatives of the sister genera *Botrypus* and *Sceptridium* were used as outgroups (Pryer et al. 2004; Rothfels et al. 2015).

Assessment of Ploidy Level

The ploidy level of many described taxa was determined through chromosome counts included in the original descriptions. Ploidy of taxa not determined by chromosome counts was inferred by three methods: allozyme profiles, spore size, and flow cytometry. Visualization of co-dominant alleles through allozyme electrophoresis allows differentiation of recombinational heterozygosity that varies across individuals in a population vs. fixed heterozygosity that is constant in all plants of a taxon. Fixed heterozygosity is characteristic of allopolyploid taxa containing two homoeologous sets of chromosomes derived through polyploidization. Allozyme profiles of taxa with published chromosome counts were consistent with this interpretation. Spore size was measured as a morphological marker for cryptic specimens not examined cytologically or by allozyme profiles. Spore size in *Botrychium* taxa is consistently smaller in diploid taxa than in polyploid taxa. For example, the range of spore size in diploid *B. lunaria* and *B. neolunaria* is 33–39 µm whereas that of their allopolyploid derivative *B. yaaxudakeit* Stensvold and Farrar is 46–57 µm (Stensvold et al. 2002). Additionally, we investigated the ploidy level of several taxa using flow cytometry analysis (Dauphin et al. 2016).

DNA Extraction, PCR Amplification, and Sequencing

Total genomic DNA was extracted from leaf material dried in silica gel or from herbarium specimens using the DNeasy plant mini kit (Qiagen, Hilden, Düsseldorf, Germany) and the manufacturer's protocol. Total DNA was quantified with spectrophotometry (NanoDrop 2000, Thermo Scientific, Wilmington, Delaware), and the quality was inspected on agarose gel (1.5%).

For each taxon, the four non-coding chloroplast loci *matK* intron, $trnH^{GUG}$–*psbA* and $trnL^{UAA}$-$trnF^{GAA}$ intergenic spacers, and *rpL16* intron region were targeted based on amplification success and informativeness at intraspecific level (Taberlet et al. 1991; Sang et al. 1997; Small et al. 1998; Tate and Simpson 2003; Shaw et al. 2005; Hauk et al. 2012). Protocol and optimized polymerase chain reaction (PCR) programs were used following Dauphin et al. (2014) except for the *matK* intron. Specific forward matK_F3 (5' GGA ATC ATC ATC TTT ATG AGT TTG AGG 3') and reverse matK_R3 (5' GCA CGT ATT GTA CTT TTG TGT TTG C 3') primers targeting the *matK* intron

were designed on GenBank accessions from *Botrypus* and *Sceptridium* (Shinohara et al. 2013). This region was amplified using the following PCR program: 180 sec of initial denaturation at 94°C; 35 cycles of 94°C for 60 sec, 56°C for 45 sec, and 72°C for 60 sec; with a final extension at 72°C for 600 sec. All PCR amplifications were carried out in a 25 µL reaction volume containing 17.4 µL of ddH_2O, 5 µL of buffer, 0.5 µL of dNTP mix at 10 mM, 0.5 µL of each 10 mM primer, 0.1 µL of GoTaq G2 DNA polymerase (Promega, Madison, Wisconsin), and ~ 10 ng of DNA.

The PCR products were purified using the EZ-seq kit and loci were sequenced on an ABI3130 XL Automated Sequencer by Macrogen Europe (Amsterdam, Netherlands), using the same primer combinations as used in PCR amplifications. Sequencing reactions were carried out on 5 µL purified PCR products with 5 µL of primer at 10 mM. All regions were sequenced with both forward and reverse primers to ensure unambiguous chromatogram profiles.

Sequence Alignment

Sequences were edited and assembled with Geneious V.6.1.6 (Kearse et al. 2012). All sequences were aligned with MAFFT online (Katoh and Standley 2013) using the FFT-NS-1 settings. Indels (insertion/ deletion) present in the final alignment were coded as missing values. All sequences have been deposited in GenBank. The 83 *trnH*^{GUG}*–psbA*, *trnL*^{UAA}*-trnF*^{GAA}, and *rpL16* sequences already published (Dauphin et al. 2014) were included in the dataset to complement the geographic range of several taxa. Alignments of each locus were submitted to Dryad (doi:10.5061/dryad.tc07j).

Assessment of Haplotype Number per Geographical Area

The four cpDNA regions were analyzed as a single concatenated-locus because the chloroplast genome of ferns is basically uniparentally inherited with no recombination (Gastony et al. 1992; Vogel et al. 1998). Outgroups were excluded from this analysis and indels were considered as single mutation events (Simmons and Ochoterena 2000). The number of haplotypes was calculated in DnaSP V.5.10 (Librado and Rozas 2009) for the major clades Lanceolatum (LAN), Lunaria (LUN), and Simplex-Campestre (SIC) distributed across the entire geographical range of the genus. Geographical areas were defined to cover species distribution on five continents: Asia (ASI) [defined as Eurasia east of the Ural and Caucasus Mountains], Europe (EUR), North America (NAM), Oceania (OCE) [as Australia, Tasmania, and New Zealand), and South America (SAM) [as Patagonia and the Falkland Islands) (Fig. 2). Although *B. lunaria* occurs in North Africa in the Atlas Mountains of Morocco, we have so far been unable to procure material. Each taxon distribution was coded with multiple geographical areas when it had an inter-continental distribution.

Phylogenetic Analyses

Phylogenetic relationships were simultaneously investigated using maximum likelihood (ML) in RAxML V.8 (Stamatakis 2014) and Bayesian inference (BI) in both BEAST V.1.8 (Drummond et al. 2012) and MrBayes V.3.2 (Ronquist et al. 2012). The optimal substitution model was assessed for

each locus with the program jModelTest V.2 (Darriba et al. 2012) according to the AIC (Akaike 1973) and BIC (Posada and Buckley 2004) scores (Table 2). The data were partitioned and the substitution models were unlinked between the four cpDNA loci.

ML was run with the optimal substitution model (GTR+G) implementable in RAxML V.8 and conducted with 1,000 bootstrap replicates. *Botrypus virginianus* was fixed as the outgroup to root the topology. Bootstrap support values (BS) were reported on the phylogenetic tree (Fig. 3A-C).

BI were conducted with two optimal substitution models, HKY + G for *matK* and *rpL16* and GTR+G for $trnL^{UAA}$–$trnF^{GAA}$ and $trnH^{GUG}$–*psbA*. Two independent runs of 15 million MCMC generations were completed with each including three heated and one cold chain, and with uniform priors and a chain temp of 0.1. Trees were sampled every 1,000 generations and a burn-in of 33% was set to discard the first 5,000 trees. A 50% majority-rule consensus (MRC) tree of the 10,000 remaining trees was built based on which posterior probabilities (PP) were reported (Fig. 3A-C).

Analyses were independently conducted on the concatenated data partition and on each locus to check the congruence between loci. Incongruence between the tree topology of the concatenated data partition and the topology of each locus was visually inspected with the packages ape and distory in R (R Development Core Team 2015). All phylogenetic analyses were run on the Cyberinfrastructure for Phylogenetic Research (CIPRES) (Miller et al. 2010).

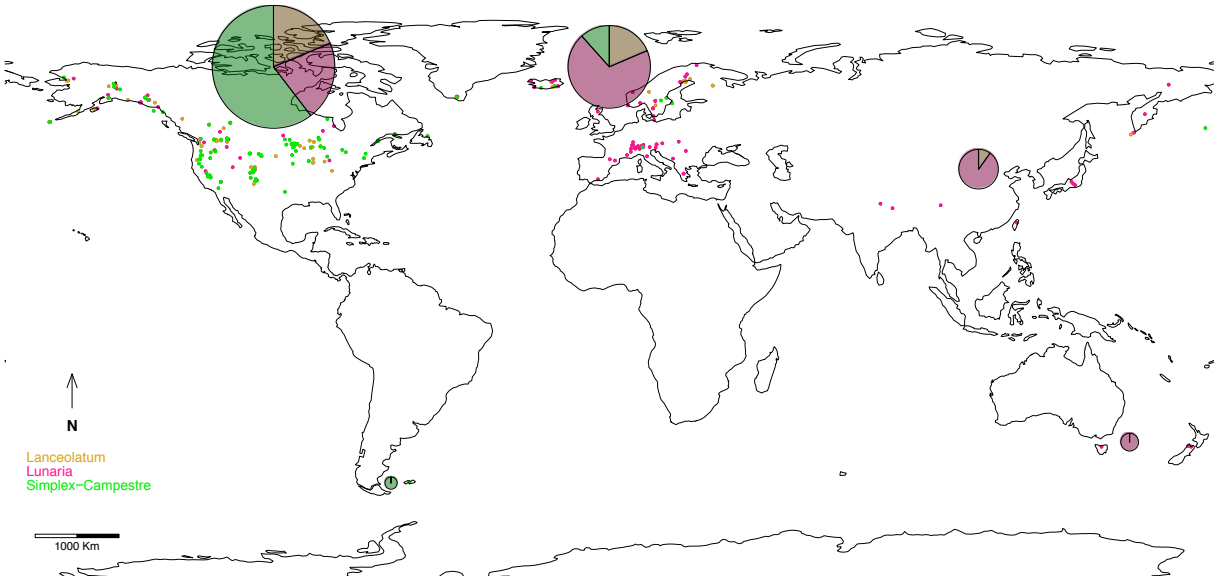


Fig. 2. Sampling and geographical distribution of haplotype numbers and their relative proportion for each major clade: Lanceolatum (gold), Lunaria (pink), and Simplex-Campestre (green). Number of haplotypes and their frequencies per taxon are given in Appendix 2.

Results

Chloroplast Characteristics

The combined plastid data set consisted of 2,385 aligned nucleotides, of which 515 (21.6%) sites were variable, 389 (16.3%) were parsimony-informative, and 126 (5.3%) were single nucleotide polymorphisms for a total of 368 individuals analyzed including outgroups (Table 2). Within *Botrychium* s. s., the two intergenic spacers $trnL^{UAA}$–$trnF^{GAA}$ and $trnH^{GUG}$–*psbA* were the most variable regions with 13.7% and 12.8%, whereas the introns *matK* and *rpL16* had the lowest variability with 9.1% and 9.7% respectively. From the three major clades, the LAN clade had the lowest variability with 22 variable sites and 16 parsimony-informative sites (0.9% and 0.7%), the LUN clade had 77 and 53 parsimony-informative sites (3.2% and 2.2%), and the SIC clade contained 137 and 102 parsimony-informative sites (5.7% and 4.3%) respectively. One sequence of *B. watertonense* W. H. Wagner (DF2896) for the *rpL16* locus is missing because PCR amplification repeatedly failed.

Number of Haplotypes per Geographical Area

A total of 118 cpDNA haplotypes was found, with 19 from the LAN clade, 59 from LUN and 62 from SIC (Table 3). The number of haplotypes was greatest in North America (96), followed by Europe (43), Asia (10), Oceania (2), and South America (1). While the SIC clade was the most diversified in North America, the LUN clade has its highest genetic diversity in Europe and Asia with 30 and 9 haplotypes respectively (Fig. 2). Details of haplotype frequencies are given in supplementary material (Appendix 2).

Phylogenetic Analyses

Topologies and support values (BS and PP) between ML and BI analyses were highly congruent (Fig. 3A-C), as well as between the different loci although their phylogenetic resolution varied. The optimal substitution models implementable were identified from the jModelTest ranking as HKY +G for *matK* (AIC = 4,672.3 and BIC = 7,988.7) and *rpL16* (AIC = 6,111.9 and BIC = 9,518.8) and GTR + G for $trnL^{UAA}$–$trnF^{GAA}$ (AIC = 3,916.9 and BIC = 6,830.6) and $trnH^{GUG}$–*psbA* (AIC = 5,710.5 and BIC = 8,981.1). The best tree produced by ML analysis (ln = – 7,758.150721) and the MRC tree depicted an appropriate phylogenetic resolution for both relationships at inter- and intraspecific level (Fig. 3A-C). A simplified global molecular phylogeny of maternal clades is presented with diploid and polyploid taxa based on the most frequent haplotype per taxon (Fig. 4).

Our phylogenetic tree resolved three major clades LAN (ML-BS/BI-PP = 100/1.00) (Fig. 3A), SIC (78/1.00) (Fig. 3B) and LUN (100/1.00) (Fig. 3C) as well as 15 subclades with a PP > 0.95: Alaskense (73/1.00), Boreale GRL (51/0.99), Campestre (100/1.00), Crenulatum (84/1.00), Farrarii (97/1.00), Hesperium (53/0.96), Lunaria 3 (99/1.00), Lunaria 4 (93/1.00), Michiganense Central NA (51/0.97), Neolunaria (56/0.99), Pallidum (100/1.00), Pseudopinnatum (40/0.98), Pumicola (99/1.00), Simplex (100/1.00), and Tunux (74/1.00). In addition, four poorly supported subclades were identified

Lunaria 1 (64/0.93), Lunaria 2 (41/0.62), Matricariifolium (41/0.75), and Pinnatum (38/0.92) (Fig. 3A-C).

There is a large polytomy in the LAN clade with a single haplotype shared by three diploids (*B. angustisegmentum*, *B. lanceolatum* "green", and *B. lanceolatum* "red") and six allopolyploids (*B. boreale* J. Milde, *B. echo* W. H. Wagner, *B. lanceolatum* "red" × *B. pallidum 1*, *B. lanceolatum* "red" × *B. pallidum 2*, *B. michiganense*, and *B. watertonense*). One taxon was monophyletic, *B. hesperium* (Maxon & R. T. Clausen) W. H. Wagner & Lellinger, and four taxa were embedded with all their representatives in subclades that included diploids, *B. alaskense* W. H. Wagner & J. R. Grant var. *alaskense*, *B. alaskense* var. *salchaketense* var. nov. ined., *B. pinnatum* H. St. John, and *B. pseudopinnatum* although their differentiation is relatively recent as indicated by short branches. In contrast, three polyploid taxa were paraphyletic: *B. boreale* fell into both the Boreale GRL subclade and the Lanceolatum polytomy, *B. michiganense* into three positions (the Matricariifolium, Michiganense Central NA subclade and the polytomy), and *B. watertonense* into the Lanceolatum polytomy and into the Farrarii subclade of the SIC clade (Fig. 3B).

TABLE 1. *Botrychium* taxa included in this study with their placement into diploid clades and, for allopolyploids, maternal diploid clades and subclades and probable maternal taxa as determined by chloroplast haplotypes. [1]Name published or in preparation; [2]Recommend for taxonomic recognition or change in existing taxonomy; [3]Recommend for further studies; [4]Two origins through hybridization between the same parents with reciprocal maternal parentage; [5]Based on our study as well as previous works using allozyme profiles.

Taxa	Diploid clade	Subclade or polytomy	Distribution	Ploidy level (x)	Probable maternal progenitor[5]
B. alaskense W. H. Wagner & J. R. Grant var. *alaskense*[1]	Lanceolatum	Alaskense	North America	4	*B. lanceolatum* "green"
B. alaskense W. H. Wagner & J. R. Grant var. *salchaketense* J. R. Grant, var. nov. ined.[1]	Lanceolatum	Alaskense	North America	4	*B. lanceolatum* "green"
B. angustisegmentum (Pease & A. H. Moore) Fernald[1]	Lanceolatum	Pseudopinnatum, Lanceolatum polytomy	North America	2	
B. ascendens W. H. Wagner[1]	Simplex−Campestre	Campestre	North America	4	*B. campestre* var. *lineare*
B. boreale × *B. lunaria 4*[3]	Lanceolatum	Lanceolatum polytomy	Europe	6	*B. boreale*
B. boreale J. Milde[1]	Lanceolatum	Boreale GRL, Lanceolatum polytomy	Europe North America	4	*B. lanceolatum* "green"
B. campestre var. *lineare* (W. H. Wagner) Farrar[1]	Simplex−Campestre	Campestre	North America	2	
B. campestre W. H. Wagner & Farrar var. *campestre*[1]	Simplex−Campestre	Campestre	North America	2	
B. crenulatum W. H. Wagner[1]	Lunaria	Crenulatum	North America	2	
B. dusenii (Christ) Alston[1,2]	Simplex−Campestre	Campestre	South America	4	*B. campestre*
B. echo W. H. Wagner[1]	Lanceolatum	Lanceolatum polytomy	North America	4	*B. lanceolatum* "red"
B. farrarii Legler & Popovich, sp. nov. ined.[1]	Simplex−Campestre	Farrarii	North America	2	
B. furculatum Popovich & Farrar, sp. nov. ined.[1]	Simplex−Campestre	Farrarii	North America	4	*B. farrarii* sp. nov. ined.
B. gallicomontanum Farrar & Johnson-Groh[1]	Simplex−Campestre	Pallidum	North America	4	*B. pallidum*
B. hesperium (Maxon & R.T.Clausen) W. H. Wagner & Lellinger[1]	Lanceolatum	Hesperium	North America	4	*B. lanceolatum* "red"
B. lanceolatum (Gmel.) Ångstr.[1] "green morphotype/genotype"	Lanceolatum	Alaskense, Boreale GRL, Lanceolatum polytomy, Pinnatum	Asia, Europe, North America	2	
B. lanceolatum (Gmel.) Ångstr.[1] "red morphotype/genotype"	Lanceolatum	Lanceolatum polytomy	Europe, North America	2	
B. lanceolatum "red" × *B. pallidum 1*[3]	Lanceolatum	Lanceolatum polytomy	North America	4	*B. lanceolatum* "red"
B. lanceolatum "red" × *B. pallidum 2*[3]	Lanceolatum	Lanceolatum polytomy	North America	4	*B. lanceolatum* "red"
B. lunaria (L.) Sw. var. *lunaria*[1]	Lunaria	Lunaria 1	Europe	2	

B. lunaria (L.) Sw. var. *melzeri* Stensvold & Farrar[1]	Lunaria	Lunaria 1	Europe	2	
B. lunaria 2[3]	Lunaria	Lunaria 2	Asia, Europe, North America	2	
B. lunaria 3[3]	Lunaria	Lunaria 3	Europe	2	
B. lunaria 4[3]	Lunaria	Lunaria 4	Europe	2	
B. lunaria 5[3]	Lunaria	Lunaria 1	Europe	2	
B. matricariifolium (Döll) A.Braun[1]	Lanceolatum	Matricariifolium	Europe, North America	4	*B. angustisegmentum*
B. michiganense W.H.Wagner ex A.V.Gilman, Farrar & Zika[1]	Lanceolatum	Matricariifolium, Michiganense Central NA, Lanceolatum polytomy	North America	4	Ancestral "*B. lanceolatum*"
B. minganense × *B. echo*[3]	Simplex−Campestre	Minganense	North America	4 (8?)	*B. minganense*
B. minganense Victorin[1]	Simplex−Campestre	Farrarii	North America	4	*B. farrarii* sp. nov. ined.
B. montanum W.H.Wagner[1,2]	Simplex−Campestre	Simplex	North America	2	
B. mormo W.H.Wagner[1,2]	Simplex−Campestre	Simplex	North America	2	
B. neolunaria Stensvold & Farrar[1]	Lunaria	Neolunaria	North America	2	
B. nordicum Stensvold & Farrar[1]	Lunaria	Lunaria 1	Europe	2	
B. pallidum W. H. Wagner[1]	Simplex−Campestre	Pallidum	North America	2	
B. paradoxum W. H. Wagner[1]	Simplex−Campestre	Farrarii	North America	4	*B. farrarii* sp. nov. ined.
B. pedunculosum W. H. Wagner[1]	Lanceolatum	Matricariifolium	North America	4	*B. lanceolatum* "green"
B. pinnatum H. St. John[1]	Lanceolatum	Pinnatum	North America	4	*B. lanceolatum* "green"
B. pseudopinnatum W. H. Wagner[1]	Lanceolatum	Pseudopinnatum	North America	6	*B. angustisegmentum*
B. pumicola Colville[1]	Simplex−Campestre	Pumicola	North America	2	
B. simplex 2[3]	Simplex−Campestre	Simplex	Europe, North America	2	
B. simplex E. Hitchc. var. *compositum* (Lasch) Milde[2]	Simplex−Campestre	Simplex	North America	2	
B. simplex E. Hitchc. var. *simplex*[1]	Simplex−Campestre	Simplex	North America	2	
B. spathulatum W. H. Wagner[1]	Simplex−Campestre	Campestre	North America	4	*B. campestre*

B. tenebrosum A. A. Eaton[1]	Simplex−Campestre	Simplex	Europe, North America	2	
B. tunux Stensvold & Farrar[1]	Lunaria	Tunux	Europe, North America	2	
B. watertonense W. H. Wagner[1,4]	Lanceolatum, Simplex−Campestre	Lanceolatum polytomy, Farrarii	North America	4, (8?)	*B. hesperium* *B. paradoxum*
B. yaaxudakeit Stensvold & Farrar[1]	Lunaria	Neolunaria	North America	4	*B. neolunaria*
Botrypus virginianus (L.) A.Michx[1]			North America, Europe	4	
Sceptridium multifidum (S.G.Gmelin) M.Nishida[1]			North America, Europe, Asia	2	

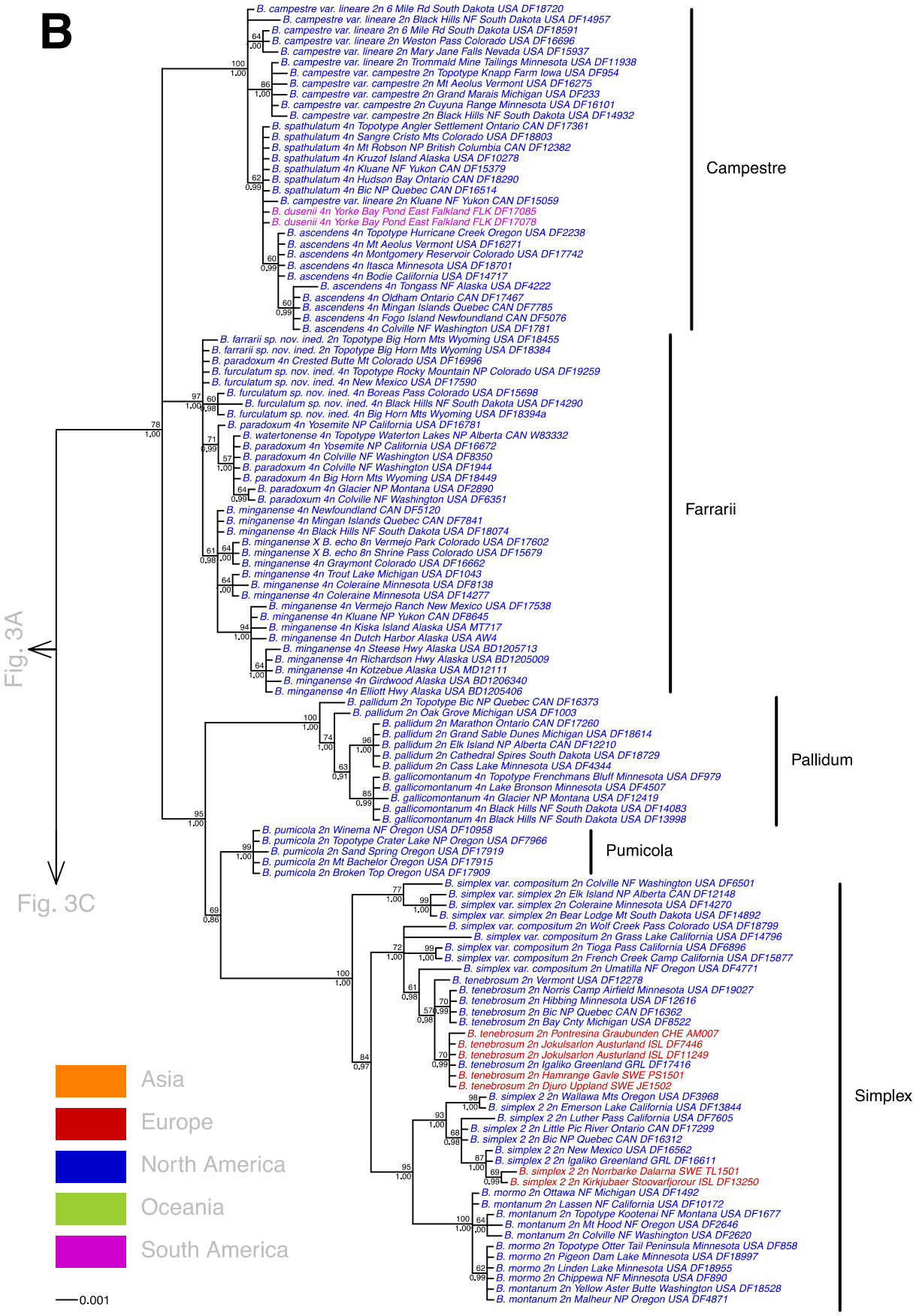
TABLE 2. Characteristics of the four cpDNA regions relative to the three major clades and evolutionary models used according to their AIC and BIC scores. Note: values in parentheses include the three outgroups, two of *Sceptridium multifidum* and one *Botrypus virginianus*; [1]Alignment with the three outgroups; [2]Parsimony informative sites; [3]Single nucleotide polymorphism; [4]Best evolutionary model selected with jModelTest2.

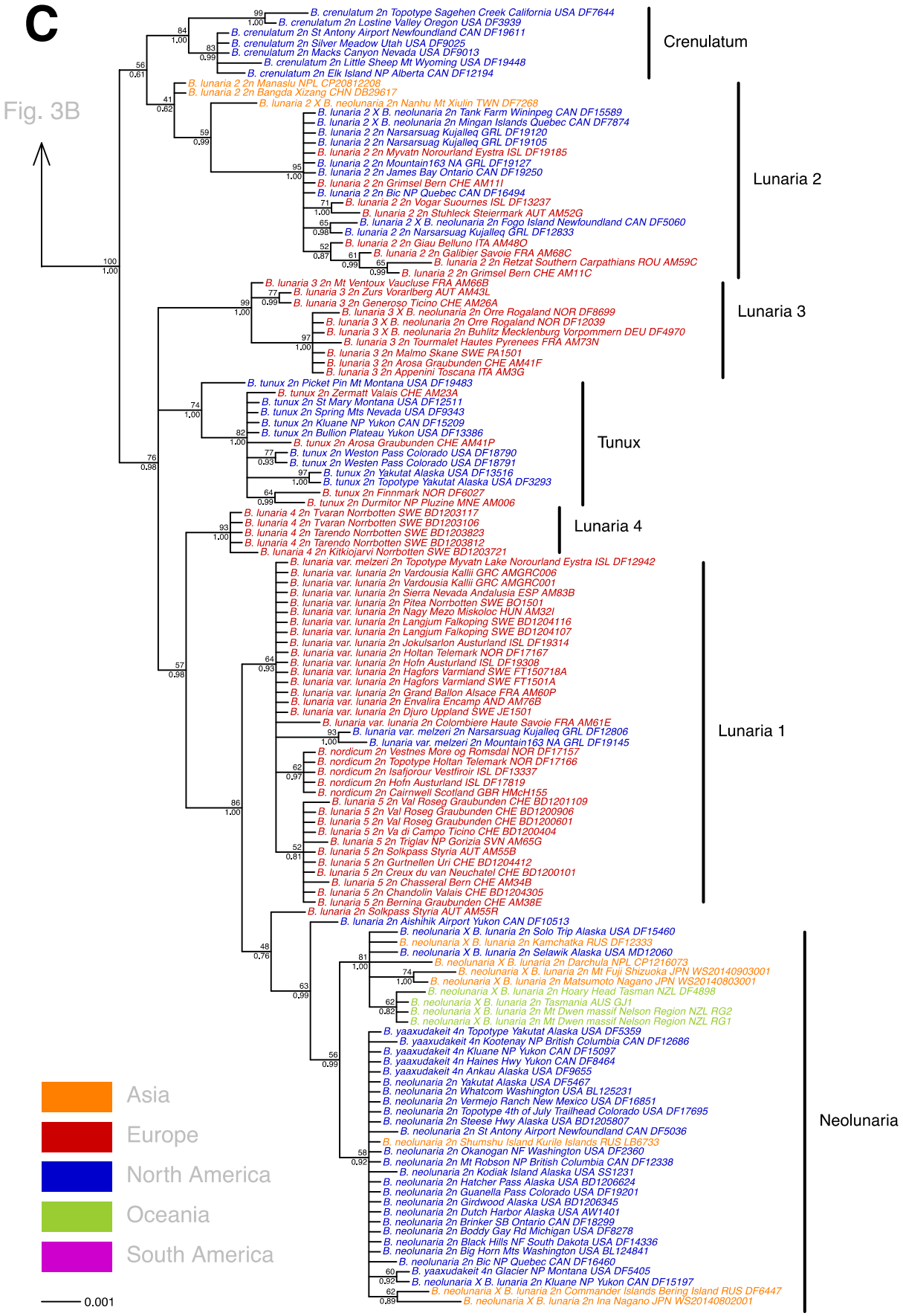
			Lanceolatum clade			Lunaria clade			Simplex-Campestre clade			Three major clades			
cpDNA region	Aligned length[1] (bp)	No. Accession	No.–% of variable sites	No–% of PI[2]	No.–% of SNP[3]	No.–% of variable sites	No–% of PI[2]	No.–% of SNP[3]	No.–% of variable sites	No–% of PI[2]	No.–% of SNP[3]	No.–% of variable sites	No–% of PI[2]	No.–% of SNP[3]	Model[4]
matK	657	365 (368)	5–0.8 (42–6.4)	4–0.6 (41–6.2)	1–0.2 (1–0.2)	17–2.6 (55–8.4)	9–1.4 (48–7.3)	8–1.2 (7–1.1)	40–6.1 (71–10.8)	36–5.8 (67–10.2)	4–0.6 (4–0.6)	60–9.1 (85–12.9)	50–7.6 (76–11.6)	10–1.5 (9–1.4)	HKY + G
rpL16	752	364 (367)	7–0.9 (142–18.9)	4–0.5 (97–12.9)	3–0.4 (45–6.0)	22–2.9 (155–20.6)	13–1.7 (109–14.5)	9–1.2 (46–6.1)	34–4.5 (163–21.7)	22–2.9 (113–15.0)	12–1.6 (50–6.6)	73–9.7 (186–24.7)	53–7.0 (134–17.8)	20–2.7 (52–6.9)	HKY + G
trnL^{UAA}– *trnF*^{GAA}	373	365 (368)	5–1.3 (57–15.3)	4–1.1 (41–11.0)	1–0.3 (16–4.3)	12–3.2 (62–16.6)	10–2.7 (47–12.6)	2–0.5 (15–4.0)	25–6.7 (69–18.5)	19–5.1 (50–13.4)	6–1.6 (19–5.1)	51–13.7 (89–23.9)	45–12.1 (72–19.3)	6–1.6 (17–4.6)	GTR + G
trnH^{GUG}– *psbA*	603	365 (368)	5–0.8 (106–17.6)	4–0.7 (56–9.3)	1–0.2 (50–8.3)	26–4.3 (127–21.1)	21–3.5 (80–13.3)	5–0.8 (47–7.8)	38–6.3 (129–21.4)	25–4.1 (77–12.8)	13-2.2 (52–8.6)	77–12.8 (155–25.7)	64–10.6 (107–17.7)	13–2.2 (48–8.0)	GTR + G
All regions	2385	365 (368)	22–0.9 (347–14.5)	16–0.7 (235–9.9)	6–0.2 (112–5.1)	77–3.2 (399–16.2)	53–2.2 (284–11.9)	24–1.0 (115–4.8)	137–5.7 (432–18.1)	102–4.3 (307–12.9)	35–1.5 (125–5.2)	261–10.9 (515–21.6)	212–8.8 (389–16.3)	49–2.1 (126–5.3)	Unlinked

TABLE 3. Number of haplotypes in the major clades relative to their geographical areas.

Geographical areas	Asia			Europe				North America				Oceania	South America	All			
Maternal clade	LAN	LUN	LAN + LUN	LAN	LUN	SIC	LAN + LUN + SIC	LAN	LUN	SIC	LAN + LUN + SIC	LUN	SIC	LAN	LUN	SIC	LAN + LUN + SIC
# *Botrychium* taxa	2	2	4	4	8	2	14	15	5	18	38	1	1	16	11	19	45
n	2	10	12	26	67	9	102	85	49	111	245	4	2	113	130	122	365
# combined haplotypes *matK* + *rpL16* + $trnL^{UAA}$—$trnF^{GAA}$ + $trnH^{GUG}$–*psbA*	1	9	10	8	30	5	43	18	20	58	96	2	1	19	59	62	118
Mean # haplotypes per taxa	0.5	4.5	2.5	2.0	3.8	2.5	3.1	1.2	4.0	3.2	2.5	2.0	1.0	1.2	5.4	3.3	2.6

Fig. 3. A–C. Bayesian tree depicting the maternal relationships among species of the LAN (A), SIC (B) and LUN (C) clade based on the four cpDNA loci. Maximum likelihood bootstrap supports (above nodes) and posterior probabilities (below nodes) are represented for each node. Clades and subclades are named on the right margin, and color indicates the geographical area of each specimen. Ploidy level of each specimen is symbolized by “2n” for diploid, and “4n”, “6n”, and “8n” for polyploid taxa.


B
B. campestre var. lineare 2n 6 Mile Rd South Dakota USA DF18720
B. campestre var. lineare 2n Black Hills NF South Dakota USA DF14957
B. campestre var. lineare 2n 6 Mile Rd South Dakota USA DF18591
B. campestre var. lineare 2n Weston Pass Colorado USA DF16696
B. campestre var. lineare 2n Mary Jane Falls Nevada USA DF15937
B. campestre var. lineare 2n Trommald Mine Tailings Minnesota USA DF11938
B. campestre var. campestre 2n Topotype Knapp Farm Iowa USA DF954
B. campestre var. campestre 2n Mt Aeolus Vermont USA DF16275
B. campestre var. campestre 2n Grand Marais Michigan USA DF233
B. campestre var. campestre 2n Cuyuna Range Minnesota USA DF16101
B. campestre var. campestre 2n Black Hills NF South Dakota USA DF14932
B. spathulatum 4n Topotype Angler Settlement Ontario CAN DF17361
B. spathulatum 4n Sangre Cristo Mts Colorado USA DF18803
B. spathulatum 4n Mt Robson NP British Columbia CAN DF12382
B. spathulatum 4n Kruzof Island Alaska USA DF10278
B. spathulatum 4n Kluane NF Yukon CAN DF15379
B. spathulatum 4n Hudson Bay Ontario CAN DF18290
B. spathulatum 4n Bic NP Quebec CAN DF16514
B. campestre var. lineare 2n Kluane NF Yukon CAN DF15059
B. dusenii 4n Yorke Bay Pond East Falkland FLK DF17085
B. dusenii 4n Yorke Bay Pond East Falkland FLK DF17078
B. ascendens 4n Topotype Hurricane Creek Oregon USA DF2238
B. ascendens 4n Mt Aeolus Vermont USA DF16271
B. ascendens 4n Montgomery Reservoir Colorado USA DF17742
B. ascendens 4n Itasca Minnesota USA DF18701
B. ascendens 4n Bodie California USA DF14717
B. ascendens 4n Tongass NF Alaska USA DF4222
B. ascendens 4n Oldham Ontario CAN DF17467
B. ascendens 4n Mingan Islands Quebec CAN DF7785
B. ascendens 4n Fogo Island Newfoundland CAN DF5076
B. ascendens 4n Colville NF Washington USA DF1781
Campestre
B. farrarii sp. nov. ined. 2n Topotype Big Horn Mts Wyoming USA DF18455
B. farrarii sp. nov. ined. 2n Topotype Big Horn Mts Wyoming USA DF18384
B. paradoxum 4n Crested Butte Mt Colorado USA DF16996
B. furculatum sp. nov. ined. 4n Topotype Rocky Mountain NP Colorado USA DF19259
B. furculatum sp. nov. ined. 4n New Mexico USA DF17590
B. furculatum sp. nov. ined. 4n Boreas Pass Colorado USA DF15698
B. furculatum sp. nov. ined. 4n Black Hills NF South Dakota USA DF14290
B. furculatum sp. nov. ined. 4n Big Horn Mts Wyoming USA DF18394a
B. paradoxum 4n Yosemite NP California USA DF16781
B. watertonense 4n Topotype Waterton Lakes NP Alberta CAN W83332
B. paradoxum 4n Yosemite NP California USA DF16672
B. paradoxum 4n Colville NF Washington USA DF8350
B. paradoxum 4n Colville NF Washington USA DF1944
B. paradoxum 4n Big Horn Mts Wyoming USA DF18449
B. paradoxum 4n Glacier NP Montana USA DF2890
B. paradoxum 4n Colville NF Washington USA DF6351
B. minganense 4n Newfoundland CAN DF5120
B. minganense 4n Mingan Islands Quebec CAN DF7841
B. minganense 4n Black Hills NF South Dakota USA DF18074
B. minganense X B. echo 8n Vermejo Park Colorado USA DF17602
B. minganense X B. echo 8n Shrine Pass Colorado USA DF15679
B. minganense 4n Graymont Colorado USA DF16662
B. minganense 4n Trout Lake Michigan USA DF1043
B. minganense 4n Coleraine Minnesota USA DF8138
B. minganense 4n Coleraine Minnesota USA DF14277
B. minganense 4n Vermejo Ranch New Mexico USA DF17538
B. minganense 4n Kluane NP Yukon CAN DF8645
B. minganense 4n Kiska Island Alaska USA MT717
B. minganense 4n Dutch Harbor Alaska USA AW4
B. minganense 4n Steese Hwy Alaska USA BD1205713
B. minganense 4n Richardson Hwy Alaska USA BD1205009
B. minganense 4n Kotzebue Alaska USA MD12111
B. minganense 4n Girdwood Alaska USA BD1206340
B. minganense 4n Elliott Hwy Alaska USA BD1205406
Farrarii
B. pallidum 2n Topotype Bic NP Quebec CAN DF16373
B. pallidum 2n Oak Grove Michigan USA DF1003
B. pallidum 2n Marathon Ontario CAN DF17260
B. pallidum 2n Grand Sable Dunes Michigan USA DF18614
B. pallidum 2n Elk Island NP Alberta CAN DF12210
B. pallidum 2n Cathedral Spires South Dakota USA DF18729
B. pallidum 2n Cass Lake Minnesota USA DF4344
B. gallicomontanum 4n Topotype Frenchmans Bluff Minnesota USA DF979
B. gallicomontanum 4n Lake Bronson Minnesota USA DF4507
B. gallicomontanum 4n Glacier NP Montana USA DF12419
B. gallicomontanum 4n Black Hills NF South Dakota USA DF14083
B. gallicomontanum 4n Black Hills NF South Dakota USA DF13998
Pallidum
B. pumicola 2n Winema NF Oregon USA DF10958
B. pumicola 2n Topotype Crater Lake NP Oregon USA DF7966
B. pumicola 2n Sand Spring Oregon USA DF17919
B. pumicola 2n Mt Bachelor Oregon USA DF17915
B. pumicola 2n Broken Top Oregon USA DF17909
Pumicola
B. simplex var. compositum 2n Colville NF Washington USA DF6501
B. simplex var. simplex 2n Elk Island NP Alberta CAN DF12148
B. simplex var. simplex 2n Coleraine Minnesota USA DF14270
B. simplex var. simplex 2n Bear Lodge Mt South Dakota USA DF14892
B. simplex var. compositum 2n Wolf Creek Pass Colorado USA DF18799
B. simplex var. compositum 2n Grass Lake California USA DF14796
B. simplex var. compositum 2n Tioga Pass California USA DF6896
B. simplex var. compositum 2n French Creek Camp California USA DF15877
B. simplex var. compositum 2n Umatilla NF Oregon USA DF4771
B. tenebrosum 2n Vermont USA DF12278
B. tenebrosum 2n Norris Camp Airfield Minnesota USA DF19027
B. tenebrosum 2n Hibbing Minnesota USA DF12616
B. tenebrosum 2n Bic NP Quebec CAN DF16362
B. tenebrosum 2n Bay Cnty Michigan USA DF8522
B. tenebrosum 2n Pontresina Graubunden CHE AM007
B. tenebrosum 2n Jokulsarlon Austurland ISL DF7446
B. tenebrosum 2n Jokulsarlon Austurland ISL DF11249
B. tenebrosum 2n Igaliko Greenland GRL DF17416
B. tenebrosum 2n Hamrange Gavle SWE PS1501
B. tenebrosum 2n Djuro Uppland SWE JE1502
B. simplex 2 2n Wallawa Mts Oregon USA DF3968
B. simplex 2 2n Emerson Lake California USA DF13844
B. simplex 2 2n Luther Pass California USA DF7605
B. simplex 2 2n Little Pic River Ontario CAN DF17299
B. simplex 2 2n Bic NP Quebec CAN DF16312
B. simplex 2 2n New Mexico USA DF16562
B. simplex 2 2n Igaliko Greenland GRL DF16611
B. simplex 2 2n Norrbarke Dalarna SWE TL1501
B. simplex 2 2n Kirkjubaer Stoovarfjorour ISL DF13250
B. mormo 2n Ottawa NF Michigan USA DF1492
B. montanum 2n Lassen NF California USA DF10172
B. montanum 2n Topotype Kootenai NF Montana USA DF1677
B. montanum 2n Mt Hood NF Oregon USA DF2646
B. montanum 2n Colville NF Washington USA DF2620
B. mormo 2n Topotype Otter Tail Peninsula Minnesota USA DF858
B. mormo 2n Pigeon Dam Lake Minnesota USA DF18997
B. mormo 2n Linden Lake Minnesota USA DF18955
B. mormo 2n Chippewa NF Minnesota USA DF890
B. montanum 2n Yellow Aster Butte Washington USA DF18528
B. montanum 2n Malheur NP Oregon USA DF4871
Simplex
Asia
Europe
North America
Oceania
South America
0.001
Fig. 3A
Fig. 3C


C
Fig. 3B
B. crenulatum 2n Topotype Sagehen Creek California USA DF7644
B. crenulatum 2n Lostine Valley Oregon USA DF3939
B. crenulatum 2n St Antony Airport Newfoundland CAN DF19611
B. crenulatum 2n Silver Meadow Utah USA DF9025
B. crenulatum 2n Macks Canyon Nevada USA DF9013
B. crenulatum 2n Little Sheep Mt Wyoming USA DF19448
B. crenulatum 2n Elk Island NP Alberta CAN DF12194
Crenulatum
B. lunaria 2 2n Manaslu NPL CP20812208
B. lunaria 2 2n Bangda Xizang CHN DB29617
B. lunaria 2 X B. neolunaria 2n Nanhu Mt Xiulin TWN DF7268
B. lunaria 2 X B. neolunaria 2n Tank Farm Winnipeg CAN DF15589
B. lunaria 2 X B. neolunaria 2n Mingan Islands Quebec CAN DF7874
B. lunaria 2 2n Narsarsuag Kujalleq GRL DF19120
B. lunaria 2 2n Narsarsuag Kujalleq GRL DF19105
B. lunaria 2 2n Myvatn Norourland Eystra ISL DF19185
B. lunaria 2 2n Mountain163 NA GRL DF19127
B. lunaria 2 2n James Bay Ontario CAN DF19250
B. lunaria 2 2n Grimsel Bern CHE AM11I
B. lunaria 2 2n Bic NP Quebec CAN DF16494
B. lunaria 2 2n Vogar Suournes ISL DF13237
B. lunaria 2 2n Stuhleck Steiermark AUT AM52G
B. lunaria 2 X B. neolunaria 2n Fogo Island Newfoundland CAN DF5060
B. lunaria 2 2n Narsarsuag Kujalleq GRL DF12833
B. lunaria 2 2n Giau Belluno ITA AM48O
B. lunaria 2 2n Galibier Savoie FRA AM68C
B. lunaria 2 2n Retzat Southern Carpathians ROU AM59C
B. lunaria 2 2n Grimsel Bern CHE AM11C
Lunaria 2
B. lunaria 3 2n Mt Ventoux Vaucluse FRA AM66B
B. lunaria 3 2n Zurs Vorarlberg AUT AM43L
B. lunaria 3 2n Generoso Ticino CHE AM26A
B. lunaria 3 X B. neolunaria 2n Orre Rogaland NOR DF8699
B. lunaria 3 X B. neolunaria 2n Orre Rogaland NOR DF12039
B. lunaria 3 X B. neolunaria 2n Buhlitz Mecklenburg Vorpommern DEU DF4970
B. lunaria 3 2n Tourmalet Hautes Pyrenees FRA AM73N
B. lunaria 3 2n Malmo Skane SWE PA1501
B. lunaria 3 2n Arosa Graubunden CHE AM41F
B. lunaria 3 2n Appenini Toscana ITA AM3G
Lunaria 3
B. tunux 2n Picket Pin Mt Montana USA DF19483
B. tunux 2n Zermatt Valais CHE AM23A
B. tunux 2n St Mary Montana USA DF12511
B. tunux 2n Spring Mts Nevada USA DF9343
B. tunux 2n Kluane NP Yukon CAN DF15209
B. tunux 2n Bullion Plateau Yukon USA DF13386
B. tunux 2n Arosa Graubunden CHE AM41P
B. tunux 2n Weston Pass Colorado USA DF18790
B. tunux 2n Westen Pass Colorado USA DF18791
B. tunux 2n Yakutat Alaska USA DF13516
B. tunux 2n Topotype Yakutat Alaska USA DF3293
B. tunux 2n Finnmark NOR DF6027
B. tunux 2n Durmitor NP Pluzine MNE AM006
Tunux
B. lunaria 4 2n Tvaran Norrbotten SWE BD1203117
B. lunaria 4 2n Tvaran Norrbotten SWE BD1203106
B. lunaria 4 2n Tarendo Norrbotten SWE BD1203823
B. lunaria 4 2n Tarendo Norrbotten SWE BD1203812
B. lunaria 4 2n Kitkiojarvi Norrbotten SWE BD1203721
Lunaria 4
B. lunaria var. melzeri 2n Topotype Myvatn Lake Norourland Eystra ISL DF12942
B. lunaria var. lunaria 2n Vardousia Kallii GRC AMGRC006
B. lunaria var. lunaria 2n Vardousia Kallii GRC AMGRC001
B. lunaria var. lunaria 2n Sierra Nevada Andalusia ESP AM83B
B. lunaria var. lunaria 2n Pitea Norrbotten SWE BO1501
B. lunaria var. lunaria 2n Nagy Mezo Miskoloc HUN AM32I
B. lunaria var. lunaria 2n Langjum Falkoping SWE BD1204116
B. lunaria var. lunaria 2n Langjum Falkoping SWE BD1204107
B. lunaria var. lunaria 2n Jokulsarlon Austurland ISL DF19314
B. lunaria var. lunaria 2n Holtan Telemark NOR DF17167
B. lunaria var. lunaria 2n Hofn Austurland ISL DF19308
B. lunaria var. lunaria 2n Hagfors Varmland SWE FT150718A
B. lunaria var. lunaria 2n Hagfors Varmland SWE FT1501A
B. lunaria var. lunaria 2n Grand Ballon Alsace FRA AM60P
B. lunaria var. lunaria 2n Envalira Encamp AND AM76B
B. lunaria var. lunaria 2n Djuro Uppland SWE JE1501
B. lunaria var. lunaria 2n Colombiere Haute Savoie FRA AM61E
B. lunaria var. melzeri 2n Narsarsuag Kujalleq GRL DF12806
B. lunaria var. melzeri 2n Mountain163 NA GRL DF19145
B. nordicum 2n Vestnes More og Romsdal NOR DF17157
B. nordicum 2n Topotype Holtan Telemark NOR DF17166
B. nordicum 2n Isafjorour Vestfiroir ISL DF13337
B. nordicum 2n Hofn Austurland ISL DF17819
B. nordicum 2n Cairnwell Scotland GBR HMcH155
B. lunaria 5 2n Val Roseg Graubunden CHE BD1201109
B. lunaria 5 2n Val Roseg Graubunden CHE BD1200906
B. lunaria 5 2n Val Roseg Graubunden CHE BD1200601
B. lunaria 5 2n Va di Campo Ticino CHE BD1200404
B. lunaria 5 2n Triglav NP Gorizia SVN AM65G
B. lunaria 5 2n Solkpass Styria AUT AM55B
B. lunaria 5 2n Gurtnellen Uri CHE BD1204412
B. lunaria 5 2n Creux du van Neuchatel CHE BD1200101
B. lunaria 5 2n Chasseral Bern CHE AM34B
B. lunaria 5 2n Chandolin Valais CHE BD1204305
B. lunaria 5 2n Bernina Graubunden CHE AM38E
B. lunaria 2n Solkpass Styria AUT AM55R
Lunaria 1
B. lunaria 2n Aishihik Airport Yukon CAN DF10513
B. neolunaria X B. lunaria 2n Solo Trip Alaska USA DF15460
B. neolunaria X B. lunaria 2n Kamchatka RUS DF12333
B. neolunaria X B. lunaria 2n Selawik Alaska USA MD12060
B. neolunaria X B. lunaria 2n Darchula NPL CP1216073
B. neolunaria X B. lunaria 2n Mt Fuji Shizuoka JPN WS20140903001
B. neolunaria X B. lunaria 2n Matsumoto Nagano JPN WS20140803001
B. neolunaria X B. lunaria 2n Hoary Head Tasman NZL DF4898
B. neolunaria X B. lunaria 2n Tasmania AUS GJ1
B. neolunaria X B. lunaria 2n Mt Dwen massif Nelson Region NZL RG2
B. neolunaria X B. lunaria 2n Mt Dwen massif Nelson Region NZL RG1
B. yaaxudakeit 4n Topotype Yakutat Alaska USA DF5359
B. yaaxudakeit 4n Kootenay NP British Columbia CAN DF12686
B. yaaxudakeit 4n Kluane NP Yukon CAN DF15097
B. yaaxudakeit 4n Haines Hwy Yukon CAN DF8464
B. yaaxudakeit 4n Ankau Alaska USA DF9655
B. neolunaria 2n Yakutat Alaska USA DF5467
B. neolunaria 2n Whatcom Washington USA BL125231
B. neolunaria 2n Vermejo Ranch New Mexico USA DF16851
B. neolunaria 2n Topotype 4th of July Trailhead Colorado USA DF17695
B. neolunaria 2n Steese Hwy Alaska USA BD1205807
B. neolunaria 2n St Antony Airport Newfoundland CAN DF5036
B. neolunaria 2n Shumshu Island Kurile Islands RUS LB6733
B. neolunaria 2n Okanogan NF Washington USA DF2360
B. neolunaria 2n Mt Robson NP British Columbia CAN DF12338
B. neolunaria 2n Kodiak Island Alaska USA SS1231
B. neolunaria 2n Hatcher Pass Alaska USA BD1206624
B. neolunaria 2n Guanella Pass Colorado USA DF19201
B. neolunaria 2n Girdwood Alaska USA BD1206345
B. neolunaria 2n Dutch Harbor Alaska USA AW1401
B. neolunaria 2n Brinker SB Ontario CAN DF18299
B. neolunaria 2n Boddy Gay Rd Michigan USA DF8278
B. neolunaria 2n Black Hills NF South Dakota USA DF14336
B. neolunaria 2n Big Horn Mts Washington USA BL124841
B. neolunaria 2n Bic NP Quebec CAN DF16460
B. yaaxudakeit 4n Glacier NP Montana USA DF5405
B. neolunaria X B. lunaria 2n Kluane NP Yukon CAN DF15197
B. neolunaria X B. lunaria 2n Commander Islands Bering Island RUS DF6447
B. neolunaria X B. lunaria 2n Ina Nagano JPN WS20140802001
Neolunaria
Asia
Europe
North America
Oceania
South America
0.001

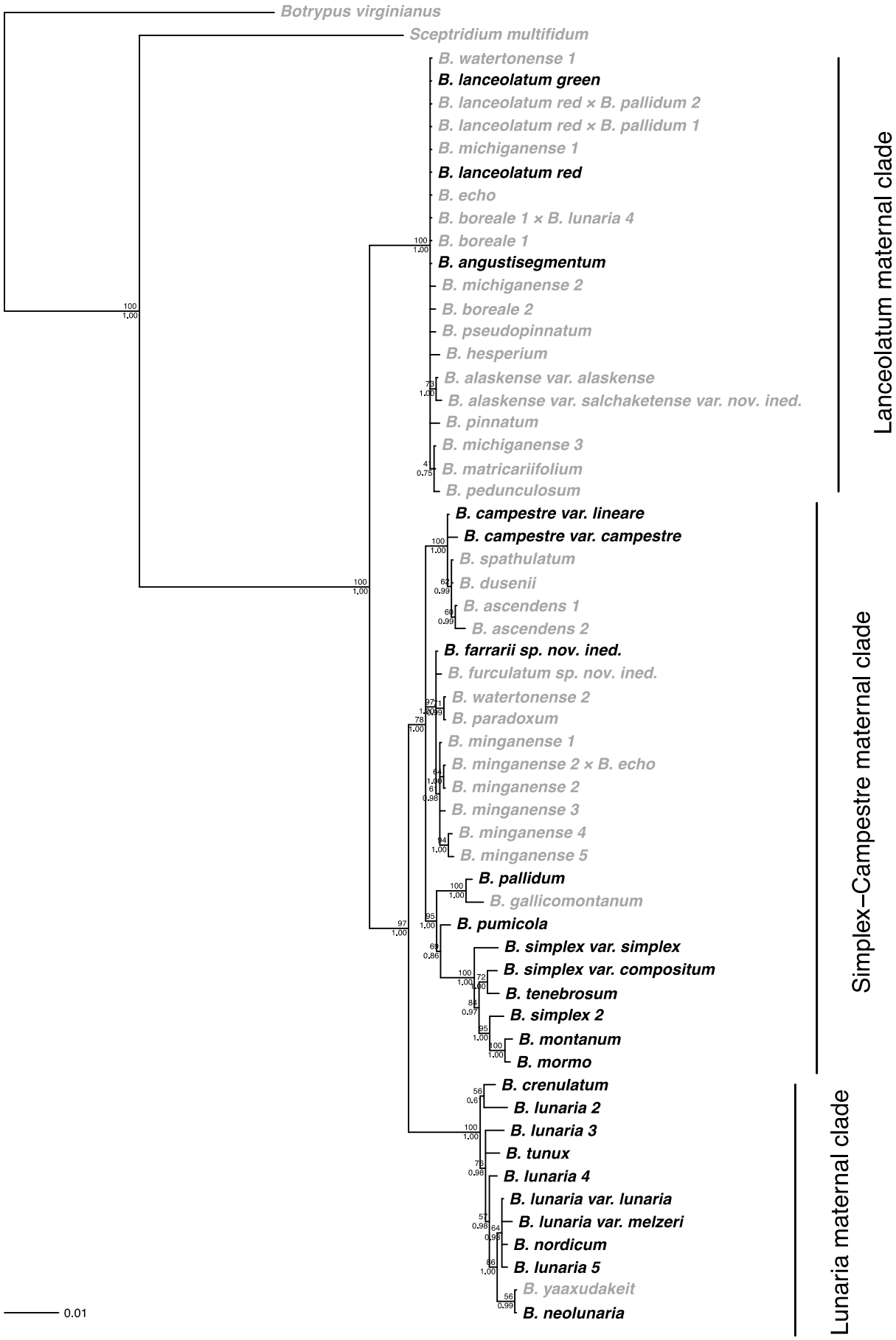

Botrypus virginianus
Sceptridium multifidum
B. watertonense 1
B. lanceolatum green
B. lanceolatum red × B. pallidum 2
B. lanceolatum red × B. pallidum 1
B. michiganense 1
B. lanceolatum red
B. echo
B. boreale 1 × B. lunaria 4
B. boreale 1
B. angustisegmentum
B. michiganense 2
B. boreale 2
B. pseudopinnatum
B. hesperium
B. alaskense var. alaskense
B. alaskense var. salchaketense var. nov. ined.
B. pinnatum
B. michiganense 3
B. matricariifolium
B. pedunculosum
B. campestre var. lineare
B. campestre var. campestre
B. spathulatum
B. dusenii
B. ascendens 1
B. ascendens 2
B. farrarii sp. nov. ined.
B. furculatum sp. nov. ined.
B. watertonense 2
B. paradoxum
B. minganense 1
B. minganense 2 × B. echo
B. minganense 2
B. minganense 3
B. minganense 4
B. minganense 5
B. pallidum
B. gallicomontanum
B. pumicola
B. simplex var. simplex
B. simplex var. compositum
B. tenebrosum
B. simplex 2
B. montanum
B. mormo
B. crenulatum
B. lunaria 2
B. lunaria 3
B. tunux
B. lunaria 4
B. lunaria var. lunaria
B. lunaria var. melzeri
B. nordicum
B. lunaria 5
B. yaaxudakeit
B. neolunaria
0.01
Lanceolatum maternal clade
Simplex–Campestre maternal clade
Lunaria maternal clade

Fig. 4. Simplified maternal tree topology including one representative per diploid and one or several per polyploid if multiple origins are suspected. Maximum likelihood bootstrap supports (above nodes) and posterior probabilities (below nodes) are represented for each node. Ploidy level of each taxon is symbolized in bold black for diploids and in gray for polyploids and outgroups. Individuals chosen are representative of the most frequent haplotype; *B. alaskense* var. *alaskense* (BD1204709), *B. alaskense* var. *salchaketense* var. nov. ined. (BD1206801), *B. angustisegmentum* (DF19039), *B. ascendens* 1 (DF2238), *B. ascendens* 2 (DF4222), *B. boreale* × *B. lunaria 4* (BD12022C), *B. boreale* 1 (BD1202604), *B. boreale* 2 (DF8863), *B. campestre* var. *campestre* (DF14932), *B. crenulatum* (DF9025), *B. dusenii* (DF17078), *B. echo* (DF13047), *B. farrarii* sp. nov. ined. (DF18384), *B. furculatum* sp. nov. ined. (DF18394a), *B. gallicomontanum* (DF12419), *B. hesperium* (DF13071), *B. lanceolatum* "green" (DF8956), *B. lanceolatum* "red" (DF13038), *B. lanceolatum* "red" × *B. pallidum 1* (DF17272), *B. lanceolatum* "red" × *B. pallidum 2* (DF2878), *B. campestre* var. *lineare* (DF18720), *B. lunaria* var. *lunaria* (JE1501), *B. lunaria 2* (DF19250), *B. lunaria 3* (AM26A), *B. lunaria 4* (BD1203823), *B. lunaria 5* (BD1204305), *B. lunaria* var. *melzeri* (DF12806), *B. matricariifolium* (DF19017), *B. michiganense* 1 (DF8268), *B. michiganense* 2 (DF18145), *B. michiganense* 3 (DF16329), *B. minganense* 1 (DF18074), *B. minganense* 2 (DF16662), *B. minganense* 2 × *B. echo* (DF17602), *B. minganense* 3 (DF14277), *B. minganense* 4 (DF17538), *B. minganense* 5 (BD1205009), *B. montanum* (DF2620), *B. mormo* (DF18997), *B. neolunaria* (DF17695), *B. nordicum* (DF13337), *B. pallidum* (DF16373), *B. paradoxum* (DF8350), *B. pedunculosum* (DF5332), *B. pinnatum* (DF2776), *B. pseudopinnatum* (DF18301), *B. pumicola* (DF7966), *B. simplex 2* (DF16312), *B. simplex* var. *simplex* (DF12148), *B. simplex* var. *compositum* (DF15877), *B. spathulatum* (DF17361), *B. tenebrosum* (DF12278), *B. tunux* (DF15209), *B. watertonense* 1 (DF2896), *B. watertonense* 2 (W83332), *B. yaaxudakeit* (DF5359), *Botrypus virginianus* (BD1206330), *Sceptridium multifidum* (BD1205311).

Discussion

In this discussion, we first investigate the phylogenetic relationships of diploid taxa and the maternal origin of allopolyploids within each of the three major clades LAN, LUN, and SIC, and focus on new morphologically cryptic taxa and multiple independent hybridization events. Then, we present several perspectives for further studies on the asymmetry in the maternal parentage of allopolyploids and their phylogeographical history that could reveal complex routes of migration on a broad scale.

Lanceolatum Clade (LAN)

Within *Botrychium* s. s., the first divergence separates the Lanceolatum clade that comprises 17 taxa of which three are diploids and 14 are polyploids with maternal lineage from one of the three diploids (Table 1). Although LAN is the second most diverse clade overall, it has a large poorly resolved polytomy among the diploids (Fig. 3A; Fig. 4). This lack of phylogenetic resolution extends across large geographical areas, including inter-continental disjunctions, while some divergence in Europe is suggested in several clades. We deduce that this lack of differentiation results from recent evolution and migrations, including long-range dispersal, sometimes at the inter-continental scale with migrations from North America to Europe. This large polytomy is also presumably due to recent allopolyploidization events involving one of the three LAN diploids which are differentiated by allozyme analysis (Stensvold 2008). However, more informative nuclear sequence data are critically needed to improve our understanding of these complex phylogenetic relationships.

The Lanceolatum clade includes six resolved and one poorly supported subclade that include derived allopolyploids with LAN maternal parentage (Fig. 3A). Several of these corroborate maternal lineages from particular diploids as inferred from morphology, geography and allozyme analysis: *B. boreale* and *B. pinnatum* parentage by *B. lanceolatum* "green", and *B. pseudopinnatum* parentage by *B. angustisegmentum*. The allotetraploids, *B. matricariifolium* and *B. pedunculosum* W. H. Wagner, cluster in a poorly supported clade without indication of specific diploid parentage, although allozyme profiles strongly suggest *B. angustisegmentum* as the parent of *B. matricariifolium* (Gilman et al. 2015), and *B. lanceolatum* "green" as the parent of *B. pedunculosum* (Farrar 2011). The allotetraploid *B. hesperium* forms a clade without association to a specific diploid, although allozyme evidence suggests *B. lanceolatum* "red" as its diploid parent (Gilman et al. 2015). A small clade of allotetraploid *B. michiganense* (Michiganense Central NA) suggests some differentiation in this taxon; whereas most collections remain embedded in the large polytomy, none indicate a specific diploid parentage. Similarly, the allotetraploid, *B. echo*, the hexaploid, *B. boreale* × *B. lunaria 4* (Dauphin et al. 2016), and the two putative allopolyploids *B. lanceolatum* "red" × *B. pallidum 1* and *B. lanceolatum* "red" × *B. pallidum 2* are embedded in the large polytomy without indication of specific parentage, although allozyme profiles suggest *B. lanceolatum* "red" parentage in all three polyploid taxa (Farrar 2011).

Further analysis of allozyme data from *B. michiganense* (Gilman et al. 2015) gives insight into recent evolution of the LAN clade. This allopolyploid, resulting from hybridization of a LAN diploid and *B. pallidum* W. H. Wagner from the SIC clade, contains LAN-derived alleles now segregated into diploids *B. lanceolatum* "red" and *B. angustisegmentum*, suggesting that the origin of *B. michiganense* may predate differentiation of these two LAN diploids. The range of *B. michiganense* in North America encompasses the range of both western *B. lanceolatum* "red" and eastern *B. angustisegmentum*, and all occurrences are within recently glaciated areas. Furthermore, in our analysis, plants of *B. michiganense* appear in the Matricariifolium clade and in the polytomy as well as the Michiganense Central NA clade, suggesting multiple origins. Together, this suggests a recent differentiation of LAN diploids and multiple origins of allopolyploids, both leading to a complex phylogeny of the LAN clade and its maternally derived allopolyploids.

The appearance of the allopolyploid *B. watertonense* in the LAN clade is informative relative to the asymmetry of maternal parentage in *Botrychium* allopolyploids, since previously, only diploids of the LAN clade were known to be the maternal donor to all named allotetraploids involving LAN diploids, except this one. The parentage of *B. watertonense* is between two allotetraploids, *B. hesperium* and *B. paradoxum* W. H. Wagner (Wagner et al. 1984), and its cytotype(s) could be multiple (see Table 1). Its appearance in the LAN clade is presumably through the maternal parentage of *B. hesperium* for which *B. lanceolatum* "red" is the maternal parent. However, *B. watertonense* also appears in the SIC clade, clustering with *B. paradoxum*, for which *B. farrarii* sp. nov. ined. is the

maternal parent. Our two samples of *B. watertonense* are from Waterton Lakes NP in Canada and adjoining Glacier NP in Montana. Apparently, factors promoting maternal parentage by LAN species may be less effective when present in only one of the four genomes involved.

Simplex–Campestre clade (SIC)

The second major lineage is the most diversified with 20 taxa of which nine are polyploids (Table 1; Fig. 4). With a center of diversity in North America, the SIC clade includes multiple cladogenesis events between and among taxa with five strongly supported subclades (Fig. 3B). It also includes the largest haplotype number in North America (Table 3; Fig. 2). This clade also occurs in Europe (*B. simplex 2* and *B. tenebrosum*), South America (*B. dusenii*), and possibly Asia (Kholia 2012).

Within the Campestre subclade, *B. campestre* [including var. *lineare*] (Farrar and Gilman, 2017) is the sole diploid, accompanied by three allotetraploid species, *B. ascendens* W. H. Wagner (with two potential origins labelled as "1" and "2" in Fig. 4), *B. dusenii*, and *B. spathulatum*. The specimen (DF15059) of *B. campestre* var. *lineare* (W. H. Wagner) Farrar from Yukon (CAN) clustering with *B. spathulatum* and *B. dusenii* is from a small, single genotype population with no allelic distinctions from genetically diverse var. *lineare* populations in the southern Rocky Mountains and likely just one of many occurrences of this and other genetic subsets derived from southern populations (Farrar and Gilman, 2017). The close genetic relationship between varieties *campestre* and *lineare* cautions interpretations of maternal parentage, but Zika and Farrar (2009) found alleles in *B. ascendens* not present in var. *campestre*, and Wagner and Wagner (1990), concluded from morphological and genetic data (contributed by Farrar) that var. *campestre* was the most likely maternal parent of *B. spathulatum*. Studies by Meza-Torres et al. (2017) revealed South American *B. dusenii*, which has likely migrated from North America to South America by long distance dispersal, to be an allotetraploid with a genomic contribution from *B. campestre* var. *campestre*. Our results failed to separate *B. dusenii* from *B. spathulatum* and we recommend treatment of *B. dusenii* and *B. spathulatum* as conspecific.

The Farrarii subclade has a single diploid, *B. farrarii* sp. nov. ined., as the unambiguous maternal parent of the allopolyploid *B. furculatum* sp. nov. ined., *B. minganense*, *B. paradoxum*, and through *B. paradoxum*, the maternal parent of at least one population of four-genome *B. watertonense* (Table 1; Fig. 4). It is interesting to note that *B. minganense* is depicted into several subclades, possibly due to multiple allopolyploid origins in different geographical areas or a differentiation followed by migration after the polyploidization event(s). Further testing of geographic groups could be profitable.

Our tree also includes two specimens from Colorado (USA) identified through morphology and allozyme analysis as probably originating through hybridization between *B. minganense* × *B. echo*. These plants cluster with Colorado *B. minganense* (Fig. 3B; Fig. 4), indicating parentage by that taxon. Because allotetraploid *B. echo* (*B. lanceolatum* "red" × *B. campestre* var. *lineare*) contains the maternal genome of *B. lanceolatum* "red", *B. minganense* × *B. echo* constitutes a second case where

the taxon carrying the "*lanceolatum* genome" is not the maternal donor in allopolyploidization. In both cases (the other being *B. watertonense*) four-genome allopolyploids result from hybridization between allotetraploids.

The Pallidum subclade is comprised of two taxa, the diploid, *B. pallidum*, and the allopolyploid *B. gallicomontanum* with *B. pallidum* as maternal donor. Curiously, an additional seven allopolyploids have actually been formed by *B. pallidum* as paternal donor (*B. furculatum* sp. nov. ined., *B. hesperium*, *B. lanceolatum* "red" × *B. pallidum 1*, *B. lanceolatum* "red" × *B. pallidum 2, B. matricariifolium*, *B. michiganense*, and *B. peduculosum*) (Farrar 2011; Gilman et al. 2015; Williams et al. 2016), which again supports the asymmetric contribution in parental donors forming allopolyploids.

The Pumicola clade includes only *B. pumicola* Colville, a species peculiar in its restricted geographical distribution and habitat preference on unconsolidated volcanic pumice in intermountain basins, and well-drained glacial moraines on volcanic mountains in southern Oregon and northern California (Farrar 2011). Two haplotypes, haplotype 1 with a frequency of 0.80 and haplotype 2 of 0.20 (see Appendix 2), constitute the current diversity found in these five populations. This low genetic diversity corroborates the hypothesis of a recent bottleneck occurring in this taxon through periodic volcanic creation and obliteration of its habitat.

The strongly supported Simplex subclade has six diploid taxa, three of which are monophyletic (*B. simplex* var. *simplex, B. simplex 2*, and *B. tenebrosum*), and three that are paraphyletic (*B. montanum, B. mormo,* and *B. simplex* var. *compositum* (Lasch) Milde) (Fig. 3B). *Botrychium simplex 2* and *B. tenebrosum,* DNA-identified for the first time from Europe, each form two monophyletic clades that are strongly supported by statistical values. *Botrychium tenebrosum* appears to be the Alpine taxon described as *B. simplex* in many regional floras and further study is necessary to determine the distributions of *B. tenebrosum* and *B. simplex 2* as well as the presence or absence of varieties *compositum* and *simplex* in Europe. Early divergent positions of North American populations of *B. simplex 2* and *B. tenebrosum* suggest European origins from North America, but the relationships are complicated.

The earliest diverging lineage of the Simplex subclade is *B. simplex* var. *simplex* and var. *compositum*, which occur in eastern North America, extending westward across the Great Plains to the Rocky Mountains. *Botrychium simplex* var. *compositum* has a characteristically "composite" or a ternately divided trophophore, and occurs in western North America. *Botrychium tenebrosum* has an inter-continental distribution ranging from Eastern North America to Europe, which includes four distinct haplotypes with variable frequencies (Appendix 2). *Botrychium simplex 2* also occurs in both North America and Europe, but its plants are large and fleshy occuring on highly calcareous seeps in California, and initially by Farrar (2012) as a putative new species based on allozyme genetics,

morphology and habitat. Eastern North American representatives of this clade are similar morphologically to western plants, including their larger size and greater ternation of basal pinnae than in *B*. *simplex* var. *simplex*, but they lack much of the allelic diversity and distinctive alleles of that taxon, perhaps due to founder effect. The paraphyly of *B*. *montanum* and *B*. *mormo* found here corroborates the low genetic differentiation between populations of both taxa investigated through allozyme genetics (Farrar 2011). Further study is required to determine whether characters other than range and habitat warrant continued taxonomic recognition of both taxa.

Three diploid *Botrychium* species, *B*. *campestre*, *B*. *pallidum* and *B*. *pumicola*, all in the SIC clade, are capable of vegetative reproduction via gemmae produced in leaf axils of their underground stems (Farrar and Johnson-Groh 1990). The distribution of these species in the SIC clade suggests that gemma production could be a synapomorphy of the clade, which has been subsequently lost in the Farrarii and Simplex subclades.

Lunaria Clade (LUN)

A major finding of our study is the considerable genetic diversity identified among taxa within the LUN clade. This third major lineage encompasses 11 taxa, including ten diploids and a single polyploid (*B*. *yaaxudakeit*) distributed in seven strongly supported subclades (Fig. 3C). The LUN clade is widespread in the northern hemisphere and disjunct to Oceania (Fig. 2). Our dense sampling in the major European mountain chains (Alps, Appenines, Balkans, Black Forest, Carpathians, Greece, Jura, Massif Central, Pyrenées, Sierra Nevada, and Vosges), North America, and Scandinavia, as well as partial sampling from Asia, supports its diversity center in continental Europe where different haplotypes with deep phylogenetic divergences are found within sympatric populations (Maccagni et al. 2017). Nevertheless, the geographical origin of the common ancestor of all representatives of this clade remains unknown and further investigations in the Atlas Mountains in North Africa, and the major mountain ranges in Asia are essential to resolve this issue.

The Crenulatum subclade is restricted to North America with two geographical clusters, one in the Sierra, Cascade and coastal mountains of the west coast, and the second distributed discontinuously from the Rocky Mountains to Newfoundland in eastern North America (Fig. 3C). Genetic similarity of *B*. *crenulatum* W. H. Wagner to the North American clade of *B*. *lunaria* led Stensvold (2008) to consideration of *B*. *crenulatum* as a variety of *B*. *lunaria*. However, closer examination of Stensvold's data relative to our results of multiple clades within traditional *B*. *lunaria* indicate that the haplotype divergence between *B*. *crenulatum* and the European clade containing the type locality of *B*. *lunaria*, shows a much higher differentiation than between it and the North American clade. Consequently, Farrar (2012) and Stensvold and Farrar (2016) accept the species status of *B*. *crenulatum*.

Lunaria 2 is a subclade geographically widely distributed across northern North America to the European Alps, the Carpathian Mountains, Taiwan, and the Himalayan Mountains (Nepal, Tibet). The deeply divergent phylogenetic placement of specimens from Asia may provide data relevant to the origin of the Lunaria 2 clade, and possibly to the entire LUN clade. Further sampling throughout mountainous, steppe and boreal areas of Asia is critically needed.

The Lunaria 3 subclade is exclusively European and sister to the Tunux subclade (Fig. 3C). The clade is centered in the European Alps, and the three samples from northern Europe were identified by Stensvold (2008) as containing alleles of both *B. lunaria* and *B. neolunaria* and thus possibly derived by introgression between these two species. Representatives from this subclade have a morphology similar to *B. lunaria*, yet ecologically tend to occur on limestone soils in grasslands dominated by *Sesleria* or *Elymus* (alkaline soil) whereas typical *B. lunaria* tends to occur on those dominated by *Nardus* (acidic soil) (Maccagni et al. 2017).

The Tunux subclade includes a single taxon, *B. tunux* Stensvold and Farrar, which is here reported for the first time in the Alps and Balkans. This discovery extends its geographical area over several thousand kilometers southward from its previously known localities in Scandinavia. Two of its subgroups depict a geographical clustering, one from Alaska (Topotype) and another from Colorado, while the specimen from Montenegro shows greatest genetic similarity to the specimen from Norway.

The strongly supported Lunaria 4 is exclusively restricted to northern Sweden and the five specimens from three sites constituting this clade share a morphology of largely incised pinnae distinct from *B. lunaria* var. *lunaria*. This observation merits further investigation for possible recognition of *B. lunaria 4* as a new taxon. Interestingly, other *B. lunaria* specimens from southern Scandinavia with typical *lunaria* morphology occur within the Lunaria 1 and Lunaria 3 subclades, indicating the presence of several distinct taxa in Scandinavia.

The Lunaria 1 subclade includes three taxa (*B. lunaria* var. *lunaria*, *B. lunaria* var. *melzeri*, and *B. nordicum*) and putatively a fourth (*B. lunaria 5*), all occurring in Europe and Greenland. *Botrychium lunaria* var. *lunaria* occurs from Iceland to Scandinavia, the European Alps, Balkans and Greece. This clade, containing most of our samples from Sweden, likely includes the type locality of *B. lunaria*, although the exact locality of the type specimen for the species is unknown (Jonsell and Jarvis 1994). The taxon *B. nordicum* forms a subclade and its identification is reported for the first time from much further south in Scotland.

Within the Neolunaria subclade, we found two distinct lineages, the first with a broad distribution in North America (including the Topotype) and northern Asia, the second from Alaska to Kamchatka, Japan, Nepal, Tasmania and New Zealand. *Botrychium neolunaria* as described by Stenvold and Farrar (2016) has remarkably low genetic variability with several unique alleles that have not been detected outside of North America. Testing plants from Oceania, Stensvold (2008) reported plants

from Alaska, Japan, and New Zealand had combined alleles from both *B. lunaria* var. *lunaria* (Lunaria 2) and *B. neolunaria* (indicated as "X" in Fig. 3C), presumably having been derived through hybridization in Alaska. Thus, it is likely that all plants in Oceania are introgressed with maternal parentage by *B. neolunaria* (Stensvold and Farrar, 2016).

The single interclade allopolyploid in the LUN clade, *B. yaaxudakeit*, derived through hybridization between *B. neolunaria* and *B. lunaria* var. *lunaria*, clusters with the broad North American group of *B. neolunaria*, indicating maternal parentage by *B. neolunaria*. Our results suggest two formations of *B. yaaxudakeit*, one in Alaska and a second in Montana.

Cryptic Groups and Polyploidy Implications

Our results support the taxonomic delineation of nearly all *Botrychium* taxa previously described on the basis of morphology, chloroplast, and nuclear data. These results have also revealed several new morphologically cryptic taxa as well as substantial genetic diversity hidden within common species. In total, 47 taxa are identified in this new molecular phylogeny based on multiple accessions per taxon, and sampled throughout their known geographical distribution.

Our study also strongly corroborates previously hypothesized maternal parentage of allopolyploid taxa, and provides support for multiple independent hybridization events generating the same allopolyploid taxon. Strong asymmetry in maternal parentage of allopolyploids by certain diploid species raises questions and may provide clues to understanding of how the process of polyploidy takes place within natural populations. With most allopolyploids originating through interclade hybridization, our data also reveals possible bias against polyploid formation through close intraclade hybridization. Sexual development of *Botrychium* gametophytes, including pheromonal interactions (Atallah and Banks 2015; Tanaka et al. 2014), and chromosome pairing behavior may be fruitful in explaining these asymmetric patterns.

Our broad sampling, combined with phylogenetic analyses, also provides evidence of long-distance dispersal ability of *Botrychium* taxa on intra- and inter-continental scales, as well as probable directions of dispersal. Vectors of migration undoubtedly involve wind dispersal of spores and probable amphitropical bird migrations from the Northern to Southern Hemisphere (Farrar and Stensvold in press). Transcontinental migrations patterns in the northern hemisphere are likely intimately associated with Pleistocene glaciations. Further studies are needed to retrace the phylogeography history of the genus, especially in Asia where the taxonomic diversity remains little known.

Acknowledgments

We greatly thank members of the Evolutionary Botany lab at the Université de Neuchâtel (Olivier Bachmann, Marco Fracassetti, Julien Vieu, and Guillaume Wos), and members of the Rothfels lab at the University of California at Berkeley (Forrest Freund, Ingrid Jordon-Thaden, and Carl Rothfels) for their relevant ideas, criticisms and suggestions on this work. We are also grateful to collectors and colleagues worldwide for their precious help in collecting material this study: Alaska (Mike Duffy MD, Stacy Studebaker SS, and Abi Woodbridge AW), Hungary (Jozsef Sulyok), Japan (Wataru Shinohara WS), Romania (Claudia Danau), Russia (Peter Zika LB, MT), Scotland (Heather McHaffie HMcH), New Zealand (Shannel Courtney RG), Slovenia (Amadej Trnkoczy), Sweden (Pontus Axén PA, Joakim Ekman JE, Tomas Ljung TL, Birgitta Öster BO, Peter Ståhl PS, Frida Turner FT, and Sture Westerberg SW), Tasmania (Greg Jordan GJ), and in Washington state (Ben Legler BL). The abbreviations for these collectors corresponds to the collection number in Table 1. These collections complement the majority of collections made by Alessio Maccagni (AM), Benjamin Dauphin (BD), and Don Farrar (DF). We also acknowledge the Centre national de données et d'informations sur la flore de Suisse (Info Flora) for providing precise locality information of species in Switzerland. We are thankful to Mary Endress (Universität Zürich) and several anonymous reviewers for their helpful comments and constructive criticism. This study was supported by a grant from the Swiss National Science Foundation – 31003A_156456.

In compliance with ethical standards, this research was funded by Swiss National Science Foundation grant 31003A_156456 to J.R. Grant. The authors declare that they have no conflict of interest.

Literature cited

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267–281 in *Proceedings of the second international symposium on information theory*. eds. B.N. Petrov and F. Csaki. Budapest, Hungary: Akademiai Kiado.

Atallah, N. M. and J. A. Banks. 2015. Reproduction and the pheromonal regulation of sex type in fern gametophytes. *Frontiers in Plant Science* 6: 1–6.

Ballard, H. E. and K. J. Sytsma. 2000. Evolution and biogeography of the woody Hawaiian violets (*Viola*, Violaceae): Arctic origins, herbaceous ancestry and bird dispersal. *Evolution* 54: 1521–1532.

Bickford, D., D. J. Lohman, N. S. Sodhi, P. K. Ng, R. Meier, K. Winker, K. I. Krista, and I. Das. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22: 148–155.

Bozukov, V., B. Tsenov, and M. Vatsev. 2010. A first find of *Botrychium* (Ophioglossaceae) in Bulgarian palaeomacroflora. *Comptes Rendus de l'Académie Bulgare des Sciences* 63: 889–892.

Clausen, R. T. 1938. A monograph of the Ophioglossaceae. *Memoirs of the Torrey Botanical Club* 19: 1–177.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods* 9: 772–772.

Dauphin, B., J. R. Grant, and P. Mráz. 2016. Ploidy level and genome size variation in the homosporous ferns *Botrychium* sl (Ophioglossaceae). *Plant Systematics and Evolution* 302: 575–584.

Dauphin, B., J. Vieu, and J. R. Grant. 2014. Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* s.s., Ophioglossaceae). *American Journal of Botany* 101: 128–140.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.

Farrar, D. R. 2011 (last updated 15 Dec. 2011). Systematics and taxonomy of the genus *Botrychium*. http://www.public.iastate.edu/~herbarium/botrychium/Moonwort-Systematics.pdf. (accessed 01 May 2016).

Farrar, D. R. 2012. Ophioglossaceae. Pp. 115–117 in *The Jepson manual vascular plants of California*. eds. B. G. Baldwin, D. H. Goldman, D. J. Keil, R. Patterson, T. J. Rosatti, and D. H. Wilken. Berkeley: University of California Press.

Farrar, D. R. and A.V. Gilman. 2017. Relationships in the *Botrychium campestre* (Ophioglossaceae) complex. *Brittonia* 69: 265–275.

Farrar, D. R., A.V. Gilman, and R. C. Moran. 2017. Ophioglossaceae. Online in *Gleason & Cronquist manual of the vascular plants of the northeastern United States and Adjacent Canada*. eds. R. F. C. Naczi, J. R. Abbott, and Collaborators. New York: New York Botanical Garden.

Farrar, D. R. and C. L. Johnson-Groh. 1990. Subterranean sporophytic gemmae in moonwort ferns, *Botrychium* subgenus *Botrychium*. *American Journal of Botany* 77: 1168–1175.

Farrar, D. R. and C. L. Johnson-Groh. 1991. A new prairie moonwort (*Botrychium* subgenus *Botrychium*) from northwestern Minnesota. *American Fern Journal* 81: 1–6.

Farrar, D. R. and S. J. Popovich. 2012. Ophioglossaceae. Pp. 24–35 in *Colorado flora-eastern slope – a field guide to the vascular plants*. eds. W. A. Weber and R. C. Wittman. Boulder: University Press of Colorado.

Farrar, D. R. and M. C. Stensvold. In press. Observations on bipolar disjunctions of moonwort ferns (Botrychium, Ophioglossaceae). *American Journal of Botany* XX: XX–XX.

Gastony, G. J. and G. Yatskievych. 1992. Maternal inheritance of the chloroplast and mitochondrial genomes in cheilanthoid ferns. *American Journal of Botany* 79: 716–722.

Gilman, A. V., D. R. Farrar, and P. F. Zika. 2015. *Botrychium michiganense* sp. nov. (Ophioglossaceae), a new North American moonwort. *Journal of the Botanical Research Institute of Texas* 9: 295–309.

Haroni, S. A., M. Alifriqui, and V. Simonneaux. 2009. Recent dynamics of the wet pastures at Oukaimeden plateau (High Atlas mountains, Morroco). *Biodiversity and Conservation* 18: 167–189.

Hauk, W. D. 1995. A molecular assessment of relationships among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). *American Fern Journal* 85: 375–394.

Hauk, W. D., L. Kennedy, and H. M. Hawke. 2012. A phylogenetic investigation of *Botrychium* s.s. (Ophioglossaceae): Evidence from three plastid DNA sequence datasets. *Systematic Botany* 37: 320–330.

Hauk, W. D., C. R. Parks, and M. W. Chase. 2003. Phylogenetic studies of Ophioglossaceae: Evidence from *rbcL* and *trnL-F* plastid DNA sequences and morphology. *Molecular Phylogenetics and Evolution* 28: 131–151.

Jonsell, B. and C. E. Jarvis. 1994. Lectotypifications of Linnaean names for Flora Nordica Vol. 1 (Lycopodiaceae – Papaveraceae). *Nordic Journal of Botany* 14: 145–164.

Kato, M. 1987. A phylogenetic classification of Ophioglossaceae. *Garden's Bulletin Singapore* 40: 1–14

Katoh, K. and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Mentjies, and A. Drummond. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.

Kholia, B. S. 2012. *Botrychium simplex* E. Hitchcock: A new moonwort for the Indian Himalayan Mountains. *American Fern Journal* 102: 86–90.

Lewis, L. R., E. Behling, H. Gousse, E. Qian, C. S. Elphick, J. F. Lamarre, J. Bêty, J. Liebezeit, R. Rozzi, and B. Goffinet. 2014. First evidence of bryophyte diaspores in the plumage of transequatorial migrant birds. *PeerJ* 2: e424.

Librado, P. and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.

Maccagni, A., C. Parisod, and J. R. Grant. 2017. Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. *Alpine Botany* 127: 185–196.

Meza-Torres, E. I. and M. S. Ferrucci. 2013. Ophioglossaceae: An exomorphological approach with emphasis in South American species. *Indian Fern Journal* 30: 192–205.

Meza-Torres, E. I., M. C. Stensvold, D. R. Farrar, and M. S. Ferrucci. 2017. Circumscription of the South American moonwort *Botrychium* (Ophioglossaceae). *Plant Biosystems* 151: 1–11.

Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop* (GCE). IEEE: 1–8.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

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

Piñeiro, R., M. Popp, K. Hassel, D. Listl, K. B. Westergaard, K. I. Flatberg, H. K. Stenoien, and C. Brochmann. 2012. Circumarctic dispersal and long- distance colonization of South America: The moss genus *Cinclidium*. *Journal of Biogeography* 39: 2041–2051.

Posada, D. and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53: 793–808.

PPG I. 2016. A community-derived classification for extant lycopods and ferns. *Journal of Systematic and Evolution* 54: 563–603.

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

R Development Core Team. 2015 (last updated 10 December 2015). R: A language and environment for statistical computing. Version 3.2. R Foundation for Statistical Computing. Vienna. https://www.r-project.org (accessed 01 January 2016).

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Rothfels, C. J., F. W. Li, E. M. Sigel, L. Huiet, A. Larsson, D. O. Burge, M. Ruhsam, M. Deyholos, D. E. Soltis, N. C. Stewart, S. W. Shaw, L. Pokorny, T. Chen, C. dePamphilis, L. DeGironimo, L. Chen, X. Wei, X. Sun, P. Korall, D. W. Stevenson, S. W. Graham, G. K-S. Wong, and K. M. Pryer. 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *American Journal of Botany* 102: 1089–1107.

Rothwell, G. W. and R. A. Stockey.1989. Fossil Ophioglossales in the Paleocene of Western North America. *American Journal of Botany* 76: 637–644.

Sandoz, F. A. 2016. Influence of ecological factors on arbuscular mycorrhizal fungi (AMF) communities of *Botrychium lunaria*. Master's Dissertation, University of Neuchâtel, Switzerland.

Sang, T., D. Crawford, and T. Stuessy. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120–1136 .

Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling, and R. L. Small. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.

Shinohara, W., N. Nakato, Y. Yatabe-Kakugawa, T. Oka, J. K. Kim, N. Murakami, and N. Sahashi. 2013. The use of *matK* in Ophioglossaceae phylogeny and the determination of *Mankyua* chromosome number shed light on chromosome number evolution in Ophioglossaceae. *Systematic Botany* 38: 564–570.

Schönswetter, P., R. Elven, and C. Brochmann. 2008. Trans-Atlantic dispersal and large-scale lack of genetic structure in the circumpolar, arctic-alpine sedge *Carex bigelowii* s.l. (Cyperaceae). *American Journal of Botany* 95: 1006–1014.

Simmons, M. P. and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: 369–381.

Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: Choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* 85: 1301–1315.

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

Stensvold, M. C. 2008. A taxonomic and phylogeographic study of the *Botrychium lunaria* complex. Dissertation. Ames: Iowa State University.

Stensvold, M. C. and D. R. Farrar. 2016. Genetic diversity in the worldwide *Botrychium lunaria* (Ophioglossaceae) complex, with new species and new combinations. *Brittonia* 1–28.

Stensvold, M. C., C. L. Johnson-Groh, and D. R. Farrar. 2002. Two new species of moonworts (*Botrychium* subg. *Botrychium*) in Alaska. *American Fern Journal* 92: 150–160.

Sun, B. Y., T. G. Baek, Y. D. Kim, and C. S. Kim. 2009. Phylogeny of the family Ophioglossaceae with special emphasis on genus *Mankyua*. *Korean Journal of Plant Taxonomy* 39: 135–142.

Sun, B. Y., M. H. Kim, C. H. Kim, and C. W. Park. 2001. *Mankyua* (Ophioglossaceae): A new fern genus from Cheju Island, Korea. *Taxon* 50: 1019-1024.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

Tanaka, J., K. Yano, K. Aya, K. Hirano, S. Takehara, E. Koketsu, R. L. Ordonio, S. H. Park, M. Nakajima, M. Ueguchi-Tanaka, and M. Matsuoka. 2014. Antheridiogen determines sex in ferns via a spatiotemporally split gibberellin synthesis pathway. *Science* 346: 469–473.

Tate, J. A. and B. B. Simpson. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* 28: 723–737.

Victorin, M. 1927. Sur un *Botrychium* nouveau de la flore Americaine et ses rapports avec le *B. simplex*. *Transactions Royal Canada Institute* 5: 319–340

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

Wagner, F. S. 1993. Chromosomes of North American grapeferns and moonworts (Ophioglossaceae: *Botrychium*). *Contributions from the University of Michigan Herbarium* 19: 83–92.

Wagner, W. H. and J. R. Grant. 2002. *Botrychium alaskense*, a new moonwort from the Interior of Alaska. *American Fern Journal* 92: 164–170.

Wagner, W. H. and L. P. Lord. 1956. The morphological and cytological distinctness of *Botrychium minganense* and *B. lunaria* in Michigan. *Bulletin of the Torrey Botanical Club* 83: 261–280.

Wagner, W. H. and J. T. Mickel. 2004. *Botrychium tolucense*. Pp. 164 in *Pteridophytes of Mexico*. eds. J. T. Mickel and A. R. Smith. New York: Mem. New York Bot. Gard.

Wagner, W. H. and F. S. Wagner. 1990. Notes on the fan-leaflet group of moonworts in North America with descriptions of two new members. *American Fern Journal* 80: 73–81.

Wagner, W. H., F. S. Wagner, C. Haufler, and J. K. Emerson. 1984. A new nothospecies of moonwort (Ophioglossaceae, *Botrychium*). *Canadian Journal of Botany* 62: 629–634.

Williams, E. W., D. R. Farrar, and D. Henson. 2016. Cryptic speciation in allotetraploids: Lessons from the *Botrychium matricariifolium* complex. *American Journal of Botany* 103: 740–753.

Williams, E. W. and D. M. Waller. 2012. Phylogenetic placement of species within the genus *Botrychium* s.s. (Ophioglossaceae) on the basis of plastid sequences, amplified fragment length polymorphisms, and flow cytometry. *International Journal of Plant Sciences* 173: 516–531.

Zika, P. F. and D. R. Farrar. 2009. *Botrychium ascendens* W.H. Wagner (Ophioglossaceae) in Newfoundland and notes on its origin. *American Fern Journal* 99: 249–259.

Appendix 1

Locality data for *Botrychium* specimens included in this study with collection date, voucher number, herbarium acronym, and Genbank accessions for *matK*, *rpL16*, *trnH-psbA*, and *trnL-trnF* region. New accessions are indicated with an asterisk and a dash indicates missing data.

Outgroup: *Botrypus virginianus* (L.) A.Michx, USA, Alaska, 2012, BD1206330, NEU, KY191410*, KF700540, KY192273*, KF700455; *Sceptridium multifidum* (S.G.Gmelin) M.Nishida, USA, Alaska, 2012, BD1205311, NEU, KY191411*, KF700542, KF700625, KF700457; *Sceptridium multifidum* (S.G.Gmelin) M.Nishida, USA, Alaska, 2012, BD1205503, NEU, KY191412*, KF700463, KF700626, KF700458; ***Botrychium* s.s.**: *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, USA, Alaska, 2012, BD1204709, NEU, KY191051*, KF700462, KF700546, KF700380; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, USA, Alaska, 2012, BD1205308, NEU, KY191052*, KF700461, KF700545, KF700379; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, USA, Alaska, 2012, BD1205703, NEU, KY191050*, KY191413*, KF700547, KF700381; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, USA, Alaska, 2012, BD1206705, NEU, KY191048*, KF700460, KF700544, KF700378; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, USA, Alaska, 2004, DF11430, ISC, KY191045*, KY191414*, KY191987*, KY191698*; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, CAN, Yukon, 2007, DF15116, ISC, KY191047*, KY191415*, KY191988*, KY191699*; *B*. *alaskense* W.H.Wagner & J.R.Grant var. alaskense, RUS, Kurile Islands, 2014, LB6732, NEU, KY191049*, KY191416*, KY191989*, KY191700*; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *alaskense*, USA, Alaska, 2012, MD12077, NEU, KY191046*, KF700459, KF700543, KF700377; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *salchaketense* J.R.Grant, var. nov. ined., USA, Alaska, 2012, BD1206801, NEU, KY191053*, KF700464, KF700548, KF700382; *B*. *alaskense* W.H.Wagner & J.R.Grant var. *salchaketense* J.R.Grant, var. nov. ined., USA, Alaska, 2012, BD1206803, NEU, KY191054*, KF700465, KF700549, KF700383; *B*. *angustisegmentum* (Pease & A.H.Moore) Fernald, CAN, Alberta, 2005, DF12205, ISC, KY191057*, KY191417*, KY191990*, KY191701*; *B*. *angustisegmentum* (Pease & A.H.Moore) Fernald, USA, Minesota, 2012, DF18959, ISC, KY191058*, KF700474, KF700558, KF700392; *B*. *angustisegmentum* (Pease & A.H.Moore) Fernald, USA, Wisconsin, 2012, DF19039, ISC, KY191055*, KF700473, KF700557, KF700391; *B*. *angustisegmentum* (Pease & A.H.Moore) Fernald, USA, Michigan, 2002, DF7951, ISC, KY191059*, KY191418*, KY191991*, KY191702*; *B*. *angustisegmentum* (Pease & A.H.Moore) Fernald, USA, Ohio, 2002, DF7988, ISC, KY191056*, KY191419*, KY191992*, KY191703*; *B*. *ascendens* W.H.Wagner, USA, California, 2006, DF14717, ISC, KY191060*, KY191420*, KY191993*, KY191704*; *B*. *ascendens* W.H.Wagner, USA, Vermont, 2008, DF16271, ISC, KY191066*, KY191421*, KY191994*, KY191705*; *B*. *ascendens* W.H.Wagner, CAN, Ontario, 2009, DF17467, ISC, KY191067*, KY191422*, KY191995*, KY191706*; *B*. *ascendens* W.H.Wagner, USA, Colorado, 2009, DF17742, ISC, KY191065*, KY191423*, KY191996*, KY191707*; *B*. *ascendens* W.H.Wagner, USA, Washington, 1997, DF1781, ISC, KY191061*, KY191424*, KY191997*, KY191708*; *B*. *ascendens* W.H.Wagner, USA, Minnesota, 2011, DF18701, ISC, KY191063*, KY191425*, KY191998*, KY191709*; *B*. *ascendens* W.H.Wagner, USA, Oregon, 1997, DF2238, ISC, KY191069*, KY191426*, KY191999*, KY191710*; *B*. *ascendens* W.H.Wagner, USA, Alaska, 1999, DF4222, ISC, KY191068*, KY191427*, KY192000*, KY191711*; *B*. *ascendens* W.H.Wagner, CAN, Newfoundland, 2001, DF5076, ISC, KY191062*, KY191428*, KY192001*, KY191712*; *B*. *ascendens* W.H.Wagner, CAN, Quebec, 2002, DF7785, ISC, KY191064*, KY191429*, KY192002*, KY191713*; *B*. *boreale* J.Milde, SWE, Norrbotten, 2012, BD1202204, NEU, KY191078*, KF700469, KF700553, KF700387; *B*. *boreale* J.Milde, SWE, Norrbotten, 2012, BD1202604, NEU, KY191075*, KF700468, KF700552, KF700386; *B*. *boreale* J.Milde, SWE, Norrbotten, 2012, BD1202702, NEU, KY191071*, KF700467, KF700551, KF700385; *B*. *boreale* J.Milde, SWE, Norrbotten, 2012, BD1202805, NEU, KY191080*, KF700470, KF700554, KF700388; *B*. *boreale* J.Milde, SWE, Norrbotten, 2012, BD1203001, NEU, KY191070*, KF700466, KF700550, KF700384; *B*. *boreale* J.Milde, SWE, Norrbotten, 2012, BD1203903, NEU, KY191081*, KF700471, KF700555, KF700389; *B*. *boreale* J.Milde, NOR, Trondelag, 2006, DF14841, ISC, KY191079*, KY191430*, KY192003*, KY191714*; *B*. *boreale*

J.Milde, ISL, Norourland Eystra, 2011, DF18771, ISC, KY191076*, KY191431*, KY192004*, KY191715*; *B*. *boreale* J.Milde, GRL, Kujalleq, 2003, DF8863, ISC, KY191077*, KY191432*, KY192005*, KY191716*; *B*. *boreale* J.Milde, RUS, Karelia, 2013, KG1, NEU, KY191073*, KY191433*, KY192006*, KY191717*; *B*. *boreale* J.Milde, RUS, Karelia, 2013, KG2, NEU, KY191074*, KY191434*, KY192007*, KY191718*; *B*. *boreale* J.Milde, SWE, Gavle, 2015, PS1502, NEU, KY191072*, KY191435*, KY192008*, KY191719*; *B*. *boreale* × *B*. *lunaria 4*, SWE, Norrbotten, 2012, BD12022C, NEU, KY191082*, KF700472, KF700556, KF700390; *B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, USA, Minnesota, 2005, DF11938, ISC, KY191136*, KY191498*, KY192071*, KY191786*; *B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, USA, South Dakota, 2007, DF14957, ISC, KY191133*, KY191499*, KY192072*, KY191787*; *B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, CAN, Yukon, 2007, DF15059, ISC, KY191134*, KY191500*, KY192073*, KY191788*;*B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, USA, Nevada, 2007, DF15937, ISC, KY191135*, KY191501*, KY192074*, KY191789*; *B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, USA, Colorado, 2008, DF16696, ISC, KY191137*, KY191502*, KY192075*, KY191790*; *B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, USA, South Dakota, 2011, DF18591, ISC, KY191131*, KY191503*, KY192076*, KY191791*; *B*. *campestre* var. *lineare* (W.H.Wagner) Farrar, USA, South Dakota, 2011, DF18720, ISC, KY191132*, KY191504*, KY192077*, KY191792*; *B*. *campestre* W.H.Wagner & Farrar var. *campestre*, USA, South Dakota, 2007, DF14932, ISC, KY191083*, KY191436*, KY192009*, KY191720*; *B*. *campestre* W.H.Wagner & Farrar var. *campestre*, USA, Minnesota, 2008, DF16101, ISC, KY191084*, KY191437*, KY192010*, KY191721*; *B*. *campestre* W.H.Wagner & Farrar var. *campestre*, USA, Vermont, 2008, DF16275, ISC, KY191086*, KY191438*, KY192011*, KY191722*; *B*. *campestre* W.H.Wagner & Farrar var. *campestre*, USA, Michigan, 1987, DF233, ISC, KY191085*, KY191439*, KY192012*, KY191723*; *B*. *campestre* W.H.Wagner & Farrar var. *campestre*, USA, Iowa, 1995, DF954, ISC, KY191087*, KY191440*, KY192013*, KY191724*; *B*. *crenulatum* W.H.Wagner, CAN, Alberta, 2005, DF12194, ISC, KY191088*, KY191447*, KY192020*, KY191731*; *B*. *crenulatum* W.H.Wagner, USA, Wyoming, 2012, DF19448, ISC, KY191089*, KY191448*, KY192021*, KY191732*; *B*. *crenulatum* W.H.Wagner, CAN, Newfoundland, 2013, DF19611, ISC, KY191093*, KY191449*, KY192022*, KY191733*; *B*. *crenulatum* W.H.Wagner, USA, Oregon, 1999, DF3939, ISC, KY191090*, KY191450*, KY192023*, KY191734*; *B*. *crenulatum* W.H.Wagner, USA, California, 2002, DF7644, ISC, KY191094*, KY191451*, KY192024*, KY191735*; *B*. *crenulatum* W.H.Wagner, USA, Nevada, 2003, DF9013, ISC, KY191091*, KY191452*, KY192025*, KY191736*; *B*. *crenulatum* W.H.Wagner, USA, Utah, 2003, DF9025, ISC, KY191092*, KY191453*, KY192026*, KY191737*; *B*. *dusenii* (Christ) Alston, FLK, East Falkland, 2008, DF17078, ISC, KY191095*, KY191454*, KY192027*, KY191738*; *B*. *dusenii* (Christ) Alston, FLK, East Falkland, 2008, DF17085, ISC, KY191096*, KY191455*, KY192028*, KY191739*; *B*. *echo* W.H.Wagner, USA, Colorado, 2005, DF13047, ISC, KY191101*, KY191456*, KY192029*, KY191740*; *B*. *echo* W.H.Wagner, USA, Colorado, 2005, DF13093, ISC, KY191097*, KY191457*, KY192030*, KY191741*; *B*. *echo* W.H.Wagner, USA, New Mexico, 2008, DF16241, ISC, KY191098*, KY191458*, KY192031*, KY191742*; *B*. *echo* W.H.Wagner, USA, Colorado, 2008, DF17029, ISC, KY191100*, KY191459*, KY192032*, KY191743*; *B*. *echo* W.H.Wagner, USA, Colorado, 2009, DF17698, ISC, KY191099*, KY191460*, KY192033*, KY191744*; *B*. *farrarii* Legler & Popovich, sp. nov. ined., USA, Wyoming, 2010, DF18384, ISC, KY191323*, KY191461*, KY192034*, KY191745*; *B*. *farrarii* Legler & Popovich, sp. nov. ined., USA, Wyoming, 2010, DF18455, ISC, KY191324*, KY191462*, KY192035*, KY191746*; *B*. *furculatum* Popovich & Farrar, sp. nov. ined., USA, South Dakota, 2006, DF14290, ISC, KY191103*, KY191463*, KY192036*, KY191747*; *B*. *furculatum* Popovich & Farrar, sp. nov. ined., USA, Colorado, 2007, DF15698, ISC, KY191104*, KY191464*, KY192037*, KY191748*; *B*. *furculatum* Popovich & Farrar, sp. nov. ined., USA, New Mexico, 2009, DF17590, ISC, KY191105*, KY191465*, KY192038*, KY191749*; *B*. *furculatum* Popovich & Farrar, sp. nov. ined., USA, Wyoming, 2010, DF18394a, ISC, KY191102*, KY191466*, KY192039*, KY191750*; *B*. *furculatum* Popovich & Farrar, sp. nov. ined., USA, Colorado, 2012, DF19259, ISC, KY191106*, KY191467*, KY192040*, KY191751*; *B*. *gallicomontanum* Farrar & Johnson-Groh, USA, Montana, 2005, DF12419, ISC, KY191109*, KY191468*, KY192041*, KY191752*; *B*. *gallicomontanum* Farrar & Johnson-Groh, USA, South Dakota, 2006, DF13998, ISC, KY191107*, KY191469*, KY192042*, KY191753*; *B*. *gallicomontanum* Farrar & Johnson-Groh, USA, South Dakota, 2006,

DF14083, ISC, KY191108*, KY191470*, KY192043*, KY191754*; *B*. *gallicomontanum* Farrar & Johnson-Groh, USA, Minnesota, 2000, DF4507, ISC, KY191110*, KY191471*, KY192044*, KY191755*; *B*. *gallicomontanum* Farrar & Johnson-Groh, USA, Minnesota, 1995, DF979, ISC, KY191111*, KY191472*, KY192045*, KY191756*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, Montana, 2004, DF10488, ISC, KY191115*, KY191473*, KY192046*, KY191757*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, Colorado, 2005, DF13071, ISC, KY191117*, KY191474*, KY192047*, KY191758*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, Colorado, 2006, DF14457, ISC, KY191113*, KY191475*, KY192048*, KY191759*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, Colorado, 2007, DF15803, ISC, KY191112*, KY191476*, KY192049*, KY191760*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, New Mexico, 2008, DF16236, ISC, KY191118*, KY191477*, KY192050*, KY191761*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, Idaho, 2010, DF18178, ISC, KY191116*, KY191478*, KY192051*, KY191762*; *B*. *hesperium* (Maxon & R.T.Clausen) W.H.Wagner & Lellinger, USA, Washington, 2000, DF4485, ISC, KY191114*, KY191479*, KY192052*, KY191763*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2012, AW3, NEU, KY191386*, KF700476, KF700560, KF700394; *B*. *lanceolatum* (Gmel.) Ångstr. "green", SWE, Norrbotten, 2012, BD1202311, NEU, KY191399*, KF700486, KF700570, KY191767*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2012, BD1205330, NEU, KY191395*, KF700485, KF700569, KF700401; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2012, BD1206209, NEU, KY191401*, KF700487, KF700571, KY191768*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2012, BD1206630, NEU, KY191388*, KF700478, KF700562, KF700396; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2004, DF10543, ISC, KY191387*, KY191483*, KY192056*, KY191769*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", CAN, Yukon, 2004, DF11767, ISC, KY191391*, KY191484*, KY192057*, KY191770*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, New Mexico, 2009, DF17744, ISC, KY191396*, KY191485*, KY192058*, KY191771*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", ISL, Austurland, 2011, DF18750, ISC, KY191389*, KY191486*, KY192059*, KY191772*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", ISL, Norourland Eystra, 2012, DF19171, ISC, KY191393*, KF700482, KF700566, KY191773*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", ISL, Austurland, 2012, DF19337, ISC, KY191390*, KF700480, KF700564, KY191774*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Washington, 1997, DF2760, ISC, KY191385*, KY191487*, KY192060*, KY191775*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Idaho, 2001, DF5417, ISC, KY191394*, KY191488*, KY192061*, KY191776*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", CAN, British Columbia, 2003, DF8338, ISC, KY191400*, KY191489*, KY192062*, KY191777*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Oregon, 2003, DF8956, ISC, KY191397*, KY191490*, KY192063*, KY191778*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", RUS, Kurile Islands, 2014, LB6734, NEU, KY191398*, KY191491*, KY192064*, KY191779*; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2012, MD12076, NEU, KY191384*, KF700475, KF700559, KF700393; *B*. *lanceolatum* (Gmel.) Ångstr. "green", USA, Alaska, 2012, SS1233, NEU, KY191392*, KF700481, KF700565, KF700398; *B*. *lanceolatum* "red" × *B*. *pallidum 1*, CAN, Ontario, 2009, DF17272, ISC, KY191312*, KY191480*, KY192053*, KY191764*; *B*. *lanceolatum* "red" × *B*. *pallidum 1*, CAN, Ontario, 2010, DF18269, ISC, KY191313*, KY191481*, KY192054*, KY191765*; *B*. *lanceolatum* "red" × *B*. *pallidum 2*, USA, Montana, 1998, DF2878, ISC, KY191383*, KY191482*, KY192055*, KY191766*; *B*. *lanceolatum* (Gmel.) Ångstr. "red", CHE, Graubunden, 2012, BD1201501, NEU, KY191128*, KF700488, KF700572, KF700402; *B*. *lanceolatum* (Gmel.) Ångstr. "red", CHE, Graubunden, 2012, BD1201502, NEU, KY191129*, KF700489, KF700573, KF700403; *B*. *lanceolatum* (Gmel.) Ångstr. "red", USA, Alaska, 2012, BD1205343, NEU, KY191127*, KF700484, KF700568, KF700400; *B*. *lanceolatum* (Gmel.) Ångstr. "red", USA, Alaska, 2012, BD1206310, NEU, KY191121*, KF700477, KF700561, KF700395; *B*. *lanceolatum* (Gmel.) Ångstr. "red", CAN, Yukon, 2004, DF11117, ISC, KY191124*, KY191492*, KY192065*, KY191780*; *B*. *lanceolatum* (Gmel.) Ångstr. "red", USA, Colorado, 2005, DF13038, ISC, KY191119*, KY191493*, KY192066*, KY191781*; *B*. *lanceolatum* (Gmel.) Ångstr. "red", USA, New Mexico, 2008, DF16253, ISC, KY191130*, KY191494*, KY192067*, KY191782*; *B*. *lanceolatum* (Gmel.) Ångstr. "red", ISL, Austurland, 2011, DF18735, ISC, KY191122*, KY191495*, KY192068*, KY191783*; *B*. *lanceolatum* (Gmel.) Ångstr. "red", GRL, Kujalleq, 2012, DF19093,

ISC, KY191125*, KF700483, KF700567, KF700399; *B*. *lanceolatum* (Gmel.) Ångstr. "red", ISL, Austurland, 2012, DF19326, ISC, KY191123*, KF700479, KF700563, KF700397; *B*. *lanceolatum* (Gmel.) Ångstr. "red", GRL, Kujalleq, 2003, DF8873, ISC, KY191126*, KY191496*, KY192069*, KY191784*; *B*. *lanceolatum* (Gmel.) Ångstr. "red", USA, Alaska, 2003, DF9865, ISC, KY191120*, KY191497*, KY192070*, KY191785*; *B*. *lunaria* (L.) Sw. var. *lunaria*, HUN, Miskoloc, 2014, AM32I, NEU, KY191170*, KY191535*, KY192109*, KY191823*; *B*. *lunaria* (L.) Sw. var. *lunaria*, AUT, Styria, 2014, AM55R, NEU, KY191173*, KY191536*, KY192110*, KY191824*; *B*. *lunaria* (L.) Sw. var. *lunaria*, FRA, Alsace, 2014, AM60P, NEU, KY191162*, KY191537*, KY192111*, KY191825*; *B*. *lunaria* (L.) Sw. var. *lunaria*, FRA, Haute Savoie, 2014, AM61E, NEU, KY191159*, KY191538*, KY192112*, KY191826*; *B*. *lunaria* (L.) Sw. var. *lunaria*, AND, Encamp, 2014, AM76B, NEU, KY191161*, KY191539*, KY192113*, KY191827*; *B*. *lunaria* (L.) Sw. var. *lunaria*, ESP, Andalusia, 2014, AM83B, NEU, KY191172*, KY191540*, KY192114*, KY191828*; *B*. *lunaria* (L.) Sw. var. *lunaria*, GRC, Kallii, 2015, AMGRC001, NEU, KY191174*, KY191541*, KY192115*, KY191829*; *B*. *lunaria* (L.) Sw. var. *lunaria*, GRC, Kallii, 2015, AMGRC006, NEU, KY191175*, KY191542*, KY192116*, KY191830*; *B*. *lunaria* (L.) Sw. var. *lunaria*, SWE, Falkoping, 2012, BD1204107, NEU, KY191168*, KF700496, KF700579, KF700410; *B*. *lunaria* (L.) Sw. var. *lunaria*, SWE, Falkoping, 2012, BD1204116, NEU, KY191169*, KF700497, KF700580, KF700411; *B*. *lunaria* (L.) Sw. var. *lunaria*, SWE, Norrbotten, 2015, BO1501, NEU, KY191171*, KY191543*, KY192117*, KY191831*; *B*. *lunaria* (L.) Sw. var. *lunaria*, CAN, Yukon, 2004, DF10513, ISC, KY191158*, KY191544*, KY192118*, KY191832*; *B*. *lunaria* (L.) Sw. var. *lunaria*, NOR, Telemark, 2009, DF17167, ISC, KY191166*, KY191545*, KY192119*, KY191833*; *B*. *lunaria* (L.) Sw. var. *lunaria*, ISL, Austurland, 2012, DF19308, ISC, KY191165*, KF700493, KF700577, KF700407; *B*. *lunaria* (L.) Sw. var. *lunaria*, ISL, Austurland, 2012, DF19314, ISC, KY191167*, KF700494, KF700578, KF700408; *B*. *lunaria* (L.) Sw. var. *lunaria*, SWE, Varmland, 2015, FT1501A, NEU, KY191163*, KY191546*, KY192120*, KY191834*; *B*. *lunaria* (L.) Sw. var. *lunaria*, SWE, Varmland, 2015, FT150718A, NEU, KY191164*, KY191547*, KY192121*, KY191835*; *B*. *lunaria* (L.) Sw. var. *lunaria*, SWE, Uppland, 2015, JE1501, NEU, KY191160*, KY191548*, KY192122*, KY191836*; *B*. *lunaria* (L.) Sw. var. *melzeri* Stensvold & Farrar, GRL, Kujalleq, 2005, DF12806, ISC, KY191203*, KY191557*, KY192131*, KY191845*; *B*. *lunaria* (L.) Sw. var. *melzeri* Stensvold & Farrar, ISL, Norourland Eystra, 2005, DF12942, ISC, KY191204*, KY191558*, KY192132*, KY191846*; *B*. *lunaria* (L.) Sw. var. *melzeri* Stensvold & Farrar, GRL, NA, 2012, DF19145, ISC, KY191202*, KF700499, KF700582, KF700413; *B*. *lunaria 2*, CHE, Bern, 2014, AM11C, NEU, KY191142*, KY191505*, KY192078*, KY191793*; *B*. *lunaria 2*, CHE, Bern, 2014, AM11I, NEU, KY191143*, KY191506*, KY192079*, KY191794*; *B*. *lunaria 2*, ITA, Belluno, 2014, AM48O, NEU, KY191141*, KY191507*, KY192080*, KY191795*; *B*. *lunaria 2*, AUT, Steiermark, 2014, AM52G, NEU, KY191152*, KY191508*, KY192081*, KY191796*; *B*. *lunaria 2*, ROU, Southern Carpathians, 2014, AM59C, NEU, KY191151*, KY191509*, KY192082*, KY191797*; *B*. *lunaria 2*, FRA, Savoie, 2014, AM68C, NEU, KY191140*, KY191510*, KY192083*, KY191798*; *B*. *lunaria 2*, NPL, Manaslu, 2008, CP20812208, GH, KY191145*, KY191511*, KY192084*, KY191799*; *B*. *lunaria 2*, CHN, Xizang, 2000, DB29617, GH, KY191138*, KY191512*, KY192085*, KY191800*; *B*. *lunaria 2*, GRL, Kujalleq, 2005, DF12833, ISC, KY191148*, KY191513*, KY192086*, KY191801*; *B*. *lunaria 2*, ISL, Suournes, 2005, DF13237, ISC, KY191153*, KY191514*, KY192087*, KY191802*; *B*. *lunaria 2*, CAN, Winnipeg, 2007, DF15589, ISC, KY191157*, KY191515*, KY192088*, KY191803*; *B*. *lunaria 2*, CAN, Quebec, 2008, DF16494, ISC, KY191139*, KY191516*, KY192089*, KY191804*; *B*. *lunaria 2*, GRL, Kujalleq, 2012, DF19105, ISC, KY191149*, KF700501, KF700584, KF700415; *B*. *lunaria 2*, GRL, Kujalleq, 2012, DF19120, ISC, KY191150*, KF700502, KF700585, KF700416; *B*. *lunaria 2*, GRL, 2012, DF19127, ISC, KY191146*, KF700498, KF700581, KF700412; *B*. *lunaria 2*, ISL, Norourland Eystra, 2012, DF19185, ISC, KY191147*, KF700500, KF700583, KF700414; *B*. *lunaria 2*, CAN, Ontario, 2012, DF19250, ISC, KY191144*, KY191517*, KY192090*, KY191805*; *B*. *lunaria 2*, CAN, Newfoundland, 2001, DF5060, ISC, KY191154*, KY191518*, KY192091*, KY191806*; *B*. *lunaria 2*, TWN, Xiulin, 2002, DF7268, ISC, KY191156*, KY191519*, KY192092*, KY191807*; *B*. *lunaria 2*, CAN, Quebec, 2002, DF7874, ISC, KY191155*, KY191520*, KY192093*, KY191808*; *B*. *lunaria 3*, CHE, Ticino, 2014, AM26A, NEU, KY191178*, KY191521*, KY192094*, KY191809*; *B*. *lunaria 3*, ITA, Toscana, 2014, AM3G, NEU,

KY191176*, KY191522*, KY192095*, KY191810*; *B. lunaria 3*, CHE, Graubunden, 2014, AM41F, NEU, KY191177*, KY191523*, KY192096*, KY191811*; *B. lunaria 3*, AUT, Vorarlberg, 2014, AM43L, NEU, KY191182*, KY191524*, KY192097*, KY191812*; *B. lunaria 3*, FRA, Vaucluse, 2014, AM66B, NEU, KY191180*, KY191525*, KY192098*, KY191813*; *B. lunaria 3*, FRA, Hautes Pyrenees, 2014, AM73N, NEU, KY191181*, KY191526*, KY192099*, KY191814*; *B. lunaria 3*, NOR, Rogaland, 2005, DF12039, ISC, KY191184*, KY191527*, KY192100*, KY191815*; *B. lunaria 3*, DEU, Mecklenburg Vorpommern, 2003, DF4970, ISC, KY191183*, KY191528*, KY192101*, KY191816*; *B. lunaria 3*, NOR, Rogaland, 2003, DF8699, ISC, KY191185*, KY191529*, KY192102*, KY191817*; *B. lunaria 3*, SWE, Skane, 2015, PA1501, NEU, KY191179*, KY191530*, KY192103*, KY191818*; *B. lunaria 4*, SWE, Norrbotten, 2012, BD1203106, NEU, KY191189*, KF700505, KF700588, KF700419; *B. lunaria 4*, SWE, Norrbotten, 2012, BD1203117, NEU, KY191190*, KF700506, KF700589, KF700420; *B. lunaria 4*, SWE, Norrbotten, 2014, BD1203721, NEU, KY191186*, KF700495, KY192104*, KF700409; *B. lunaria 4*, SWE, Norrbotten, 2012, BD1203812, NEU, KY191187*, KF700503, KF700586, KF700417; *B. lunaria 4*, SWE, Norrbotten, 2012, BD1203823, NEU, KY191188*, KF700504, KF700587, KF700418; *B. lunaria 5*, CHE, Bern, 2014, AM34B, NEU, KY191193*, KY191531*, KY192105*, KY191819*; *B. lunaria 5*, CHE, Graubunden, 2014, AM38E, NEU, KY191191*, KY191532*, KY192106*, KY191820*; *B. lunaria 5*, AUT, Styria, 2014, AM55B, NEU, KY191196*, KY191533*, KY192107*, KY191821*; *B. lunaria 5*, SVN, Gorizia, 2014, AM65G, NEU, KY191197*, KY191534*, KY192108*, KY191822*; *B. lunaria 5*, CHE, Neuchatel, 2012, BD1200101, NEU, KY191194*, KF700491, KF700575, KF700405; *B. lunaria 5*, CHE, Ticino, 2012, BD1200404, NEU, KY191198*, KF700507, KF700590, KF700421; *B. lunaria 5*, CHE, Graubunden, 2012, BD1200601, NEU, KY191199*, KF700508, KF700591, KF700422; *B. lunaria 5*, CHE, Graubunden, 2012, BD1200906, NEU, KY191200*, KF700509, KF700592, KF700423; *B. lunaria 5*, CHE, Graubunden, 2012, BD1201109, NEU, KY191201*, KF700510, KF700593, KF700424; *B. lunaria 5*, CHE, Valais, 2012, BD1204305, NEU, KY191192*, KF700490, KF700574, KF700404; *B. lunaria 5*, CHE, Uri, 2012, BD1204412, NEU, KY191195*, KF700492, KF700576, KF700406; *B. matricariifolium* (Döll) A.Braun, CHE, Uri, 2012, BD1201601, NEU, KY191209*, KF700513, KF700596, KF700427; *B. matricariifolium* (Döll) A.Braun, SWE, Falkoping, 2012, BD1204005, NEU, KY191206*, KF700511, KF700594, KF700425; *B. matricariifolium* (Döll) A.Braun, SWE, Falkoping, 2012, BD1204102, NEU, KY191211*, KF700515, KF700598, KF700429; *B. matricariifolium* (Döll) A.Braun, USA, Michigan, 1995, DF1189, ISC, KY191214*, KY191549*, KY192123*, KY191837*; *B. matricariifolium* (Döll) A.Braun, CAN, Alberta, 2005, DF12184, ISC, KY191208*, KY191550*, KY192124*, KY191838*; *B. matricariifolium* (Döll) A.Braun, USA, Michigan, 1995, DF1292, ISC, KY191213*, KY191551*, KY192125*, KY191839*; *B. matricariifolium* (Döll) A.Braun, USA, Michigan, 1996, DF1437, ISC, KY191215*, KY191552*, KY192126*, KY191840*; *B. matricariifolium* (Döll) A.Braun, USA, Wisconsin, 1996, DF1463, ISC, KY191207*, KY191553*, KY192127*, KY191841*; *B. matricariifolium* (Döll) A.Braun, CAN, Quebec, 2008, DF16389, ISC, KY191205*, KY191554*, KY192128*, KY191842*; *B. matricariifolium* (Döll) A.Braun, USA, Minnesota, 2012, DF19017, ISC, KY191210*, KF700514, KF700597, KF700428; *B. matricariifolium* (Döll) A.Braun, USA, Minnesota, 2012, DF19067, ISC, KY191212*, KY191555*, KY192129*, KY191843*; *B. matricariifolium* (Döll) A.Braun, USA, Michigan, 1987, DF465, ISC, KY191216*, KY191556*, KY192130*, KY191844*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, CAN, Alberta, 2005, DF12087, ISC, KY191218*, KY191559*, KY192133*, KY191847*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, USA, Montana, 2005, DF12394, ISC, KY191221*, KY191560*, KY192134*, KY191848*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, USA, South Dakota, 2007, DF14903, ISC, KY191220*, KY191561*, KY192135*, KY191849*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, CAN, Quebec, 2007, DF16329, ISC, KY191217*, KY191562*, KY192136*, KY191850*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, USA, Minnesota, 2010, DF18145, ISC, KY191222*, KY191563*, KY192137*, KY191851*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, USA, Washington, 1997, DF1872, ISC, KY191219*, KY191564*, KY192138*, KY191852*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, USA, Michigan, 2003, DF8268, ISC, KY191224*, KY191565*, KY192139*, KY191853*; *B. michiganense* W.H.Wagner ex A.V.Gilman, Farrar & Zika, USA, Washington, 2003, DF8363, ISC, KY191223*, KY191566*, KY192140*,

KY191854*; *B*. *minganense* × *B*. *echo*, USA, Colorado, 2007, DF15679, ISC, KY191241*, KY191567*, KY192141*, KY191855*; *B*. *minganense* × *B*. *echo*, USA, Colorado, 2009, DF17602, ISC, KY191242*, KY191568*, KY192142*, KY191856*; *B*. *minganense* Victorin, USA, Alaska, 2012, AW4, NEU, KY191228*, KF700516, KF700599, KF700430; *B*. *minganense* Victorin, USA, Alaska, 2012, BD1205009, NEU, KY191237*, KF700520, KF700603, KF700434; *B*. *minganense* Victorin, USA, Alaska, 2012, BD1205406, NEU, KY191229*, KF700517, KF700600, KF700431; *B*. *minganense* Victorin, USA, Alaska, 2012, BD1205713, NEU, KY191238*, KF700521, KF700604, KF700435; *B*. *minganense* Victorin, USA, Alaska, 2012, BD1206340, NEU, KY191230*, KF700518, KF700601, KF700432; *B*. *minganense* Victorin, USA, Michigan, 1995, DF1043, ISC, KY191239*, KY191569*, KY192143*, KY191857*; *B*. *minganense* Victorin, USA, Minnesota, 2006, DF14277, ISC, KY191226*, KY191570*, KY192144*, KY191858*; *B*. *minganense* Victorin, USA, Colorado, 2008, DF16662, ISC, KY191231*, KY191571*, KY192145*, KY191859*; *B*. *minganense* Victorin, USA, New Mexico, 2009, DF17538, ISC, KY191240*, KY191572*, KY192146*, KY191860*; *B*. *minganense* Victorin, USA, South Dakota, 2010, DF18074, ISC, KY191225*, KY191573*, KY192147*, KY191861*; *B*. *minganense* Victorin, CAN, Newfoundland, 2001, DF5120, ISC, KY191236*, KY191574*, KY192148*, KY191862*; *B*. *minganense* Victorin, CAN, Quebec, 2002, DF7841, ISC, KY191235*, KY191575*, KY192149*, KY191863*; *B*. *minganense* Victorin, USA, Minnesota, 2003, DF8138, ISC, KY191227*, KY191576*, KY192150*, KY191864*; *B*. *minganense* Victorin, CAN, Yukon, 2003, DF8645, ISC, KY191233*, KY191577*, KY192151*, KY191865*; *B*. *minganense* Victorin, USA, Alaska, 2012, MD12111, NEU, KY191234*, KF700519, KF700602, KF700433; *B*. *minganense* Victorin, USA, Alaska, 2014, MT717, NEU, KY191232*, KY191578*, KY192152*, KY191866*; *B*. *montanum* W.H.Wagner, USA, California, 2003, DF10172, ISC, KY191244*, KY191579*, KY192153*, KY191867*; *B*. *montanum* W.H.Wagner, USA, Montana, 1997, DF1677, ISC, KY191247*, KY191580*, KY192154*, KY191868*; *B*. *montanum* W.H.Wagner, USA, Washington, 2010, DF18528, ISC, KY191248*, KY191581*, KY192155*, KY191869*; *B*. *montanum* W.H.Wagner, USA, Washington, 1997, DF2620, ISC, KY191243*, KY191582*, KY192156*, KY191870*; *B*. *montanum* W.H.Wagner, USA, Oregon, 1997, DF2646, ISC, KY191246*, KY191583*, KY192157*, KY191871*; *B*. *montanum* W.H.Wagner, USA, Oregon, 2000, DF4871, ISC, KY191245*, KY191584*, KY192158*, KY191872*; *B*. *mormo* W.H.Wagner, USA, Michigan, 1996, DF1492, ISC, KY191251*, KY191585*, KY192159*, KY191873*; *B*. *mormo* W.H.Wagner, USA, Minnesota, 2012, DF18955, ISC, KY191250*, KY191586*, KY192160*, KY191874*; *B*. *mormo* W.H.Wagner, USA, Minnesota, 2012, DF18997, ISC, KY191252*, KY191587*, KY192161*, KY191875*; *B*. *mormo* W.H.Wagner, USA, Minnesota, 1993, DF858, ISC, KY191253*, KY191588*, KY192162*, KY191876*; *B*. *mormo* W.H.Wagner, USA, Minnesota, 1994, DF890, ISC, KY191249*, KY191589*, KY192163*, KY191877*; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2014, AW1401, NEU, KY191260*, KY191593*, KY192167*, KY191881*; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2012, BD1205807, NEU, KY191275*, KF700528, KF700611, KF700442; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2012, BD1206345, NEU, KY191261*, KF700523, KF700606, KF700437; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2012, BD1206624, NEU, KY191263*, KF700525, KF700608, KF700439; *B*. *neolunaria* Stensvold & Farrar, USA, Washington, 2012, BL124841, NEU, KY191255*, KF700522, KF700605, KF700436; *B*. *neolunaria* Stensvold & Farrar, USA, Washington, 2012, BL125231, NEU, KY191279*, KF700529, KF700612, KF700443; *B*. *neolunaria* Stensvold & Farrar, NPL, Darchula, 2012, CP1216073, GH, KY191259*, KY191594*, KY192168*, KY191882*; *B*. *neolunaria* Stensvold & Farrar, RUS, Kamchatka, 2005, DF12333, ISC, KY191283*, KY191595*, KY192169*, KY191883*; *B*. *neolunaria* Stensvold & Farrar, CAN, British Columbia, 2005, DF12338, ISC, KY191270*, KY191596*, KY192170*, KY191884*; *B*. *neolunaria* Stensvold & Farrar, USA, South Dakota, 2006, DF14336, ISC, KY191256*, KY191597*, KY192171*, KY191885*; *B*. *neolunaria* Stensvold & Farrar, CAN, Yukon, 2007, DF15197, ISC, KY191284*, KY191598*, KY192172*, KY191886*; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2007, DF15460, ISC, KY191285*, KY191599*, KY192173*, KY191887*; *B*. *neolunaria* Stensvold & Farrar, CAN, Quebec, 2008, DF16460, ISC, KY191254*, KY191600*, KY192174*, KY191888*; *B*. *neolunaria* Stensvold & Farrar, USA, New Mexico, 2008, DF16851, ISC, KY191278*, KY191601*, KY192175*, KY191889*; *B*. *neolunaria* Stensvold & Farrar, USA, Colorado, 2009, DF17695, ISC, KY191277*, KY191602*, KY192176*, KY191890*; *B*. *neolunaria* Stensvold & Farrar, CAN, Ontario, 2010, DF18299, ISC, KY191258*,

KY191603*, KY192177*, KY191891*; *B*. *neolunaria* Stensvold & Farrar, USA, Colorado, 2012, DF19201, ISC, KY191262*, KF700524, KF700607, KF700438; *B*. *neolunaria* Stensvold & Farrar, USA, Washington, 1997, DF2360, ISC, KY191271*, KY191604*, KY192178*, KY191892*; *B*. *neolunaria* Stensvold & Farrar, NZL, Tasman, 2001, DF4898, ISC, KY191282*, KY191605*, KY192179*, KY191893*; *B*. *neolunaria* Stensvold & Farrar, CAN, Newfoundland, 2001, DF5036, ISC, KY191274*, KY191606*, KY192180*, KY191894*; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2001, DF5467, ISC, KY191280*, KY191607*, KY192181*, KY191895*; *B*. *neolunaria* Stensvold & Farrar, RUS, Bering Island, 2002, DF6447, ISC, KY191281*, KY191608*, KY192182*, KY191896*; *B*. *neolunaria* Stensvold & Farrar, USA, Michigan, 2003, DF8278, ISC, KY191257*, KY191609*, KY192183*, KY191897*; *B*. *neolunaria* Stensvold & Farrar, AUS, Tasmania, 2013, GJ1, NEU, KY191276*, KY191610*, KY192184*, KY191898*; *B*. *neolunaria* Stensvold & Farrar, RUS, Kurile Islands, 2014, LB6733, NEU, KY191273*, KY191611*, KY192185*, KY191899*; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2012, MD12060, NEU, KY191272*, KF700527, KF700610, KF700441; *B*. *neolunaria* Stensvold & Farrar, NZL, Nelson Region, 2014, RG1, NEU, KY191267*, KY191612*, KY192186*, KY191900*; *B*. *neolunaria* Stensvold & Farrar, NZL, Nelson Region, 2015, RG2, NEU, KY191268*, KY191613*, KY192187*, KY191901*; *B*. *neolunaria* Stensvold & Farrar, USA, Alaska, 2012, SS1231, NEU, KY191265*, KF700526, KF700609, KF700440; *B*. *neolunaria* Stensvold & Farrar, JPN, Nagano, 2014, WS20140802001, MAK, KY191264*, KY191590*, KY192164*, KY191878*; *B*. *neolunaria* Stensvold & Farrar, JPN, Nagano, 2014, WS20140803001, MAK, KY191266*, KY191591*, KY192165*, KY191879*; *B*. *neolunaria* Stensvold & Farrar, JPN, Shizuoka, 2014, WS20140903001, MAK, KY191269*, KY191592*, KY192166*, KY191880*; *B*. *nordicum* Stensvold & Farrar, ISL, Vestfiroir, 2005, DF13337, ISC, KY191288*, KF700530, KF700613, KF700444; *B*. *nordicum* Stensvold & Farrar, NOR, More og Romsdal, 2007, DF17157, ISC, KY191290*, KF700531, KF700614, KF700445; *B*. *nordicum* Stensvold & Farrar, NOR, Telemark, 2009, DF17166, ISC, KY191289*, KY191614*, KY192188*, KY191902*; *B*. *nordicum* Stensvold & Farrar, ISL, Austurland, 2009, DF17819, ISC, KY191287*, KY191615*, KY192189*, KY191903*; *B*. *nordicum* Stensvold & Farrar, GBR, Scotland, 2015, HMcH155, NEU, KY191286*, KY191616*, KY192190*, KY191904*; *B*. *pallidum* W.H.Wagner, USA, Michigan, 1995, DF1003, ISC, KY191296*, KY191617*, KY192191*, KY191905*; *B*. *pallidum* W.H.Wagner, CAN, Alberta, 2005, DF12210, ISC, KY191293*, KY191618*, KY192192*, KY191906*; *B*. *pallidum* W.H.Wagner, CAN, Quebec, 2008, DF16373, ISC, KY191297*, KY191619*, KY192193*, KY191907*; *B*. *pallidum* W.H.Wagner, CAN, Ontario, 2009, DF17260, ISC, KY191295*, KY191620*, KY192194*, KY191908*; *B*. *pallidum* W.H.Wagner, USA, Michigan, 2011, DF18614, ISC, KY191294*, KY191621*, KY192195*, KY191909*; *B*. *pallidum* W.H.Wagner, USA, South Dakota, 2011, DF18729, ISC, KY191292*, KY191622*, KY192196*, KY191910*; *B*. *pallidum* W.H.Wagner, USA, Minnesota, 2000, DF4344, ISC, KY191291*, KY191623*, KY192197*, KY191911*; *B*. *paradoxum* W.H.Wagner, USA, California, 2008, DF16672, ISC, KY191304*, KY191624*, KY192198*, KY191912*; *B*. *paradoxum* W.H.Wagner, USA, California, 2008, DF16781, ISC, KY191305*, KY191625*, KY192199*, KY191913*; *B*. *paradoxum* W.H.Wagner, USA, Colorado, 2008, DF16996, ISC, KY191302*, KY191626*, KY192200*, KY191914*; *B*. *paradoxum* W.H.Wagner, USA, Wyoming, 2010, DF18449, ISC, KY191298*, KY191627*, KY192201*, KY191915*; *B*. *paradoxum* W.H.Wagner, USA, Washington, 1997, DF1944, ISC, KY191299*, KY191628*, KY192202*, KY191916*; *B*. *paradoxum* W.H.Wagner, USA, Montana, 1998, DF2890, ISC, KY191303*, KY191629*, KY192203*, KY191917*; *B*. *paradoxum* W.H.Wagner, USA, Washington, 2002, DF6351, ISC, KY191300*, KY191630*, KY192204*, KY191918*; *B*. *paradoxum* W.H.Wagner, USA, Washington, 2003, DF8350, ISC, KY191301*, KY191631*, KY192205*, KY191919*; *B*. *pedunculosum* W.H.Wagner, USA, Oregon, 2004, DF11066, ISC, KY191310*, KY191632*, KY192206*, KY191920*; *B*. *pedunculosum* W.H.Wagner, USA, Washington, 2011, DF18840, ISC, KY191307*, KY191633*, KY192207*, KY191921*; *B*. *pedunculosum* W.H.Wagner, USA, Oregon, 2001, DF5332, ISC, KY191311*, KY191634*, KY192208*, KY191922*; *B*. *pedunculosum* W.H.Wagner, USA, Idaho, 2001, DF5413, ISC, KY191309*, KY191635*, KY192209*, KY191923*; *B*. *pedunculosum* W.H.Wagner, USA, Washington, 2002, DF6489, ISC, KY191308*, KY191636*, KY192210*, KY191924*; *B*. *pedunculosum* W.H.Wagner, USA, Alaska, 2012, MD12078, NEU, KY191306*, KF700512, KF700595, KF700426; *B*. *pinnatum* H.St.John, USA, Alaska, 2012, AW2,

NEU, KY191314*, KF700532, KF700615, KF700446; *B. pinnatum* H.St.John, USA, Alaska, 2012, BD1206619, NEU, KY191315*, KF700534, KF700617, KF700448; *B. pinnatum* H.St.John, USA, Alaska, 2004, DF11437, ISC, KY191319*, KY191637*, KY192211*, KY191925*; *B. pinnatum* H.St.John, CAN, British Columbia, 2005, DF12373, ISC, KY191320*, KY191638*, KY192212*, KY191926*; *B. pinnatum* H.St.John, CAN, Yukon, 2007, DF15126, ISC, KY191316*, KY191639*, KY192213*, KY191927*; *B. pinnatum* H.St.John, USA, Colorado, 2010, DF18204, ISC, KY191321*, KY191640*, KY192214*, KY191928*; *B. pinnatum* H.St.John, USA, Washington, 1997, DF2776, ISC, KY191322*, KY191641*, KY192215*, KY191929*; *B. pinnatum* H.St.John, USA, Oregon, 1999, DF3925, ISC, KY191318*, KY191642*, KY192216*, KY191930*; *B. pinnatum* H.St.John, USA, Alaska, 2012, SS1232, NEU, KY191317*, KF700535, KF700618, KF700449; *B. pseudopinnatum* W.H.Wagner, CAN, Ontario, 2010, DF18301, ISC, KY191327*, KY191643*, KY192217*, KY191931*; *B. pseudopinnatum* W.H.Wagner, CAN, Ontario, 2010, DF18314, ISC, KY191328*, KY191644*, KY192218*, KY191932*; *B. pseudopinnatum* W.H.Wagner, CAN, Ontario, 1989, DF689, ISC, KY191325*, KY191645*, KY192219*, KY191933*; *B. pseudopinnatum* W.H.Wagner, CAN, Ontario, 1989, DF693, ISC, KY191326*, KY191646*, KY192220*, KY191934*; *B. pumicola* Colville, USA, Oregon, 2004, DF10958, ISC, KY191333*, KY191647*, KY192221*, KY191935*; *B. pumicola* Colville, USA, Oregon, 2009, DF17909, ISC, KY191329*, KY191648*, KY192222*, KY191936*; *B. pumicola* Colville, USA, Oregon, 2009, DF17915, ISC, KY191330*, KY191649*, KY192223*, KY191937*; *B. pumicola* Colville, USA, Oregon, 2009, DF17919, ISC, KY191331*, KY191650*, KY192224*, KY191938*; *B. pumicola* Colville, USA, Oregon, 2002, DF7966, ISC, KY191332*, KY191651*, KY192225*, KY191939*; *B. simplex 2*, ISL, Stoovarfjorour, 2005, DF13250, ISC, KY191340*, KY191652*, KY192226*, KY191940*; *B. simplex 2*, USA, California, 2005, DF13844, ISC, KY191338*, KY191653*, KY192227*, KY191941*; *B. simplex 2*, CAN, Quebec, 2008, DF16312, ISC, KY191337*, KY191654*, KY192228*, KY191942*; *B. simplex 2*, USA, New Mexico, 2008, DF16562, ISC, KY191343*, KY191655*, KY192229*, KY191943*; *B. simplex 2*, GRL, Greenland, 2008, DF16611, ISC, KY191339*, KY191656*, KY192230*, KY191944*; *B. simplex 2*, CAN, Ontario, 2009, DF17299, ISC, KY191341*, KY191657*, KY192231*, KY191945*; *B. simplex 2*, USA, Oregon, 1999, DF3968, ISC, KY191345*, KY191658*, KY192232*, KY191946*; *B. simplex 2*, USA, California, 2002, DF7605, ISC, KY191342*, KY191659*, KY192233*, KY191947*; *B. simplex 2*, SWE, Dalarna, 2015, TL1501, NEU, KY191344*, KY191660*, KY192234*, KY191948*; *B. simplex* E.Hitchc. var. *compositum* (Lasch) Milde, USA, California, 2006, DF14796, ISC, KY191348*, KY191441*, KY192014*, KY191725*; *B. simplex* E.Hitchc. var. *compositum* (Lasch) Milde, USA, California, 2007, DF15877, ISC, KY191347*, KY191442*, KY192015*, KY191726*; *B. simplex* E.Hitchc. var. *compositum* (Lasch) Milde, USA, Colorado, 2011, DF18799, ISC, KY191351*, KY191443*, KY192016*, KY191727*; *B. simplex* E.Hitchc. var. *compositum* (Lasch) Milde, USA, Oregon, 2000, DF4771, ISC, KY191350*, KY191444*, KY192017*, KY191728*; *B. simplex* E.Hitchc. var. *compositum* (Lasch) Milde, USA, Washington, 2002, DF6501, ISC, KY191346*, KY191445*, KY192018*, KY191729*; *B. simplex* E.Hitchc. var. *compositum* (Lasch) Milde, USA, California, 2002, DF6896, ISC, KY191349*, KY191446*, KY192019*, KY191730*; *B. simplex* E.Hitchc. var. *simplex*, CAN, Alberta, 2005, DF12148, ISC, KY191336*, KY191661*, KY192235*, KY191949*; *B. simplex* E.Hitchc. var. *simplex*, USA, Minnesota, 2006, DF14270, ISC, KY191335*, KY191662*, KY192236*, KY191950*; *B. simplex* E.Hitchc. var. *simplex*, USA, South Dakota, 2007, DF14892, ISC, KY191334*, KY191663*, KY192237*, KY191951*; *B. spathulatum* W.H.Wagner, USA, Alaska, 2004, DF10278, ISC, KY191355*, KY191664*, KY192238*, KY191952*; *B. spathulatum* W.H.Wagner, CAN, British Columbia, 2005, DF12382, ISC, KY191356*, KY191665*, KY192239*, KY191953*; *B. spathulatum* W.H.Wagner, CAN, Yukon, 2007, DF15379, ISC, KY191354*, KY191666*, KY192240*, KY191954*; *B. spathulatum* W.H.Wagner, CAN, Quebec, 2008, DF16514, ISC, KY191352*, KY191667*, KY192241*, KY191955*; *B. spathulatum* W.H.Wagner, CAN, Ontario, 2009, DF17361, ISC, KY191358*, KY191668*, KY192242*, KY191956*; *B. spathulatum* W.H.Wagner, CAN, Ontario, 2010, DF18290, ISC, KY191353*, KY191669*, KY192243*, KY191957*; *B. spathulatum* W.H.Wagner, USA, Colorado, 2011, DF18803, ISC, KY191357*, KY191670*, KY192244*, KY191958*; *B. tenebrosum* A.A.Eaton, CHE, Graubunden, 2015, AM007, KY191368*, KY191671*, KY192245*, KY191959*; *B. tenebrosum* A.A.Eaton, ISL, Austurland, 2004, DF11249, ISC, KY191365*, KY191672*, KY192246*, KY191960*; *B. tenebrosum* A.A.Eaton,

USA, Vermont, 2005, DF12278, ISC, KY191369*, KY191673*, KY192247*, KY191961*; *B*. *tenebrosum* A.A.Eaton, USA, Minnesota, 2005, DF12616, ISC, KY191363*, KY191674*, KY192248*, KY191962*; *B*. *tenebrosum* A.A.Eaton, CAN, Quebec, 2008, DF16362, ISC, KY191360*, KY191675*, KY192249*, KY191963*; *B*. *tenebrosum* A.A.Eaton, GRL, Greenland, 2009, DF17416, ISC, KY191364*, KY191676*, KY192250*, KY191964*; *B*. *tenebrosum* A.A.Eaton, USA, Minnesota, 2012, DF19027, ISC, KY191367*, KY191677*, KY192251*, KY191965*; *B*. *tenebrosum* A.A.Eaton, ISL, Austurland, 2002, DF7446, ISC, KY191366*, KY191678*, KY192252*, KY191966*; *B*. *tenebrosum* A.A.Eaton, USA, Michigan, 2003, DF8522, ISC, KY191359*, KY191679*, KY192253*, KY191967*; *B*. *tenebrosum* A.A.Eaton, SWE, Uppland, 2015, JE1502, NEU, KY191361*, KY191680*, KY192254*, KY191968*; *B*. *tenebrosum* A.A.Eaton, SWE, Gavle, 2015, PS1501, NEU, KY191362*, KY191681*, KY192255*, KY191969*; *B*. *tunux* Stensvold & Farrar, MNE, Pluzine, 2015, AM006, NEU, KY191372*, KY191682*, KY192256*, KY191970*; *B*. *tunux* Stensvold & Farrar, CHE, Valais, 2014, AM23A, NEU, KY191382*, KY191683*, KY192257*, KY191971*; *B*. *tunux* Stensvold & Farrar, CHE, Graubunden, 2014, AM41P, NEU, KY191370*, KY191684*, KY192258*, KY191972*; *B*. *tunux* Stensvold & Farrar, USA, Montana, 2005, DF12511, ISC, KY191377*, KY191685*, KY192259*, KY191973*; *B*. *tunux* Stensvold & Farrar, USA, Yukon, 2005, DF13386, ISC, KY191371*, KY191686*, KY192260*, KY191974*; *B*. *tunux* Stensvold & Farrar, USA, Alaska, 2012, DF13516, ISC, KY191381*, KF700538, KF700621, KF700452; *B*. *tunux* Stensvold & Farrar, CAN, Yukon, 2012, DF15209, ISC, KY191374*, KF700536, KF700619, KF700450; *B*. *tunux* Stensvold & Farrar, USA, Colorado, 2012, DF18790, ISC, KY191380*, KF700537, KF700620, KF700451; *B*. *tunux* Stensvold & Farrar, USA, Colorado, 2011, DF18791, ISC, KY191379*, KY191687*, KY192261*, KY191975*; *B*. *tunux* Stensvold & Farrar, USA, Montana, 2012, DF19483, ISC, KY191375*, KY191688*, KY192262*, KY191976*; *B*. *tunux* Stensvold & Farrar, USA, Alaska, 1998, DF3293, ISC, KY191378*, KY191689*, KY192263*, KY191977*; *B*. *tunux* Stensvold & Farrar, NOR, Finnmark, 2001, DF6027, ISC, KY191373*, KY191690*, KY192264*, KY191978*; *B*. *tunux* Stensvold & Farrar, USA, Nevada, 2003, DF9343, ISC, KY191376*, KY191691*, KY192265*, KY191979*; *B*. *watertonense* W.H.Wagner, USA, Montana, 1998, DF2896, ISC, KY191402*, —, KY192266*, KY191980*; *B*. *watertonense* W.H.Wagner, CAN, Alberta, 1983, W83332, MICH, KY191403*, KY191692*, KY192267*, KY191981*; *B*. *yaaxudakeit* Stensvold & Farrar, CAN, British Columbia, 2005, DF12686, ISC, KY191408*, KY191693*, KY192268*, KY191982*; *B*. *yaaxudakeit* Stensvold & Farrar, CAN, Yukon, 2006, DF15097, ISC, KY191407*, KF700539, KF700622, KF700453; *B*. *yaaxudakeit* Stensvold & Farrar, USA, Alaska, 2001, DF5359, ISC, KY191409*, KY191694*, KY192269*, KY191983*; *B*. *yaaxudakeit* Stensvold & Farrar, USA, Montana, 2001, DF5405, ISC, KY191405*, KY191695*, KY192270*, KY191984*; *B*. *yaaxudakeit* Stensvold & Farrar, CAN, Yukon, 2003, DF8464, ISC, KY191406*, KY191696*, KY192271*, KY191985*; *B*. *yaaxudakeit* Stensvold & Farrar, USA, Alaska, 2003, DF9655, ISC, KY191404*, KY191697*, KY192272*, KY191986*.

Frequency of haplotypes per taxon.

Taxa	n	No. combined haplotypes *matK* + *rpL16* + *trnL*^{UAA}—*trnF*^{GAA} + *trnH*^{GUG}—*psbA*																							
		Overall # combined haplotypes	H1	Freq. H1	H2	Freq. H2	H3	Freq. H3	H4	Freq. H4	H5	Freq. H5	H6	Freq. H6	H7	Freq. H7	H8	Freq. H8	H9	Freq. H9	H10	Freq. H10	H11	Freq. H11
B. alaskense W. H. Wagner & J.R.Grant var. *alaskense*	8	1	8	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. alaskense W. H. Wagner & J.R.Grant var. *salchaketense* J.R.Grant. var. nov. ined.	2	1	2	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. angustisegmentum (Pease & A. H. Moore) Fernald	5	5	1	0.20	1	0.20	1	0.20	1	0.20	1	0.20	–	–	–	–	–	–	–	–	–	–	–	–
B. ascendens W. H. Wagner	10	3	4	0.40	5	0.50	1	0.10	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. boreale × *B. lunaria 4*	1	1	1	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. boreale J. Milde	12	2	1	0.08	11	0.92	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. campestre var. lineare (W. H. Wagner) Farrar	7	6	1	0.14	1	0.14	1	0.14	1	0.14	2	0.29	1	0.14	–	–	–	–	–	–	–	–	–	–
B. campestre W. H. Wagner & Farrar var. campestre	5	5	1	0.20	1	0.20	1	0.20	1	0.20	1	0.20	–	–	–	–	–	–	–	–	–	–	–	–
B. crenulatum W. H. Wagner	7	5	1	0.14	3	0.43	1	0.14	1	0.14	1	0.14	–	–	–	–	–	–	–	–	–	–	–	–
B. dusenii (Christ) Alston	2	1	2	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. echo W. H. Wagner	5	1	5	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. farrarii Legler & Popovich, sp. nov. ined.	2	2	1	0.50	1	0.50	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. furculatum Popovich & Farrar, sp. nov. ined.	5	3	1	0.20	2	0.40	2	0.40	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. gallicomontanum Farrar & Johnson-Groh	5	2	1	0.20	4	0.80	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. hesperium (Maxon & R. T. Clausen) W. H. Wagner &Lellinger	7	3	2	0.29	4	0.57	1	0.14	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. lanceolatum "red" × *B. pallidum 1*	2	1	2	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. lanceolatum "red" × *B. pallidum 2*	1	1	1	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. lanceolatum (Gmel.) Ångstr. "green"	18	6	1	0.61	1	0.06	1	0.06	1	0.06	3	0.17	1	0.06	–	–	–	–	–	–	–	–	–	–
B. lanceolatum (Gmel.) Ångstr. "red"	12	5	8	0.67	1	0.08	1	0.08	1	0.08	1	0.08	–	–	–	–	–	–	–	–	–	–	–	–
B. lunaria (L.) Sw. var. *lunaria*	18	4	5	0.83	1	0.06	1	0.06	1	0.06	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. lunaria (L.) Sw. var. *melzeri* Stensvold & Farrar	3	3	1	0.33	1	0.33	1	0.33	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. lunaria 2	20	11	1	0.05	9	0.45	1	0.05	1	0.05	2	0.10	1	0.05	1	0.05	1	0.05	1	0.05	1	0.05	1	0.05
B. lunaria 3	10	5	2	0.20	5	0.50	1	0.10	1	0.10	1	0.10	–	–	–	–	–	–	–	–	–	–	–	–

B. lunaria 4	5	2	4	0.80	1	0.20	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. lunaria 5	11	6	2	0.18	1	0.09	1	0.09	5	0.45	1	0.09	1	0.09	–	–	–	–	–	–	–	–	–	–
B. matricariifolium (Döll) A. Braun	12	2	1	0.92	1	0.08	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. michiganense W. H. Wagner ex A. V. Gilman, Farrar &Zika	8	3	2	0.25	4	0.50	2	0.25	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. minganense Victorin	16	11	3	0.19	1	0.06	1	0.06	2	0.13	2	0.13	1	0.06	2	0.13	1	0.06	1	0.06	1	0.06	1	0.06
B. minganense × *B. echo*	2	1	2	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. montanum W. H. Wagner	6	5	1	0.17	1	0.17	2	0.33	1	0.17	1	0.17	–	–	–	–	–	–	–	–	–	–	–	–
B. mormo W. H. Wagner	5	2	1	0.20	4	0.80	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. neolunaria Stensvold & Farrar	20	5	6	0.80	1	0.05	1	0.05	1	0.05	1	0.05	–	–	–	–	–	–	–	–	–	–	–	–
B. nordicum Stensvold & Farrar	5	1	5	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. pallidum W. H. Wagner	7	3	5	0.71	1	0.14	1	0.14	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. paradoxum W. H. Wagner	8	4	4	0.50	1	0.13	1	0.13	2	0.25	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. pedunculosum W. H. Wagner	6	2	2	0.33	4	0.67	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. pinnatum H. St. John	9	2	3	0.33	6	0.67	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. pseudopinnatum W. H. Wagner	4	2	3	0.75	1	0.25	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. pumicola Colville	5	2	4	0.80	1	0.20	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. simplex 2	9	6	2	0.22	2	0.22	1	0.11	1	0.11	2	0.22	1	0.11	–	–	–	–	–	–	–	–	–	–
B. simplex E. Hitchc. var. *compositum* (Lasch) Milde	6	5	1	0.17	2	0.33	1	0.17	1	0.17	1	0.17	–	–	–	–	–	–	–	–	–	–	–	–
B. simplex E. Hitchc. var. *simplex*	3	3	1	0.33	1	0.33	1	0.33	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. spathulatum W. H. Wagner	7	1	7	1.00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. tenebrosum A. A. Eaton	11	4	4	0.36	1	0.09	5	0.45	1	0.09	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. tunux Stensvold & Farrar	13	8	4	0.31	1	0.08	1	0.08	1	0.08	1	0.08	2	0.15	2	0.15	1	0.08	–	–	–	–	–	–
B. watertonense W. H. Wagner	2	2	1	0.50	1	0.50	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
B. yaaxudakeit Stensvold & Farrar	6	3	1	0.17	4	0.67	1	0.17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Chapter 2

Ploidy Level and Genome Size Variation in the Homosporous Ferns *Botrychium* s.l. (Ophioglossaceae)

Benjamin Dauphin[1], Jason Grant[1] and Patrik Mráz[2,3]

[1]Laboratoire de botanique évolutive, Institut de Biologie, Université de Neuchâtel, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

[2]Ecology & Evolution Unit, Department of Biology, Université de Fribourg, Chemin du Musée 10, 1700 Fribourg, Switzerland

[3]Herbarium and Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ–12801 Praha, Czech Republic

This chapter was published in *Plant Systematics and Evolution* 302: 575–584.

Abstract

Recent cytological and molecular studies have investigated genome size variation and evolution in the homosporous ferns, but representatives of the Ophioglossaceae were largely overlooked, despite their evolutionary singularity. Flow cytometry analysis was performed on 41 individuals of eight species of the genera *Botrychium (B.), Botrypus*, and *Sceptridium* to estimate their ploidy level variation. In a subset of individuals, we also estimated the absolute genome size and corresponding C-values. Additionally, a classical chromosome count was made on the recently described species *B. alaskense*. Ploidy level and new genome size records were determined for *Botrychium alaskense*, *B. boreale, B. lanceolatum, B. "*neolunaria" ined., *B. pinnatum*, *Botrypus virginianus* and *Sceptridium multifidum*. In addition, we confirmed the genome size of *B. matricariifolium*, *B. minganense* and *B. lunaria*. Two of the three major sub-clades of *Botrychium* differ slightly in their averaged homoploid genome size (subclade Lanceolatum, 24.72 ± 0.40 pg; subclade Lunaria, 27.51 ± 0.47 pg). Flow cytometry and chromosome counting confirmed that *B. alaskense* is a tetraploid. A new hexaploid cytotype, putatively formed through an autopolyploidization from the sympatric tetraploid cytotype, was detected in a single individual of *B. boreale*. This is only the second report of hexaploidy in the genus *Botrychium* and our data highlight the potential to find other ploidy levels within other *Botrychium* species. Interestingly, no difference within the monoploid genome sizes was detected between ploidy levels, thus supporting the hypothesis of genome size stability after polyploidization and rejecting the scenario of genome downsizing.

Keywords: *Botrychium*; chromosome; homosporous ferns; flow cytometry; genome size; polyploidy.

Introduction

With the largest genome sizes and highest chromosome number in the plant kingdom, the homosporous ferns, including the Ophioglossaceae family, have a unique evolutionary history with deep divergences from over 300 mya (Pryer et al. 2004; Leitch et al. 2005). Belonging to this fern group, *Botrychium* s.l. (sensu latu, in the broad sense) (or grapeferns) includes five more narrowly circumscribed genera, namely, *Botrychium* (*B*. hereafter) s.s. (sensu strictu, in the narrow sense) (Milde) Clausen, *Botrypus* (L.) Michx, *Japanobotrychium* Masam, *Osmundopteris* (Milde) Clausen, and *Sceptridium* (Lyon) Clausen (Clausen 1938; Kato 1987; Hauk et al. 2003; Shinohara et al. 2013). Representatives of these genera differ mainly in sterile leaf blade size and shape. Their morphological differentiation is further supported by phylogenetic analyses (Hauk et al. 2012; Shinohara et al. 2013; Dauphin et al. 2014). Here, we focus on the cryptic species of *Botrychium* s.s. and its two sister genera *Botrypus* and *Sceptridium*, which have the same base of 45 chromosomes (Paris et al. 1989; Wagner 1993).

The largest genus within *Botrychium* s.l. is *Botrychium* s.s. or the moonworts, which include 30 species (Hauk et al. 2012; Dauphin et al. 2014), of which 28 have known ploidy levels. Twelve species were reported to be diploid (2n = 2x = 90), 15 tetraploid (2n = 4x = 180) and one species, *B*. *pseudopinnatum*, hexaploid (2n = 6x = 270) (Wagner 1993; Farrar 2011). However, the ploidy levels have not yet been reported for the two species *B*. *alaskense* and *B*. *boreale*. Based on these differing ploidy levels, it is obvious that polyploidy has played a key role in the speciation of *Botrychium* s.s. (Wood et al. 2009). Most importantly, the polyploid speciation in *Botrychium* has tightly been associated with interspecific hybridization, as suggested by morphological (Wagner and Lord 1956; Wagner and Grant 2002), karyological (Wagner and Lord 1956; Wagner 1993), allozyme (Hauk and Haufler 1999; Zika and Farrar 2009; Farrar 2011) and molecular phylogenetics data (Hauk et al. 2003; Williams and Waller 2012; Dauphin et al. 2014).

Despite the importance of polyploidization in the evolution of *Botrychium* s.s., records on ploidy level variation within the genus are rather scarce (Table 1), mainly because of methodological limits concerning (i) chromosome counting (up to 270 chromosomes for hexaploid specimens); and (ii) problematic long-term cultivation of specimens under experimental conditions because of their dependence on arbuscular mycorrhizae fungi (Kovác et al. 2007; Winther and Friedman 2007). However, the relatively recent introduction of flow cytometry into plant research may resolve the first methodological issue (Doležel et al. 2007). Flow cytometric analyses estimate ploidy levels of samples by measuring an absolute amount of the DNA in isolated nuclei; therefore, when used appropriately, this method might provide an important insight into genome size variation in plants. Absolute genome size is considered an important taxonomic criterion that may help to discriminate closely related taxa and infer their evolutionary histories (Leitch et al. 2005). To date, genome size has been measured in 12 species of *Botrychium* s.l. (Table 1) and consequently, it is still unknown for the majority of

species. Given the high chromosome number per nuclei in *Botrychium* s.l., their holoploid genome was considered to be either intermediate or large by Leitch et al. (2005), varying from a 2C value = 22.90 ± 1.82 pg in *Botrypus virginianus* to 53.68 ± 5.05 pg in *B. minganense* (Williams and Waller 2012). Importantly, when comparing the holoploid genome sizes in closely related genera of *Botrychium* s.l., Williams and Waller (2012) found that the genome of *Botrypus* (1Cx = 5.73 pg) was 2.3 times smaller than that of *Botrychium* s.s. (overall mean 1Cx = 13.36 pg), which supported their recognition as distinct genera.

The present study aimed to assess the ploidy levels and genome size variation from 41 specimens of eight species of *Botrychium, Botrypus* and *Sceptridium* collected in the Europe and North America. Using flow cytometry and the classical karyological approach, these data can be used to retrace genome evolution in the Ophioglossaceae which is the plant family with highest chromosome counts observed to date.

Materials and methods

Sampling

Living plants, including their below-ground parts, were dug up together with surrounding soil and co-occurring vegetation in Switzerland, Sweden and the United States (Alaska) during the summer of 2012. Plants were then potted and maintained at the botanical gardens of the Université de Neuchâtel and the Université de Fribourg in Switzerland until the analyses were conducted. After analysis, vouchers were deposited in the Herbarium of the Université de Neuchâtel (NEU) (except for the accession BD022C). All accessions sampled are listed in Table 2.

Chromosome counting

In the field, root tips of several individuals of *Botrychium alaskense* were taken, washed with distilled water and pre-treated with a bromo-naphthalene solution saturated in the aqueous phase for three hours. Fixation was performed using 10 mL of fixative solution containing 75% absolute ethanol, 24% of glacial acetic acid, 5 drops of carmine acetic acid and 1 drop of iron acetate (Sharma and Sharma 1980). The coloration step involved fixing root tips in a porcelain dish filled with carmine acetic acid and 10 ml of iron acetate with gentle heating for 2 minutes without ebullition. The crushed root tips and young cells in mitosis were viewed using a Leica Leitz microscope. We referred to previous chromosome counting for the remaining species analyzed in this study (Table 1).

Table 1. Published chromosome counts/DNA-ploidy level estimations of *Botrychium* s.s. with their geographical origins.

Taxa	Country	Nb.	Chromosome no. (2n) / Ploidy level	2C-value ± SD (pg) error	References
B. ascendens W.H.Wagner	USA	1	180/4x	-	Wagner (1993)
B. campestre W.H.Wagner & D.Farrar	USA	2	90/2x	-	Wagner (1993); Wagner and Wagner (1990)
B. crenulatum W.H.Wagner	USA	1	90/2x	-	Wagner (1993)
B. echo W.H.Wagner	USA	1	180/4x	51.63	Wagner (1993); Williams and Waller (2012)
B. gallicomontanum D.Farrar & Johnson-Groh	USA	1	Unknown/4x	-	Wagner (1993)
B. hesperium (Maxon & R.T.Clausen) W.H.Wagner & Lellinger	USA	1	180/4x	47.16	Wagner (1993); Williams and Waller (2012)
B. lanceolatum (S.G.Gmelin) Angström ssp. *angustisegmentum* (Pease & A.H.Moore) R.T.Clausen	USA	2	90/2x	-	Löve and Löve (1976); Wagner (1993)
B. lanceolatum (S.G.Gmelin) Angström ssp. *lanceolatum* Pease & A.H.Moore	ITA; GRL; USA	4	90/2x	29.56±0.95	Fabbri (1963); Löve and Löve (1976); Dalgaard (1989); Wagner (1993); Williams and Waller (2012)
B. lunaria (L.) Swartz	ITA; GEO; GRL; RUS;	5	90-60/2x	29.44±3.23	Zhukova et al. (1976);Dalgaard (1989); Wagner (1993); Gagnidze et al. (1998); Peruzzi et al. (2003); Williams and Waller (2012)
B. matricariifolium (Döll) A.Braun	USA	2	180/4x	50.82±1.22	Löve and Löve (1976); Wagner (1993); Williams and Waller (2012)
B. michiganense (W.H.Wagner) A.V.Gilmand, D.R.Farrar & P.F.Zika	USA	-	-/4x	46.64±2.85	Williams and Waller (2012)
B. minganense M.Victorin	USA	2	180/4x	53.68±5.05	Löve and Löve (1976); Wagner (1993); Williams and Waller (2012)
B. montanum W.H.Wagner	USA	1	90/2x	28.19	Wagner (1993); Williams and Waller (2012)
B. mormo W.H.Wagner	USA	1	90/2x	-	Wagner (1993)
B. pallidum W.H.Wagner	USA	2	90/2x	24.05±1.49	Wagner and Wagner (1990); Wagner (1993); Williams and Waller (2012)
B. paradoxum W.H.Wagner	USA	1	180/4x	-	Wagner (1993)
B. pedunculosum W.H.Wagner	USA	1	180/4x	-	Wagner (1993)
B. pinnatum H.St.John	ITA; USA	2	180/4x	-	Fabbri (1963); Wagner (1993)
B. pseudopinnatum W.H.Wagner	USA	2	270/6x	-	Wagner and Wagner (1990); Wagner (1993)
B. pumicola (Underw) Coville	USA	1	90/2x	-	Wagner (1993)
B. simplex E.Hitchc	USA	1	90/2x	22.05	Wagner (1993); Williams and Waller (2012)
B. spathulatum W.H.Wagner	USA	2	180/4x	53.07±5.43	Wagner and Wagner (1990); Wagner (1993)
B. watertonense W.H.Wagner	USA	1	180/4x	-	Wagner (1993)
Botrypus virginianus (L.) Michaux	USA	1	180/4x	22.90±1.82; 20.44±0.22	Wagner (1993); Bainard et al. (2011); Williams and Waller (2012)
Sceptridium multifidum (S.G.Gmelin) M.Nishida ex Tagawa	USA	1	90/2x	-	Wagner (1993)

Note. Nb = number of localities of each *Botrychium* species that was analyzed for chromosome counts; GEO: Georgia; GRL: Greenland; ITA: Italy; RUS: Russia; USA: United States.

Table 2. Botrychium s.l. samples analyzed using flow cytometry and chromosome squashes

Species	Accession	Country	Locality	Coordinates
Botrychium alaskense W.H.sWagner & J.R.Grant var. *alaskense*	BD047AC	USA	Alaska (Salcha River)	N 64°28'05.3'; W 146°44'58.1"
Botrychium alaskense W.H.Wagner & J.R.Grant var. *alaskense*	BD053BEG	USA	Alaska (320 miles, Richardson Highway)	N 64°25'41.0"; W 146°53'50.3"
Botrychium alaskense W.H.Wagner & J.R.Grant var. *alaskense*	BD068BC	USA	Alaska (322 miles, Richardson Highway)	N 64°26'47.1"; W 146°54'34.4"
Botrychium boreale J.Milde	BD022AD	SWE	Norrbotten province (Siknäs)	N 65°45'30.97"; W 22°40'03.22"
Botrychium boreale J.Milde	BD022C	SWE	Norrbotten province (Siknäs)	N 65°45'30.97"; W 22°40'03.22"
Botrychium lanceolatum (S.G.Gmelin) Ångström ssp. *lanceolatum*	BD022EF	SWE	Norrbotten province (Siknäs)	N 65°45'30.97"; W 22°40'03.22"
Botrychium lanceolatum (S.G.Gmelin) Ångström ssp. *lanceolatum*	BD034C	SWE	Norrbotten province (Pissiniemi)	N 67°34'57"; E 23°33'19"
Botrychium lanceolatum (S.G.Gmelin) Ångström ssp. *lanceolatum*	BD053K	USA	Alaska (320 miles, Richardson Highway)	N 64°25'41.0"; W 146°53'50.3"
Botrychium lanceolatum (S.G.Gmelin) Ångström ssp. *lanceolatum*	BD063DDX	USA	Alaska (86 miles, Seward Highway)	N 60°54'25.6"; W 149°05'01.1"
Botrychium lunaria (L.) Swartz	BD014A	CH	Graubünden canton (Prontresina, Val Roseg)	N 46°26'00.24"; E 9°51'46.84"
Botrychium lunaria (L.) Swartz	BD016B	CH	Uri canton (Gurtnellen)	N 46°43'47.89"; E 8°36'30.95"
Botrychium lunaria (L.) Swartz	BD035A	SWE	Norrbotten province (Kätkesuando)	N 68°06'59"; E 23°20'19"
Botrychium matricariifolium (Döll) A.Braun	BD016A	CH	Uri canton (Gurtnellen)	N 46°43'47.89"; E 8°36'30.95"
Botrychium minganense M.Victorin	BD063AD	USA	Alaska (86 miles, Seward Highway)	N 60°54'25.6"; W 149°05'01.1"
Botrychium minganense M.Victorin	BD065A	USA	Alaska (88 miles, Seward Highway)	N 60°55'07.9"; W 149°07'50.8"
Botrychium "neolunaria M.Stensvold" ined.	BD058AC	USA	Alaska (8 miles, Chena Hot Springs Road)	N 64°53'03.5"; W 147°21'45.6"
Botrychium "neolunaria M.Stensvold" ined.	BD063LR	USA	Alaska (86 miles, Seward Highway)	N 60°54'25.6"; W 149°05'01.1"
Botrychium pinnatum H.St.John	BD061AI	USA	Alaska (Anchorage, the "Dome")	N 61°10'13.6"; W 149°39'18.6"
Botrychium pinnatum H.St.John	ABI1	USA	Alaska (Dutch harbor)	N 53°53'58"; W 166°33'05"
Botrychium pinnatum H.St.John	L1245316	USA	Washington state (Colville National Forest)	N48°75'12.05"; W117°19'67.46"
Botrypus virginianus (L.) Michx.	BD063Q	USA	Alaska (86 miles, Seward Highway)	N 60°54'25.6"; W 149°05'01.1"
Sceptridium multifidum (S.G.Gmelin) M. Nishida ex Tagawa	BD034A	SWE	Norrbotten province (Pissiniemi)	N 67°34'57"; E 23°33'19"

USA: United States of America; SWE: Sweden; CH: Switzerland. Vouchers were deposited at the Herbarium of the University of Neuchâtel (NEU), Switzerland.

Flow cytometry

A Partec CyFlow SL flow cytometer (Partec GmBH, Münster, Germany) equipped with a green 532-nm laser was used to assess ploidy level in 41 accessions of *Botrychium* (Table 3). The samples were prepared from fresh leaf tissue by chopping them with razor blades in 1 ml of general-purpose buffer (0.5 mM spermine, 30 mM sodium citrate, 20 mM 4-morpholine propane sulfonate, 80 mM KCl, 20 mM NaCl, 0.5% Triton X-100, pH = 7.0) with subsequent incubation for 5 minutes, according to Loureiro et al. (2007). Samples were then filtered through a green Partec filter (40 μm) and stained with a 30-μL mixture of propidium iodide and RNase (ratio 1:1, both of initial concentration of 1%). *Secale cereale* cv. Dankovské (2C = 16.19 pg, Doležel et al. 1998) was used as an internal standard for all species and *Vicia faba* ssp. *faba* var. *equine* cv. Inovec (2C = 26.90 pg) (Doležel et al. 1992; Doležel et al. 1998) was used only for *Sc*. *multifidum* and all polyploid species (see Table 3). After initial ploidy level screening, we estimated the absolute genome size in 15 individuals selected from the fittest plants and those that provided enough material for repeated measurements. To isolate and stain nuclei, we used the same protocol as above. Two internal standards were used: *Secale cereale* cv. Dankovské and *Vicia faba* ssp. *faba* var. *equine* cv. Inovec. Holoploid (2C-value) and monoploid genome size (1Cx-value) were subsequently

calculated according to the method of Greilhuber et al. (2005). Each plant was analyzed once a day on three consecutive days to minimize instrumental errors (Doležel et al. 2007), except for four plants (both individuals of *B*. *lanceolatum* ssp. *lanceolatum* and *B*. "neolunaria" ined.), that were analyzed only twice with *Vicia faba* as the standard. Histograms were accumulated at a flow rate of approximately 10–30 particles per second for a total count of 1500 (ploidy level estimations) and 5000 (absolute genome size estimations) nuclei. Coefficient of variations (CV) of the peaks of internal standards ranged from 1.78% to 6.47%, with an average value of 3.58%; and the CV of peaks of measured samples varied between 1.96% and 5.46%, with an average value of 3.61%. Measurements exceeding 3% divergence between independent runs were usually discarded and the sample was re-analyzed. Repeated measurements were averaged to obtained mean 2C-values per accession.

Statistical analyses

The estimated genome sizes of this study were averaged per species with 11 published records in Williams and Waller (2012) (Table 1). In total, five diploid and nine tetraploid species of *Botrychium* s.s. were considered in a combined dataset. Monoploid genome size was assessed with the holoploid genome size divided by the ploidy level. Normality assumptions were tested on residuals of a linear model with the Shapiro–Wilk normality test using the package Stats in the R statistical software (R Development Core Team 2015). Overall mean genome size comparison between ploidy levels was performed following the Welch t-test for both holo- and monoploid genome size. The graphical representation of quantiles was plotted with ploidy level as the clustering factor.

Results

Ploidy level in Botrychium s.l.

The ploidy levels for 41 accessions belonging to eight species were inferred by flow cytometry analyses and are listed in Table 3. Among them, three species were confirmed to be diploid, four species were tetraploid and for one species, *B. boreale*, we found both tetra- and hexaploid plants. The hexaploid record is only the second one reported in the genus *Botrychium* (see Discussion). Tetraploidy is confirmed here for the first time in both *B. boreale* and *B. alaskense*. In addition to flow cytometric estimations, the tetraploid status of *B. alaskense* was corroborated by a classical chromosome counting (2n = 4x = ca 180), based on x = 45 (Wagner 1993) (Fig. 1).

Genome size variation in Botrychium s.l.

The absolute genome size values (2C) and their derived 1Cx values determined for 15 accessions belonging to nine taxa are listed in Table 4. Genome size estimations were fairly similar when two different standards were used (*Secale cereale* and *Vicia faba*), with the maximum divergence (3.54%) recorded in *B. alaskense* (BD053G). The 2C-values in diploid *Botrychium* species ranged from 24.72 ± 0.40 pg to 27.51 ± 0.60 pg (Table 4; Fig. 2.a). However, in the closely related diploid *Sceptridium multifidum*, we recorded the smallest genome size with 16.11 ± 0.06 pg (Table 4; Fig. 2.b). In tetraploids, the 2C-values measured against *Secale* and *Vicia* standards ranged from 49.69 ± 0.91 to 52.07 ± 1.6 pg and from 50.51 ± 0.50 to 52.14 ± 1.16 pg, respectively. The smallest genome size (2C = 18.93 ± 0.26 pg) among tetraploids was recorded in *Botrypus virginianus* (Table 4). The genome

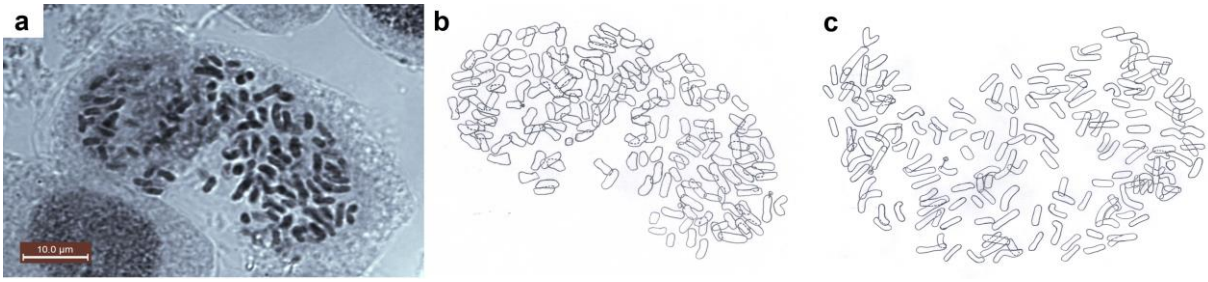


Fig. 1. Mitotic metaphase chromosomes of tetraploid (2n=4x=180) Botrychium alaskense (a and b: plant BD1204805 and c: plant BD04803; Salcha River population, Alaska, U.S.A.; N 64°28'05.3", W 146°44'58.1"). The drawings were made by Philippe Küpfer and Benjamin Dauphin.

Table 3. Ploidy level variation in *Botrychium* species analyzed with *Secale cereale* (2C-value = 16.19 pg) as an internal standard.

Species	# plants/ # sites	Mean fluorescence ratio	Min - Max fluorescence ratio range	SD	Estimated ploidy level
Botrychium lanceolatum ssp. *lanceolatum*	10/4	1.46	1.42 - 1.50	0.52	2
Botrychium lunaria	3/3	1.69	1.65 - 1.74	0.75	2
Botrychium "neolunaria" ined. [1]	4/2	1.67	1.56 - 1.76	1.55	2
Botrychium alaskense	8/3	2.97	2.86 - 3.05	1.21	4
Botrychium boreale	2/1	3.10	3.04 - 3.16	1.37	4
Botrychium matricariifolium	1/1	2.38	-	-	4
Botrychium minganense	4/2	2.73	2.67 - 2.80	1.58	4
Botrychium pinnatum	9/3	2.96	2.88 - 3.10	1.10	4
Botrychium boreale [2]	1/1	4.55	-	-	6

[1]The North American plants of the *B. lunaria* complex have been proposed as a distinct, yet unpublished, taxon;

[2]Analyzed using *Vicia faba* because of the large genome size.

size of one hexaploid individual of *B. boreale* (BD022C)was estimated as 79.09 ± 4.02 pg with *Secale*, and 76.52 ± 1.31 pg with *Vicia* as standards, respectively (Table 4; Fig. 2.f).

Holo- and monoploid genome size comparison in Botrychium s.s.

For the holoploid genome size, median values differed greatly between ploidy levels, with 27.14 pg for diploid and 50.82 pg for tetraploid species, whereas they were very similar between ploidy levels for monoploid genome sizes: 13.57 pg for diploids and 12.71 pg for tetraploids (Fig. 3). Additionally, holoploid genome sizes varied significantly ($p < 0.001$) between ploidy levels, with a 2C-value mean of 25.98 pg for diploids and 50.65 pg for tetraploids (Fig. 3.a). In contrast, monoploid genome sizes were not significantly different between diploids and tetraploids ($p = 0.64$), displaying very similar means of 12.99 pg and 12.66 pg for diploid and tetraploid species, respectively (Fig. 3.b).

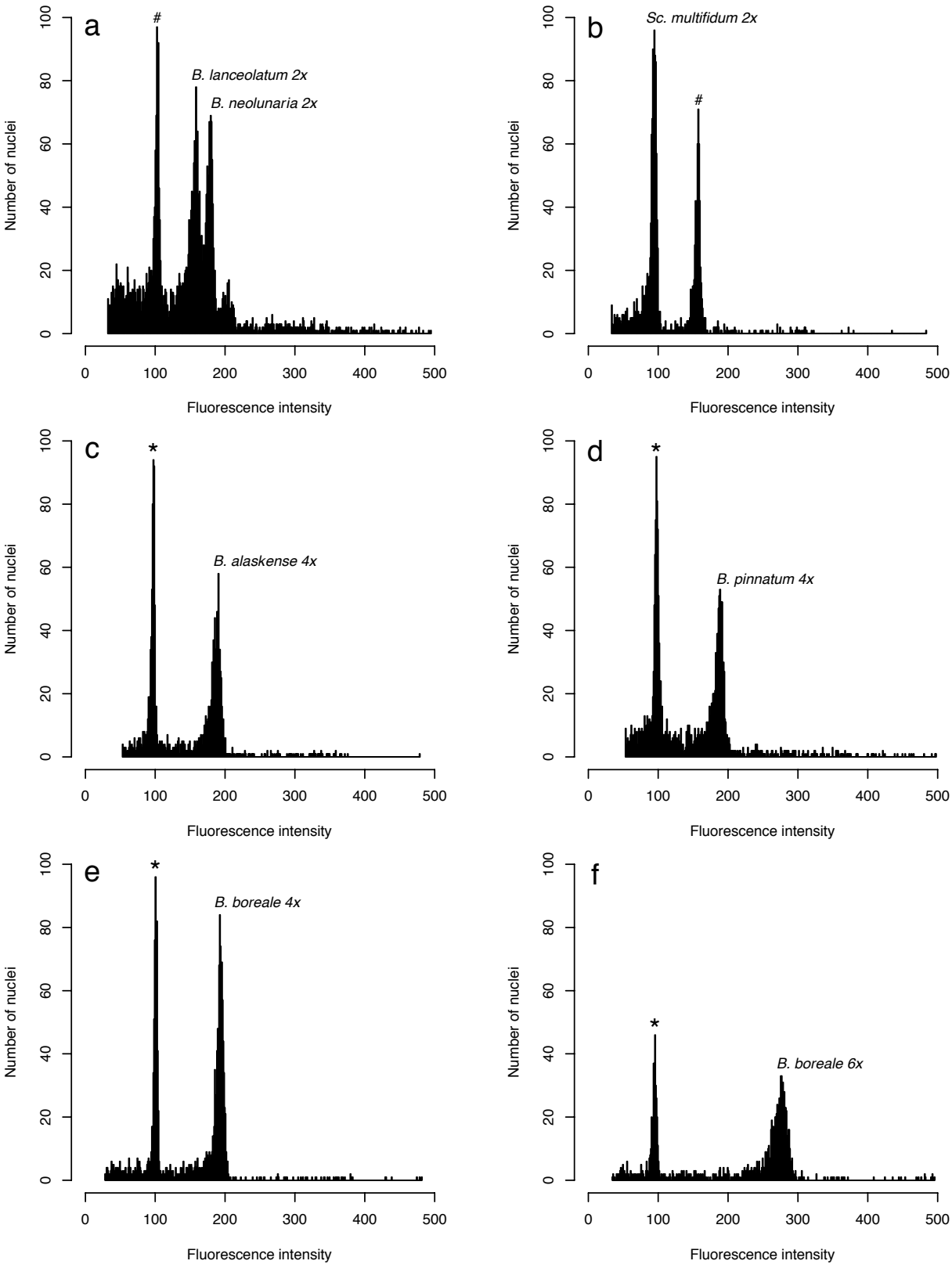


Fig. 2. Histograms of the absolute DNA content of Botrychium s.l.: a) B. lanceolatum 2x and B. "neolunaria" 2x; b) Sc. multifidum 2x; c) B. alaskense 4x; d) B. pinnatum 4x; e) B. boreale 4x, and f) B. boreale 6x. The two internal standards are symbolized by "#" for Secale cereale and "*" for Vicia faba.

Table 4. Absolute genome size values of 15 accessions of seven species of the genera *Botrychium*, *Botrypus* and *Sceptridium*, compared with two internal standards (*Secale cereale* and *Vicia faba*).

Species	Accession	Ploidy level (x)	2C-value *Secale* (pg) with (± SD)	1Cx-value *Secale* (pg)	2C-value *Vicia* (pg) with (± SD)	1Cx-value *Vicia* (pg)
Sceptridium multifidum	BD034A	2	NM	NM	16.11 (± 0.06)	8.06
Botrypus virginianus	BD063Q	4	NM	NM	18.93 (± 0.26)	4.73
Botrychium lanceolatum ssp. *lanceolatum*	BD063T	2	24.62 (± 0.37)	12.31	NM	NM
Botrychium lanceolatum ssp. *lanceolatum*	BD063U	2	24.81 (± 0.49)	12.41	NM	NM
Mean values for *B. lanceolatum*			24.72 (± 0.40)	12.36	-	-
*Botrychium "*neolunaria*"* ined.[1]	BD058A	2	27.26 (± 0.17)	13.63	NM	NM
*Botrychium "*neolunaria*"* ined.[1]	BD058B	2	27.75 (± 0.60)	13.87	NM	NM
Mean values for *B. "*neolunaria*"* ined.[1]			27.51 (± 0.47)	13.75	-	-
Botrychium alaskense	BD053E	4	50.77 (± 0.49)	12.69	51.96 (± 0.32)	12.99
Botrychium alaskense	BD053G	4	50.36 (± 1.01)	12.59	52.21 (± 0.52)	13.05
Botrychium alaskense	BD053B	4	50.54 (± 0.78)	12.63	52.06 (± 0.12)	13.01
Botrychium alaskense	BD068B	4	50.91 (± 0.60)	12.73	51.04 (± 0.10)	12.76
Mean values for *B. alaskense*			50.64 (± 0.67)	12.66	51.82 (± 0.52)	12.95
Botrychium pinnatum	BD061D	4	50.20 (± 0.88)	12.55	50.58 (± 0.80)	12.64
Botrychium pinnatum	BD061H	4	49.17 (± 0.72)	12.29	50.45 (± 0.31)	12.61
Mean values for *B. pinnatum*			49.69 (± 0.91)	12.42	50.51 (± 0.50)	12.63
Botrychium boreale	BD022A	4	52.70 (± 0.47)	13.17	52.63 (± 0.16)	13.16
Botrychium boreale	BD022B	4	51.43 (± 0.61)	12.86	51.65 (± 0.33)	12.91
Botrychium boreale	BD022C	6	79.09 (± 4.02)	13.18	76.52 (± 1.31)	12.75
Mean values for *B. boreale*			-	13.07	-	12.94
Overall mean values for diploid species of *Botrychium* s.s.			26.11 (± 1.52)	13.06	-	-
Overall mean values for tetraploid species of *Botrychium* s.s.			50.80 (± 1.14)	12.72	51.49 (± 0.81)	12.84

NM: not measured;[1]The North American genotype of the *B. lunaria* complex has been proposed as a distinct, yet not described, species.

Discussion

Ploidy level and genome size variation in Botrychium s.l. *and related species*

Our results confirmed the previously reported ploidy levels of six species (*B. lanceolatum, B. lunaria, B. "*neolunaria" ined., *B. matricariifolium*, *B. minganense*, and *B. pinnatum*), that were mostly based on exact chromosome counts (Wagner 1955; Wagner and Lord 1956; Wagner and Wagner 1986; Wagner et al. 1990; Wagner 1993, Stensvold 2007), or more recently, by flow cytometry on specimens obtained exclusively from North America (Bainard et al 2011; Bai et al 2012; Williams and Waller 2012). Here, we analyzed specimens from the Alps and Scandinavia, as well as North America. In addition, we determined the tetraploid level of *B. alaskense* and *B. boreale* for the first time, and found intraspecific variation in the latter taxon, showing both tetra- and hexaploid cytotypes. The hexaploid record is only the second one in the genus *Botrychium*, and the

first outside North America. Previously, hexaploidy had only been reported in one population of *B. pseudopinnatum* from the northwest shore of Lake Superior (Ontario, Canada) (Wagner et al. 1990).

The absolute genome sizes were assessed for the first time in four species *B. alaskense*, *B. boreale*, *B. pinnatum* and *Sceptridium multifidum* (see Table 4). For three species (*B. lanceolatum*, *B.* "neolunaria" ined. and *Botrypus virginianus*) with previously published genome sizes (Bainard et al. 2011; Williams and Waller 2012), we found very similar values, despite different internal standards, protocols and types of material (fresh *vs*. dried) used. Exceptions occurred when the genome size of the standard species was too different from that of the species being analyzed, as was the case with Zea mays and *Botrypus virginianus* with 22.9 ± 1.8 pg (Williams and Waller 2012; in our study 2C = 18.93 ± 0.26 pg), or *B. lanceolatum* with 30.44 ± 0.9 pg (Williams and Waller 2012; in our study 2C = 24.72 ± 0.40 pg). Conversely, when *Allium cepa* was used as an internal standard (Bainard et al. 2011), the difference between the reported (2C = 20.44 pg) and our genome size estimate (2C = 18.93 ± 0.26 pg) for *Botrypus virginianus* was 7.4%. Thus, with an appropriate internal standard (i.e. with the genome size value close to that of measured samples), genome size can be estimated on material dried in silica gel, which however exhibits higher coefficient of variation (around 5%) (Williams and Waller 2012). This method may even be preferable, since *Botrychium* is hardly ever cultivated because of its obligate arbuscular mycorrhizal fungal symbionts (Clausen 1938), in addition to the problem of extracting living specimens in populations of rare or endangered species (Casson et al. 2002).

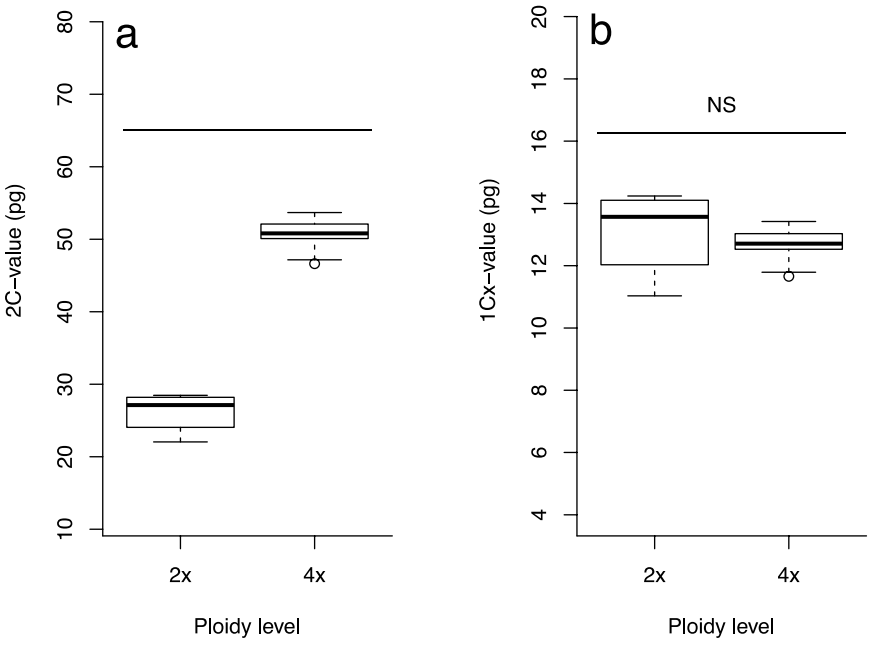


Fig. 3. Averaged 2C (a) and 1Cx (b) values of Botrychium s.s. per ploidy level, including published genome sizes in Williams and Waller (2012). It includes five diploid and nine tetraploid species. The significance of the statistical test (t-test) is given for both holoploid and monoploid genome sizes.

The newly recognized species *B. "*neolunaria*"* ined. shows almost identical genome size (± 0.02 of the mean fluorescence ratio, Table 3) to *B. lunaria*, supporting their close relationship in the cryptic Lunaria group, which has been corroborated by phylogenetic inferences based on plastid regions, where both taxa are positioned with the allopolyploid *B. yaaxudakeit* (*B. lunaria* × *B. "*neolunaria*"* ined.) in the Lunaria clade (Stensvold 2008; Dauphin et al. 2014). Furthermore, differences in monoploid genome sizes between *Botrychium* s.s. (1Cx = 12.83 pg), *Botrypus* (1Cx = 4.73 pg) and *Sceptridium* (1Cx = 8.06 pg) were consistent with their phylogenetic placement within distant clades (Hauk et al. 2003).

Absolute genome size variation and its relationships to auto- and allopolyploid events

Based on its morphology *Botrychium boreale* has long been considered a tetraploid species (Hultén 1968; William 1996). We confirmed this hypothesis and even found intraspecific variation in ploidy level. At the Siknäs site (Norrbotten, Sweden), we found one hexaploid individual intermixed (the minimum distance between 4x – 6x plants was 5 cm) within the typical tetraploid population of *B. boreale*, raising questions about its origin. The following arguments support its autopolyploid origin from co-occurring allotetraploid *B. boreale*. First, the hexaploid plant was morphologically indistinguishable from the co-occurring tetraploids. Second, it had almost exactly 1.5-fold the genome size of tetraploid *B. boreale* plants (1.52- and 1.47-fold with *Secale* and *Vicia*, respectively, as the standards), which suggested a fusion of reduced (2N) and unreduced gametes (4N) (see below). Third,

hexaploid and tetraploid plants (BD12022C) from this locality shared the same plastid haplotype (Dauphin et al. 2014). These data suggested that both polyploidization mechanisms (auto- and allo-) could operate in *Botrychium*, although allopolyploidization seems to be more frequent in this genus (Farrar 2011). Nevertheless, because allopolyploidy is more easily detected than autopolyploidy (Soltis et al. 2007), the latter hybridization process could be underestimated in *Botrychium*. As the fusion of reduced and unreduced gametes is the most common pathway of polyploidization (Ramsey and Schemske 1998), we expect that this process (fusion of 2N and 4N gametes) was also involved in the formation of the hexaploid plant. The hybridogeneous origin of *B. boreale* could have contributed to this pathway by increasing the frequency of unreduced gamete formation, which is 13.5-times higher in allo- ("outcrossing taxa") than in autopolyploid species ("selfing taxa") (Ramsey and Schemske 1998).

Based on allozyme data and morphology, *B. alaskense*, *B. pinnatum* and *B. boreale* have been suggested to be allotetraploids that arose from hybridization between the two diploid species, *B. lanceolatum* and *B. lunaria* / *B.* "neolunaria" ined. (Wagner 1993; Farrar 2011), which diverged ca. 5 mya (Stensvold 2008). The small, but still distinguishable, difference in their genome sizes (2C = 24.72 ± 0.40 pg in *B. lanceolatum,* 2C = 27.51 ± 0.47 pg in *B. lunaria*, Fig. 2.a), tetraploid *B. boreale* (2C = 52.07 ± 0.50 pg) showed almost an exact addition of both putative genomes. Thus, our genome size data supports the hybrid hypothesis proposed by Wagner (1993) and Farrar (2011). In contrast, the two remaining tetraploids had a slightly lower amount of nuclear DNA than expected under a hybrid scenario and complete additivity: 2C = 50.64 and 2C = 49.69 pg for *B. alaskense* and *B. pinnatum*, respectively.

Genome size stability

Despite significant variation between ploidy levels of holoploid genome sizes in *Botrychium* s.s., no difference was detected for monoploid genome size (Fig. 3). Yet, we had hypothesized genomic mechanisms leading to a size reduction in polyploid genomes compared to diploids, as has often been reported in plants (Leitch and Bennett 2004). Interestingly, genome downsizing seems to be ineffective or absent in this genus, which could explain why the genome sizes of *Botrychium* s.l. taxa are among the highest in vascular plants. Our study thus corroborates the pattern found in other groups of ferns (Henry et al. 2014). Accordingly, a low activity of transposable elements (Brandes et al. 1997) combined with a high retention rate of chromosomes (Barker 2013) may lead to an unprecedented conservation in genome sizes of homosporous ferns.

Conclusions

This study reports new ploidy levels and genome size estimates in the genus *Botrychium*. Our data confirm the predominance of polyploid taxa in the genus, and in addition to allopolyploidy, flow cytometric analysis suggests that autopolyploidization may have been involved in the origin of the hexaploid cytotype of *B. boreale*. Applications of efficient flow cytometry methods and more detailed sampling have revealed higher ploidy level variation in the genus than had previously been thought. Intermediate patterns in genome sizes in *B. alaskense*, *B. boreale* and *B. pinnatum* support the hypothesis that their allopolyploid origin arose between two diploid taxa, *B. lanceolatum* and *B. lunaria* / *B.* "neolunaria" ined. Finally, no difference in monoploid genome size was detected between ploidy levels indicating that strong genomic mechanisms must allow the stability of those genomes among the largest land plants.

Acknowledgments

This work was supported by the Fonds Marguerite Wüthrich and A. Matthey-Dupraz at the Université de Neuchâtel, the Fonds Dr. Joachim de Giacomi of the Swiss Academy of Natural Sciences, the Bourse de voyages of the Swiss Academy of Natural Sciences, and the Fonds de donations of the Université de Neuchâtel. The authors thank Lennart Stenberg for assistance with fieldwork in Sweden, and Anne-Catherine Pasche and Philippe Küpfer for help with laboratory work.

References

Bai C, Alverson WS, Follansbee A, Waller DM (2012) New reports of nuclear DNA content for 407 vascular plant taxa from the United States. Ann Bot 110: 1623–1629. doi: 10.1093/aob/mcs222

Bainard JD, Henry T, Bainard LD, Newmaster SG (2011) DNA content variation in monilophytes and lycophytes: Large genomes that are not endopolyploid. Chromosom Res 19: 763–775. doi: 10.1007/s10577-011-9228-1

Barker MS (2013) Karyotype and genome evolution in pteridophytes. In Plant genome Divers vol 2, Phys Struct Behav Evol plant genomes. Edited by Leitch IJ, Greilhuber J, Doležel J, Wendel JF. Wien: Springer. pp 245-253. doi: 10.1007/978-3-7091-1160-4

Bennett MD, Leitch IJ (1997) Nuclear DNA amounts in angiosperms: 583 new estimates. Ann Bot 80: 169–196. doi: 10.1006/anbo.1997.0415

Bennetzen JL, Ma J, Devos KM (2005) Mechanisms of recent genome size variation in flowering plants. Ann Bot 95: 127–132. doi: 10.1093/aob/mci008

Brandes A, Heslop-Harrison JS, Kamm A, Kubis S, Doudrick RL, Schmidt T (1997) Comparative analysis of the chromosomal and genomic organization of Ty1-copia-like retrotransposons in pteridophytes, gymnosperms and angiosperms. Plant Mol Biol 33: 11-21. doi: 10.1023/A:1005797222148

Casson JI, Shackleford I, Parker I, Schultz J (2002) Conservation approach for goblin fern (*Botrychium mormo* W.H. Wagner). USDA Forest Service, Eastern Region, Milwaukee, Wisconsin, USA.

Clausen RT (1936) A monograph of the Ophioglossaceae. Mem Torrey Bot Club 19: 1–177.

Dalgaard V (1989) Additional chromosome numbers in vascular plants from the Disko Bugt area (west Greenland). Willdenowia 19: 199–213.

Dauphin B, Vieu J, Grant JR (2014) Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* s.s., Ophioglossaceae). Am J Bot 101: 128–40. doi: 10.3732/ajb.1300154

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák MA, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann Bot 82: 17–26.

Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc 2: 2233–2244. doi:10.1038/nprot.2007.310

Doležel J, Sgorbati S, Lucretti S (1992) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiol Plant 85: 625–631. doi: 10.1111/j.1399-3054.1992.tb04764.x

Fabbri F (1963) Primo supplement alle Tavole Cromosomiche delle Pteridophyta di Alberto Chiarugi. Caryologia 16: 237–335.

Farrar DR (2011) Systematics and taxonomy of genus *Botrychium*, pp 1–29. Available at: http: //www.public.iastate.edu/~herbarium/botrychium.html

Flavell RB, Bennett MD, Smith JB, Smith DB (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genet 12: 257–269.

Gagnidze RI, Tsn G, Danelia IM, Churadze MV (1998) Chromosome numbers of the species of the Georgian flora. Bot Zhurn 83: 143–147.

Greilhuber J, Doležel J, Lysák MA, Bennett MD (2005) The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. Ann Bot 95: 255–260. doi: 10.1093/aob/mci019

Hauk WD (1995) A molecular assessment of relationships among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). Am Fern J 85: 375–394. doi: 10.2307/1547816

Hauk WD, Parks CR, Chase MW (2003) Phylogenetic studies of Ophioglossaceae: evidence from rbcL and trnL-F plastid DNA sequences and morphology. Mol Phylogenet Evol 28: 131–151. doi: 10.1016/S1055-7903(03)00032-0

Hauk WD, Kennedy L, Hawke HM (2012) A phylogenetic investigation of *Botrychium* ss (Ophioglossaceae): Evidence from three plastid DNA sequence datasets. Syst Bot 37: 320–330. doi: 10.1600/036364412X635377

Henry TA, Bainard JD, Newmaster, SG (2014) Genome size evolution in Ontario ferns (Polypodiidae): evolutionary correlations with cell size, spore size, and habitat type and an absence of genome downsizing. Genome 57: 555–566. doi: 10.1139/gen-2014-0090

Hultén E (1968) Flora of Alaska and neighboring territories – A manual of the vascular plants, Stanford University Press, pp 40.

Kato M (1987) A phylogenetic classification of Ophioglossaceae. Gard. Bull. Singapore 40: 1–14.

Kellogg EA, Bennetzen JL (2004) The evolution of nuclear genome structure in seed plants. Am J Bot 91: 1709–1725. doi: 10.3732/ajb.91.10.1709

Kovács GM, Balázs T, Pénzes Z (2007) Molecular study of arbuscular mycorrhizal fungi colonizing the sporophyte of the eusporangiate rattlesnake fern (*Botrychium virginianum*, Ophioglossaceae). Mycorrhiza 17: 597–605. doi: 10.1007/s00572-007-0137-2

Leitch IJ, Bennett MD (2004) Genome downsize in polyploids plants. Biol J Linn Soc 82: 651–663. doi: 10.1111/j.1095-8312.2004.00349.x

Leitch IJ, Soltis DE, Soltis PS, Bennett MD (2005) Evolution of DNA amounts across land plants (Embryophyta). Ann Bot 95: 207–17. doi: 10.1093/aob/mci014

Loureiro J, Rodriguez E, Doležel J, Santos C (2007) Two new nuclear isolation buffers for plant DNA flow cytometry: A test with 37 species. Ann Bot 100: 875–888. doi: 10.1093/aob/mcm152

Löve A, Löve D (1976) IOPB chromosome number reports LIII. Taxon 25: 483–500.

Lysák M, Doležel J (1998) Estimation of nuclear DNA content in *Sesleria* (Poaceae). Caryologia 51: 123–132. doi: 10.1080/00087114.1998.10589127

Obermayer R, Leitch IJ, Hanson L, Bennett MD (2002) Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. Ann Bot 90: 209–217. doi: 10.1093/aob/mcf167

Paris CA, Wagner FS, Wagner WH (1989) Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. Am Fern J 79: 46–54. doi: 10.2307/1547159

Peruzzi L, Cesca G, Puntillo D (2003) *Isoëtes* (Isoetaceae), *Ophioglossum* and *Botrychium* (Ophioglossaceae) in Calabria (S Italy): more karyological and taxonomical data. Caryologia 56: 359–363. doi: 10.1080/00087114.2003.10589345

Pryer KM, Schuettpelz E, Wolf PG, Schneider H, Smith AR, Cranfill R (2004) Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. Am J Bot 91: 1582–1598. doi: 10.3732/ajb.91.10.1582

R Development Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna.

Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu Rev Ecol Syst 29: 467–501.

Ramsey J, Schemske DW (2002) Neopolyploidy in flowering plants. Annu Rev Ecol Syst 33: 589–639.

Sharma AK, Sharma A (1980) Chromosome Techniques – Theory and Practice, Butterwort, pp 354–355.

Shinohara W, Nakato N, Yatabe-Kakugawa Y, Oka T, Kim JK, Murakami N, Noda H, Sahashi N (2013) The use of *matK* in Ophioglossaceae phylogeny and the determination of Mankyua chromosome number shed light on chromosome number evolution in Ophioglossaceae. Syst Bot 38: 564–570. doi: 10.1600/036364413X670232

Soltis DE, Soltis PS, Bennett MD, Leitch IJ (2003) Evolution of genome size in the angiosperms. Am J Bot 90: 1596–1603. doi: 10.3732/ajb.90.11.1596

Soltis, DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband, BC, Judd WS (2007) Autopolyploidy in angiosperms: have we grossly underestimated the number of species? Taxon 56: 13–30. doi: 10.2307/25065732

Stensvold MC (2008) A taxonomic and phylogeographic study of the *Botrychium lunaria* complex. PhD Thesis, University of Iowa, Iowa, pp 142.

Veselý P, Bureš P, Šmarda P, Pavlíček T (2012) Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? Ann Bot 109: 65–75. doi: 10.1093/aob/mcr267

Wagner FS (1993) Chromosomes of North American grapeferns and moonworts (Ophioglossaceae: *Botrychium*). Contrib from Univ Michigan Herb 19: 83–92.

Wagner WH (1955) Cytotaxonomic observations on North American ferns. Rhodora 57: 219–240.

Wagner WH, Lord LP (1956) The morphological and cytological distinctness of *Botrychium minganense* and *B. lunaria* in Michigan. Bull Torrey Bot Club 83: 261–280.

Wagner WH, Wagner FS (1980) Polyploidy in pteridophytes, Springer University Press, pp 199–214.

Wagner WH, Wagner FS (1981) New species of moonworts, *Botrychium* subg. *Botrychium* (Ophioglossaceae) from North America. Am Fern J 71: 20–30. doi: 10.2307/1546672

Wagner WH, Wagner FS (1983) Two moonworts of the Rocky Mountains; *Botrychium hesperium* and a new species formerly confused with it. Am Fern J 73: 53–62. doi: 10.2307/1546534

Wagner WH, Wagner FS (1986) Three new species of moonworts (*Botrychium* subg. *Botrychium*) endemic in western North America. Am Fern J 76: 33–47. doi: 10.2307/1547557

Wagner WH, Wagner FS (1990) Notes on the fan-leaflet group of moonworts in North America with descriptions of 2 new members. Am Fern J 80: 73–81. doi: 10.2307/1547172

Wagner WH, Grant JR (2002) *Botrychium alaskense*, a new moonwort from the interior of Alaska. Am Fern J 92: 164–170. doi: 10.1640/0002-8444(2002)092[0164:BAANMF]2.0.CO;2

William JC (1996) Flora of the Yukon Territory, NRC Research Press, pp 33.

Williams EW, Waller DM (2012) Phylogenetic placement of species within the genus *Botrychium* s.s. (Ophioglossaceae) on the basis of plastid sequences, amplified fragment length polymorphisms, and flow cytometry. Int J Plant Sci 173: 516–531. doi: 10.1086/664711

Winther JL, Friedman WE (2008) Arbuscular mycorrhizal associations in Lycopodiaceae. New Phytol 177: 790–801. doi: 10.1111/j.1469-8137.2007.02276.x

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH (2009) The frequency of polyploid speciation in vascular plants. P Natl Acad Sci 106: 13875–13879. doi: 10.1073/pnas.0811575106

Zhukova PG, Petrovsky VV (1976) Chromosome numbers in some plant species from West Chukotka. II. Bot Zhurn 61: 963–969 (in Russian). [ЖуковаП.Г., Петровский В.В. 1976. Хромосомные числа некоторых растений Западной Чукотки. II // Ботанический журнал. 1976. Т. 61, No 7. С. 963–969].

Zika PF, Farrar DR (2009) *Botrychium ascendens* WH Wagner (Ophioglossaceae) in Newfoundland and notes on its origin. Am Fern J 99: 249–259. doi: 10.1640/0002-8444-99.4.249

Chapter 3

Rapid Allopolyploid Radiation of Moonwort Ferns (*Botrychium*; Ophioglossaceae) revealed by Pacbio Sequencing of Homologous and Homeologous Nuclear Regions

Benjamin Dauphin[1,2,4], Jason R. Grant[1], Donald R. Farrar[3] and Carl J. Rothfels[2]

[1] Laboratoire de Génétique Évolutive, Université de Neuchâtel, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

[2] Department of Integrative Biology, University of California at Berkeley, 3040 Valley Life Sciences Building, Berkeley, CA 94720–3140, USA

[3] Department of Ecology, Evolution, and Organismal Biology, Iowa State University, 353 Bessey Hall, Ames, IA 50011–1020, USA

This chapter is submitted to *Molecular phylogenetics and Evolution*.

Abstract

Polyploidy is a major speciation process in vascular plants, and is postulated to be particularly important in shaping the diversity of extant ferns. However, limitations in the availability of bi-parental markers for ferns have greatly limited phylogenetic investigation of polyploidy in this group. With a large number of allopolyploid species, the genus *Botrychium* is a classic example in ferns where recurrent polyploidy is postulated to have driven frequent speciation events. Here, we use PacBio sequencing and the PURC bioinformatics pipeline to capture all homeologous or allelic copies of four long (~1kb) low-copy nuclear regions from a sample of 45 specimens from the cosmopolitan genus *Botrychium* (representing 37 taxa: 25 diploids and 20 polyploids), and three outgroups. This sample includes most currently recognized *Botrychium* species in Europe and North America, and the majority of our specimens were genotyped with co-dominant nuclear allozymes to ensure species identification. We analyzed the sequence data using maximum likelihood (ML) and Bayesian inference (BI) concatenated-data ("gene tree") approaches to explore the relationships among *Botrychium* species. Finally, we estimated divergence times among *Botrychium* lineages and inferred the multi-labeled polyploid species tree showing the origins of the polyploid taxa, and their relationships to each other and to their diploid progenitors. We found strong support for the monophyly of the major lineages within *Botrychium* and identified most of the parental donors of the polyploids; these results largely corroborate earlier morphological and allozyme-based investigations. Each polyploid had at least two distinct homeologs, indicating that all sampled polyploids are likely allopolyploids (rather than autopolyploids). Our divergence-time analyses revealed that these allopolyploid lineages originated recently—within the last two million years—and thus that the genus has undergone a recent radiation, correlated with multiple independent allopolyploidizations across the phylogeny. Also, we found strong parental biases in the formation of allopolyploids, with individual diploid species participating multiple times as either the maternal or paternal donor (but not both). Finally, we discuss the role of polyploidy in the evolutionary history of *Botrychium* and the interspecific reproductive barriers possibly involved in these parental biases.

Keywords: Ferns; low-copy makers; PacBio; polyploid network; PURC; reticulate evolution.

Introduction

Polyploidy is a major speciation process in plants (Stebbins, 1950; Soltis and Soltis, 2009; Rothfels and Otto, 2016), particularly within ferns, which have rates of polyploid speciation approximately twice those of angiosperms (Otto and Whitton, 2000; Wood et al., 2009). The fern genus *Botrychium* (Ophioglossaceae) is a classic example where recurrent inter-specific hybridization between diploid progenitors has been postulated to have formed an allopolyploid complex (Gilman et al., 2015; Williams et al., 2016; see also, e.g. *Myriopteris*, Grusz et al. 2009; *Dryopteris*, Sessa et al., 2012; *Cystopteris*, Rothfels et al., 2014). *Botrychium* is the largest genus of the Botrychioideae subfamily, containing approximately 48 taxa (24 diploids and 24 polyploids), of which 35 are currently accepted and 13 remain under investigation (PPG I 2016; Dauphin et al., 2017). *Botrychium* species are small perennial plants, colloquially known as moonworts, and have a relatively simple form constituted by a common stalk dichotomously divided into a trophophore and a sporophore (Clausen, 1938). Bisexual gametophytes of *Botrychium* are underground (Wagner et al., 1985), non photosynthetic, and endomycorrhized by fungi species of Glomeromycota (Winther and Friedman, 2007). The genus has a broad circumtemperate to circumboreal distribution, primarily in the northern hemisphere, where species occurs in various specific habitats, for example in perennial herbaceous meadows of mountainous regions or along older roadsides and utility corridors.

Allopolyploidy in *Botrychium* was first hypothesized based on taxa with morphological characters intermediate between their putative parents (Wagner and Wagner, 1981, 1983, 1986). In contrast, several polyploids are superficially indistinguishable from diploid taxa, and were first identified by karyological studies that determined their polyploid status and chromosome number (Wagner and Lord, 1956; Wagner and Wagner, 1993). The hypotheses of polyploidy (and specifically allopolyploidy) were then refined and greatly elaborated by 15 years of investigative work by D.R. Farrar (see Farrar 2011) who characterized the genetic profiles of nearly all *Botrychium* taxa in North America (see also Hauk and Haufler, 1999) based on 22 co-dominant enzyme loci. These data enabled the reconstruction of the bi-parental origin of allopolyploids and the proposal of new hybridization hypotheses (Stensvold et al., 2002; Zika and Farrar, 2009; Gilman et al., 2015). This genetic fingerprinting has since been complemented by other molecular tools, such as amplified fragment length polymorphism (AFLP), to provide new insights on the genetic diversity within specific allopolyploid complexes (Williams and Waller, 2012; Williams et al. 2016).

Previous molecular phylogenies have mostly focused on the relationships of maternal lineages among *Botrychium* taxa (Hauk et al., 2012; Williams and Waller, 2012; Dauphin et al., 2014; Dauphin et al., 2017). The genus comprises an estimated 48 taxa divided among three major clades: Lanceolatum (3 diploids, 14 polyploids), Lunaria (10 diploids, 1 polyloid), and Simplex-Campestre (11 diploids, 9 polyploids; Dauphin et al., 2017). These studies additionally revealed various morphologically cryptic taxa (including both diploids and polyploids) and provided a phylogenetic

foundation that helped elucidate diagnostic morphological characters (Stensvold and Farrar, 2017; Williams et al., 2016; Meza-Torres et al., 2017; Farrar and Gilman, 2017). However, despite the recent advances in our understanding of the maternal phylogenetic relationships among *Botrychium* taxa, phylogenetic information about the paternal lines through nuclear sequencing is lacking.

Low-copy nuclear regions, which provide information on both maternal and paternal evolutionary histories, are crucial for phylogenetic studies of polyploid groups. Unfortunately, there are no universal primer sets targeting a single-copy nuclear gene available for ferns, probably due to limited genetic resources, and the deep divergences within ferns; for example, Ophioglossidae, the subclass to which *Botrychium* belongs (PPG I 2016), diverged from its extant sister group approximately 350 mya (Rothfels et al., 2015). However, with recent advances in sequencing technologies, pteridologists have new opportunities to target variable nuclear regions in their group of interest (Rothfels et al., 2013). In particular, it is now possible to investigate reticulate evolution of polyploid groups by simultaneously sequencing all homeologs present within individual accessions without cloning, and thus to capture and characterize the bi-parental copies of a targeted gene (Rothfels et al., 2017).

In this study, we used the PacBio sequencing approach, in conjunction with the PURC analysis pipeline (Rothfels et al., 2017), to investigate the nuclear phylogenetic relationships of *Botrychium* diploid species and the reticulate histories of allopolyploids, inferring their origins, and characterizing the temporal scale of allopolyploidy events. Moreover, we investigated the combination of diploid in the formation of allopolyploids and their preferences to be either maternal or paternal progenitors. Finally, we discuss our results in a broader context about the role of polyploidy in generating new species and the possible reproductive barriers involved in the selection of parental contributors of allopolyploids.

Materials and methods

Plant material and DNA extraction

We analyzed 45 specimens of *Botrychium* (25 diploid and 20 polyploid) and three outgroup accessions from the related genera *Botrypus* and *Sceptridium*; in total we sampled 37 distinct named *Botrychium* taxa (Table 1). Our sampling includes representatives of each of the three major clades of *Botrychium* (Lanceolatum, Lunaria, and Simplex-Campestre; Hauk et al., 2012), and encompasses the phylogenetic diversity found in each of them, as inferred from plastid loci (Dauphin et al., 2014). Importantly, 35 of the 45 *Botrychium* specimens included were additionally assayed for 22 allozyme loci to ensure species identification and to connect our sequence-based results with the existing wealth of allozyme data. All vouchers were deposited at either the herbarium of the University of Neuchâtel (NEU) or at the Ada Hayden Herbarium (ISC) at Iowa State University.

Genomic DNA was extracted from herbarium specimens or silica-dried material using the DNeasy Plant Mini Kit according to the manufacturer's protocol (Qiagen, Hilden, Düsseldorf, Germany). Quantity and quality of total DNA was inspected with spectrophotometry (NanoDrop 2000, Thermo Scientifc, Wilmington, USA), fluorescence (Qubit Fluorometric Quantification, Thermo Scientifc, Wilmington, USA), and on agarose gels (1%).

Choice of candidate genes, PCR amplification tests and optimization

Primer sets were designed from transcriptome data from *Botrypus* and *Sceptridium*, targeting intron-containing regions approximately 1kb in length for 10 single-copy nuclear markers (*ApPEFP_C*, *CRY2*, *CRY4*, *DET1*, *gapCpSh*, *IBR3*, *pgiC*, *SQD1*, *TPLATE*, and *transducin*; Rothfels et al., 2013) using the program Primer3 (Koressaar and Remm 2007), as implemented in Geneious v.6.1 (Kearse et al., 2012; http://www.geneious.com). PCR amplification was carried out in 25 µl reactions using 12.5 µl of GoTaq G2 colorless master mix (Promega, Madison, Wisconsin, USA), 1 mM of $MgCl_2$, 10 µM of each primer, and 5 ng of DNA. For these amplification tests, material included representatives of each major clade of *Botrychium*, and two outgroups (*Botrypus virginianus* and *Sceptridium dissectum*). Approximately 4–10 primer combinations were tested on gradient PCR for each locus with standard PCR conditions: initial denaturation at 94°C for 180 sec, 35 cycles of 94°C for 60 sec, 48–65°C for 45 sec, and 72°C for 90 sec, with a final extension at 72°C for 600 sec. The highest annealing temperature in the PCR gradient was adjusted for each primer set, taking the lower T_M of the two primers, subtracting 5°C, and decreasing up to T_M –14°C. Amplified products were visualized on 1% agarose gel at 100 V for 90 min to confirm amplification success.

Sanger sequencing and variability tests

To assess phylogenetic informativeness at the species level for each marker, we sequenced a subset of diploid species using an ABI3130 XL Automated Sequencer by Macrogen Europe (Amsterdam, Netherlands), using the same amplification primers. We roughly estimated the variability of each locus although several chromatograms were ambiguous with multiple peaks. For the *SQD1* region, the largest band (~800bp) was extracted using the QIAquick Gel Extraction kit and manufacturer's guidelines (Qiagen, Hilden, Düsseldorf, Germany), and purified before sequencing.

Pool preparation and the PacBio sequencing of homologous and homeologous copies

Four regions were both consistently amplifiable and informative at the species level in *Botrychium*: *ApPEFP_C*, *CRY2cA*, *CRY2cB*, and *transducin* (Table 2). For each of these regions we designed barcoded forward primers using the first 48 primers provided by Pacific Biosciences and followed the data-generation protocol of Rothfels et al. (2017). Briefly, each sample was individually barcoded with a unique 14-bp sequence linked to the 5' end of the forward primer (reverse primers were not barcoded). PCR products were inspected on a 1% agarose gel and band intensity was scored by eye according to six categories: blank, very weak, weak, medium, strong, and very strong. These

band strengths were taken to correspond approximately to concentrations of 5, 10, 15, 20, 30, and 50 ng/μl respectively (Rothfels et al., 2017). Amplicons were pooled in approximately equal quantities based on these concentration estimates (doubled for tetraploid taxa), keeping a constant targeted sequencing coverage of about 36X for the target sequences, for a total of 414 potential targeted sequences. The pooled amplicons were cleaned with an AMPure XP bead purification as recommended by Pacific Biosciences (2016).

We submitted 150 μl of the cleaned pool for library preparation using the P6-C4 chemistry and the SMRTbell™ Template Prep Kit, which includes a repair-ends and ligation reaction, and a purification step (Pacific Biosciences, 2016). Sequencing was conducted on a single PacBio SMRT cell, based on Circular Consensus Sequencing (CCS) technology (Travers et al., 2010) using the PacBio RS II platform (Pacific Biosciences, 2015). Library preparation and sequencing were performed at the Sequencing and Genomic Technologies Core Facility of the Duke University Center for Genomic and Computational Biology (Durham, North Carolina, USA).

Bioinformatics pipeline

We processed the raw CCS reads with the Pipeline for Untangling Reticulate Complexes (PURC; Rothfels et al., 2017), which relies on four other dependencies: BLAST+ v.2.4 (Camacho et al., 2009); CUTADAPT v.1.3 (Martin, 2011); MUSCLE v.3.8.31 (Edgar, 2004); and USEARCH v.8.1 (Edgar, 2010). Thus, we converted the fastq file to fasta and removed all raw CCS reads that were <600 bases long or had more than five expected errors using USEARCH's fastq_filter command (Edgar, 2010). Then, PURC was run on the filtered CCS to de-multiplex the reads, remove primer and barcode sequences, annotate the reads with locus and accession names, remove chimeras, correct sequencing and PCR errors through an iterative clustering approach, and infer final alignments for each locus (Rothfels et al., 2017).

To explore the effectiveness of PURC in inferring the true biological sequences from the PacBio CCS reads, we tested ten different clustering regimes (Supplemental Table 1). The first five regimes (a, b, c, d, and e) followed the default UCHIME settings (Edgar, 2011), required a minimum cluster size of five reads for a cluster to be retained, and differed in their sequence-similarity cutoffs for two sequences to be clustered together (Supplemental Table 1). The second five regimes (f, g, h, i, and j) used the same clustering parameters as for the first five, but with a more stringent chimera detection parameters (abundance_skew, minh, xn, and dn; Edgar, 2011; Supplemental Table 1). For each locus, the alignments from all regimes were merged and a phylogeny inferred from that master alignment using the program AliView v.1.18 (Larsson 2014). Inspection of these phylogenies revealed clear cross-regime consensuses as to what the true biological sequences were; results from those regimes that differed from this consensus were clearly due to, e.g., failures of that regime to capture all the chimeras present or to erroneous lumping together of distinct sequences (see Rothfels et al. 2017). The

final sequences were submitted in GenBank and final alignments of each locus were deposited on Dryad.

Combining homeologs of nuclear loci

All 48 specimens included in this study were previously analyzed in a global plastid phylogeny of *Botrychium* (Dauphin et al., 2017). Because the chloroplast genome is maternally inherited in ferns (Gastony and Yatskievych, 1992; Vogel et al., 1998; Guillon and Raquin, 2000), we deduced for each polyploid taxon the maternal homeolog (labeled "A") as the one that matched with its plastid phylogenetic placement. The second homoelog copy (labeled "B") is therefor considered to be paternal. For the concatenated data analyses the labeled homeologs (A and B) were treated as independent accessions.

Nuclear gene tree inference

We analyzed our four nuclear loci with maximum likelihood (ML) and Bayesian inference (BI) using RAxML v.8 (Stamatakis 2014) and MrBayes v.3.2 (Ronquist et al., 2012), respectively, to infer individual gene trees and the tree from the concatenated data. For each locus we implemented the optimal substitution model (Table 3), based on BIC scores (Posada and Buckley 2004) calculated with the program jModelTest v.2 (Darriba et al., 2012). We ran ML tree searches using a randomized stepwise addition maximum parsimony starting tree and gamma-distributed site rates, and we performed 1,000 bootstrap replicates on the best-scoring ML tree with a bootstrap random number seed. For the Bayesian inference, parameters were unlinked among partitions (loci) and each locus was permitted its own average rate. We implemented the same substitution models as for the ML analyses, except we used a GTR instead of a HKY for the *CRY2cA* locus (Table 3). We kept all other priors at their default values and we ran two independent runs of 10 million MCMC generations, each including three heated and one cold chain, a chain temp of 0.1, sampling trees every 1,000 generations, and a burn-in of 25%. A phylogram summarizing the 7,500 remaining trees was produced, on which PP were reported (Fig. 1).

Divergence time estimates

To estimate divergence times within *Botrychium*, we analyzed the concatenated nuclear data (partitioned by locus, substitution models as above) under an uncorrelated lognormal relaxed clock model using BEAST 2 v.2.4 (Bouckaert et al., 2014). Site models and clock models were unlinked among partitions, and a birth-death tree prior was applied. We set broad priors on the clock rate with an exponential distribution (mean of 10.0) for *ucldMean.c* and a gamma distribution ($\alpha = 0.5396$, $ß = 0.3819$) for *ucldStdev.c*. We constrained the age of three well-supported nodes, using two macrofossils, one attributable to *Botrypus* (Rothwell and Stockey, 1989) and the other to *Sceptridium* (Bozukov et al., 2010), and one secondarily derived age estimate for the divergence of the Lanceolatum and Lunaria clades (Stensvold, 2008). Using the fossils as minimum age of the clades,

we applied a uniform prior distribution on the root (divergence *Botrypus-Sceptridium+Botrychium*) with a lower bound of 57.0 million years (myr), and a uniform prior for the divergence *Sceptridium-Botrychium* with a lower bound of 23.0 myr. We set a normal distribution for the divergence time of the Lanceolatum-Lunaria clades, with a mean of 5 myr and a sigma of 2.0 to keep broad priors, and we constrained the calibrated clades as monophyletic to facilitate the convergence of analysis.

We ran four independent analyses of this model for 100 million generations, with parameters sampled every 1,000 generations. We inspected convergence in Tracer v.1.6 (Drummond and Rambaut, 2007), and we excluded the first 10 million generations from each of the runs as burn-in; the effective sample size (ESS) of each parameter was greater than 300. We then summarized our post burn-in samples with TreeAnnotator v.2.4 (Bouckaert et al., 2014) to generate the maximum clade credibility (MCC) chronogram depicting the divergence time estimates with 95% highest posterior density (HPD) intervals.

Multi-labeled species tree inference

To investigate more deeply the polyploidization events, we inferred a multi-labeled species tree (or "genome tree") using AlloPPNet (Jones et al., 2013). Because AlloPPNet is limited to diploid and tetraploid taxa, for this analysis we removed accessions of the hexaploid *B. pseudopinnatum* (DF18314) and kept the 47 specimens; 25 diploid, 19 tetraploid, and three outgroups for a total of 36 taxa, with one or two specimens for each taxon.

We used the R (R Core Team 2015) script and AlloPPNet R codes (AlloppDT_5beastxml_toplevel.r and AlloppDT_6beastxml_bits.r) available online (Jones 2014) to generate a XML file implementable in BEAST v.1.8.2 (Drummond and Rambaut 2007). To facilitate convergence of the analysis, we manually added a starting tree (obtained from the previous phylogenetic analysis excluding the *B. pseudopinnatum* accession) and constrained the monophyly of the three major clades (Lanceolatum, Lunaria, and Simplex-Campestre). We partitioned nuclear data by locus and kept the default HKY substitution model and strict molecular clock for each gene.

We ran two independent analyses of this model for 500 million generations, sampling parameters every 10,000 generations. We inspected consistency of parameters among the two runs using Tracer v.1.6 (Drummond and Rambaut 2007) to ensure convergence and adequate ESS values. Then, we generated a MCC tree with a burn-in of 10% using TreeAnnotator v.2.4 (Bouckaert et al., 2014). All phylogenetic analyses were run on the Cyberinfrastructure for Phylogenetic Research portal (CIPRES; Miller et al., 2010).

Results

Nuclear loci and PURC regimes comparison

Of the over 100 primer sets tested, only three consistently amplified long nuclear fragments in single copy (*ApPEFP_C* and *transducin*) or had similarly-size copies that were undetectable when viewing on agarose gel (*CRY2cA* and *CRY2cB*; Table 2). From the PacBio run, we obtained 60,151 CCS reads, of which 33,417 were > 600 bp long and had fewer than five expected errors (Supplementary Fig. 1). Cleaning the dataset, PURC detected and removed 180 interlocus concatemers, 10,316 sequences with failed barcodes, 970 sequences with more than two barcodes, and 85 sequences that were unclassifiable (contaminants that did not match any of the reference sequences).

In our PURC regime comparisons, we recovered similar numbers of inferred alleles between stringent (f, g, h, i, and j) and less stringent regimes (a, b, c, d, and e) for *ApPEFP_C* and *transducin*, but different numbers for *CRY2cA* and *CRY2cB*. Manual inspection of alignments showed that the less stringent regimes more consistently recovered the expected number of alleles, especially in cases where the CCS coverage was low (e.g., between 5 and 10X). Therefore, we retained the inferences from regime d as the most appropriate for *ApPEFP_C*, *CRY2cA*, and *transducin*, and regime c for *CRY2cB*, although some exceptions were applied for low CCS coverage sequences, based on manual inspections of the alignments.

Because intra-gametophytic selfing and homozygous populations are common in *Botrychium* (Hauk and Haufler, 1999; Stensvold and Farrar, 2017), we expected most diploids to have a single sequence for each locus, the tetraploids to have two sequences (fixed heterozygosity via allopolyploidy), and the single hexaploid to have three sequences. In general, our results matched expectations for the diploids (Table 1), although the tetraploids frequently had fewer sequences than expected (Table 1). In total, we obtained 23 different alleles for *ApPEFP_C*, 32 for *CRY2cA*, 26 for *CRY2cB*, and 33 for *transducin* (Table 3).

Nuclear phylogeny of the concatenated data

Our nuclear dataset contained four loci with alignment lengths ranging from 915 to 1,764 base pairs (bp), for a total of 4,878 aligned bp for 48 accessions (including outgroups). The *transducin* locus had the most variable sites and *ApPEFP_C* was the most conserved region (Table 3).

Topologies of the ML and BI trees inferred from the concatenated nuclear data were congruent, with 56% of branches strongly supported (BV > 70 and PP > 0.95; Fig. 1). We found strong support for the three major clades (Lanceolatum: 100/1.00; Lunaria: 99/1.00; Simplex-Campestre:100/1.00; Fig. 1).

Table 1 Allele counts and coverage per taxon.

Taxa	Accession	Ploidy	Locus			
			ApPEFP_C	*CRY2cA*	*CRY2cB*	*transducin*
B. alaskense var. *alaskense*	BD1204709	4	2 (48, 34)	2 (50, 42)	1 (70)	2 (40, 19)
B. alaskense var. *salchaketense*	BD1206801	4	2 (21, 17)	2 (26, 24)	1 (40)	2 (49, 45)
B. angustisegmentum	DF19039	2	1 (40)	1 (43)	1 (85)	1 (49)
B. ascendens	DF17467	4	2 (15, 6)	2 (22, 17)	1 (22)	2 (19, 8)
B. boreale	BD1203001	4	2 (24, 24)	2 (30, 25)	1 (47)	2 (14, 9)
B. campestre var. *campestre*	DF954	2	1 (23)	1 (57)	1 (44)	1 (31)
B. campestre var. *lineare*	DF16696	2	1 (54)	1 (10)	1 (5)	1 (17)
B. crenulatum	DF19611	2	1 (75)	1 (24)	NA	1 (73)
B. dusenii	DF17085	4	2 (53, 47)	2 (49, 39)	1 (5)	2 (36, 27)
B. echo	DF17698	4	2 (49, 42)	2 (60, 30)	2 (53, 43)	2 (100, 56)
B. farrarii	DF18455	2	1 (58)	1 (13)	1 (13)	1 (7)
B. furculatum	DF19259	4	1 (56)	1 (106)	1 (157)	2 (48, 36)
B. gallicomontanum	DF14083	4	1 (6)	1 (5)	1 (5)	2 (63, 53)
B. hesperium	DF18178	4	2 (80, 53)	2 (37, 34)	2 (73, 46)	2 (66, 52)
B. lanceolatum "green"	BD1202311	2	1 (30)	1 (25)	1 (52)	1 (25)
B. lanceolatum "green"	DF19171	2	1 (71)	1 (85)	1 (137)	1 (49)
B. lanceolatum "green"	DF19337	2	1 (48)	1 (43)	1 (62)	1 (52)
B. lanceolatum "red"	DF19326	2	1 (30)	1 (31)	1 (44)	1 (10)
B. lanceolatum "red" × *B. pallidum*	DF17272	4	2 (48, 34)	2 (31, 30)	2 (28, 25)	2 (26, 14)
B. lunaria 2	AM11I	2	1 (113)	1 (93)	1 (16)	1 (37)
B. lunaria var. *lunaria*	BD1204116	2	1 (21)	1 (10)	NA	NA
B. lunaria var. *melzeri*	DF19145	2	1 (29)	1 (35)	NA	1 (9)
B. matricariifolium	DF19017	4	2 (30, 29)	2 (8, 7)	2 (26, 7)	NA
B. matricariifolium	DF19067	4	2 (48, 25)	2 (28, 28)	2 (45, 41)	2 (20, 13)
B. michiganense	DF18145	4	2 (45, 33)	2 (42, 32)	2 (33, 33)	2 (27, 10)
B. minganense	BD1205713	4	2 (53, 41)	2 (33, 23)	1 (57)	2 (51, 42)
B. minganense	DF17538	4	2 (65, 50)	2 (38, 27)	1 (38)	2 (21, 11)
B. montanum	DF18528	2	1 (22)	1 (27)	1 (24)	1 (28)
B. mormo	DF890	2	1 (30)	1 (35)	1 (55)	1 (32)
B. neolunaria	DF18299	2	1 (23)	1 (16)	NA	1 (5)
B. neolunaria	DF19201	2	1 (58)	1 (85)	1 (14)	1 (86)
B. nordicum	DF17819	2	1 (30)	1 (25)	NA	1 (13)
B. pallidum	DF17260	2	1 (23)	1 (11)	1 (24)	1 (19)
B. pallidum	DF18729	2	1 (17)	1 (25)	1 (31)	1 (13)
B. paradoxum	DF16672	4	2 (39, 12)	1 (56)	1 (74)	2 (29, 20)
B. pedunculosum	DF5413	4	2 (80, 56)	2 (41, 34)	2 (37, 24)	2 (33, 15)
B. pinnatum	BD1206619	4	2 (78, 22)	2 (43, 23)	1 (56)	2 (58, 50)
B. pseudopinnatum	DF18314	6	3 (21, 18, 15)	3 (33, 32, 20)	2 (46, 41)	3 (57, 49, 41)
B. pumicola	DF17915	2	1 (110)	1 (22)	1 (36)	1 (60)
B. simplex 2	DF17299	2	1 (27)	1 (16)	1 (17)	1 (8)
B. simplex var. *compositum*	DF18799	2	1 (65)	1 (18)	1 (11)	1 (40)
B. spathulatum	DF18803	4	2 (43, 35)	1 (81)	1 (8)	2 (58, 32)
B. tenebrosum	AM1502	2	1 (66)	1 (73)	1 (44)	1 (57)
B. tenebrosum	DF17416	2	1 (60)	1 (60)	1 (44)	1 (30)
B. tunux	AM23A	2	1 (22)	1 (55)	1 (13)	1 (97)
Botrypus virginianus	BD1206330	4	1 (52)	1 (49)	2 (101, 48)	1 (9)
Sceptridium multifidum	AM1503	2	1 (5)	1 (5)	NA	1 (80)
Sceptridium multifidum	BD1205311	2	1 (52)	1 (40)	1 (61)	1 (49)

NA, no data available since PCR amplfication failed multiple times.

Table 2 Primer sets and PCR conditions.

Protein region	Primer names (Forward, Reverse)	Sequence (5' - 3') (Forward, Reverse)	PCR Program[1]	Expected length (bp)
ApPEFP_C	76AP_F2, 559AP_R2	TCATTGGGTGGGTGCAGGACTG, CTGCTGGAAGTGCAGTTATTCT	659035	900
CRY2	1749C2_F5, 4483C2_R4	TGGGAAGTCAATGATGATCAAGG, GCCKTACATACTCTCCATGAGG	589035	1150
transducin	1904TR_F4, 1826TR_R6	GAAAGAGGTAGAAGCTCATTCTGG, TGAGCAAACAARGACCGGTGGC	609035	1300

[1]First two digits is the annealing temperature (°C), the following two is elongation time (in seconds), and the two last is the number of cycles.

Divergence time estimates

All parameters from the divergence-time analyses had post-burnin pooled ESSs > 3200. The mean estimate for the *Botrypus* - *Sceptridium* divergence time was 60.20 myr (HPD, 57.00–71.3 myr), that for the *Sceptridium* - *Botrychium* divergence was 36.19 myr (HPD, 23–53.34 myr), and that between the Lanceolatum and Lunaria clades was 8.51 myr (HPD, 5.96–11.08 myr; Fig. 2).

Multi-labeled species tree

In our allopolyploid network inference, the majority of parameters were estimated with ESS values > 300 for both combined runs, with some exceptions for the *apspnetwork* statistics. We infered a mean of 11 allopolyploidy events (HPD, 10–12). In total, 45,000 trees were retained and used to generate the MCC tree (highest log clade credibility = –94.707745; Fig. 3).

Table 3 Datasest characteristics.

Locus	Aligned length in bp.	No. Specimens	Missing data[1]	No. Alleles	No. Parsimony information sites	No. Variable sites	No. SNPs	Best model	lnL	BIC scores	Implemented model in RaxML	Implemented model in MrBayes	Implemented model in BEAST2
ApPEFP_C	905 (915)	45 (48)	0%	21 (23)	30 (66)	41 (116)	11 (50)	HKY+G	−2048.2941	5044.4186	HKY+G	HKY+G	HKY+G
CRY2cA	1037 (1115)	45 (48)	0%	30 (32)	38 (93)	51 (138)	13 (45)	TIM3+G	−2481.5967	5270.7011	HKY+G	GTR+G	TN93+G
CRY2cB	1066 (1084)	40 (42)	13%	23 (26)	31 (65)	48 (132)	17 (65)	HKY+I	−2498.3041	5758.3452	HKY+G	HKY+I	HKY+I
transducin	1221 (1764)	43 (46)	4%	30 (33)	77 (117)	96 (259)	19 (142)	HKY+G	−4026.9818	9063.1344	HKY+G	HKY+G	HKY+G
All loci	4229 (4878)	45 (48)	4%	104 (114)	176 (341)	236 (645)	60 (302)				Unliked	Unliked	Unliked
Overall in %					4.16 (6.99)	5.58 (13.22)	1.41 (6.19)						

Note: outgroups are considered in parentheses; [1]For each locus, a specimen was counted as missing data when no allele or homeolog was successfully sequenced.

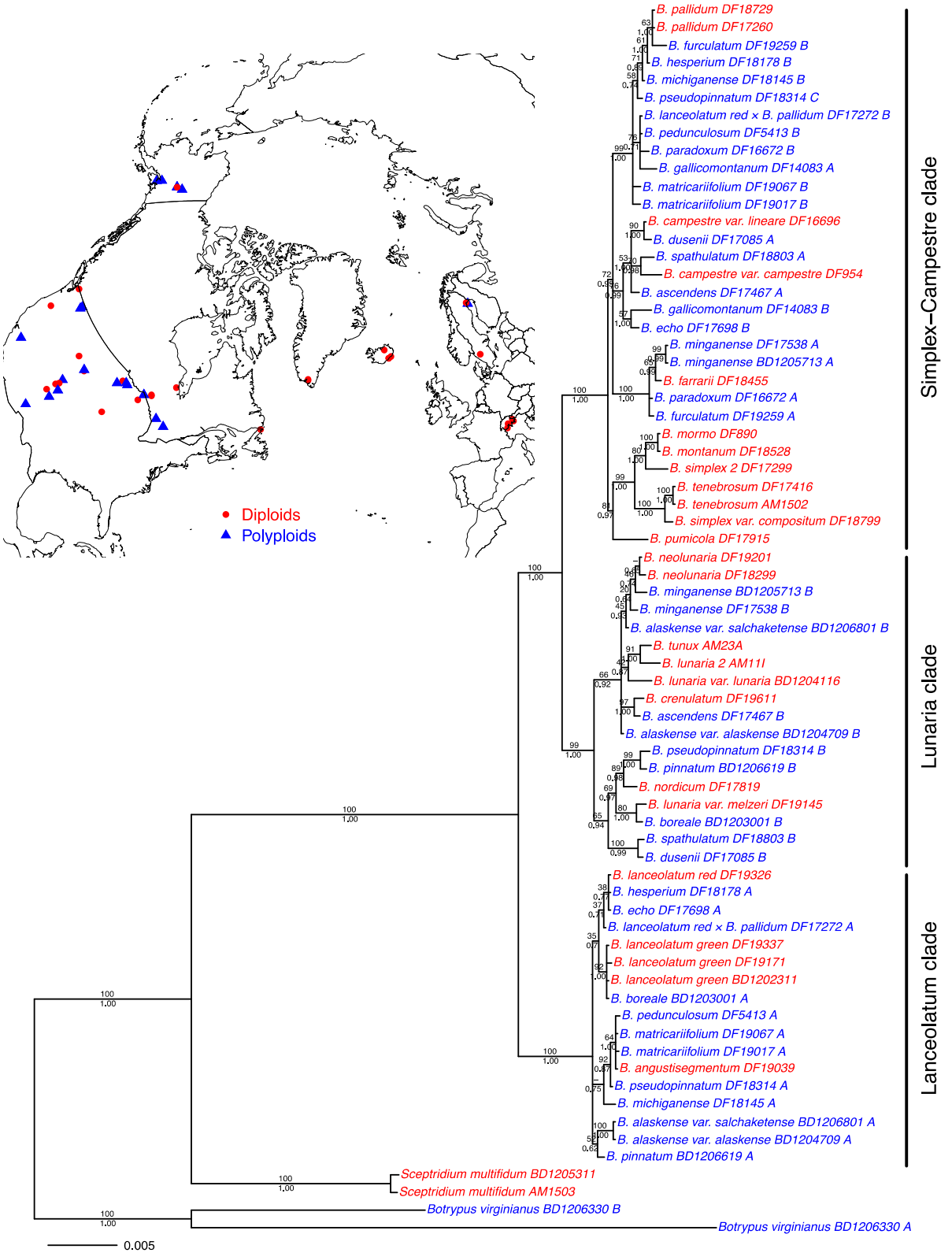


Fig. 1 Bayesian phylogram of the multi-labeled gene tree from the concatenated nuclear data. Maximum likelihood bootstrap support values are depicted above branches, and posterior probabilities are below branches. Diploid species are shown in red and the polyploids in blue. For the polyploids, each subgenome (that from the maternal progenitor and that from the paternal progenitor) appears on the tree individually: “A” for the maternal subgenome, “B” for the paternal subgenome, and “C” for the additional subgenome of the single hexaploid. The map shows the collection sites of the sampled accessions, with red circles for diploids and blue triangles for polyploids.

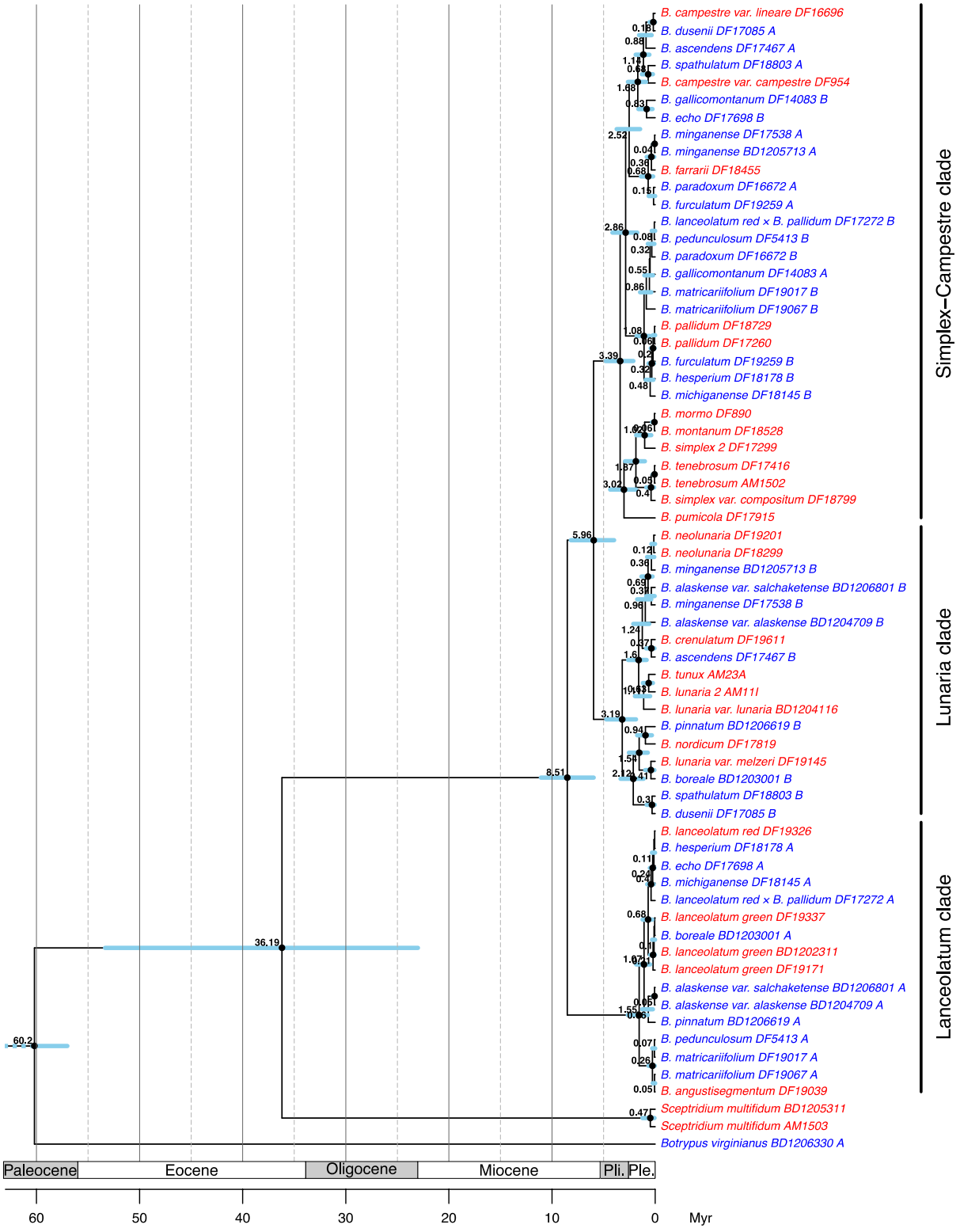


Fig. 2 Time-calibrated *Botrychium* phylogeny from the concatenated nuclear data. Black dots indicate a posterior probability > 0.95 and node ages are specified on the left side of nodes. The color scheme and subgenome designations are as in Figure 1 and the geological time scale follows the International Commission on Stratigraphy (2012; http://www.stratigraphy.org/index.php/ics-chart-timescale).

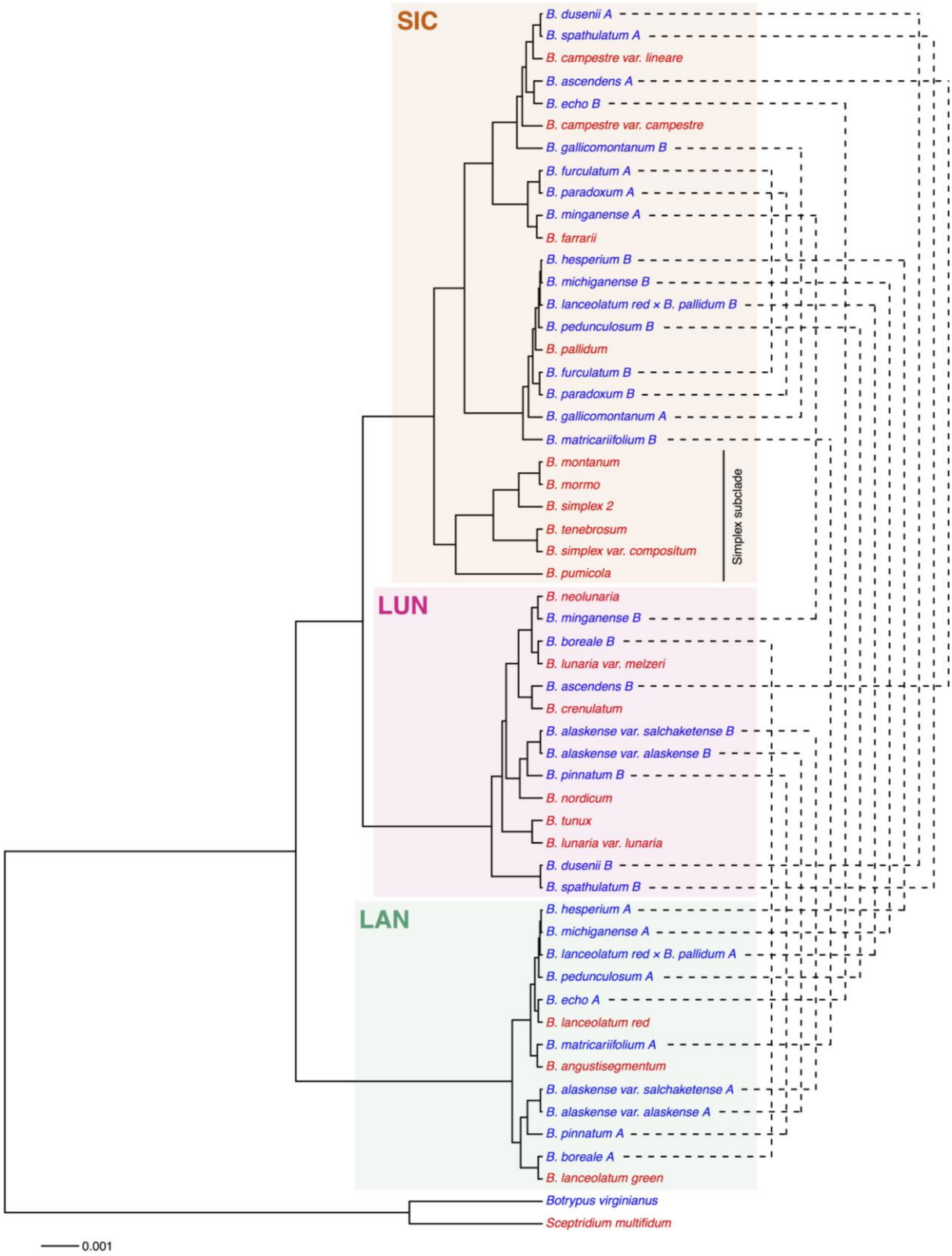


Fig. 3 The multi-labeled maximum clade credibility species tree inferred by AlloPPNet from our nuclear data. Dashed lines on the right-hand side connect subgenomes present in individual polyploid taxa. The color scheme and subgenome designations are as in Figure 1.

Discussion

Generating low-copy nuclear data for polyploid complexes

PacBio sequencing combined with the bioinformatics pipeline PURC (Pipeline for Untangling Reticulate Complexes; Rothfels et al. 2017) open news perspectives to study groups with polyploid taxa and limited genomic resources. As a first result, we provided primer sets targeting four long nuclear loci for *Botrychium*, *Botrypus*, and *Sceptridium*, which now enables investigations of the evolutionary history of their nuclear genomes. Our success for generating low-copy nuclear data was primarily due to the availability of transcriptomes of species closely related to the group of interest (Rothfels et al., 2013), which made possible primer design. Based on a single PacBio run, we sequenced 45 specimens of *Botrychium* (25 diploids, 19 allotetraploids, and one hexaploid) and three outgroups, in a much faster and more economical way than is possible using conventional cloning and Sanger sequencing. The use of PURC greatly improved the detection of PCR-mediated chimeras and simultaneously retrieved with success the homologs of diploids and homeologs of allopolyploids.

Corroborating plastid-based inferences of relationship among diploid species

A first relevant outcome is the concordance of our nuclear phylogeny with earlier inferences based on plastid data for the relationships among the three major clades of *Botrychium*, with the first divergence separating the Lanceolatum (LAN) clade from the two other major lineages (the Lunaria [LUN] and Simplex-Campestre clade [SIC]; Hauk et al., 2012; Williams and Waller, 2012; Dauphin et al., 2014; Dauphin et al., 2017). The deep phylogenetic divergences among the genera *Botrychium*, *Botrypus*, and *Sceptridium*, are also consistent with those inferred in the plastid phylogeny of Ophioglossaceae (Hauk et al., 2003).

However, within the three clades themselves, there are several cases of plastid-nuclear incongruence among the diploids. In the LAN clade, the two diploids *B. angustisegmentum* and *B. lanceolatum* "green" are supported in two distinct clades, while *B. lanceolatum* "red" is embedded in a poorly supported group (Fig. 1). In contrast, in our species tree (Fig. 3), *B. angustisegmentum* and *B. lanceolatum* "red" form a monophyletic group, which is inconsistent with allozyme-based genetic distance that supports *B. lanceolatum* "red" as the early divergent diploid lineage of the LAN clade (Stensvold and Farrar, 2017). Because of weak differentiation among the LAN diploids, presumably due to a recent speciation events, these phylogenetic relationships remain unclear (Figs. 1–3).

Within the LUN clade, all diploid species are differentiated in our nuclear phylogeny, including recently published cryptic taxa. Compared to earlier plastid phylogeneties, we found a greater than expected divergence between the varieties of *B. lunaria* (var. *lunaria* and *melzeri*), and between *B. lunaria* and *B. nordicum* (Fig. 3), which is congruent with their genetic distance based on the allozyme allele frequencies and supports the species status of *B. nordicum* (Stensvold and Farrar, 2017).

Plastid and nuclear genomes generate a concordant phylogenetic signal for the SIC clade, with an exception for the placement of *B. pallidum*. In the plastid phylogenetic reconstruction, *B. pallidum* is the sister group to the rest of the Simplex clade (ML = 95; PP = 1.00; see Fig. 4 in Dauphin et al., 2017), whereas that species is embedded within the Campestre clade in our nuclear phylogeny. Both the concatenated-data tree and the species tree support the clade formed by the four diploids *B. campestre* var. *campestre*, *B. campestre* var. *lineare*, *B. farrarii*, and *B. pallidum* (Figs. 1–3). This result is in agreement with morphological data, where *B. pallidum* displays a tendency to form two lobes on the basal pinna, typical of the Campestre clade (Wagner and Wagner, 1990).

Simplex clade free of polypoids

Within the SIC clade, the Simplex clade comprises six taxa (*B. montanum*, *B. mormo*, *B. pumicola*, *B. simplex* var. *compositum*, *B. simplex 2*, and *B. tenebrosum*), none of which form any allopolyploids (Figs. 1–3). Yet, the Simplex clade is widely distributed in North America, and also occurs in several regions of Europe and its species are found with other diploid taxa in natural populations, which should allow inter-specific hybridization (Abbott, 2017). At least two pre-zygotic barriers may explain in part these limitations to inter-specific hybridization. First, the bisexual gametophytes of *Botrychium* taxa may have different phenological cycles so that non-synchrony of gametangial maturity for members of the Simplex clade could possibly be involved, and needs further investigation. The second scenario is that cytogenetic incompatibilities prevent allopolyploidy between diploids of the Simplex clade and the other members of *Botrychium*. Supporting this idea is the significantly smaller genome size of *B. simplex* in comparison with all other species of the genus (except *B. pallidum*; Williams and Waller, 2012), suggesting that differences in their chromosomal structure may prevent pairing in meiosis. Interestingly, several sterile hybrids (*B. lunaria* × *B. simplex*) were reported between diploid taxa of SIC and LUN, but these cytogenetic incompatibilities appear permeable at the intra-clade level and the diploid taxa of the Simplex clade do show the capability to produce viable homoploid hybrids (between *B. simplex* var. *compositum* and *B. simplex 2*) within natural populations (Wagner, 1980, 1991; Wagner and Beitel, 1985; Wagner and Wagner, 1988; D.R. Farrar, personal communication). An analogous case of homoploid hybrids is found in the LUN clade where there are several introgressed genotypes between *B. lunaria* and *B. neolunaria* (Stensvold, 2008), but surprinsingly, these two closely related species have also formed an allopolyploid, *B. yaaxudakeit* (Stensvold et al., 2002; Stensvold and Farrar, 2017). Further works are critical to elucidate the mechanisms underlying their reproductive barriers at conspecific level.

Confirming parentage of allopolyploids

Reticulation is extensive in *Botrychium* and involves diploid progenitors among all three major clades. On the basis of our phylogenetic results, we identified most of the parental donors of polyploids, which largely corroborate earlier morphology and allozyme-based investigations (Figs. 3–

4). Over half (12 of 20) of the diploid taxa contributed to at least one polyploid, and most show a clear bias in favor of being either a maternal or paternal progenitor. Most of diploid donors contributed to two or three allotetraploids but some are much more prolific, such as *B. pallidum*, which is the progenitor of eighteen allotetraploids. *Botrychium pallidum's* prolific polyploid formation may in part be explained by its widespread ancestral occurrence in North America where it likely occurred sympatrically with both eastern and western diploid species of *Botrychium*. Among the allotetraploids, our results support the parentage hypotheses of *B. ascendens* (Zika and Farrar, 2009), *B. matricariifolium* (Williams et al., 2016), *B. michiganense* (Gilman et al., 2015), and the majority of the remaining unpublished taxa, although some discrepancies (symbolized as dashed lines in Fig. 4) between our results and allozyme-based genetic profiles (Farrar, 2011) exist, as described below.

Several allopolyploidy hypotheses presented remain tentative (Fig. 4), pending inclusion of putative new diploid taxa that may prove to be more probable diploid parents than those currently indicated. Specifically, unrepresented in our study is a diploid North American "*B. lunaria* var. *lunaria"* that appears to be distinct from European *B. lunaria* (Dauphin et al., 2017; Maccagni et al., 2017; Stensvold and Farrar, 2017). Consequently, and in consideration of the currently restricted occurrence of *B. nordicum* to Iceland and Norway, the *B. nordicum* parentage of the North American *B. pinnatum*, *B. pseudopinnatum, B. spathulatum,* and the South American *B. dusenii,* are questionable. Additionally, a putative new diploid form of *B. paradoxum* that may prove parental to allotetraploid *B. paradoxum* has recently been detected (D.R. Farrar, personal communication). Parentages in question due to non-inclusion of these candidate diploid taxa are indicated (dashed lines in Fig. 4) and further study with key diploid additions can be expected to further clarify these ambiguous parentages of *Botrychium* allopolyploids.

Predominance of allopolyploidy

In general, an allopolyploid taxon is easier to identify in the field than an autopolyploid which might be expected to have no or few morphological characters distinct from its parent. Thus, while we cannot refute the presence of undetected autopolyploidy, our nuclear data demonstrate that all our sampled polyploids are allopolyploids, usually with parents from different major clades (Lanceolatum, Lunaria, and Simplex-Campestre; Fig. 3). *Botrychium* species are known to maintain gametophytic selfing as the predominant mode of reproduction (Hauk and Haufler, 1999; Stensvold and Farrar, 2017), which produces strict homozygotes in a single generation, and drastically decreases the allele diversity at the individual level (Haufler et al., 2016; Sessa et al., 2016). With a fixed heterozygosity, the allopolyploids may have an evolutionary advantage relative to their diploid progenitors, including greater resilience under ecological changes (Hegarty and Hiscock, 2009; Abbott et al., 2013).

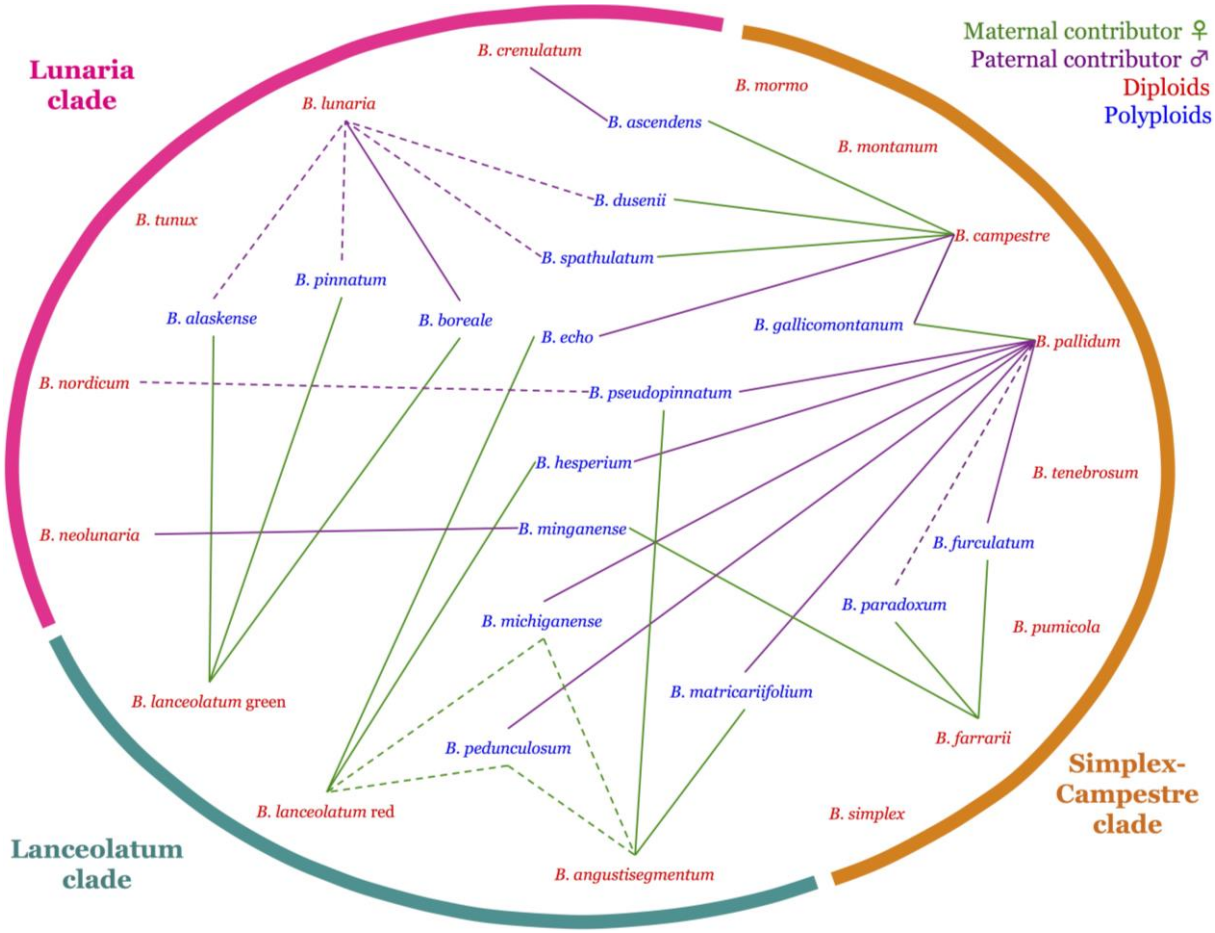


Fig. 4 Reticulation diagram, showing the origins of the sampled polyploids. Maternal contributors of polyploids are shown as green lines and their paternal contributors as violet lines. Uncertainties of parental progenitors—cases where our results either conflict with earlier hypotheses (Farrar, 2011; Gilman et al., 2015; Meza-Torres et al., 2017) or could be erroneous due to unsampled diploids (Stensvold and Farrar, 2017)—are represented with dashed lines.

Recent rapid radiation in an ancient lineage

Probably the major breakthrough of our study is the discovery of the rapid and recent radiation of allopolyploid taxa in *Botrychium*. Our divergence time estimates provide the first chronology of hybridization events within *Botrychium*, in which all presently known allopolyploids have arisen in the last two million years (Fig. 2). This timescale could be even shorter, given the probability, with greater sampling, of detecting diploids that may be more closely related to allopolyploids than those inferred in our phylogenetic analysis (e.g., the paternal donor of *B. dusenii* and *B. spathulatum* in Fig. 3). In the more conservative case, with the oldest ages considered (Fig. 2), *Botrychium* allopolyploids are young in the context of Ophioglossaceae, which have been diverging for ~160 my from the other fern lineages (Pryer et al., 2004; Rothfels et al., 2015). Comprised of 50% polyploids, speciation via allopolyploidy plays an important role in the genus, as similarly reported in many other fern lineages (*Asplenium*, Perrie and Brownsey, 2005; *Pteris*, Chao et al., 2012; *Dryopteris*, Sessa et al., 2012; *Cystopteris*, Rothfels et al., 2014).

Parental biases in the formation of allopolyploids

The allopolyploids in our dataset are not randomly formed from the pool of diploid donors—instead some diploids are over-represented in the formation of polyploids. Among the diploid progenitors of polyploids, there is an additional tendency for individual species to be either the maternal or paternal progenitor, but not both. A striking example is the three diploids of the LAN clade, which have contributed to fourteen allopolyploid taxa, but always as the maternal donor (Dauphin et al., 2017). In contrast, the prolific *B. pallidum* has generated nine allopolyploids as a paternal donor and only one as the maternal contributor. Similar patterns of non-random parental contributions have been inferred in *Asplenium* (Perrie et al., 2010), and *Dryopteris* (Sessa et al., 2012), while more reciprocal patterns (individual diploids functioning as both the maternal and paternal parents of polyploids) were reported for the polyploid *Astrolepis integerrima* (Beck et al., 2012) and *Polypodium hesperium* (Sigel et al., 2014).

Conclusion

PacBio sequencing technology in conjunction with the use of the PURC bioinformatic pipeline provides new insight into phylogenetics of *Botrychium* and the major role played by allopolyploidy in the diversification of this genus. Our study presents a timescale of inter-specific hybridization events and recent and rapid radiation of allopolyploids in the last two million years. Our findings also confirm a predominance of allopolyploidy, with no evidence of autopolyploidy in our accessions. Distinctive features of allopolyploid formation in *Botrychium* include strong biases for individual diploid species to be either the maternal or paternal contributors to polyploids, a tendency for parental diploids to be from different major clades, and unequal participation among species and clades, with particular species in certain clades parenting most allopolyploids.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Swiss National Science Foundation (grant: 31003A_156456) and its associated mobility grant, and by the University of California Berkeley. The authors thank Fay-Wei Li, Alessio Maccaggni, Kathleen Pryer, and Eric Schuettpelz for sharing herbarium specimens. We also thank Ingrid Jordon-Thaden and Lydia Smith for their assistance with laboratory work, and the staff at Duke University's Sequencing and Genomic Technologies Core Facility.

References

Abbott, R.J., 2017. Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. J. Syst. Evol. 55, 238–258.

Abbott, R.J., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., et al., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246.

Beck, J.B., Allison, J.R., Pryer, K.M., Windham, M.D., 2012. Identifying multiple origins of polyploid taxa: A multilocus study of the hybrid cloak fern (*Astrolepis integerrima*; Pteridaceae). Am. J. Bot. 99, 1857–1865.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.

Bozukov, V., Tsenov, B., Vatsev, M., 2010. A first find of *Botrychium* (Ophioglossaceae) in Bulgarian palaeomacroflora. C. R. Acad. Bulg. Sci. 63, 889–892.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 1–9.

Chao, Y.S., Dong, S.Y., Chiang, Y.C., Liu, H.Y., Chiou, W.L., 2012. Extreme multiple reticulate origins of the *Pteris cadieri* complex (Pteridaceae). Int. J. Mol. Sci. 13, 4523–4544.

Clausen, R.T., 1938. A monograph of the Ophioglossaceae. Mem. Torrey Bot. Club 19, 1–177.

Darriba, D., Taboada, G.L., Doallo, Posada, R.D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772–772.

Dauphin, B., Vieu, J., Grant, J.R., 2014. Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* s.s., Ophioglossaceae). Am. J. Bot. 101, 128–140.

Dauphin, B., Farrar, D.R., Maccagni, A., Grant, J.R., 2017 A worldwide molecular phylogeny provides new insight on cryptic diversity within the moonworts (*Botrychium* s.s., Ophioglossaceae). Syst. Bot. 42, XX–XX.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.

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

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 26, 2460–2461.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 27, 2194–2200.

Farrar, D.R., 2011 (last updated 15 Dec. 2011). Systematics and taxonomy of the genus *Botrychium*. http://www.public.iastate.edu/~herbarium/botrychium/Moonwort-Systematics.pdf. (accessed 01 May 2017).

Farrar, D.R., Gilman, A.V., 2017. Relationships in the *Botrychium campestre* (Ophioglossaceae) complex. Brittonia 69, 265–275.

Gastony, G.J., Yatskievych, G., 1992. Maternal inheritance of the chloroplast and mitochondrial genomes in cheilanthoid ferns. Am. J. Bot. 79, 716–722.

Gilman, A.V., Farrar, D.R., Zika, P.F., 2015. *Botrychium michiganense* sp. nov. (Ophioglossaceae), a new North American moonwort. J. Bot. Res. Inst. Texas 9, 295–309.

Grusz, A.L., Windham, M.D., Pryer, K.M., 2009. Deciphering the origins of apomictic polyploids in the *Cheilanthes yavapensis* complex (Pteridaceae). Am. J. Bot 96, 1636–1645.

Guillon, J.M., Raquin, C., 2000. Maternal inheritance of chloroplasts in the horsetail *Equisetum variegatum* (Schleich.). Curr. Genet. 37, 53–56.

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

Hauk, W.D., Haufler, C.H., 1999. Isozyme variability among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). Am. J. Bot. 86, 614–633.

Hauk, W.D., Kennedy, L., Hawke, H.M., 2012. A phylogenetic investigation of *Botrychium* s.s. (Ophioglossaceae): Evidence from three plastid DNA sequence datasets. Syst. Bot. 37, 320–330.

Hegarty, M., Hiscock, S., 2007. Polyploidy: doubling up for evolutionary success. Curr. Biol. 17, R927–R929.

Jones, G., 2014. Manual for using AlloppNET for diploids and allotetraploids. MANUAL, 1–10. URL: https://sites.google.com/site/touchingthedata/software/allopolyploids (accessed 16.03.01).

Jones, G., Sagitov, S., Oxelman, B., 2013. Statistical inference of allopolyploid species networks in the presence of incomplete lineage sorting. Syst. Biol. 62, 467–478.

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A.J., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.

Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30, 3276–3278.

Maccagni, A., Parisod, C., Grant, J.R., 2017. Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. *Alpine Botany* 127, 185–196.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 1–10.

Meza-Torres, E.I., Stensvold, M.C., Farrar, D.R., Ferrucci, M.S., 2017. Circumscription of the South American moonwort *Botrychium* (Ophioglossaceae). Plant Biosystems 151, 1–11.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE). IEEE, 1–8.

Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Annu. Rev. Genet. 34, 401–437.

Pacific Biosciences 2015. Revolutionize genomics with SMRT Sequencing. URL: http://pacificbiosciences.com/brochure 1–9 (accessed 16.03.01).

Pacific Biosciences 2016. Procedure checklist – amplicon template preparation and sequencing. URL http://www.pacb.com/wp-content/uploads/Procedure-Checklist-Amplicon-Template-Preparation-and-Sequencing.pdf (accessed 16.03.01).

Perrie, L.R., Brownsey, P.J., 2005. Insights into the biogeography and polyploid evolution of New Zealand *Asplenium* from chloroplast DNA sequence data. Am. Fern J. 95, 1–21.

Perrie, L.R., Shepherd, L.D., De Lange, P.J., Brownsey, P.J., 2010. Parallel polyploid speciation: distinct sympatric gene- pools of recurrently derived allo- octoploid *Asplenium* ferns. Mol. Ecol. 19, 2916–2932.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

PPG I, 2016. A community-derived classification for extant lycopods and ferns. J. Syst. Evol. 54, 563–603.

R Development Core Team. 2015 (last updated 15.01.10). R: a language and environment for statistical computing. Version 3.2. R Foundation for Statistical Computing. Vienna. https://www.rproject.org (accessed 16.01.01).

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Rothfels, C.J., Johnson, A.K., Windham, M.D., Pryer, K.M., 2014. Low-copy nuclear data confirm rampant allopolyploidy in the Cystopteridaceae (Polypodiales). Taxon 63, 1026–1036.

Rothfels, C.J., Larsson, A., Li, F.W., Sigel, E.M., Huiet, L., Burge, D.O., Ruhsam, M., Graham, S.W., Stevenson, D.W., Wong, G.K., Korall, P., 2013. Transcriptome-mining for single-copy nuclear markers in ferns. PLoS One 8, e76957.

Rothfels, C.J., Li, F.W., Sigel, E.M., Huiet, L., Larsson, A., Burge, D.O., Ruhsam, M., Deyholos, M., Soltis, D.E., Stewart, C.N., Shaw, S.W., 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. Am. J. Bot. 102, 1089–1107.

Rothfels, C.J., Otto, S.P., 2016. Polyploid speciation. In: Kliman R.M., editor. Encyclopedia of Evolutionary Biology. Oxford: Academic Press. p. 317–326.

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

Rothwell, G.W., Stockey, R.A., 1989. Fossil Ophioglossales in the Paleocene of Western North America. Am. J. Bot. 76, 637–644.

Sessa, E.B., Testo, W.L., Watkins, J.E., 2016. On the widespread capacity for, and functional significance of, extreme inbreeding in ferns. New Phytol. 211, 1108–1119.

Sessa, E.B., Zimmer, E.A., Givnish, T.J., 2012. Reticulate evolution on a global scale: A nuclear phylogeny for New World *Dryopteris* (Dryopteridaceae). Mol. Phylogenet. Evol. 64, 563–581.

Sigel, E.M., Windham, M.D., Pryer, K.M., 2014. Evidence for reciprocal origins in *Polypodium hesperium* (Polypodiaceae): A fern model system for investigating how multiple origins shape allopolyploid genomes. Am. J. Bot 101, 1476–1485.

Soltis, P.S., Soltis, D.E., 2009. The role of hybridization in plant speciation. Annu. Rev. Plant Biol. 60, 561–588.

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

Stebbins, G.L., 1950. Variation Evolution in Plants. New York: Columbia Univ. Press

Stensvold, M.C., 2008. A taxonomic phylogeographic study of the *Botrychium lunaria* complex. Dissertation, Iowa State University, Ames, Iowa, U.S.A.

Stensvold, M.C., Farrar, D.R., 2017. Genetic diversity in the worldwide *Botrychium lunaria* (Ophioglossaceae) complex, with new species and new combinations. Brittonia 69, 148–175.

Stensvold, M.C., Farrar, D.R., Johnson-Groh, C., 2002. Two new species of moonworts (*Botrychium* subg. *Botrychium*) from Alaska. Am. Fern J. 92, 150–160.

Travers, K.J., Chin, C.S., Rank, D.R., Eid, J.S., Turner, S.W., 2010. A flexible and efficient template format for circular consensus sequencing and SNP detection. Nucleic Acids Res. 38, 1–8.

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

Wagner, F.S., 1993. Chromosomes of North American grapeferns and moonworts (Ophioglossaceae: *Botrychium*). Contrib. Univ. Mich. Herb. 19, 83–92.

Wagner, W.H., 1980. A probable new hybrid grapefern *Botryhium matricariifolium* x *simplex*, from Central Michigan. The Michigan Botanist 19: 31–36.

Wagner, W.H., 1991. New examples of the moonwort hybrid, *Botrychium matricariifolium* x *simplex* (Ophioglossaceae). The Canadian Field Naturalist 105: 91–94.

Wagner, W.H., Beitel, J.M., 1985. Evidence for interspecific hybridization in pteridophytes with subterranean mycoparasitic gametophytes. P. Roy. Soc. Edinb. B. 86: 273–281.

Wagner, W.H., Lord, L.P., 1956. The morphological and cytological distinctness of *Botrychium minganense* and *B. lunaria* in Michigan. B. Torrey Bot. Club 83, 261–280.

Wagner, W.H., Wagner, F.S., 1981. New species of moonworts, *Botrychium* subg. *Botrychium* (Ophioglossaceae), from North America. Am. Fern J. 71, 20–30.

Wagner, W.H., Wagner, F.S., 1983. Two moonworts of the Rocky Mountains; *Botrychium hesperium* and a new species formerly confused with it. Am. Fern J. 73, 53–62.

Wagner, W.H., Wagner, F.S., 1986. Three new species of moonworts (*Botrychium* subg. *Botrychium*) endemic in western North America. Am. Fern J. 76, 33–47.

Wagner, W.H., Wagner, F.S., 1988. Detecting *Botrychium* hybrids in the Lake Superior region. The Michigan Botanist 27, 75–80.

Wagner, W.H., Wagner, F.S., 1990. Notes on the fan-leaflet group of moonworts in North America with descriptions of two new members. Am. Fern J. 80, 73–81.

Wagner, W.H., Wagner, F.S., Beital, J.M., 1985. Evidence for interspecific hybridization in pteridophytes with subterranean mycoparasitic gametophytes. Proc. Roy. Soc. Edin. 86B, 273–281.

Williams, E.W., Waller, D.M., 2012. Phylogenetic placement of species within the genus *Botrychium* s.s. (Ophioglossaceae) on the basis of plastid sequences, amplified fragment length polymorphisms, and flow cytometry. Int. J. Plant Sci. 173, 516–531.

Williams, E.W., Farrar, D.R., Henson, D., 2016. Cryptic speciation in allotetraploids: Lessons from the *Botrychium matricariifolium* complex. Am. J. Bot. 103, 740–753.

Winther, J.L., Friedman, W.E., 2007. Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). Am. J. Bot. 94, 1248–1255.

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., Rieseberg, L.H., 2009. The frequency of polyploid speciation in vascular plants. P. Natl. Acad. Sci. USA 106, 13875–13879.

Zika, P.F., Farrar, D.R., 2009. *Botrychium ascendens* WH Wagner (Ophioglossaceae) in Newfoundland and notes on its origin. Am. Fern J. 99, 249–259.

Supplementary Materials

Table S1 PURC clustering regimes run in our study

Regime	Cluster thresholds				UCHIME settings			
a	0.995	0.995	0.995	0.995	2.0	0.28	8.0	1.4
b	0.995	0.995	0.995	0.990	2.0	0.28	8.0	1.4
c	0.995	0.995	0.990	0.990	2.0	0.28	8.0	1.4
d	0.997	0.995	0.995	0.995	2.0	0.28	8.0	1.4
e	0.997	0.997	0.997	0.997	2.0	0.28	8.0	1.4
f	0.995	0.995	0.995	0.995	1.1	0.20	3.0	0.5
g	0.995	0.995	0.995	0.990	1.1	0.20	3.0	0.5
h	0.995	0.995	0.990	0.990	1.1	0.20	3.0	0.5
i	0.997	0.995	0.995	0.995	1.1	0.20	3.0	0.5
j	0.997	0.997	0.997	0.997	1.1	0.20	3.0	0.5

PURC (the four cluster thresholds) settings are explained in Rothfels et al. (2016) and UCHIME (abskew, minh, xn, and dn) settings in the USEARCH manual available online at http://www.drive5.com/usearch/manual/UCHIME_score.html

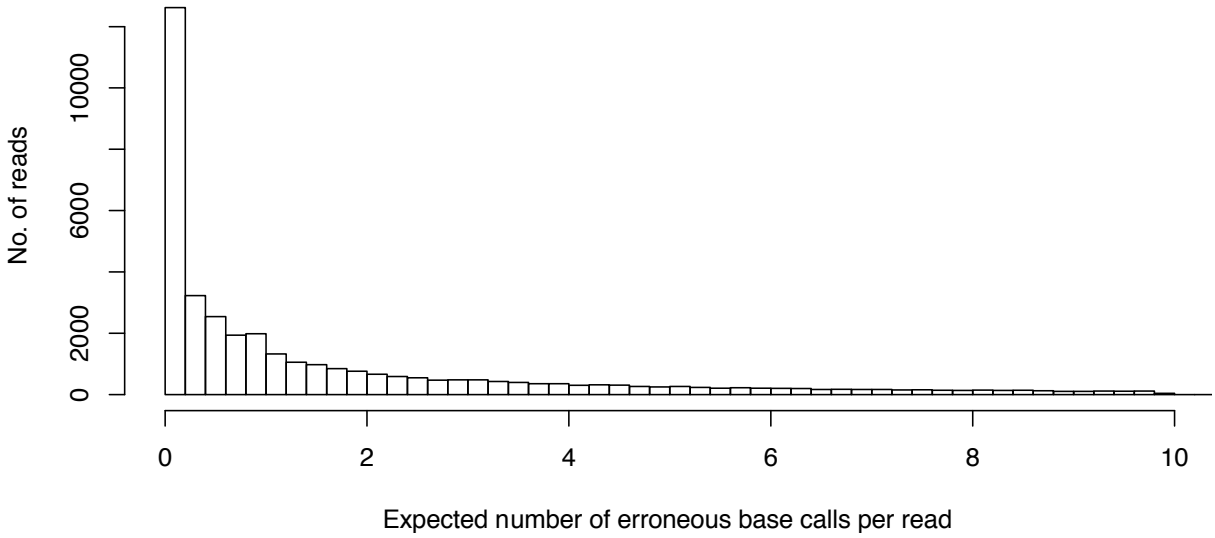


Fig. S1 Estimates of the expected errors of our PacBio CCS reads using the R script available in the Dryad Digital Repository: doi: 10.5061/dryad.dj82k (Rothfels et al. 2016).

Table S2 Summary of the PURC results for the ten compared regimes

	Locus		
Regime	ApPEFP	CRY2	TRANS
a	67 (47.1 ± 25.6)	128 (35 ± 21.4)	64 (39.8 ± 24.2)
b	67 (47.1 ± 25.6)	124 (36.2 ± 23.2)	64 (39.8 ± 24.2)
c	67 (95.4 ± 53.2)	129 (56.2 ± 41.7)	66 (56.7 ± 34.8)
d	**67 (42.5 ± 25.3)**	**127 (31.3 ± 20.7)**	**64 (37.1 ± 22.7)**
e	59 (19.5 ± 11)	116 (20.3 ± 13.9)	58 (17.4 ± 10.4)
f	66 (46 ± 25.9)	121 (35.6 ± 22.2)	64 (39.8 ± 24.3)
g	66 (46 ± 25.9)	119 (36.2 ± 23.5)	64 (39.8 ± 24.3)
h	67 (85.8 ± 45.4)	120 (53.5 ± 39.7)	66 (55.5 ± 33.4)
i	67 (39.4 ± 24.2)	116 (31 ± 20.2)	63 (36 ± 22.5)
j	59 (19.5 ± 11)	108 (20.4 ± 14.7)	56 (17 ± 10.5)

Total number of the final consensus sequences inferred for each regime and allele, followed by the mean coverage and the ± standard deviation of the coverage.

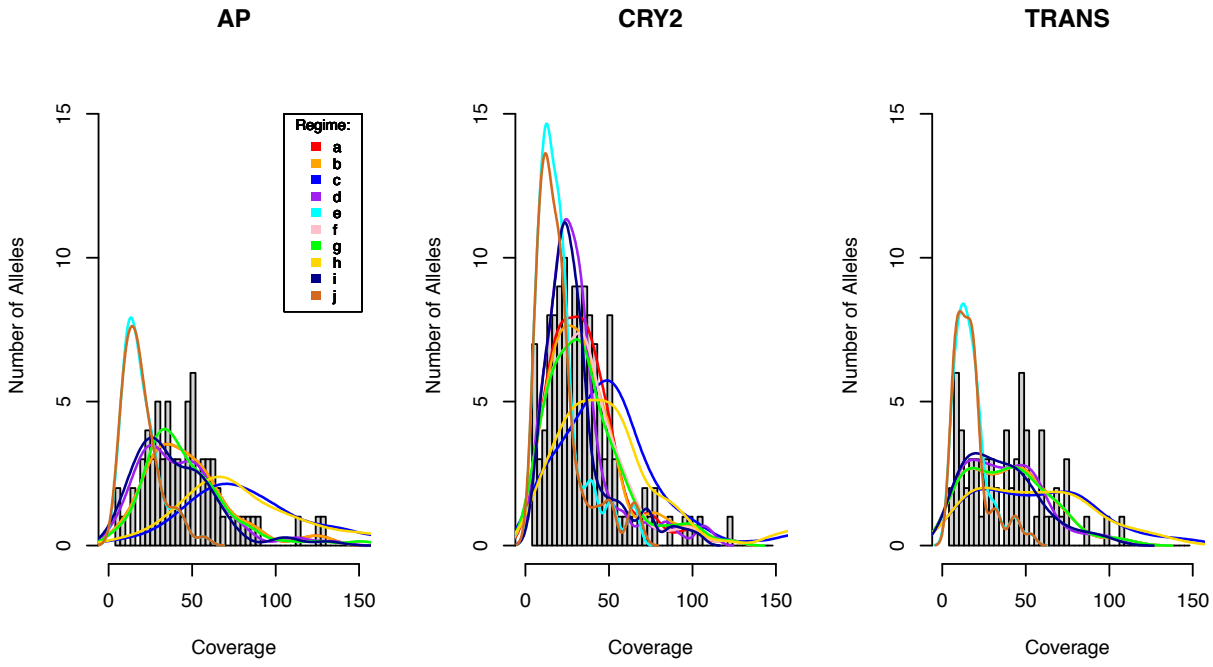


Fig. S2 Coverage of the inferred allele sequences with PURC using the R script available in the Dryad Digital Repository: doi: 10.5061/dryad.dj82k (Rothfels et al. 2016).

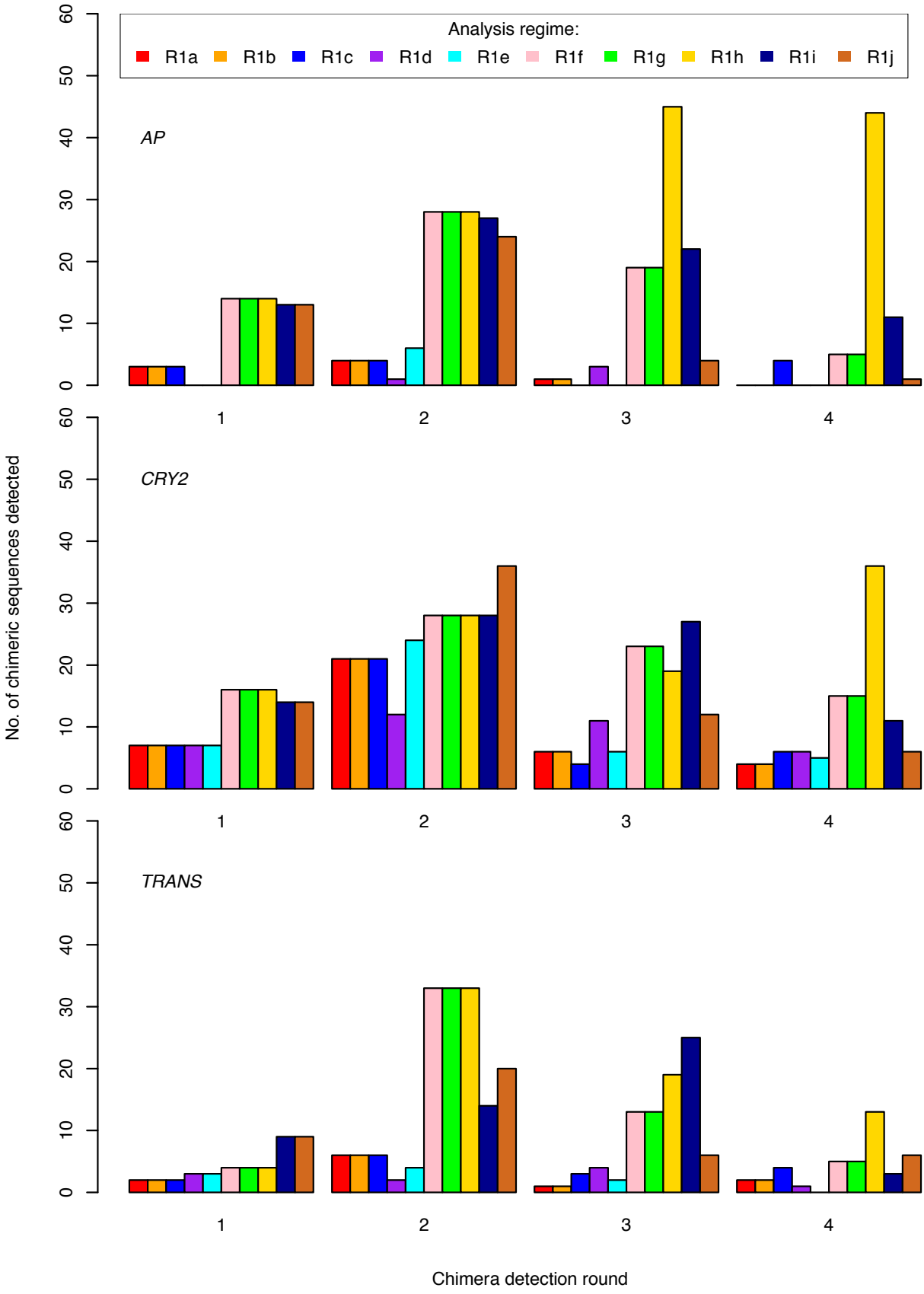


Fig. S3 Number of chimeras sequences found in our dataset using the R script available in the Dryad Digital Repository: doi: 10.5061/dryad.dj82k (Rothfels et al. 2016).

Chapter 4

Dispersal and Diversification in Ancestral Vascular Plants

B. Dauphin[1], J. R. Grant[1], M. C. Stensvold[2], A. Maccagni[1,3], J. Nason[4], D. R. Farrar[4]

[1]Université de Neuchâtel, Laboratoire de Génétique Évolutive, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

[2]P.O. Box 1042, Sitka, Alaska, 99835, U. S. A.

[3]Departement Umweltwissenschaften, Forschungsgruppe Pflanzenökologie, Universiät Basel, Hebelstrasse 1, 4056 Basel, Switzerland

[4]Iowa State University, Department of Ecology, Evolution, and Organismal Biology, 2200 Osborn Dr., Room 251 Bessey Hall, Ames, Iowa 50011–4009, U. S. A.

This chapter is in preparation for its submission.

Abstract

Key evolutionary novelties have enabled vascular plants to conquer terrestrial habitats and diversify to build up a large portion of the tree of life. Although the speciation mechanisms of flowering plants are well documented, the major biological processes driving diversification of early vascular plants—dispersal ability and suitable mating systems—remain largely unknown. Here, we report how the shift from sporophytic outcrossing to gametophytic selfing coupled with the dispersal capability of spores supports long-distance colonization events that ultimately have led to the speciation of highly variable and strongly genetically depauperate species. Illustrating the versatility of the life cycle of extant species, our results provide new insights on speciation mechanisms that allowed the evolution of early vascular plants over 450 million years ago.

Introduction

More than 450 million year ago, early vascular plants appeared and spread to terrestrial environments (Kenrick and Crane 1997). They rapidly explored new ecological niches and diversified in suitable habitats to form over 300,000 extant species. Key evolutionary innovations such as dispersal by airborne spores and mixed breeding systems enabled this conquest, and today, several representatives of early divergent vascular plants subsist and are the witnesses of this very ancient period. Indeed, phylogenetic studies have shown the Ophioglossales (*Botrychium*, *Ophioglossum,* and related taxa), and their sister group, the Psilotales (*Psilotum* and *Tmesipteris*), to be among the early divergent lineages of ferns, themselves sister to the seed plants (Spermatophyta) (Pryer et al. 2004; Rothfels et al. 2015; PPG I 2016). In seed plants, dispersal units (seeds) of diploid plants are basically diploid, thus carrying a "built-in" level of genetic diversity. In diploid ferns, dispersal units are usually haploid spores, having been produced by meiosis in the diploid plant. The haploid spore, upon germination in a suitable environment, grows into a free-living haploid gametophyte plant that is genetically capable of producing both male and female gametangia (Haufler et al. 2016). It is thus capable, upon fusion of egg and sperm, of producing a diploid, sporophyte plant, even in isolation. However, that egg and sperm of this isolated gametophyte are both produced by mitosis, and thereby are genetically identical. Growth experiments have illustrated deleterious effects that might be expected from this extreme form of inbreeding (Klekowski 1982), but surprisingly, many fern species actually have the capability of producing a variable percentage of pure homozygous sporophytes through gametophytic selfing (Sessa et al. 2016).

Being part of the earliest divergent fern lineages, plants of the genus *Botrychium* are peculiar among vascular plants with their inbreeding behavior as well as morphological characters (Fig. 1B). Studies on *Botrychium* have shown that most of its taxa usually reproduce through gametophytic selfing (Hauk and Haufler 1999). Based on population genetic data derived through enzyme electrophoresis, populations typically contain genetic diversities among the lowest recorded for vascular plants (Ranker and Haufler 2008). The same studies show a lack of allelic heterozygosity, very little genetic variability at all among plants of a population, and very little diversity among populations. It was therefore concluded from this evidence that genotypes have arisen that are both devoid of a genetic load of deleterious alleles, but nevertheless suited to their current environment. However, comparative genetic analyses among species show differentiation equal to that recorded for sexually outcrossing species (Dauphin et al. 2017), which leads us to the hypothesis that the lack of diversity within species is the derived state that has arisen from an ancestral polymorphic and outcrossing taxon. A first indication supporting this hypothesis of out-crossing ancestors was found in a study of the species *Botrychium lineare*, in which populations sampled at high elevations in the Rocky Mountains were remarkably genetically diverse and highly heterozygous (Farrar and Gilman 2017). Further investigations revealed that high-elevation populations of *B. lineare* contained all the

alleles of its long-suspected sister species *B. campestre* distributed in the Great Plains and in eastern North America at low elevation, which displayed depauperate genetic variability. Hence, it seems that both taxa were derived from an outbreeding ancestor which guided our later assessment of *Botrychium* evolution and its possible broader relevance to reproductive behavior of early vascular plants using the worldwide distributed *Botrychium lunaria* complex as a model (Fig. 1A-B) (Stensvold and Farrar 2017).

Materials and methods

To investigate population genetics and mating systems in *Botrychium lunaria*, we sampled 485 individuals from 17 populations across its geographical range in the Alps and Jura Mountains in Switzerland, with an average of 29 individuals per population (Fig. 2A-B). We selected sampling locations based on biogeographical area, topography, soil characteristics, and the known demographical history of plants at these sites during the Quaternary climatic oscillations, with both historical contexts to be either a migratory route or a potential glacial refuge (table S2). We conducted electrophoresis on 22 gene loci of 10 enzyme systems (Stensvold and Farrar 2017) and calibrated the scoring of alleles with reference specimens for genotyping. We removed rare alleles (overall frequencies <0.05) to avoid any inflation of polymorphism (Hartl and Clark 1998) and excluded monomorphic loci. Additionally, we included genetic data of all other taxa of the *B. lunaria* complex for comparison of genetic diversities and genotypes frequencies derived at the conspecific level (table S1, Stensvold and Farrar 2017).

We first assessed the allele richness and the percentage of polymorphic loci to characterize the genetic diversity within populations, and we calculated the levels of heterozygosity among populations, testing 11 loci for departure from Hardy-Weinberg equilibrium (HWE). The degree of inbreeding within populations was also estimated based on the expected and observed proportion of heterozygotes, which is relevant to evaluate the prevalence for gametophytic and sporophytic selfing. We then characterized the genetic structure among populations using both genotype- and frequency-based methods, and combined multivariate with Bayesian clustering analyses. We performed a principal component analysis (PCA) (Hotelling 1933) based on a covariance matrix to explore the genotypic variation across all individuals, and we carried out a discriminant analysis of principal components (DAPC) (Jombart et al. 2010) to detect weak differentiation among populations by minimizing the variance within populations and maximizing it between populations. Additionally, to help in elucidating the spatial pattern of admixture populations, we used STRUCTURE (Pritchard et al. 2000) to infer the overall genetic structure (K = 1–20) under an admixture model with correlated allele frequencies (Falush et al. 2003). Also, we estimated Nei's (1978) unbiased genetic identity (*GI*) and genetic distance (*D*) to determine the degree of differentiation between populations, and we

inferred an un-rooted neighbor-joining tree to illustrate the genetic distances among populations and between taxa of the species complex. Isolation-by-distance (IBD) was tested for both, the Nei's genetic distance and the conditional genetic distance (cGD) calculated from a Population Graph (Dyer and Nason 2004). We examined the relationship between matrices of genetic distances and pairwise-population geographical distances using 1000 random permutations based on a Mantel's test (Rousset 1997).

Table 1 Comparison of genetic diversity between seed plants, ferns, genus *Botrychium*, *B. lunaria* complex, and the Swiss Alps *B. lunaria* analyzed in this study as measured by percent polymorphic loci and average number of alleles per locus.

Taxon group	Percent polymorphic loci	Average number of alleles per locus	Reference
Seed plants	mean 58.9	mean 2.29	Hamrick & Godt, 1990
Ferns	0.00 – 80, mean 36	1.00 – 2.8, mean 1.65	Li & Haufler, 1999 (data for 32 taxa)
Botrychium	0.00 – 16.7, mean 7.44	1.00 – 1.2, mean 1.10	Farrar, 1998 (data for nine diploid species)
	0.00 – 21, mean 7.30	1.00 – 1.21, mean 1.07	Hauk & Haufler, 1999 (data for five diploid species)
B. lunaria complex	5 – 35, mean 20	1.05 – 1.45, mean 1.20	Stensvold & Farrar, 2016 (data for six diploid taxa)
B. lunaria in Swiss Alps	43 – 100, mean 78.7	1.56 – 2.63, mean 2.18	This study

Table 2 Genetic diversity and level of heterozygosity among taxa of *B. lunaria* complex.

Taxon	Nb. Ind.[1]	Nb Pop.[2]	P[3]	A/L[4]	H_E[5]	H_O[6]	Departure[7]	F_{IS}[8]
B. lunaria in Swiss Alps	485	17	76.99	2.18	0.45	0.40	-0.05	0.208
B. lunaria var. *lunaria*	471	63	33.30	1.30	0.47	0.02	-0.45	0.957
B. lunaria var. *melzeri*	71	13	0.00	1.00	NA[9]	NA[9]	NA[9]	NA[9]
B. neolunaria	534	125	10.00	1.10	0.50	0.00	-0.50	0.994
B. crenulatum	115	45	15.00	1.30	0.50	0.00	-0.50	1.000
B. tunux	333	35	25.00	1.20	0.26	0.00	-0.26	1.000
B. nordicum	48	3	10.00	1.10	0.41	0.00	-0.41	1.000

Note: [1]Number of individuals analyzed, [2]Number of populations sampled, [3]Percentage of polymorphic loci, [4]Number of alleles per loci, [5]Expected proportion of heterozygosity, [6]Observed proportion of heterozygosity, [7]Departure from expected heterozygosity, [8]Inbreeding coefficient as $F_{IS} = 1 - (H_O / H_E)$, [9]Non applicable since a single homozygous genotype occurs with no variability.

Results

Genetic variation within and among Swiss populations was similar to what is observed in outbreeding seed plants and notably, significantly higher than in other *Botrychium* taxa (Table 1; table S1). We resolved 16 polymorphic loci from 9 enzyme systems and identified 54 alleles overall. The number of alleles per locus varied from 1.56 to 2.63 and the percentage of polymorphic loci ranged from 42.86 to 100 among populations (table S3). A majority of alleles found in the widely distributed taxa of the *B. lunaria* complex occurs in Swiss *B. lunaria* populations, which constitute the relictual genetic pool of a polymorphic ancestral taxon.

In our analysis, we found a considerable level of heterozygosity (H_O = 0.40) in comparison with all taxa of the *B. lunaria* complex (Fig. 1D, Table 2). No deviations from HWE were detected for all loci ($P < 0.05$) in the majority of populations (table S4), and the mean value of departure from expected heterozygosity was very close to a random mating (mean departure = – 0.05; Table 2). Also, the average inbreeding coefficient throughout the loci was low (F_{IS} = 0.208, tables S5 and S6), contrasting with the results of all other taxa of the *B. lunaria* complex that displayed an extreme case of inbreeding populations (from 0.957 to 1.000; Table 2). Together, these results support the predominance of an outcrossing mating system within the Swiss populations of *B. lunaria* and corroborate previous findings highlighted in its closely related species *B. campestre* var. *lineare* (Farrar and Gilman 2017).

Based on our PCA, the first two axes summarize 24.54% of the full variance but populations appeared largely undifferentiated with a comparable genotypic variation among populations (fig. S1). Despite its discriminant ability, the DAPC failed to differentiate populations (Fig. 2C), which was consistent with high *GI* values (mean *GI* = 0.908) (table S7). Genetic distance among populations was usually low (mean *D* = 0.099, table S7), except for three populations (ARO, GOP, and VDC), of which ARO clearly appeared as the most genetically distant in the un-rooted neighbor-joining tree (fig. S2). Focusing at the individual level, our Bayesian clustering analyses showed a large proportion of admixture among individuals that were not assignable to one of the three clusters from the optimal *K* = 3 (Fig. 2D). Also, these Bayesian inferences displayed a high level of admixture between populations, unrelated to their geographical locations, whereas previous studies reported a strong spatial structure with few inter-population migrants in closely related taxa of the complex (Stensvold and Farrar 2017).

Our Population Graph analysis displayed at least two barriers to gene flow (Fig. 2E), with a pivotal role for the Chasseron (CHO) population in the Jura, and no apparent topographic feature enforces a limit to population connectivity (fig. S3). A pattern of IBD was found based on the *cGD* (*P* = 0.011, *r* = 0.27), of which the genetic differentiation takes place in distant patches discontinuous along geographic distance (Fig. 2F and fig. S4), what is expected in these Alpine populations

separated by heterogenous habitats and major mountainous massifs ranging up to 4000 meters of elevation. These results of population genetic structure are coherent with the maintain of an outcrossing mating system in recently derived populations from a genetically variable ancestral taxon, likely native to the Alps.

Within the *B. lunaria* complex, species relationships illustrated in our un-rooted neighbor-joining tree (Fig. 1C) was congruent with the phylogenetic reconstruction based on nuclear genes (Dauphin et al. in review), in which taxa are segregated in two main clades, and where the two species *B. neolunaria* and *B. nordicum* embedded the most divergent positions.

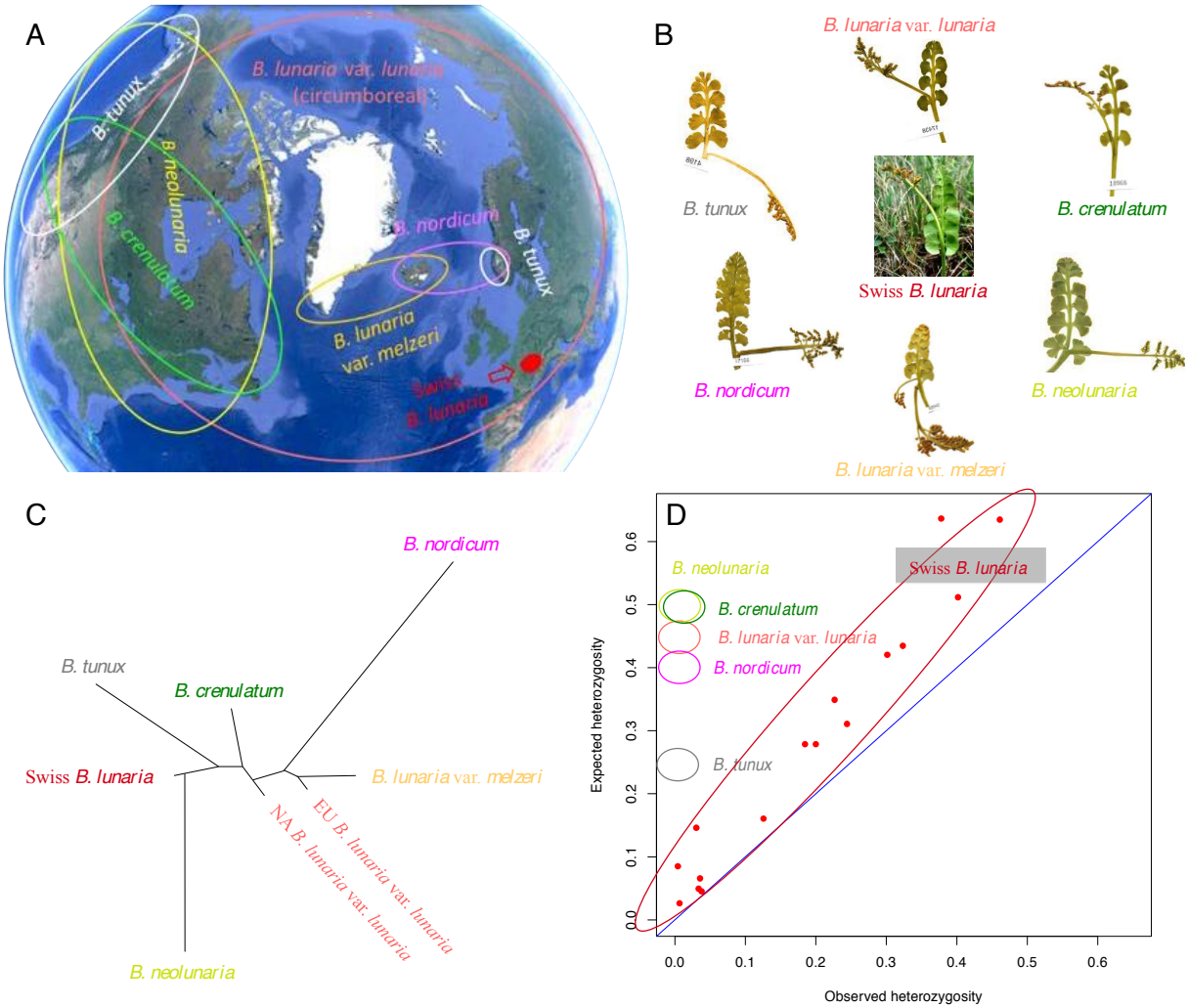


Fig. 1. Worldwide distribution of the *Botrychium lunaria* complex. (A) Geographical range of the taxa of the *B*. *lunaria* complex in northern hemisphere. (B) Morphological diversity of taxa of the *B*. *lunaria* complex. (C) Un-rooted neighbor-joining tree illustrating the genetic distances among taxa. (D) Comparison of levels of heterozygosity and departure from Hardy-Weinberg equilibrium among taxa.

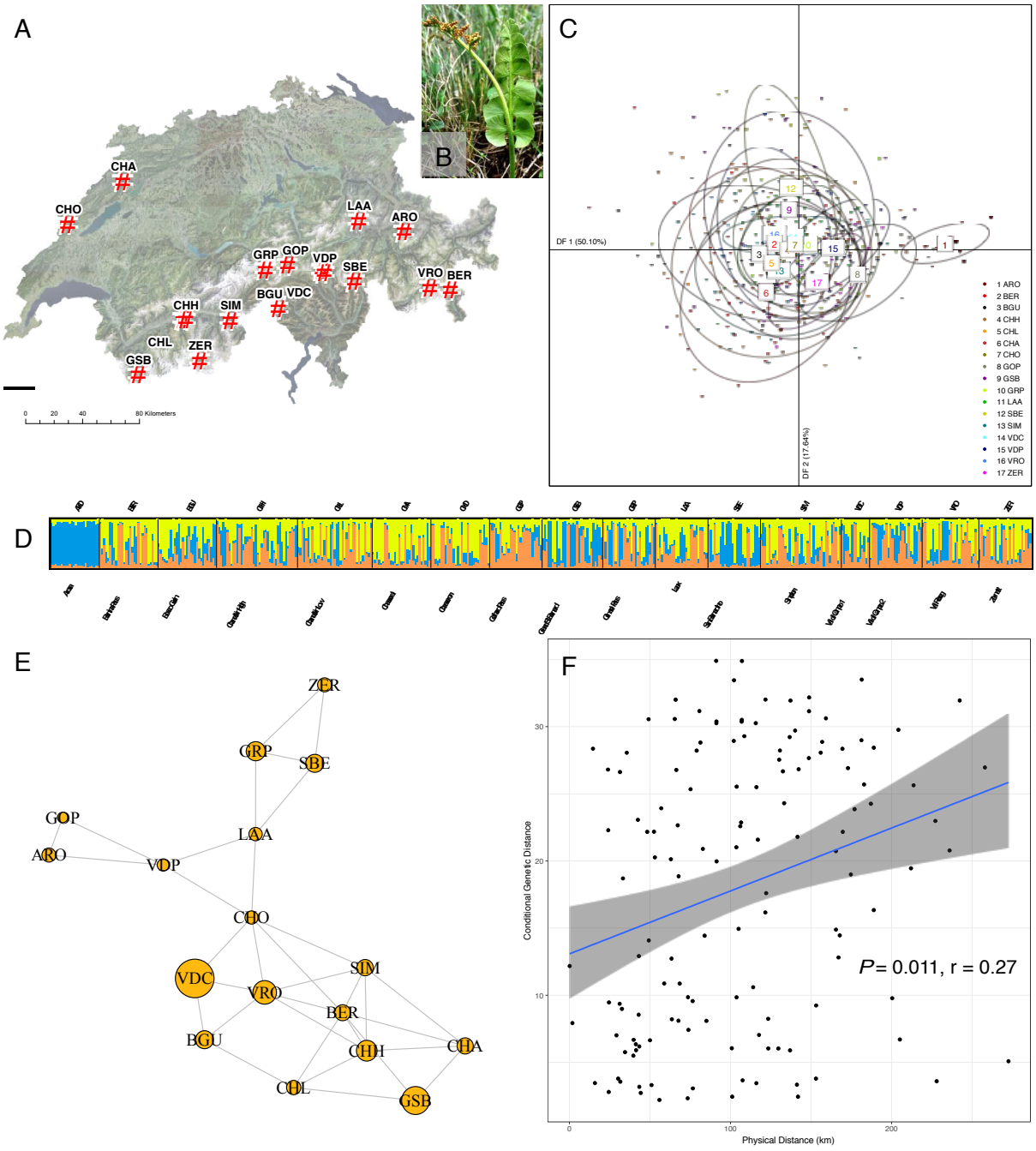


Fig. 2. Genetic structure of the Swiss *B. lunaria* populations. (A) *Botrychium lunaria* populations sampled in Swiss Alps. (B) Represents the enigmatic morphological characters of *Botrychium lunaria*; a specimen from a Swiss Alps population. (C) Discriminant analysis of principal components maximizing the variance between populations for the detection of weak genetic differentiation. The two first discriminant functions are represented along axes. (D) Admixture bar plots of posterior probability of individual assignment for $K = 3$. Each bar corresponds to an individual of *B. lunaria* and colours in each bar represent the admixture proportion of each cluster value. Names above and codes below the plot refer to the locations. (E) Population Graph representing the genetic relationships and the barriers to gene flow among the Swiss Alps *B. lunaria* populations. (F) Correlation between conditional genetic distances and physical distances, with confidence interval and mantel test.

Discussion

Our empirical findings of Swiss *B. lunaria* populations provide new insights on the dispersal and diversification mechanisms driving the evolution of early vascular plants. For a long time, population genetic studies focusing on Ophioglossaceae have reported extreme cases of inbreeding, with pure homozygous populations harboring few polymorphic loci, which have highlighted the advantages of a restricted genetic load by purifying deleterious alleles. A striking result is the unprecedented genetic diversity found in a representative of Ophioglossaceae, and the strength of the sporophytic outcrossing maintained in Alps *B. lunaria* populations. In contrast to all other taxa of the species complex, these results provide new insights on the alternation of mating systems in closely related species, and support that shifts to selfing are likely driven by environmental conditions and the geographical origins of taxa. Within another major lineage of *Botrychium*, a previous study has shown an analogous pattern, where high-elevation populations of *B. campestre* in Rocky Mountains, have displayed the highest genetic diversity and recombination rate compared to populations occurring at low elevations (Farrar and Gilman 2017). In both outcrossing cases, populations occurred on steep alpine slopes that may foster sperm movement between underground bisexual gametophytes via sub-surface water flow.

Spore dispersal by wind is known to be an important feature driving fern species demography, rather than the vicariance events (Wolf et al. 2001), and several cases of long-range dispersal at intercontinental scales are documented for the worldwide *B. lunaria* species complex (Dauphin et al. 2017; Stensvold and Farrar in review). The newly isolated migrants arose from a single or few spores can maintain a viable population through reiterative gametophytic selfing in suited habitats. Our study brings a further nuance with spore dispersal at shorter distances and presents evidence of a weakly variable population resulting from a founder event. These plants of the Arosa (ARO) locality, occurring on a thin soil with permanent permafrost, show a strong fixation of alleles with the lowest percentage of polymorphism (42.86%, table S3, Fig. 2D). This recently established population after the last glacial oscillations provides an empirical example of population differentiation associated with genetic drift, and complements to another scale our understanding of the consequences of spore dispersal in taxon diversification.

Climatic oscillations of the Quaternary period had a significant impact on species distribution in the Alps, especially 23,500 years ago during the last glacial maximum (LGM), where ice and permanent snow almost fully covered the massifs. In this context, *B. lunaria* populations and the alpine flora in general have survived following one of the two scenarios, either by the migration to peripheral areas of Alps ("*tabula rasa*", Chodat and Pampanini 1902), mostly in southern lands, or by an *in situ* survival in restricted free-ice locations of central Alps ("*nunataks*", Brockmann-Jerosch and Brockmann-Jerosch 1926). Based on the distribution of allele diversity, and the connectivity of distant populations with a diffuse topographic pattern (fig. S3), likely the Swiss *B. lunaria* has experienced both scenarios, persistence of populations (ZER, GRP, LAA, SBE) *in situ* with the high allele richness

as a signature, and a post-glacial recolonization from southerly refugia that coincides with the shared ancestral polymorphism forming one of the three phylogroups (Fig. 2E). One of the highest-diversity populations was located in the central Alps in Zermatt (ZER), which confirms the historical importance of this region as hotspot of diversity and refuge for several alpine species during the LGM (Stehlik 2000; Wohlgemut 2002). The life cycle of *Botrychium* confers a survival ability in cold and desiccant habitats, where its endomycorrhizal fungi provide it carbohydrate nutrition as well as water and minerals. Also, the mating system versatility of *B. lunaria* allows to overcome unfavorable conditions for outcrossing, where sparse populations may persisted in *nunataks* through gametophytic selfing without loss of genetic diversity.

Our study provides new insights for understanding the biological mechanisms involved in speciation of the first vascular plants. A better understanding of their evolutionary history is crucial for predicting the fate of these early divergent vascular plants and their resilience capacity in context of global changes.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

B.D. and J.R.G. were funded by the Swiss National Science Foundation (SNSF grant 31003A_156456). D.F, J.N, and M.S were supported by Iowa State University, and B.D., J.R.G., and A.M. by the University of Neuchâtel.

References

1. P. Kenrick, P. R. Crane, The origin and early evolution of plants on land. *Nature* **389**, 33–39 (1997).

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

3. C. J. Rothfels, F. W. Li, E. M. Sigel, L. Huiet, A. Larsson, D. O. Burge, M. Ruhsam, M. Deyholos, D. E. Soltis, C. N. Stewart, S. W. Shaw, The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *Am. J. Bot.* **102**, 1089–1107 (2015).

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

5. E. J. Klekowski, Genetic load and soft selection in ferns. *Heredity* **49**, 191–197 (1982).

6. E. B. Sessa, W. L. Testo, J. E. Watkins, On the widespread capacity for, and functional significance of, extreme inbreeding in ferns. *New Phytol.* **211**, 1108–1119 (2016).

7. W. D. Hauk, C. H. Haufler, Isozyme variability among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). *Am. J. Bot.* **86**, 614–633 (1999).

8. T. A. Ranker, C. H. Haufler, *Biology and evolution of ferns and lycophytes* (Cambridge University Press, Cambridge, ed. 1, 2008), pp. 113–119.

9. B. Dauphin, D. R. Farrar, A. Maccagni, J. R. Grant, A worldwide molecular phylogeny provides new insight on cryptic diversity within the moonworts (*Botrychium* s.s., Ophioglossaceae). *Syst. Bot.* **42**, XX–XX (2017).

10. D. R. Farrar, A. V. Gilman, Relationships in the *Botrychium campestre* (Ophioglossaceae) complex. *Brittonia* **69**, 265–275 (2017).

11. M. C. Stensvold, D. R. Farrar, Genetic diversity in the worldwide *Botrychium lunaria* (Ophioglossaceae) complex, with new species and new combinations. *Brittonia* **69**, 148–175 (2017).

12. D. L. Hartl, A. G. Clark, *Principles of population genetics* (Sinauer Associates, Sunderland, ed. 3, 1998), pp. 54–57.

13. H. Hotelling, Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417–441 (1933).

14. T. Jombart, S. Devillard, F. Balloux, Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94 (2010).

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

16. D. Falush, M. Stephens, J. K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).

17. M. Nei, Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590 (1978).

18. R. J. Dyer, J. D. Nason, Population graphs: the graph theoretic shape of genetic structure. *Mol*. *Ecol* **13**, 1713–1727 (2004).

19. F. Rousset, Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**, 1219–1228 (1997).

20. P. G. Wolf, H. Schneider, T. A. Ranker, Geographic distributions of homosporous ferns: does dispersal obscure evidence of vicariance? *J*. *Biogeogr* **28**, 263–270 (2001).

21. I. Stehlik, Nunataks and peripheral refugia for alpine plants during quaternary glaciation in the middle part of the Alps. *Bot*. *Helv*. **110**, 25–30 (2000).

22. T. Wohlgemut, Alpine plant species richness in the Swiss Alps: diversity hot spots reconsidered. *Mém*. *Soc*. *Bot*. *Genève* **3**, 63–74 (2002).

Supplementary Materials

Materials

Case study and sampling

Botrychium Sw. is a genus of the fern family Ophioglossaceae with 35 accepted taxa divided among three major lineages; Lanceolatum, Lunaria, and Simplex-Campestre (Hauk et al. 2003; PPG I 2016; Dauphin et al. 2017). *Botrychium* species are small perennial plants, commonly known as moonworts, and have a relatively simple morphology constituted by an annual common stalk dichotomously divided into a trophophore and a sporophore (Clausen 1938).

The *B. lunaria* complex (L.) Sw. forms a monophyletic group of six morphologically distinct diploid taxa; *B. crenulatum*, *B. lunaria* var. *lunaria*, *B. lunaria* var. *melzeri*, *B. neolunaria*, *B. nordicum*, and *B. tunux*. It has a circumtemperate and circumboreal distribution in the northern hemisphere, with some disjunctions in Oceania (Stensvold 2008; Stensvold and Farrar 2017). Several cases of long-range dispersal through wind and possibly bird migrations were reported for *B. crenulatum*, *B. neolunaria*, and *B. tunux*, which provide evidence of the capability of these fern species to spread out at inter-continental scales (Stensvold 2008; Dauphin et al. 2017; Stensvold and Farrar in review). The species complex includes a single allotetraploid, *B. yaaxudakeit*, that arose from interspecific hybridization between the North American genotype of *B. lunaria* var. *lunaria* and *B. neolunaria* (Stensvold et al. 2002).

Botrychium lunaria var. *lunaria* is widespread in Europe where it occurs on alpine meadows and disturbed deglaciated areas, rarely below the tree line. A phylogeographical study has shown its high haplotype diversity in European Alps and provided evidence of at least two southern peripheral refugia for the survival of *B. lunaria*'s ancestor during the Quaternary glaciations (Maccagni et al. 2017). Also, authors have characterized three main genetic groups in Alps populations based on cpDNA but no morphological or ecological differences were discernable for supporting any correlation between genetics and character traits.

Swiss Alps sampling took place in 15 locations across the biogeographical range of *B. lunaria* var. *lunaria* (Fig. 2A), with an altitudinal range from 1552 to 2644 meters (table S2), and two additional localities (CHA and CHO) from the Jura Mountains were sampled as a distinct geographical area. The population sizes, approximated from the number of emergent sporophytes, varied from 80 to over 1000 individuals, although it is difficult to assess the effective population sizes without a quantification of mature underground gametophytes. In total, 485 sporophytes were randomly collected across the population distribution, and fresh leaf samples were stored in cold conditions during transport up to laboratory. All genotyped specimens were individually numbered and vouchers were deposited at the Ada Hayden Herbarium (ISC) at Iowa State University for morphological study.

Molecular data

Allozyme data were generated following the Stensvold and Farrar's procedure (2017). As soon as possible after collection, 1 cm of the common stalk tissue was cut off and crushed in a buffer solution containing phosphate-polyvinylpyrrolidone (Cronn et al. 1997). Lysate was filtered and stored up to 12 months in microcentrifuge tube at –80°C until the electrophoresis run. Enzyme electrophoresis was carried out in horizontal 12% starch gel using the three buffer systems 7, 9 and 11 from Soltis et al. (1983) to reveal allelic variants on 22 loci of eleven enzyme systems by differentiating migration patterns. Thus, loci of those enzyme systems were resolved for aspartate aminotransferase (AAT) and triose-phosphate isomerase (TPI) with the buffer system 7, for malate dehydrogenase (MDH), phosphoglucomutase (PGM), 6-phosphogluconate dehydrogenase (6-PGD) and phospho-glucoisomerase (PGI) with the buffer system 9 and, for aconitase (ACN), diaphorase (DIA), isocitrate degydrogenase (IDH) and shikimate dehydrogenase (SKDH) with the buffer system 11 (for more details refer to Stenvold and Farrar 2017). The stain recipes were used following Soltis et al. (1983) and enzymatic activity reflected by stained bands was visualized on the starch gel containing each time identical sets of 26 loaded samples.

Scoring of alleles was calibrated with reference *Botrychium* specimens characterized by their known banding patterns for all loci and scoring accuracy was assessed by comparing gels run multiple times. Some specimens were run twice on the same gel to remove any uncertainties and allele variants for each locus were numbered by their migration order starting with the number 1 for the biggest allozyme weight.

Data were assembled in a matrix with individuals in row and allelic variants in column. We considered a polymorphic locus when a least two alleles were identified and we excluded rare alleles (overall frequencies <0.05) to avoid any inflation of polymorphism (Hartl and Clark 1998).

Methods

Genetic diversity and mating system characterization

We calculated summary statistics of genetic diversity within populations using POPGENE (Yeh et al. 1997); allelic frequency, percentage of polymorphic loci (% P), Shannon index, mean Nei's (1973) expected heterozygosity (H_E), observed heterozygotes on average through all loci (H_O), and the Nei's (1978) unbiased genetic identity (*GI*) and genetic distance (*D*). We calculated the levels of heterozygosity among populations and the inbreeding coefficient based on genotype frequencies ($F_{IS} = 1 - (H_O / H_E)$). The departure from Hardy-Weinberg equilibrium was independently tested for the 11 loci based on Chi-square test using POPGENE (Yeh et al. 1997).

Population genetic structure

Principal component analysis (PCA) (Hotelling 1933) was performed using the *dudi.pca* function from adegenet v.2.0.1 (Jombart and Ahmed 2011) as implemented in R (R Core Team 2017). We retained five axes including 50.22 % of the full variance and we plotted the two first axes that represent 14.46 % and 10.08 % respectively (fig. S1). Then, we investigated population differentiation using a discriminant analysis of principal components (DAPC) (Jombart et al. 2010), which minimizing the variance within populations and maximizing it between populations. We assessed the trade-off between power of discrimination and over-fitting based on the *a.score* and *optim.a.score* function (adegenet v.2.0.1) that identified seven PCs as sufficient for describing the data.

Bayesian clustering analysis was carried out using STRUCTURE v2.3.4 (Pritchard et al. 2000), without stipulate spatial prior distribution, under the admixture model with correlated allele frequencies (Falush et al. 2003). We ran 10 replicates for K = 1–20 with a burnin-in of 250 000 MCMC iterations, followed by 500 000 MCMC for each run. We inspected the convergence of the replicates through runs and diagnosed the optimal K value based on the delta K method (Evanno et al. 2005) implemented in Structure Harvester v.0.6.94 (Earl and vonHoldt 2012). Then, we averaged replicates for each value of K using CLUMPP v.1.1.2 (Jakobsson and Rosenberg 2007), which minimizes the variance across all iterations and between runs of this optimal K value.

Gene flow and pattern of isolation-by-distance

Gene flow among sampled populations was investigated within a graph theoretic framework based on genetic covariance (Dyer and Nason 2004) using the R packages gstudio and popgraph (available online at https://dyerlab.ces.vcu.edu/). In this network approach, node size represents a measure of multilocus genetic variance within populations, and these nodes are connected by edges only if they have a significant genetic covariance between populations. The resulting topology illustrates the population connectivity and reveals the "phylogroups" of the meta-population. To help in interpretation, we mapped the nodes and edges onto real space (fig. S3) using ggplot2 (Wickham 2016), to show how covariance is spatially structured in the context of the Alps massif.

We inferred the pattern of isolation-by-distance between geographical distance and the Nei's genetic distance (fig. S4A) or the conditional genetic distance (fig. S4B) calculated from our Population Graph (Dyer and Nason 2004). We tested between matrices of genetic distances and pairwise-population geographical distances using 1000 random permutations based on a Mantel's test (Rousset 1997) as implemented in the R-package vegan (Oksanen et al. 2017).

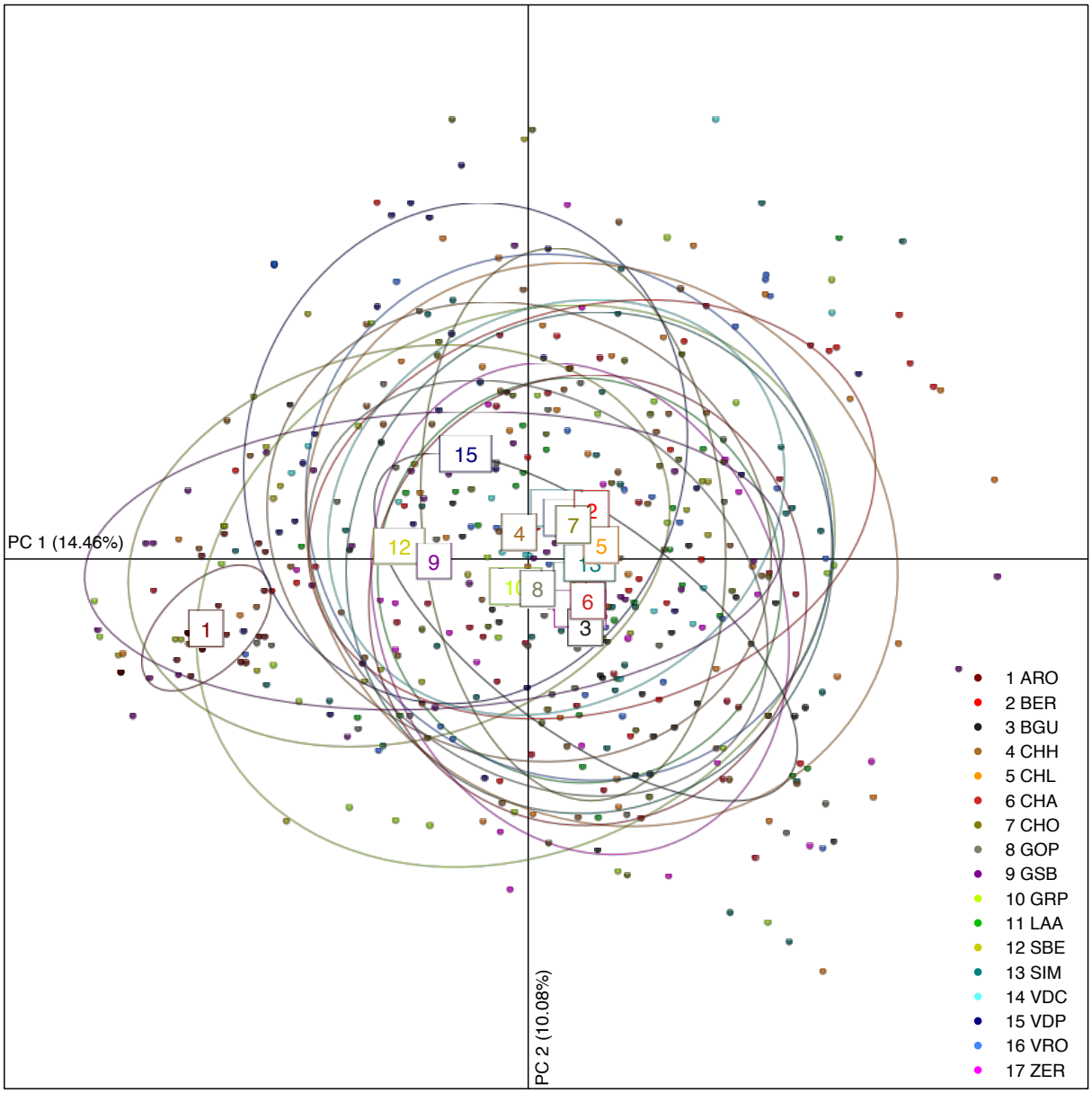


Fig. S1. Principal component analysis showing genotypic variation between individuals and populations. The two first principal components are plotted with each dot representing an individual genotyped, and colored by population.

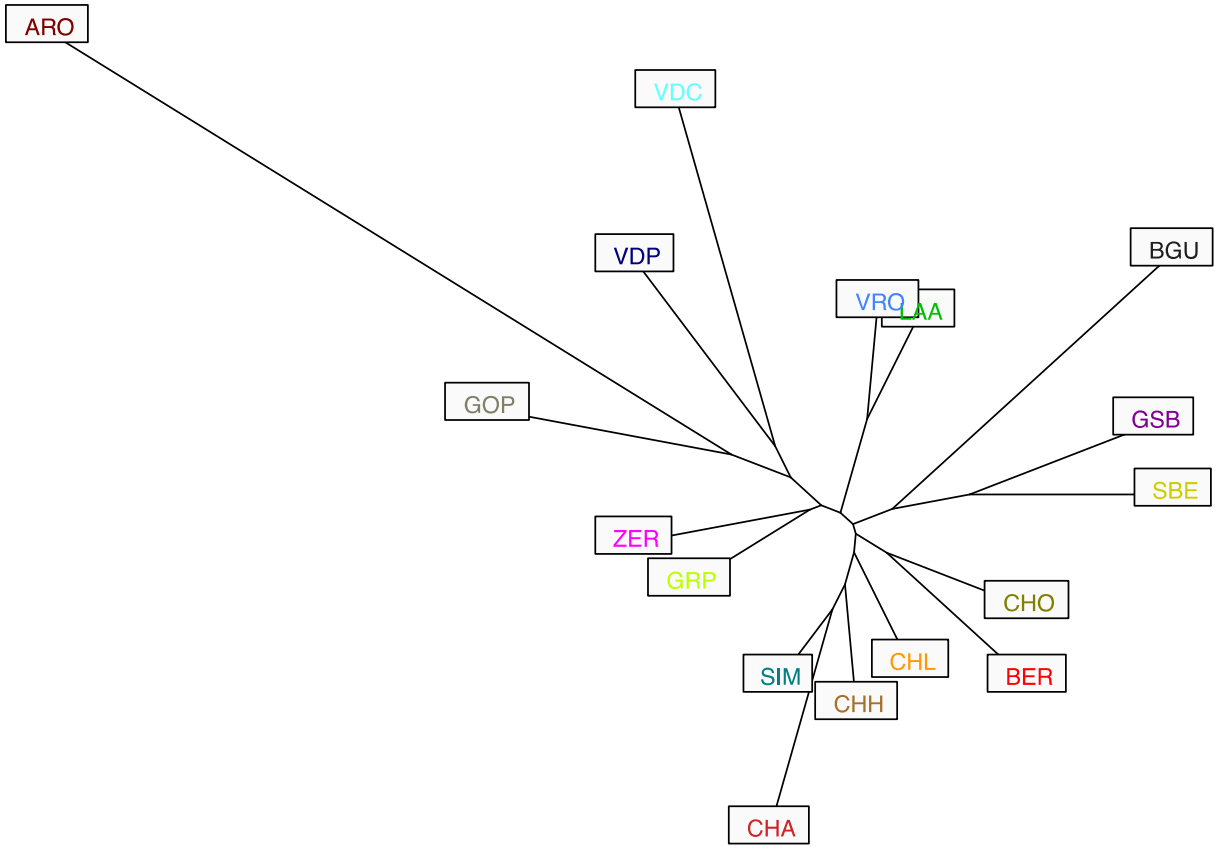


Fig. S2. Un-rooted neighbor-joining tree based on Nei's genetic distance illustrating the relationships among Swiss Alps *B*. *lunaria* populations.

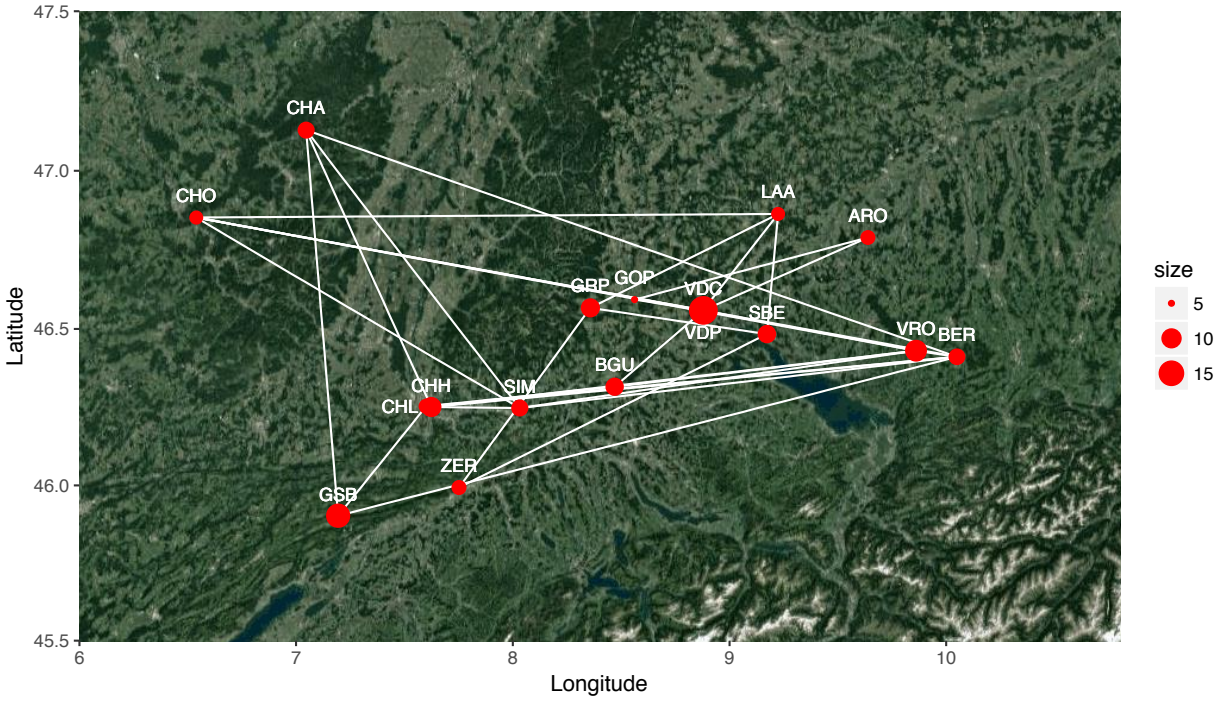


Fig. S3. Spatial Population Graph with a mapping of the nodes and edges onto real space.

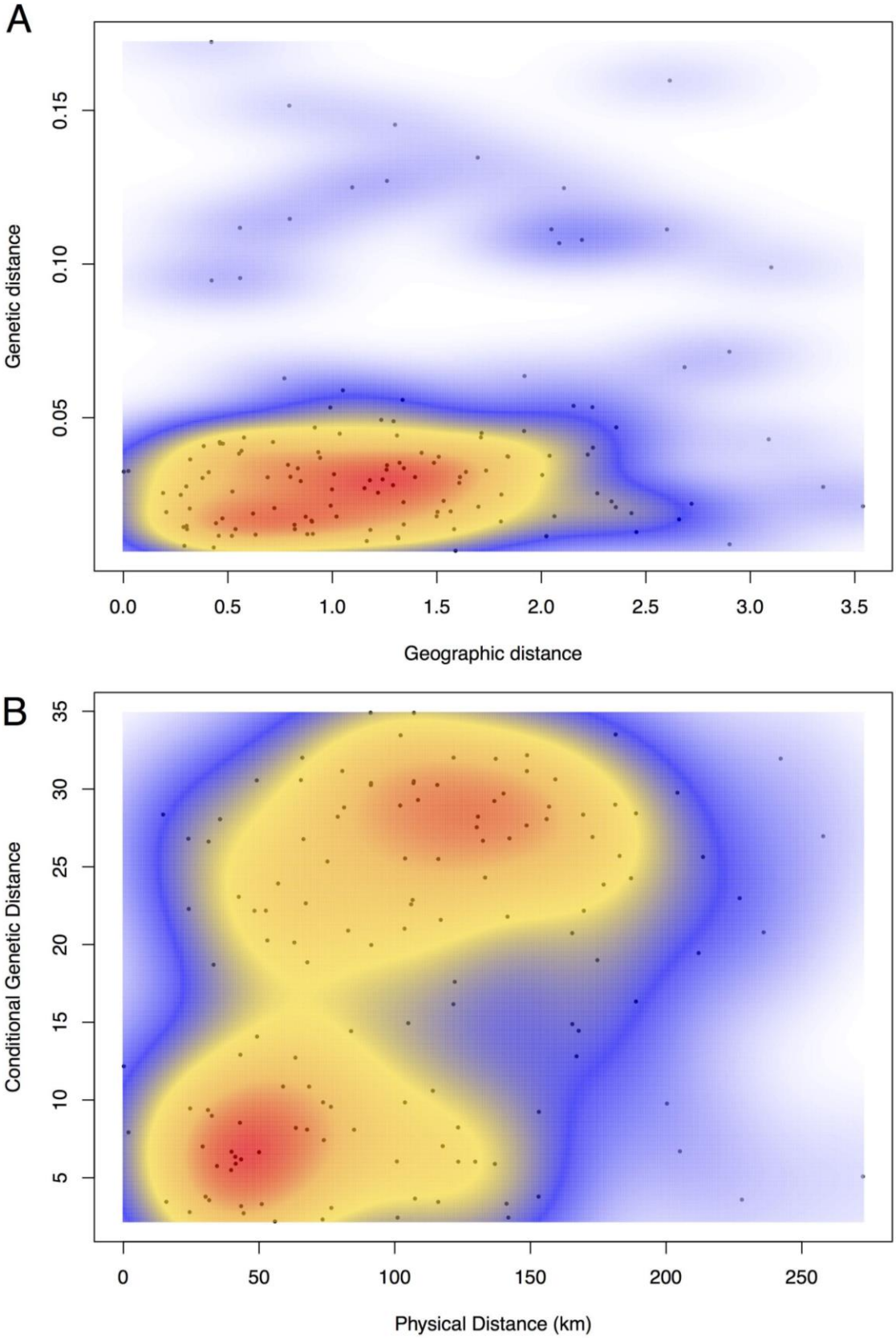


Fig. S4. Pattern of isolation-by-distance with correlations between genetic and geographic distances using a two-dimensional Kernel density estimation denoting the continuous cline and patches of genetic. (A) Represents the relation between Nei's genetic distance and geographical distances, and (B) shows the positive correlation between the conditional genetic distance and physical distances.

Table S1.

Comparison of genetic distinction between seed plants, ferns and *Botrychium* at the taxonomic levels of infraspecific groups, varieties, subspecies and interspecific taxa as measured by genetic identity (GI). A GI value of 1.00 means the plants have the same allele combination.

Taxonomic level	Seed plants	Ferns	*Botrychium*
Conspecific	0.80 – 1.00, mean 0.95[a]	0.83 – 1.0, mean 0.94[b]	0.67 – 1.00, mean 0.91[c]
		0.92 – 1.0, mean 0.96[d]	
Variety	0.71 – 0.99, mean 0.91[a]		0.67 – 1.00, mean 0.85[c]
Subspecies		0.88 – 0.94, mean 0.91[e]	0.74 – 0.79, mean 0.77[f]
Congeneric	0.35 – 0.97, mean 0.67[g]	0.09 – 0.98, mean 0.57[b]	0.18 – 0.80, mean 0.48[b]
		0.19 – 0.92, mean 0.55[c]	0.05 – 0.84, mean 0.41[c]

[a]Gottleib 1981; Crawford 1983, 1985; [b]Farrar, 2005; [c]Hauk & Haufler, 1999 (*B. simplex*); [d]Suter et al., 2000 (European *Asplenium*); [e]Kelloff et al., 2002 (*Athyrium*); [f]Stensvold & Farrar, 2017 (*B. lanceolatum* subsp. *angustisegmentum*, *B. lanceolatum* subsp. *lanceolatum* (green form), subsp. *lanceolatum* (red form); [g]Crawford 1983.

Table S2.

Details of Swiss Alps *B*. *lunaria* populations.

Pop. Nb.	Code	Locality	Longitude[1]	Latitude[1]	Vouchers	Geographical area	Biogeographical area[2]	Elevation (m)[3]
1	ARO	Arosa	9.63885	46.78957	BD160501-BD160530	Alps	Eastern Central Alps	2644
2	BER	Bernina	10.04977	46.41120	AM1-AM29	Alps	Southern Alps	2066
3	BGU	Bosco Gurin	8.47011	46.31574	BD161001-BD161030	Alps	Southern Alps	1911
4	CHH	Chandolin High	7.62394	46.25086	JG5672-JG5714; BD1204301-BD1204320	Alps	Western Central Alps	2513
5	CHL	Chandolin Low	7.60002	46.25330	JG5635-JG5671	Alps	Western Central Alps	2071
6	CHA	Chasseral	7.04625	47.12800	JG5575-JG5604	Jura	Jura	1552
7	CHO	Chasseron	6.53919	46.85266	JG5545-JG5574	Jura	Jura	1574
8	GOP	Gothard Pass	8.56179	46.59290	BD160301-BD160330	Alps	Eastern Central Alps	1788
9	GSB	Grand St. Bernard	7.19378	45.90279	JG5605-JG5634	Alps	Western Central Alps	1905
10	GRP	Grimsel Pass	8.35811	46.56677	BD160801-BD160830	Alps	Western Central Alps	2067
11	LAA	Laax	9.22446	46.86409	BD160601-BD160630	Alps	Northern Alps	2204
12	SBE	San Bernardino	9.17264	46.48341	BD160701-BD160730	Alps	Eastern Central Alps	1944
13	SIM	Simplon	8.03214	46.24785	JG5715-JG5757	Alps	Western Central Alps	2029
14	VDC	Val di Campo 1	8.87939	46.55868	BD1204501-BD1204520	Alps	Southern Alps	1965
15	VDP	Val di Campo 2	8.87715	46.55818	BD160401-BD160430	Alps	Southern Alps	1965
16	VRO	Val Roseg	9.86097	46.43056	BD1204601-BD1204620	Alps	Eastern Central Alps	2025
17	ZER	Zermatt	7.75168	45.99337	BD160901-BD160930	Alps	Western Central Alps	2534

[1]Geographic coordinates are given in decimal degree WGS84;
[2]Data from https://map.geo.admin.ch/;
[3]Variables derived with ArcMap (ESRI) from a digital elevation model at 25 meters of spatial resolution available at https://www.swisstopo.admin.ch/.

Table S3.

Genetic diversity among Swiss Alps *B. lunaria* populations.

Population	Nb. Ind.	Number of Loci	Number of polymorphic loci	% polymorphic loci	Alleles per locus — Std.	
Arosa	24	14	6	42.86	1.57	0.85
Bernina	29	16	15	93.75	2.38	0.81
Bosco Gurin	29	16	11	68.75	1.88	0.81
Chandolin High	40	16	15	93.75	2.63	1.02
Chandolin Low	37	16	16	100.00	2.56	0.89
Chasseral	29	16	11	68.75	2.06	1.00
Chasseron	29	15	12	80.00	2.13	0.83
Gothard Pass	26	14	9	64.29	2.07	1.00
Grand St. Bernard	30	16	14	87.50	2.44	0.96
Grimsel Pass	26	14	12	85.71	2.43	1.02
Laax	26	13	10	76.92	2.08	0.86
San Bernardino	26	14	12	85.71	2.36	0.93
Simplon	40	16	14	87.50	2.63	0.96
Val di Campo 1	14	16	8	50.00	1.56	0.63
Val di Campo 2	26	13	10	76.92	2.00	0.82
Val Roseg	28	16	12	75.00	2.06	0.77
Zermatt	26	14	10	71.43	2.29	1.14
Average of populations	29	15	12	76.99	2.18	0.90

Table S4.

Expected and observed heterozygosity among Swiss Alps *B. lunaria* populations based on 11 loci.

Population	Nb. Ind.	H_E[1]	H_O[2]	Departure
Arosa	24	0.105	0.050	-0.055
Bernina	29	0.351	0.293	-0.058
Bosco Gurin	29	0.295	0.242	-0.054
Chandolin High	40	0.328	0.265	-0.063
Chandolin Low	37	0.331	0.254	-0.077
Chasseral	29	0.323	0.308	-0.016
Chasseron	29	0.273	0.273	0.000
Gothard Pass	26	0.212	0.173	-0.039
Grand St. Bernard	30	0.340	0.187	-0.153
Grimsel Pass	26	0.309	0.208	-0.101
Laax	26	0.253	0.227	-0.027
San Bernardino	26	0.257	0.150	-0.107
Simplon	40	0.346	0.314	-0.031
Val di Campo 1	14	0.265	0.161	-0.104
Val di Campo 2	26	0.175	0.140	-0.035
Val Roseg	28	0.291	0.254	-0.038
Zermatt	26	0.236	0.189	-0.047
Average of populations	29	0.276	0.217	-0.059

Note: [1]Nei's (1973) expected heterozygosity; [2]Mean observed heterozygosity.

Table S5.

Expected and observed heterozygosity of Swiss Alps *B. lunaria* populations at selected two-alleles loci[1].

Locus	Nb. Ind.	% observed	p % observed	q % expected	2pq % observed	2pq Departure from expected*	F_{IS}[2]
AAT-2	407	0.10	0.90	0.17	0.14	-0.03	0.150
AAT-3	330	0.18	0.82	0.28	0.24	-0.04	0.164
DIA-1	450	0.66	0.34	0.37	0.26	-0.11	0.265
IDH-1	301	0.71	0.29	0.40	0.31	-0.08	0.231
MDH-1	313	0.96	0.04	0.07	0.06	-0.01	0.141
MDH-2	431	0.51	0.49	0.48	0.40	-0.09	0.168
PGI-2	392	0.87	0.13	0.20	0.19	-0.02	0.220
PGM-1	332	0.96	0.04	0.07	0.05	-0.02	0.275
6-PGD	407	0.70	0.30	0.39	0.30	-0.09	0.195
SKDH	473	0.88	0.12	0.19	0.11	-0.08	0.360
TPI-1	288	0.66	0.34	0.42	0.36	-0.06	0.125
Average of loci	375	0.63	0.37	0.45	0.40	-0.05	0.208

Note: [1]Approximations from 2 alleles-system derived by deletions of alleles with low frequencies (<0.05) in Swiss Alps *B. lunaria*; [2]Inbreeding coefficient as $F_{IS} = 1 - (H_O / H_E)$; *Loci in Hardy-Weinberg equilibrium (*P*-value of 0.05).

Table S6.

Departure from expected and observed heterozygosity among Swiss Alps *B*. *lunaria* populations.

Population	AAT-2	AAT-3	DIA-1	IDH-1	MDH-1	MDH-2	PGI-2	PGM-1	6-PGD	SKDH	TPI-1
Arosa	—	—	0.00	—	—	—	0.00	—	—	-0.23	—
Bernina	-0.05	-0.05	-0.12	-0.03	0.00	-0.13	-0.02	-0.05	-0.06	0.01	-0.13
Bosco Gurin	0.01	—	0.00	-0.06	0.01	-0.05	-0.06	—	-0.20	0.01	-0.01
Chandolin High	0.01	-0.10	-0.06	-0.05	0.00	0.02	0.04	-0.05	-0.06	-0.04	-0.15
Chandolin Low	0.01	-0.14	-0.11	-0.12	-0.05	-0.15	-0.04	0.00	-0.07	-0.12	-0.09
Chasseral	-0.09	0.01	0.04	-0.14	—	0.29	-0.08	—	-0.02	-0.15	0.01
Chasseron	0.01	-0.05	-0.04	0.22	—	0.00	0.05	—	-0.02	0.00	0.09
Gothard Pass	—	—	-0.12	—	—	-0.10	-0.04	0.00	-0.05	0.00	—
Grand St. Bernard	-0.05	-0.06	-0.26	-0.27	—	-0.34	-0.06	0.00	-0.18	-0.25	-0.24
Grimsel Pass	-0.07	0.01	-0.18	-0.08	-0.07	-0.18	0.00	0.00	-0.16	-0.26	—
Laax	0.00	—	-0.11	—	0.00	-0.11	—	0.01	-0.20	0.00	—
San Bernardino	-0.07	0.01	-0.36	-0.31		-0.09	0.00	-0.08	-0.07	-0.17	—
Simplon	-0.11	-0.04	0.05	0.05	0.00	-0.01	-0.09	-0.05	0.03	0.01	-0.15
Val di Campo 1	0.02	-0.22	—	—	—	-0.21	—	—	—	—	0.08
Val di Campo 2	—	—	-0.30	—	0.01	-0.10	0.00	—	-0.01	0.00	—
Val Roseg	-0.05	-0.11	-0.11	-0.10	0.00	-0.10	0.01	0.00	—	0.01	0.08
Zermatt	0.01	0.06	-0.10	—	0.00	-0.11	—	0.00	-0.19	-0.10	—
Average of populations	-0.03	-0.04	-0.11	-0.08	-0.01	-0.09	-0.02	-0.01	-0.09	-0.08	-0.05

Table S7.

Matrix of Nei's (1978) unbiased genetic identity (above the diagonal) and genetic distance (below the diagonal) for the 17 Swiss Alps *B. lunaria* populations.

Pop.	ARO	BER	BGU	CHH	CHL	CHA	CHO	GOP	GSB	GRP	LAA	SBE	SIM	VDC	VDP	VRO	ZER
ARO	—	0.7842	0.7515	0.8141	0.7970	0.7733	0.7709	0.8990	0.8110	0.8313	0.7652	0.8141	0.8026	0.7999	0.8534	0.8111	0.8155
BER	0.2430	—	0.9584	0.9881	0.9919	0.9778	0.9627	0.8892	0.9709	0.9159	0.8802	0.9055	0.9896	0.9708	0.8568	0.9846	0.9140
BGU	0.2857	0.0425	—	0.9635	0.9706	0.9626	0.9362	0.8545	0.9615	0.9014	0.8710	0.8929	0.9712	0.9062	0.8209	0.9628	0.8977
CHH	0.2056	0.0120	0.0372	—	0.9929	0.9881	0.9632	0.8944	0.9841	0.9210	0.8841	0.9130	0.9958	0.9707	0.8805	0.9844	0.9146
CHL	0.2269	0.0082	0.0298	0.0072	—	0.9869	0.9569	0.8920	0.9734	0.9281	0.8919	0.9059	0.9950	0.9735	0.8666	0.9914	0.9268
CHA	0.2571	0.0224	0.0382	0.0120	0.0131	—	0.9391	0.8672	0.9695	0.9084	0.8617	0.8883	0.9921	0.9465	0.8388	0.9675	0.9129
CHO	0.2602	0.0380	0.0659	0.0375	0.0441	0.0628	—	0.8630	0.9512	0.8812	0.9069	0.8815	0.9616	0.9370	0.9016	0.9575	0.8690
GOP	0.1065	0.1174	0.1573	0.1116	0.1143	0.1425	0.1473	—	0.8667	0.9034	0.8638	0.8666	0.9006	0.8849	0.9290	0.8903	0.9076
GSB	0.2095	0.0296	0.0393	0.0161	0.0269	0.0310	0.0500	0.1430	—	0.9080	0.8743	0.9167	0.9759	0.9542	0.8638	0.9726	0.8907
GRP	0.1848	0.0878	0.1038	0.0823	0.0746	0.0961	0.1265	0.1016	0.0965	—	0.9551	0.9818	0.9199	0.9003	0.8735	0.9235	0.9933
LAA	0.2676	0.1276	0.1381	0.1232	0.1144	0.1489	0.0977	0.1464	0.1343	0.0459	—	0.9537	0.8823	0.8834	0.9021	0.9067	0.9433
SBE	0.2057	0.0993	0.1133	0.0910	0.0988	0.1184	0.1262	0.1431	0.0870	0.0184	0.0474	—	0.8989	0.8837	0.8739	0.9140	0.9598
SIM	0.2199	0.0105	0.0292	0.0042	0.0050	0.0079	0.0392	0.1047	0.0244	0.0835	0.1252	0.1066	—	0.9641	0.8675	0.9854	0.9187
VDC	0.2233	0.0296	0.0985	0.0298	0.0269	0.0550	0.0651	0.1223	0.0468	0.1050	0.1240	0.1236	0.0365	—	0.8851	0.9802	0.8935
VDP	0.1585	0.1545	0.1974	0.1272	0.1432	0.1757	0.1036	0.0736	0.1464	0.1353	0.1031	0.1348	0.1421	0.1220	—	0.8690	0.8576
VRO	0.2094	0.0155	0.0380	0.0158	0.0087	0.0330	0.0434	0.1162	0.0278	0.0795	0.0979	0.0899	0.0147	0.0200	0.1404	—	0.9177
ZER	0.2040	0.0899	0.1080	0.0893	0.0760	0.0912	0.1404	0.0969	0.1157	0.0068	0.0583	0.0410	0.0848	0.1126	0.1536	0.0859	—

Table S8.

Matrix of Nei's (1978) unbiased genetic identity (above the diagonal) and genetic distance (below the diagonal) for all taxa of the *B. lunaria* complex.

Population	*B. neolunaria*	*B. lunaria* var. *lunaria*	Swiss *B. lunaria*	*B. crenulatum*	*B. lunaria* var. *melzeri*	*B. nordicum*	*B. tunux*	*B.* "lanceolatum"
B. neolunaria	—	0.6943	0.7809	0.6846	0.6681	0.5027	0.5899	0.2056
B. lunaria var. *lunaria*	0.3649	—	0.9333	0.9124	0.9208	0.7048	0.8208	0.3857
Swiss *B. lunaria*	0.2473	0.069	—	0.9097	0.8656	0.6241	0.8657	0.3398
B. crenulatum	0.3789	0.0917	0.0946	—	0.8148	0.5978	0.773	0.3181
B. lunaria var. *melzeri*	0.4033	0.0825	0.1443	0.2048	—	0.6382	0.7413	0.3737
B. nordicum	0.6878	0.3498	0.4715	0.5145	0.4491	—	0.5632	0.412
B. tunux	0.5277	0.1974	0.1442	0.2575	0.2993	0.5741	—	0.4338
B. "lanceolatum"	1.5817	0.9527	1.0795	1.1453	0.9842	0.8868	0.8351	—

Note: *B.* "lanceolatum" includes the *B. lanceolatum* "green", *B. lanceolatum* "red", and *B. angustisegmentum*.

References and Notes

1. W. D. Hauk, C. R. Parks, M. W. Chase, Phylogenetic studies of Ophioglossaceae: Evidence from *rbcL* and *trnL-F* plastid DNA sequences and morphology. *Mol*. *Phylogenet*. *Evol*. **28**, 131–151 (2003).
2. PPG I, A community-derived classification for extant lycopods and ferns. *J*. *Syst*. *Evol*. **54**, 563–603 (2016).
3. B. Dauphin, D. R. Farrar, A. Maccagni, J. R. Grant, A worldwide molecular phylogeny provides new insight on cryptic diversity within the moonworts (*Botrychium* s.s., Ophioglossaceae). *Syst*. *Bot*. **42**, XX–XX (2017).
4. R. T. Clausen, A monograph of the Ophioglossaceae. *Mem*. *Torrey Bot*. *Club* **19**, 1–177 (1938).
5. M. C. Stensvold, A taxonomic and phylogeographic study of the *Botrychium lunaria* complex. Dissertation, Ames, Iowa State University (2008).
6. M. C. Stensvold, D. R. Farrar, Genetic diversity in the worldwide *Botrychium lunaria* (Ophioglossaceae) complex, with new species and new combinations. *Brittonia* **69**, 148–175 (2017).
7. M. C. Stensvold, C. L. Johnson-Groh, D. R. Farrar, Two new species of moonworts (*Botrychium* subg. *Botrychium*) in Alaska. *Am*. *Fern J*. **92**, 150–160 (2002).
8. A. Maccagni, C. Parisod, J. R. Grant, Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. *Alpine Bot*. **127**, 185–196 (2017).
9. R. Cronn, R. M. Brothers, K. Klier, P. K Bretting, J. F. Wendel, Allozyme variation in domesticated annual sunflower and its wild relatives. Theor. Appl. Genet. **95**, 532–545 (1997).
10. D. E. Soltis, C. H. Haufler, D. C. Darrow, G. J. Gastony, Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. *Am*. *Fern J*. **73**, 9–27 (1983).
11. D. L. Hartl, A. G. Clark, *Principles of population genetics* (Sinauer Associates, Sunderland, ed. 3, 1998), pp. 54–57.
12. F. C. Yeh, R-C. Yang, T. Boyle, Z-H. Ye, J. X. Mao, PopGene, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton (1997).
13. M. Nei, *The theory and estimation of genetic distance* (University Hawaii Press, Honolulu, 1973), pp. 45–54.
14. M. Nei, Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590 (1978).
15. H. Hotelling, Analysis of a complex of statistical variables into principal components. *J*. *Educ*. *Psychol*. **24**, 417–441 (1933).
16. T. Jombart, I. Ahmed, adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
17. R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, (2017).

18. T. Jombart, S. Devillard, F. Balloux, Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. **11**, 94 (2010), doi.org/10.1186/1471-2156-11-94.

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

20. D. Falush, M. Stephens, J. K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).

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

22. D. A. Earl, STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv*. *Genet*. *Resour*. **4**, 359–361 (2012).

23. M. Jakobsson, N. A. Rosenberg, CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806 (2007).

24. R. J. Dyer, J. D. Nason, Population graphs: the graph theoretic shape of genetic structure. *Mol*. *Ecol* **13**, 1713–1727 (2004).

25. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. (Springer, New York, ed. 2, 2016) Springer, pp. 4-5.

26. F. Rousset, Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**, 1219–1228 (1997).

27. J. Oksanen, R. Kindt, P. Legendre, B. O'Hara, M. H. Stevens, The vegan package. *Community Ecol*. *Package* **10**, 631–637 (2007).

Chapter 5

Recent Postglacial Recolonization from *In Situ* Survival and Intense Gene Flow within Natural Populations of the Alpine Fern *Botrychium lunaria* (Ophioglossaceae) revealed by ddRAD Sequencing

Benjamin Dauphin[1], Vinciane Mossion[1], Stefanie Wirth[2], Frédéric Sandoz[1] and Jason Grant[1]

[1]Laboratoire de Génétique évolutive, Université de Neuchâtel, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

[2]Centre d'hydrogéologie et géothermie, Université de Neuchâtel, rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

This chapter is in preparation for its submission in *Alpine Botany*.

Abstract

Quaternary climatic oscillations are postulated to have deeply affected population demography of species in the Alps. While plant refugia during the last glacial maximum (LGM) are well known to have existed on the periphery of the Alps, the occurrence of such glacial refugia free of ice in the Central Alps is still debated. An representative of the alpine flora, *B. lunaria* (Ophioglossaceae) has long been thought to only have the capability for gametophytic selfing that subsequently led to its highly homozygous populations. However, a recent study has shown that this species maintains sporophytic outcrossing as its predominant mode of reproduction in the Alps, which raises thought-provoking issues about the evolutionary consequences of such a shift of mating system in the context of the LGM. Therefore, the aim of this study was to explore the consequences of the LGM on diversity and genetic structure of the outcrossing *B. lunaria* populations. We performed a ddRAD sequencing to produce genome-wide data from 186 individuals belonging to four populations in the Val d'Hérens in the Swiss Central Alps. We used the dDocent bioinformatics pipeline to generate single nucleotide polymorphisms (SNPs) for inferring the levels of heterozygosity and the population differentiation based on multivariate and Bayesian clustering analyses. We found a weak degree of deviation from the Hardy-Weinberg equilibrium (departure = 0.014) and an average observed heterozygosity (H_O) of 0.246, which supports the predominance for sporophytic outcrossing and substantial gene flow within populations (G_{IS} = 0.059). We also found a similar genetic variation among populations, with a H_A range of 1.365–1.380, and we detected a weak genetic structure among populations that coincides with a pattern of isolation-by-distance (r = 0.019, P = 0.002). The lack of population differentiation associated with the finding of high conspecific genetic diversity suggests a recent postglacial recolonization from in situ survival that sheltered a polymorphic ancestral population on favorably south-exposed slopes along steep crests of the Val d'Hérens during the LGM. Our results provide new insights not only into the mating system in alpine populations of early-divergent vascular plants, but also about how the alpine landscape could have harbored a high allele richness for well-adapted species to very cold habitats during the LGM.

Keywords: Alps, ddRAD, GBS, glacial refugia, last glacial maximum, nunataks.

Introduction

The dramatically diverse reproductive modes in seed plants have played a key role in the transmission of their genetic diversity along evolutionary history (Hamrick et al. 1996; Barrett 2002). In early-divergent vascular plants, ferns have also experienced diverse and unique life cycles that have strongly sculpted genetic variability and population genetic structure among species (Haufler et al. 2016).

Homosporous ferns have three possible sexual mating systems (Sessa et al. 2016): sporophytic outcrossing, sporophytic selfing, and gametophytic selfing. The third mating system, which results from the fusion of genetically identical egg and sperm is exclusive to homosporous ferns and lycophytes having bisexual (hermaphroditic) gametophytes (Soltis and Soltis 1987). This gametophytic selfing is an extreme form of inbreeding that produces strict homozygous sporophytes in a single generation (Klekowski and Lloyd 1968), which rapidly reduces genetic variation within populations. Although inbreeding and genetic drift are often perceived as detrimental, this breeding system likely confers key evolutionary abilities for the dispersal and colonization of homosporous lineages (Ranker and Haufler 2008; Wubs et al. 2010).

The genus *Botrychium* Sw. (Ophioglossaceae) is a classic example in homosporous ferns that shows the predominance of gametophytic selfing for the majority of species (Hauk and Haufler 1999) and sporophytic outcrossing for a restricted number of species (*B. campestre* var. *lineare* (W. H. Wagner) D. R. Farrar, Farrar and Gilman 2017; *B. lunaria* var. *lunaria* (L.) Sw., Dauphin et al. in prep). Bisexual gametophytes of *Botrychium* are underground (Wagner et al. 1985), non photosynthetic, and endomycorrhized by fungi species of Glomeromycota (Winther and Friedman 2007), which not only provide minerals and water, but also are capable of providing carbohydrates and other organic products from nearby vascular plants to which they are also connected (Berch and Kendrick 1982). These first underground development stages may foster gametophytic selfing because of the low probability that two genetically distinct gametophytes are close enough for supporting a cross-fertilization via the migration of the flagelleted sperm (~ 2 cm) into sub-surface water flow (Peck et al. 1990). Several species (*B. campestre*, *B. pallidum*, and *B. pumicola*) also have vegetative reproduction via gemmae produced in leaf axils of their underground stems, which provides an alternative for overcoming the difficulties of cross-fertilization as well as the selfing mating systems (Farrar and Johnson-Groh, 1990; Camacho, 1996). These sexual and asexual mating systems, sometimes mixed within taxa, have deep implications on population structure and genetic variation among *Botrychium* that ultimately forge the reproductive barrier and lead to the speciation.

Botrychium lunaria has a broad circumtemperate and circumboreal distribution in northern hemisphere and is widespread in European mountainous regions, where it occurs in alpine meadows, recently deglaciated areas, and human-disturbed soils (e.g. typically along ski slopes) (Farrar 2011).

Previous genetic studies investigated genetic diversity in the Lunaria species complex (Stensvold and Farrar 2017) and their phylogenetic relationships based on a worldwide sampling that showed the monophyly of Lunaria and seven well supported sub-clades, including eleven taxa of which several are under investigation (Dauphin et al. 2017). This study highlighted the large haplotypes diversity found in the Lunaria clade that suggests an ancient demographic history in main European Alps during the last glacial and inter-glacial periods beyond the Quaternary.

Two scenarios were frequently discussed about the demography of alpine species during the last glacial maximum (LGM) (Taberlet and Cheddadi 2009); either the *tabula rasa* scenario (Chodat and Pampanini 1902) in which no population persisted in Central Alps and some survived on peripheral areas that were able to recolonize the alpine landscape afterwards, or the *nunataks* scenario (Brockmann-Jerosch and Brockmann-Jerosch 1926) where populations survived in restricted free of ice locations in Central Alps that were thereafter spread out in valleys during deglaciation. A phylogeographical study of *B*. *lunaria* has revealed two peripheral refugia in Alps that have harbored alpine populations during the LGM (Maccagni et al. 2017), which supports the second scenario as the most likely although the number of specimens analyzed was limited (2 or 3) for assessing genetic diversity metrics. Palynological data attest the presence of *B*. *lunaria* in Central Alps (Fully lake) during at least the Late Quaternary, with a demographic expansion 2000 years ago characterized on that highland plateau (Finsinger and Tinner 2007). Also, it was reported that several locations in Central Alps were free of ice at moderate elevations (1800–2600 meters) during the LGM (Kelly et al. 2004; Wirsig et al. 2016), especially on the South exposed slopes where the trimline (maximal altitude of the ice sheet) reached its lower elevations (2300 m at the Gotthard pass, Hippe et al. 2014; 2290 m at the Grimsel pass, Wirsig et al. 2016; 2500 m at the Simplon pass, Dielforder and Hetzel 2014). The non-eroded crests along the Val d'Hérens (south-west exposure) (Wildi et al. 2016) display an analogous configuration with a trimline of about 2300 meters at the lower valley (Mont Noble) and 2650 meters at the upper valley (Mont Miné), which is consistent with geomorphological data (Lambiel et al. 2016).

The emergence of next-generation sequencing (NGS) has opened a new era for population genetics studies of non-model species (Ekblom and Galindo 2011). Historically, biologists used co-dominant nuclear allozymes to characterize the genetic diversity and mating system of many vascular plants (Soltis and Soltis 1989), including homosporous ferns such as Ophioglossaceae; *Botrychium* (Farrar 1998; Hauk and Haufler 1999; Stenvold and Farrar 2017), *Botrypus* (Soltis and Soltis 1986), *Ophioglossum* (Chung et al. 2012), and *Sceptridium* (Watano and Sahashi 1992; McCauley et al. 1985). These allozyme studies provided the first insights on gene flow among and within fern populations, but they were limited to a restricted number of markers (up to 11 enzyme systems) and constrained to work on fresh material with good enzyme activities. Recently, amplified fragment-length polymorphisms (AFLPs) were applied for investigating population genetic structure of

Botrychium and the genetic distances between putative new taxa (Williams and Waller 2012; Williams et al. 2016). However, by nature, AFLPs suffer, as any dominant maker, from the lack of information on levels of heterozygosity and are thus inappropriate to infer the mating system. Looking for more co-dominant nuclear loci and working on organisms with very limited genomic resources, NGS techniques were used with success to rapidly generate large-scale sequencing data from non-model plant species (Egan et al. 2012). The development of new protocols for single-nucleotide polymorphisms (SNPs) discovery allows the genotyping of a large number of individuals and populations (Peterson et al. 2012), and are therefore relevant to examine how geography and heterogeneous alpine habitats contribute to genetic structure and gene flow on a fine-scale.

In this study, we addressed key evolutionary issues for population genetics of early divergent ferns using a genome-wide sequencing for SNP genotyping (ddRAD seq). We used the broadly distributed fern *Botrychium lunaria* as a model to characterize the predominant mating system in an fern species in the Central Alps. In context of quaternary climatic oscillations, we investigated the strength of gene flow maintained among populations and the impact of the last glacial maximum on the conspecific genetic diversity at local scale.

Materials and methods

Study area and sampling design

The study was conducted during the summer of 2015 at two localities (46°12’ N; 7°29’ W and 46°05’ N; 7°32’ W) in the Val d’Hérens valley, Central Swiss Alps (Fig. 1a), with two sites sampled for each location (IT1, IT2, IIT1, and IIT2; Figs. 1b–c). The first locality is near to the entrance of the Val d’Hérens valley, connected to the main Rhone valley, whereas the second is at the end of the valley near to current glaciers. We selected both localities to be at similar elevations (2395.74 m) over an altitudinal range from 2344.34 to 2437.41 m (Table 1), with alpine meadows dominated by *Nardion strictae* (IT1 and IT2), *Caricion curvulae* (IIT1), and *Festucion variae* (IIT2) *alliance* (Braun-Blanquet 1964). The four sampling sites are hereafter called populations in genetic analyses.

We sampled dense *Botrychium lunaria* populations along altitudinal transects (steeper slopes) with two technical replicates (distant of 10 meters) per site (IT1, IT2, and IIT2), and a control site (IIT1) with a null slope, for a total of eight transects (Figs. 1b–c). We performed a modified random clustered sampling according to population density (Storfer et al. 2007) and we collected five plots (~ 2 × 2 m) per transect, distant about 8 meters, which consisted of five individuals per plot. In total, we sampled 200 specimens that we scanned (IRIS scan book 3, Louvain-La-Neuve, Belgium) on the field, dried half-plant in silica gel, and stored the remaining leaf material in liquid nitrogen for future works. We georeferenced all specimens with a differential GPS (Leica GS15, Leica Geosystems AG,

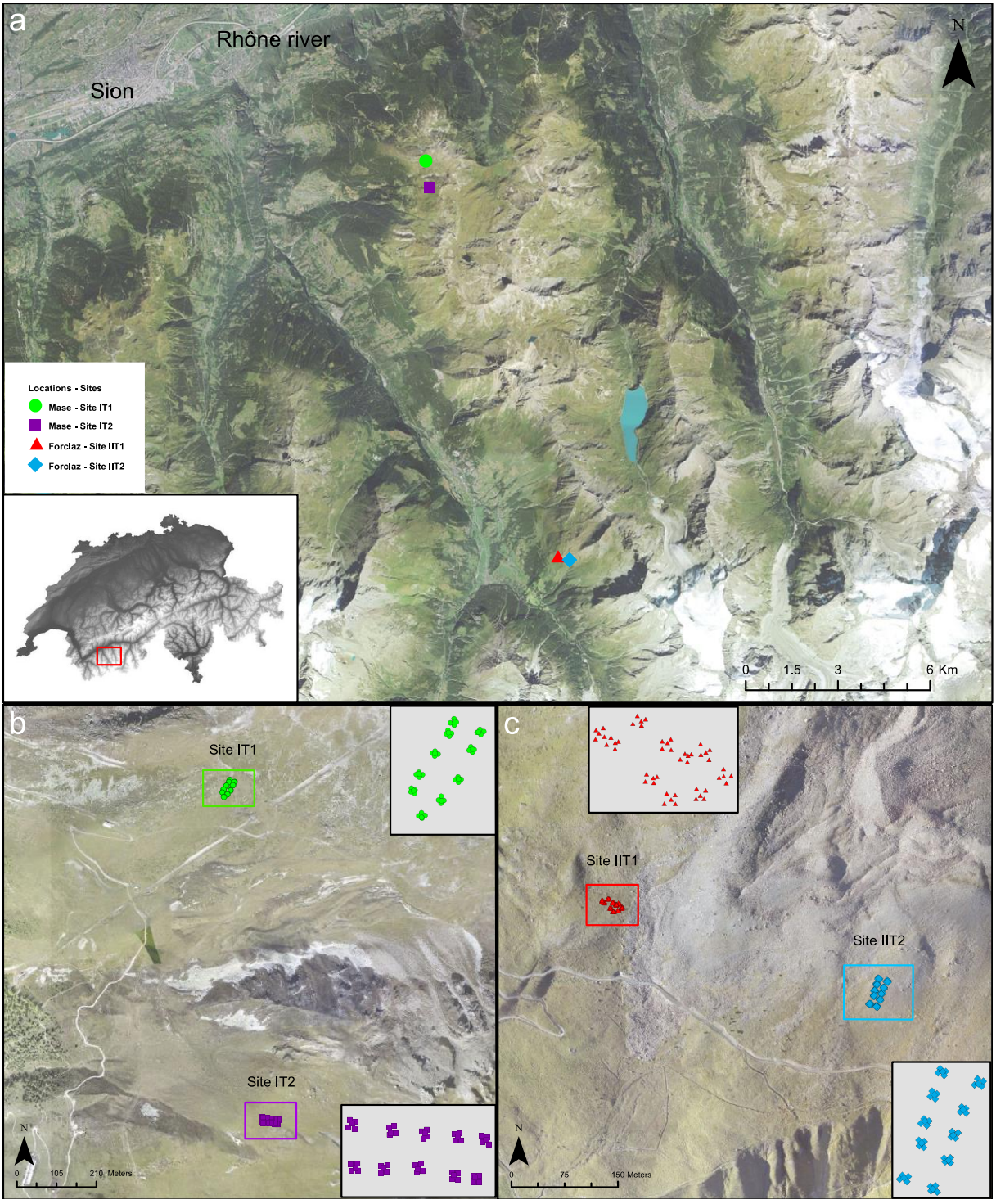


Fig. 1 Location of the *Botrychium lunaria* populations sampled in the Central Alps of Switzerland (Val d'Hérens), with a zoom on the design sampling performed at the Mase (a) and Forclaz (b) localities. Background represents the Swiss map at 25 meters of spatial resolution (www.geo.admin.ch).

Heerbrugg, Switzerland) that included a suitable horizontal and vertical accuracy of 2 cm for investigating population genetics at high spatial resolution (0.5 meters).

DNA extraction, library preparation and sequencing

Genomic DNA was extracted from fresh silica-dried material using the DNeasy Plant Mini Kit and according to the manufacture's protocol (Qiagen, Hilden, Düsseldorf, Germany). Quantity of DNAs was assessed by fluorescence (Qubit Fluorometric Quantification, Thermo Scientific, Wilmington, USA) and DNAs integrity was inspected on 1% agarose gel to ensure no degradation (Graham et al. 2015). Two double-digest restriction-associated DNA (ddRAD) libraries were prepared following Truong et al. 's protocol (2012). Briefly, 250 ng of genomic DNA was digested at 37 °C for 1 h using the restriction enzymes *Pst*I and *Mse*I (NEB), followed by ligation of the universal P7 *Mse*I adapter and tagged *Pst*I P5 adapters at 37 °C for 3 h. Samples were individually tagged with at least two different bases. PCR reaction and cycle was performed following Truong et al. (2012) with a selective amplification step using specific PCR primers containing the two additional bases GC at the 3' extremity to reduce the number of amplified fragments. Primer dimers were removed on 1.5% agarose gel using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Düsseldorf, Germany). DNA libraries were quantified with the Qubit dsDNA High Sensitivity Assay kit using the Qubit 2.0 fluorometer (ThermoFisher Scientific, Wyman, Maine, US). Subsequently, each the two pools were combined in equimolar concentrations to get two genomic libraries and purified with the Qiagen MinElute PCR Purification kit (Qiagen, Hilden, Düsseldorf, Germany). Sequencing was performed on HiSeq 2500 Illumina sequencer (v4 chemistry; San Diego, California, US) using two lanes in single-end (1 x 100 bp; Floragenex, Eugene, Oregon, US). Yeast (*Saccharomyces bayanus*) was used as a positive control.

Bioinformatics

Quality of raw reads was checked with phred scores per base using the FastQC program (Andrew 2010) and raw data files were deposited to the NCBI Sequence Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/sra/docs/submit/). We used the *dDocent* v.1.2 pipeline (Puritz et al. 2014) to identify allele variants based on six main steps and combining several bioinformatics programs. In summary, we first demultiplexed raw data using the *process_radtags* function of Stacks v.1.4 (Catchen et al. 2013), with the "rescue barcodes" option to recover reads with up to two ambiguities in the individual barcode. Then, we trimmed the demultiplexed reads using *Trimmomatic* v.0.35 (Bolger et al. 2014) to remove sequences with low quality bases (phred < 20) and a length below 36 base pairs (bp). The three other parameters "LEADING" (3), "TRAILING" (3), and "SLIDINGWINDOW" (4:15) were kept by default. Working on single-end RAD data, we performed the de novo assembly with *cd-hit-est* v.4.6.4 (Li and Godzik 2006) to cluster and assemble reads into unique reference contigs. We retained a sequence similarity threshold of 92% (value reached to

plateau) and we set a minimum coverage of 5 to retain unique sequences in individuals, which were represented in at least 10% of the total number of individual. Subsequently, we aligned the trimmed reads to the reference contigs using the MEM algorithm (Li 2013) implemented in *bwa* v.0.7.12 (Li and Durbin 2009) with default parameters; match score -A=1, mismatch score -B=4, and gap-opening penalty -O=6. After reads mapping, we identified variants sites (SNPs and indels) using *freebayes* v.1.0.2 (Garrison and Marth 2012) with default parameters; minimum mapping quality -m=5, maximum complex gap -E=3, and minimum base quality -q=5. The last main step of the pipeline consisted to SNP filtering from the resulting variant call file (vcf), which itself comprised twelve sub-steps using *VCFtools* v.0.1.15 (Danecek et al. 2011) and *vcflib* v.1.0.0 (Garrison and Marth 2012). Details about the variant filtering process are presented and the number of variant sites for each step is presented in Table S1. The final filtered vcf was converted to other program-specific input files using PGDSpider v.2.1.1 (Lischer and Excoffier 2012).

Detecting outlier and neutral loci

To detect loci not conforming to neutral expectations among populations, we carried out a genome scan based on the Bayesian method implemented in BAYESCAN v.2.1 (Foll and Gaggiotti 2008). We ran BAYESCAN with a burn-in of 50 000 iterations and 20 pilot runs of 5000 iterations each for a total of 100 000 iterations, with the thinning interval kept to 10. We set the prior odds to 100 to retain enough statistical power to detect loci under selection and to limit false positives. Thus, we characterized a locus as under selection when its $\log_{10}$ posterior odds > 1, assuming that natural selection is 10 times more likely than neutral differentiation at a locus (Foll and Gaggiotti 2008; Nielsen et al. 2009).

Then, we tested the neutral SNPs for departure from Hardy-Weinberg equilibrium (HWE) in GenoDive v.2.0b27 (Meirmans and Van Tienderen 2004), using the least squares method (Excoffier 1992; Michalakis and Excoffier 1996) with 10 000 random permutations. We removed loci that had a significant departure from HWE among populations and considered both datasets all neutral SNPs and the neutral SNPs in HWE for further genetic diversity analyses.

Genetic diversity

To investigate the mating system maintained within populations, we calculated for both datasets the expected (H_E) and observed heterozygosity (H_O), the average number of alleles (H_A) per locus, and the inbreeding coefficient (G_{IS}) within populations using GenoDive v.2.0b27 (Meirmans and Van Tienderen 2004). We estimated departure from HWE to describe the prevalence for random mating based on allele frequencies.

Population genetic structure

We characterized the genetic structure among populations using both datasets with three frequency- or genotype-based methods; pairwise differentiation, multivariate analyses without a priori

assumptions, and Bayesian clustering based on priors. These last two approaches are helpful to elucidate the spatial pattern of admixture populations and to detect hybrids among populations.

First, we estimated pairwise F_{ST} (Weir and Cockerham 1984) among populations with a significance calculated from 10 000 permutations using GenoDive v.2.0b27 (Meirmans and Van Tienderen 2004).

To explore the genotypic variation across all individuals, we performed a principal component analysis (PCA) (Hotelling 1933) based on a covariance matrix to summarize the full variance along the two first axes using the *dudi.pca* function from the adegenet v.2.0.1 package (Jombart and Ahmed 2011) implemented in R (R Core Team 2017). Then, we carried out a discriminant analysis of principal components (DAPC) (Jombart et al. 2010) to detect differentiation among populations, which minimizes the variance within populations and maximizes it between populations.

Besides, we inferred population genetic structure using Bayesian clustering analysis implemented in STRUCTURE v.2.3.4 (Pritchard et al. 2000). We set the analysis with the putative populations range of K = 1–12, under the admixture model with uncorrelated allele frequencies (Falush et al. 2003), and without a priori information of the location origin of each sample. We independently ran four times the analysis with 10 replicates for each value of K, a burn-in of 100 000 MCMC iterations, and followed by 400 000 iterations. We inspected the convergence of the replicates through runs and diagnosed the optimal K value based on the delta K method (Evanno et al. 2005) implemented in Structure Harvester v.0.6.94 (Earl and von Holdt 2012). Then, we averaged replicates for each value of K using CLUMPP v.1.1.2 (Jakobsson and Rosenberg 2007), which minimizes the variance across all iterations and between runs of that K value.

We investigated isolation by distance (IBD) within and among populations based on a Mantel test (Mantel 1967) between the genetic distance and the Euclidean geographical distance. We tested correlations based on 10 000 permutations using the *mantel.randtest* function from the ade4 v.1.7 package (Dray and Dufour 2007) implemented in R (R Core Team 2017).

Table 1 Information about localities and summary statistics for *Botrychium lunaria* populations based on 513 neutral SNPs and 348 neutral SNPs in Hardy Weinberg Equilibrium

						H_A		H_E		H_O		*Departure*		G_{IS}	
Location	Site	Latitude	Longitude	Altitude	n	Neutral	Neutral in HWE	Neutral	Neutral in HWE	Neutral	Neutral in HWE	Neutral	Neutral in HWE	Neutral	Neutral in HWE
Mase	IT1	46.20432	7.48350	2406.62	47	1.380	1.345	0.252	0.236	0.237	0.228	-0.015	-0.008	0.059	0.031
	IT2	46.19642	7.48502	2424.07	45	1.369	1.332	0.246	0.227	0.231	0.221	-0.015	-0.006	0.061	0.024
Forclaz	IIT1	46.08851	7.53906	2346.15	47	1.365	1.332	0.244	0.228	0.230	0.220	-0.014	-0.008	0.056	0.033
	IIT2	46.08741	7.54379	2406.10	47	1.366	1.328	0.244	0.225	0.230	0.219	-0.014	-0.006	0.058	0.024
Overall populations					186	1.366	1.330	0.232	0.229	0.246	0.222	0.014	-0.007	0.059	0.028

Note: n, number of individuals; H_A, average number of alleles; H_E, expected heterozygosity; H_O, observed heterozygosity; G_{IS}, inbreeding coefficient.

Results

De novo assembly, SNPs calling, and variants filtering

In total, we obtained 495 192 820 reads of 95 bp each for 190 individuals from the four sites sampled in the Val d'Hérens valley. We demultiplexed 91.5% of reads with success, recovering on average 2 385 840 reads per individual, and we discarded 2.69% of reads with the quality trimming step for a total of 11 881 042 reads dropped. Following de novo assembly, we identified 33 288 unique reference contigs on which we mapped with success 85.7% of all quality-trimmed reads. At that step, we removed 4 samples due to its high missing data (>50%). Then, we recovered a total of 96 029 raw SNPs and we identified 3 066 of them as biallelic only and genotyped in 95% of individuals (Table S1). After our additional filtering steps with stringent criteria to eliminate suspicious loci, we finally retained 343 contigs in a single vcf for a total of 513 allele variants through the 186 individuals.

Population genetic statistics

We identified no significant loci under selection and we characterized 165 loci to depart from HWE among the four populations. These unbalanced loci were removed from the neutral SNPs data set and we considered both all neutral SNPs (513) and the neutral SNPs in HWE (348) for further genetic diversity analyses.

We assessed genetic diversity within and among populations using genotype data and we found a an important variability with an overall average H_A of 1.366 (1.365–1.380) and 1.330 (1.328–1.345), an average H_E of 0.232 (0.244–0.252) and 0.229 (0.225–0.236), and an average H_O of 0.246 (0.230–0.237) and 0.222 (0.219–0.228) for the neutral SNPs and the neutral SNPs in HWE datasets, respectively (Table 1). We found a weak degree of deviation from HWE with low values for the overall average departure (0.014 and –0.007) and the inbreeding coefficient G_{IS} (0.059 and 0.028) (Table 1). Overall, our results of genetic diversity metrics were similar between both datasets (513 SNPs versus 348 SNPs, see above), with slightly lower values for the neutral SNPs in HWE (Table 1).

Table 2 Pairwise F_{ST} values (above diagonal) and *p-values* (below diagonal) for *Botrychium lunaria* populations in Val d'Hérens based on 513 neutral SNPs

		Mase		Forclaz	
Location	Site	IT1	IT2	IIT1	IIT2
Mase	IT1	–	**0.002**	**0.002**	**0.004**
	IT2	0.033	–	**0.003**	**0.003**
Forclaz	IIT1	0.021	0.003	–	**0.005**
	IIT2	< 0.001	0.004	< 0.001	–

Note: values in bold were significant at 0.05 based on 10 000 permutations.

Population genetic structure

Based on the neutral 513 SNPs dataset, pairwise F_{ST} comparisons supported weak differentiation among populations, ranging from 0.002 to 0.005, and all pairwise F_{ST} estimates were significant ($P < 0.05$) (Table 2). We found a pattern of isolation by distance (IBD) among all populations with a very low positive correlation (r = 0.019, $P = 0.002$) and among populations of the Forclaz location (r = 0.061, $P < 0.001$). However, we detected no significant pattern of IBD within populations based on these Mantel tests (Table 3).

The first axis of the principal components (PC) summarized only 1.80 % of the genotypic variation among individuals and the sum of the three first axes was low (5.17%). Here, populations appeared largely undifferentiated with a comparable genotypic diversity, which was characterized by a complete overlap among centroids and between inertia ellipses of populations (Fig. 2a).

To detect differentiation among populations using the DAPC, we retained the first 93 principal components of the PC, which included 80.1% of the variance. We saved three discriminant functions (DF) and we plotted the two first DF1 and DF2 (Fig. 2b) that represented 76.19% of the variation in the discriminant scores. The first discriminant function (DF1) separated populations within locations although a partial overlap persisted for IT1 and IT2, and the second discriminant function (DF2) segregated well the two locations (IT1-IT2 and IIT1-IIT2).

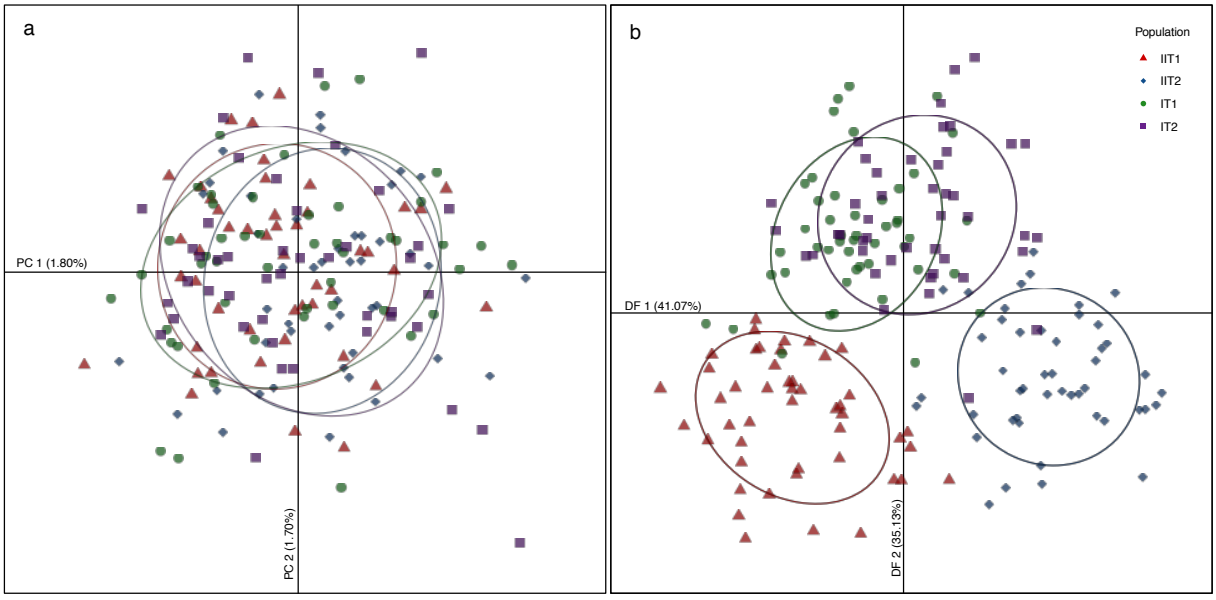


Fig. 2 Principal components analysis (PCA) (a) and discriminant analysis of principal components (DAPC) (b) showing genotypic variation between individuals and populations. Plots depict the two first principal components (PC) for the PCA and the two first discriminant functions (DF) for the DAPC. Each point represents one *Botrychium lunaria* sample.

Bayesian clustering analyses were congruent between replicates and we found the optimal $K = 5$ clusters as the most likely partitioning of individuals (Fig. 3). These distinct groups did not match to

the sampled sites that appeared to be largely admixture. Also, individuals within populations were not assignable to one of the five clusters when $K = 5$, and there was no better membership for other values of K.

Table 3 Isolation by distance within and among *Botrychium lunaria* populations in Val d'Hérens based on Mantel tests

Location	Site	r	*P*
Mase	IT1	-0.095	0.981
	IT2	-0.006	0.545
Forclaz	IIT1	-0.015	0.608
	IIT2	0.043	0.208
Among sites in Mase		0.019	0.051
Among sites in Forclaz		0.061	**< 0.001**
Among all sites		0.019	**0.002**

Note: r, regression coefficient; p-values were calculated with 10000 permutations.

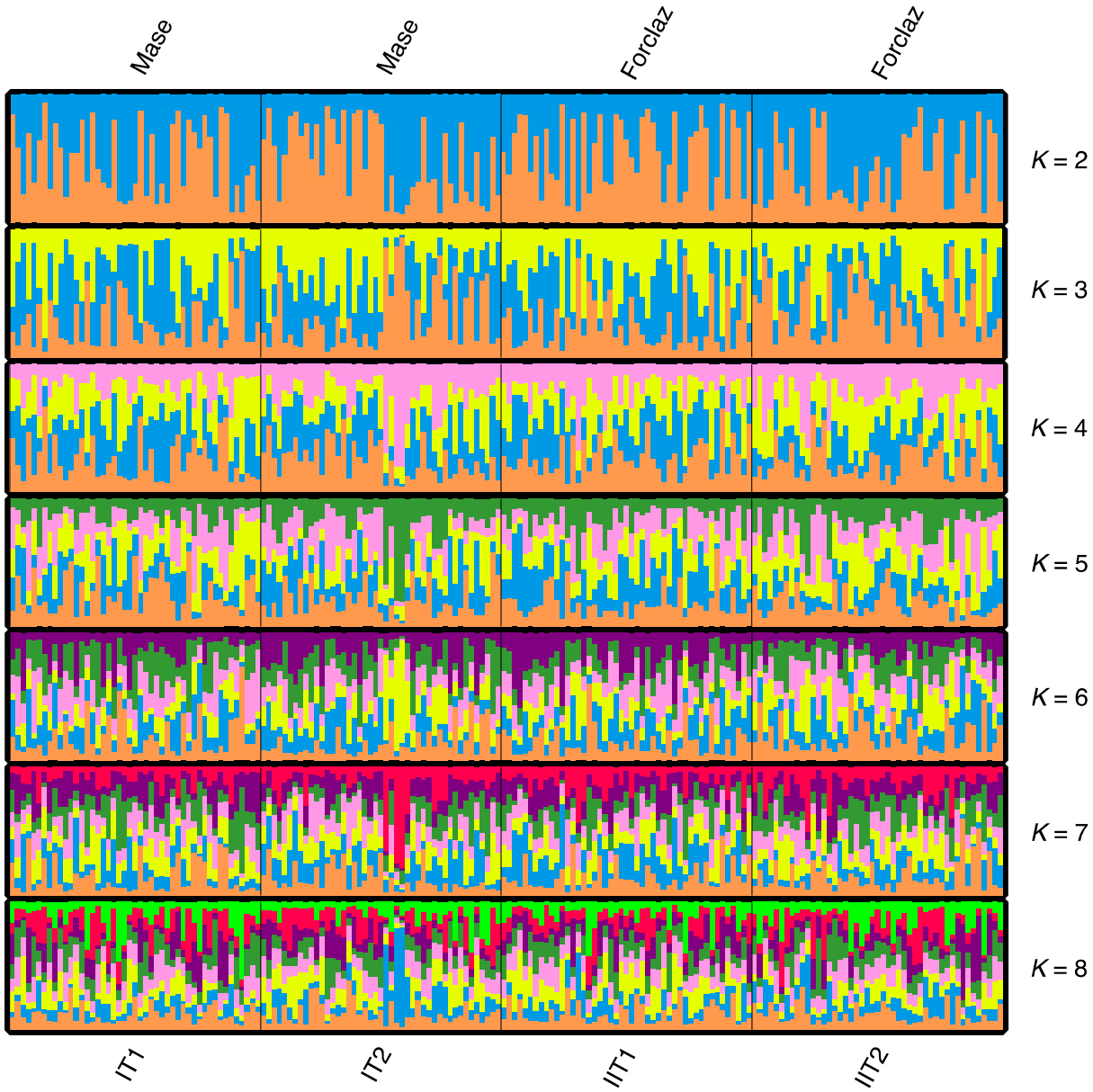


Fig. 3 Admixture bar plots of posterior probability of individual assignment for different values of *K* (2–6) based on 513 neutral SNPs, and using the program STRUCTURE. Each bar corresponds to an individual of *Botrychium lunaria*, and colors in each bar represent the admixture proportion of each cluster for each *K* value. Names above and codes below the plot refer to location and sampling site respectively.

Discussion

Genetic diversity and mating system

Our analysis of genomic data from ddRAD sequencing provides new insights on *Botrychium* population genetics. In the context of Ophioglossaceae (Ranker and Haufler 2008), we found high levels of genetic variation (H_E = 0.232) and heterozygosity (H_O = 0.246), and a moderate average number of alleles per locus (H_A = 1.366 on average) within the four *B*. *lunaria* populations analyzed in Central European Alps (Val d'Hérens, Switzerland) (Table 1). Surprisingly, those populations display very low estimates for the inbreeding coefficient (G_{IS} = 0.059) and a weak degree of deviation from Hardy-Weinberg equilibrium (Departure = 0.014), which is remarkable knowing the predominance of gametophytic selfing in *Botrychium* and more generally in Ophioglossaceae (Hauk and Haufler 1999; Soltis and Soltis 1989). For the first time in the genus, the outcrossing breeding system was reported in *B*. *campestre* var. *lineare* (Farrar and Gilman 2017) that showed up to 100% of heterozygotes in populations, which is comparable to what we have observed in the *B*. *lunaria* populations of Val d'Hérens. Our finding is the second report of sporophytic outcrossing in *B*. *lunaria* species (Dauphin et al. in prep) that was long thought as a species constituted exclusively of homozygous populations (Stenvold and Farrar 2017). It is all the more striking that these levels of heterozygosity estimated within populations can be rapidly and drastically reduced since the gametophytic selfing of homosporous ferns leads to a pure homozygous sporophyte in a single generation (Klekowski and Lloyd 1968). Hence, sporophytic outcrossing largely participates to the maintain of genetic diversity via recombination and potentially to adaptive selections taking place in heterogeneous alpine habitats (REF).

Reproductive biology and evolutionary success of Botrychium

Results of this study have broad implications in our understanding of the reproductive biology of *Botrychium* and more broadly for the life cycle of early divergent vascular plants. Circumventing the difficulties of cross-fertilization between bi-sexual underground gametophytes, environmental factors of the alpine landscape, such as the slope degree and soil porosity, may play an important role in sperm movement via the sub-surface water flow in outcrossing *Botrychium* species. The high density of mature gametophytes is intuitively a critical feature for increasing the probability of the encounter of gametes between genetically distinct individuals, in which the flagelleted sperm has a limited distance of dissemination (~ 2 cm) (Peck et al. 1990). The number of emergent sporophyte provides a relevant proxy to approximate the effective population size, with about 10-fold more mature gametophytes than above ground sporophytes. At a finer scale in our sampling, we can roughly quantify a population size of about 1000 individuals (100 emergent sporophytes plus 900 underground gametophytes) per plot of 50 × 10 m (Fig. 1a-b) and a mean density of two individuals per meter square. It gives some insights about the underground network evolving into natural populations

although further studies are needed to characterize the exact density threshold at which their breeding system shift from sporophytic outcrossing to gametophytic selfing.

Botrychium is a young genus (8.51 myr) belonging to the ancient Ophioglossaceae lineage (Rothfels et al. 2015; Dauphin et al. in review) that shows a large versatility in its reproduction modes among closely related species (Farrar and Johnson-Groh, 1990; Farrar 1998; Hauk and Haufler 1999). This genus is capable of maintaining the three sexual mating systems existing of homosporous ferns (Sessa et al. 2016) and *B. campestre* is a good example of the alternation of mating system that leads to taxa differentiation (Farrar and Gilman 2017). However, it is unclear what is the evolutionary advantage of the sporophytic outcrossing in *B. lunaria* populations of Alps while all other worldwide populations of the Lunaria complex display the gametophytic selfing as the rule? As reported from laboratory experiments by Sessa et al. (2016), the majority of fern species undergo both gametophytic selfing and sporophytic outcrossing, and few of them are confined to gametophytic selfing only. This flexibility of mating systems is postulated to have strongly contributed to fern evolution. While gametophytic selfing confers a key evolutionary innovation for the establishment of migrants in long-range dispersal via a single spore (de Groot et al. 2012), sporophytic outcrossing may lead to heterozygote superiority as well as higher genetic load due to the retention of deleterious alleles inherited (Hedrick 1987). In Alps, plant habitats are known to be heterogeneous (e.g. mosaic) and highly dynamic, with rapid shifts of vegetation due to natural and anthropic factors (Gehrig-Fasel et al. 2007). Several studies have highlighted the importance of micro-environments on genetic diversity and how the alpine landscape contributes to local adaptation of species (Holderegger et al. 2010; Manel et al. 2010). Also, heterozygous individuals may have a broader phenotypic plasticity and therefore a higher survival rate during the ice ages, which could have led to a strong positive selection on outcrossing populations along the Quaternary climatic oscillations.

Genetic structure and demographic history

Pairwise F_{ST} supported a weak but significant differentiation among the four populations (Table 2) and the principal component analysis (PCA) also showed such genetic similarities with a comparable genotypic diversity between populations (Fig. 2a). Discriminant analysis of principal components (DAPC) segregated the groups of individuals among locations and sites, with an exception for the *Mase* locality where the centroids and inertia ellipses of both sites (IT1 and IT2) were poorly discriminated (Fig. 2b). The pattern of isolation by distance (IBD) coincided with results of pairwise F_{ST} and multivariate analyses, which supports a weak genetic isolation between both locations, between both sites of the *La Forclaz* locality, but not between both sites of the *Mase* locality (Table 3) that likely form a continuous population arranged in patches (Fig. 1a). Accordingly, geographical features plays a role in isolating barrier although Bayesian clustering analyses highlighted the substantial genetic admixture within genotypes of individuals and individuals between populations (Fig. 3).

Two demographic scenarios deserve consideration regarding the weak genetic differentiation occurring among populations. Is there an intense bi-directional migration of individuals between both localities and sites or a recent population split from an ancestral variable population that has provided too little time for differentiation? The localities are thirteen kilometers apart without major high summits along the Val d'Hérens crest that could strongly hinder wind flow in northwest and southeast directions (Fig. 1a). The dispersal capability of ferns by airborne spores is well known (Wagner 1972; Crist and Farrar 1983) and several studies focusing on *Botrychium* have presented evidence of inter-continental migrations of species (Dauphin et al. 2017; Mezza-Torres et al. 2017; Stensvold and Farrar in review). However, it was shown that the local dispersal within populations of *Botrypus virginianus*, closely phylogenetically and morphologically related to *Botrychium* species, seems to be limited, with a large majority of spore lodgements noticed in the three first meters around the surveyed plants (Peck et al. 1990). Together, these studies support the episodic long dispersal events but bring a further nuance on the limited species' ability to maintain intense migration between well established populations distant of several dozens kilometers. The second and most plausible scenario, is that these *B. lunaria* populations of the Val d'Hérens valley were recently derived from an ancestral variable population, closely located to the current population occurrences, in which the sporophytic outcrossing has mitigated random genetic drift and therefore maintained the ancestral polymorphism within populations. Consistent with the genetic variation found (Table 2) and the weak IBD between populations (Table 3), differentiation among populations appears recent, and historical and geographical factors may have significantly contributed in the dynamic of contemporary populations as well as the ongoing genetic isolation in the context of the Alps postglacial recolonization (McMaster 1994; Stensvold 2007; Chung et al. 2012).

Glacial refugia in central European Alps and postglacial recolonization

Quaternary climate oscillations have deeply affected the demographic history of mountain plants (Hewitt 2004; Schönswetter et al. 2005), and unglaciated peripheral areas of the European Alps have harbored a large portion of their conspecific genetic diversity during the last glacial maximum (LGM) (Petit et al. 2003). A phylogeographical study of *B. lunaria* showed such a pattern, with at least two peripheral refugia located in southern Alps, but also it suggested two putative central refugia (Zermatt and Arosa) that could have sheltered this species during the LGM (Stehlik 2000; Maccagni et al. 2017). Adjacent to Zermatt in the southwestern part of Alps, crests of the Val d'Hérens valley were free of ice during the LGM, with a trimline of about 2300 meters at the lower valley (Mont Noble) and 2650 meters at the upper valley (Mont Miné) (Wildi et al. 2016; Lambiel et al. 2016). Thus, the *Mase* locality was slightly above the maximum elevation of the ice surfaces during the LGM while the *La Forclaz* locality was well below the trimline (Fig. 4). Several zones favorably exposed to the south insolation along the steep crests (between 2300 and 3000 meters) may have hosted *B. lunaria* by providing an habitat free of permanent snow and a suitable soil for the development of its vital

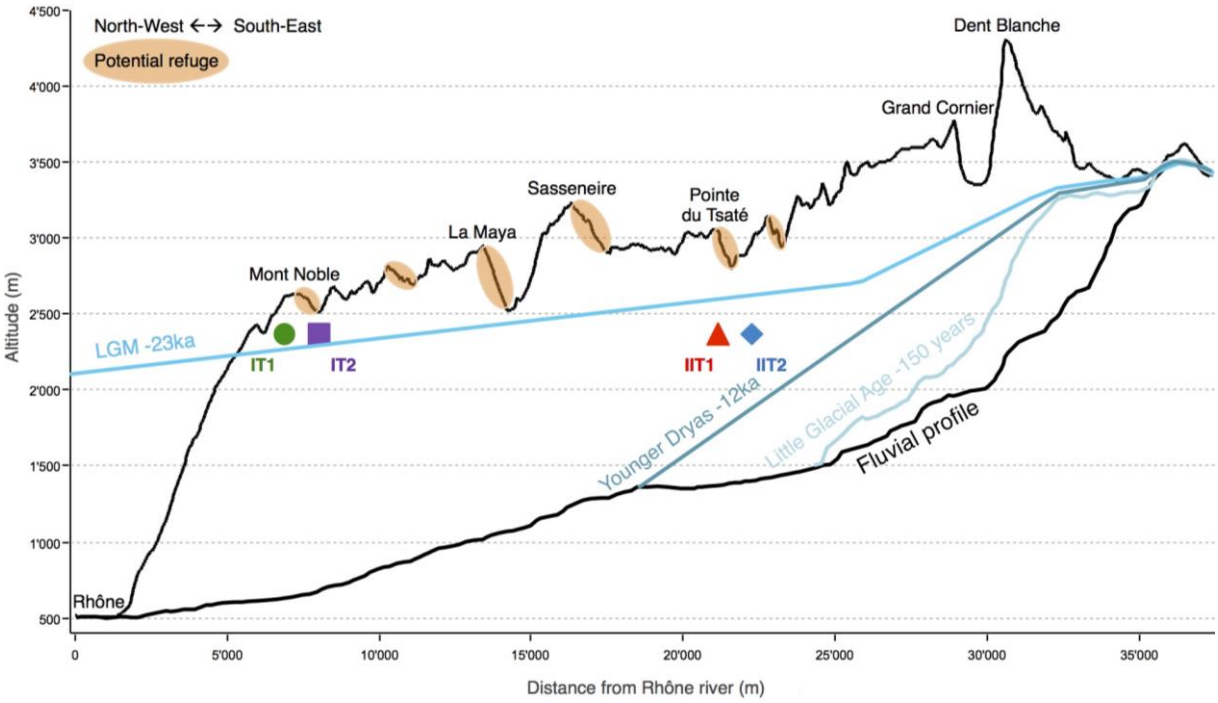


Fig. 4 Topographic profiles through the highest crests and the river bed in Val d'Hérens representing the *Botrychium lunaria* populations with the trimlines during the last glacial maximum (-23,000 years; Ehlers et al. 2011), the Younger Dryas (-12,000 years; Coutterand et al. in prep), and the little glacial age (1850; http://glaciology.ethz.ch/messnetz/?locale=fr). South exposed potential refugia are indicated by orange ovale.

arbuscular mycorrhizal fungi, at least temporarily during the growing season (from June to August) (Fig. 4). As reported for flowering plants sharing a similar habitat at high elevations (Stehlik 2000; Bettin et al. 2007; Parisod and Besnard 2007), central glacial refugia have likely sheltered *B. lunaria* and allowed in situ survival, from which the ancestral variable populations have operated a recent postglacial recolonization in its vicinity that has led to the current weak population differentiation. The lack of geographical structure in our genomic-based results supports a period of intense preglacial gene flow that has homogenized allele frequencies with the predominant outcrossing mating system.

Biological features of B. lunaria for surviving during the LGM

Climatic conditions were particularly cold during the LGM and have enforced drastic changes in the floristic composition of alpine meadows. The survival of *B. lunaria* in such central refugia implies key physiological and life cycle characteristics to maintain viable populations in favorable steep south slopes during several hundred or thousand years. As a first phenological feature, *B. lunaria* is endomycorrhized by fungi species, which provide carbohydrate nutrition as well as water and minerals, and therefore allow the maintenance of populations on thin, rocky, and poor soils in organic matter. When the season is unfavorable, for example under unusual dry conditions (Johnson-Groh 1998), the hydric stress may induce the entry into dormancy of the sporophyte, in which the annual leaf remains belowground at the bud stage. Depending on the progress of the growing season and the

habitat, species of Botrychium have an important variation of the emergence period of sporophytes (Johnson-Groh and Lee 2002), up to seven weeks of fluctuation for *B*. *mormo*. As another phenological characteristic participating to the population stability, the generation time of *B*. *lunaria* is of several decades, including about ten years for the development of the underground gametophytes and a longevity of two or three decades for the sporophytes assessed in Swiss Alps (determined from leaf scars along the underground stem). Based on these biological features, the fern species *B*. *lunaria* may have had a chance of survival in central glacial refugia during the LGM and older and colder periods in Pleistocene.

Compliance with ethical standards

Funding: JG has received a research grant from the Swiss National Science Foundation 31003A_156456.

Conflict of interest: The authors declare that they have no conflict of interest.

Declaration of authorship: BD and JG conceived and designed the research. BD and FS conducted the field work. BD performed laboratory work and analyzed the genetic data. BD and SW interpreted spatial data and glacial history of Alps. BD and JG wrote the manuscript. All authors approved the manuscript.

Acknowledgments

The authors thank Clément Duckert, Julien Vieu and Marie Zufferey for help during field work at the Val d'Hérens, Switzerland. We also thank Olivier Bachmann, Silvia Kobel, and Aria Minder for assistance with laboratory work, and Donald Farrar, Daniel Croll, and Carl Rothfels for helpful comments on earlier versions of the manuscript. This work was supported by the Swiss National Science Foundation 31003A_156456.

References

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Barrett SC (2002) The evolution of plant sexual diversity. Nat Genet 3:274–284

Berch SM, Kendrick B (1982) Vesicular-arbuscular mycorrhizae of southern Ontario ferns and fern-allies. Mycologia 74:769–776

Bettin O, Cornejo C, Edwards PJ, Holderegger R (2007) Phylogeography of the high alpine plant *Senecio halleri* (Asteraceae) in the European Alps: in situ glacial survival with postglacial stepwise dispersal into peripheral areas. Mol Ecol 16:2517–2524

De Groot GA, Verduyn B, Wubs EJ, Erkens RH, During HJ (2012) Inter-and intraspecific variation in fern mating systems after long-distance colonization: the importance of selfing. BMC Plant Biol 12:3

Dielforder A, Hetzel R (2014) The deglaciation history of the Simplon region (southern Swiss Alps) constrained by ^{10}Be exposure dating of ice-molded bedrock surfaces. Quaternary Sci Rev 84:26–38

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

Braun-Blanquet J (1964) Pflanzensoziologie. Grundzüge der Vegetationskunde, 2nd edn. Springer, Vienna

Brockmann-Jerosch H, Brockmann-Jerosch M (1926) Die Geschichte der Schweizerischen Alpen-flora. In Schröter C (ed) Das Pflanzenleben der Alpen. Raustein, Zürich. pp 1110–1215

Camacho FJ (1996) New report of subterranean sporophytic gemmae on *Botrychium pumicola*. Am Fern J 86:27–28

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Mol Ecol 22:3124–3140

Chodat R, Pampanini R (1902) Sur la distribution des plantes dans les alpes austro-orientales et plus particulièrement d'un choix de plantes des alpes cadoriques et venitiennes. Globe Rev Genevoise Géographie 41:63–132

Chung MY, López-Pujol J, Chung JM, Moon MO, Chung MG (2012) Genetic diversity in the homosporous fern *Ophioglossum vulgatum* (Ophioglossaceae) from South Korea: inference of mating system and population history. J Hered 104:263–272

Coutterand S, Reynard E, Nicoud G (in prep.) Lateglacial glacial history in the Hérens valley (Western Switzerland). XX XX:XX–XX

Crist KC, Farrar DR (1983) Genetic load and long-distance dispersal in *Asplenium platyneuron*. Can J Bot 61:1809–18

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158

Dauphin B, Farrar DR, Maccagni A, Grant JR (2017) A worldwide molecular phylogeny provides new insight on cryptic diversity within the moonworts (*Botrychium* s.s., Ophioglossaceae). Syst Bot 42:XX–XX

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Soft 22:1–20

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Egan AN, Schlueter J, Spooner DM (2012) Applications of next-generation sequencing in plant biology. Am J Bot 99:175–185

Ehlers J, Gibbard PL, Hughes PD (2011) Quaternary Glaciations – Extent and Chronology: a Closer Look. Elsevier: Amsterdam

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107:1–15

Ellenberg H (1979) Liste der Gefa Bpflanzen Mitteleuropas. Scr. Geobot. 9:1–122

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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

Farrar DR, Johnson-Groh CL (1990) Subterranean sporophytic gemmae in moonwort ferns, *Botrychium* subgenus *Botrychium*. Am J Bot 77:1168–1175

Farrar DR (1998) Population genetics of moonwort *Botrychium*. In: Berlin N, Miller P, Borovansky J, Seal US, Byers O (eds) Population and habitat viability assessment for the goblin fern (*Botrychium mormo*), The Conservation Breeding Specialist Group, Minnesota, pp 109–113

Farrar DR (2011) (last updated 15 Dec. 2011) Systematics and taxonomy of the genus *Botrychium*. http://www.public.iastate.edu/~herbarium/botrychium/Moonwort-Systematics.pdf. (accessed 01 May 2016)

Farrar DR, Gilman AV (2017) Relationships in the *Botrychium campestre* (Ophioglossaceae) complex. Brittonia 69:265–275

Finsinger W, Tinner W (2007) Pollen and plant macrofossils at Lac de Fully (2135 m asl): Holocene forest dynamics on a highland plateau in the Valais, Switzerland. The Holocene 17:1119–1127

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977–993

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. ArXiv e-print http://arxiv.org/abs/1207.3907

Gehrig-Fasel J, Guisan A, Zimmermann NE (2007) Tree line shifts in the Swiss Alps: Climate change or land abandonment? J Veg Sci 18:571–582

Graham CF, Glenn TC, McArthur AG, Boreham DR, Kieran T, Lance S, Manzon RG et al (2015) Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). Mol Ecol Resour 15:1304–1315

Hamrick JL, Godt MJ (1996). Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc London B Biol Sci 351:1291–1298

Haufler CH, Pryer KM, Schuettpelz E, Sessa EB, Farrar DR et al (2016) Sex and the single gametophyte: Revising the homosporous vascular plant life cycle in light of contemporary research. BioScience 66:928–937

Hauk WD, Haufler CH (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

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc London B Biol Sci 359:183–195

Hippe K, Ivy-Ochs S, Kober F, Zasadni J, Wieler R, Wacker L et al (2014) Chronology of Late glacial ice flow reorganization and deglaciation in the Gotthard Pass area, Central Swiss Alps, based on cosmogenic ^{10}Be and in situ ^{14}C. Quaternary Geochronology 19:14–26

Holderegger R, Buehler D, Gugerli F, Manel S (2010) Landscape genetics of plants. Trends Plant Sci 15:675–683

Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24:417–441

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Johnson-Groh CL (1998) Population demographics, underground ecology and phenology of *Botrychium mormo*. In Berlin N, Miller P, Borovansky J, Seal US, Byers O (eds) Population and habitat viability assessment (PHVA) for the goblin fern (*Botrychium mormo*). Apple Valley, Minnesota. pp 103–108

Johnson-Groh CL, Lee JM (2002). Phenology and demography of two species of *Botrychium* (Ophioglossaceae). Am J Bot 89:1624–1633

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94

Kelly MA, Buoncristiani JF, Schlüchter C (2004) A reconstruction of the last glacial maximum (LGM) ice-surface geometry in the western Swiss Alps and contiguous Alpine regions in Italy and France. Eclogae Geol Helv 97:57–75

Klekowski EJ, Lloyd RM (1968) Reproductive biology of the Pteridophyta: 1. General considerations and a study of *Onoclea sensibilis* L. Bot J the Linn Soc 60:315–324

Lambiel C, Maillard B, Kummert M, Reynard E (2016) Geomorphology of the Hérens valley (Swiss Alps). J Maps 12:160–172

Landolt E, Bäumler B, Erhardt A, Hegg O, Klötzli F, Lämmler W et al (2010) Flora indicativa–Ecological indicator values and biological attributes of the flora of Switzerland and the Alps. Haupt-Verlag, Bern

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv e-print https://arxiv.org/abs/1303.3997

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

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659

Lischer HE, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28:298–299

Maccagni A, Parisod C, Grant JR (2017) Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. Alpine Bot 127:185–196

Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. Mol Ecol 19:3824–3835

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

McCauley DE, Whittier DP, Reilly LM (1985) Inbreeding and the rate of self-fertilization in a grape fern, *Botrychium dissectum*. Am J Bot 72:1978–1981

McMaster RT (1994) Ecology, reproductive biology and population genetics of *Ophioglossum vulgatum* (Ophioglassaceae) in Massachusetts. Rhodora 96:259–286

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes 4:792–794

Meza-Torres EI, Stensvold MC, Farrar DR, Ferrucci MS (2017) Circumscription of the South American moonwort *Botrychium* (Ophioglossaceae). Plant Biosystems 151:1–11

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142:1061–1064

Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T et al (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC Evol Biol 9:276

Parisod C, Besnard G (2007) Glacial in situ survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigata* L.(Brassicaceae). Mol Ecol 16:2755–2767

Peck JH, Peck CJ, Farrar DR (1990) Influences of life history attributes on formation of local and distant fern populations. Am Fern J 80:126–142

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one 7:e37135

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R et al. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

Ranker TA, Haufler CH (2008) Biology and evolution of ferns and lycophytes. Cambridge University Press, Cambridge

Rothfels CJ, Li FW, Sigel EM, Huiet L, Larsson A, Burge DO et al (2015) The evolutionary history of ferns inferred from 25 low-copy nuclear genes. Am J Bot 102:1089–1107

Schönswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. Mol Ecol 14:3547–3555

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

Soltis DE, Soltis PS (1986) Electrophoretic evidence for inbreeding in the fern *Botrychium virginianum* (Ophioglossaceae). Am J Bot 73:588–592

Soltis DE, Soltis PS (1987) Polyploidy and breeding systems in homosporous Pteridophyta: a reevaluation. The Am Nat 130:219–232

Soltis DE, Soltis PS (1989) Polyploidy, breeding systems, and genetic differentiation in homosporous pteridophytes. In: Soltis DE, Soltis PS (eds) Isozymes in plant biology. Dioscorides Press, Portland pp 241–258

Stehlik I (2000) Nunataks and peripheral refugia for alpine plants during quaternary glaciation in the middle part of the Alps. Bot Helv 110:25–30

Stensvold MC (2007) A taxonomic and phylogeographic study of the *Botrychium lunaria* complex. Dissertation, Iowa State University

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

Taberlet P, Cheddadi R (2002) Quaternary refugia and persistence of biodiversity. Science 297:2009–2010

Truong HT, Ramos AM, Yalcin F, de Ruiter M, van der Poel HJ, Huvenaars KH et al (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. PLoS One 7:p.e37565

Wagner WH (1972) Disjunctions in homosporous vascular plants. Ann Missouri Bot Gard 59:203–217

Wagner WH, Wagner FS, Beital JM (1985) Evidence for interspecific hybridization in pteridophytes with subterranean mycoparasitic gametophytes. Proc Roy Soc Edin 86B:273–281

Watano Y, Sahashi N (1992) Predominant inbreeding and its genetic consequences in a homosporous fern genus, *Sceptridium* (Ophioglossaceae). Syst Bot 17:486–502

Weir BS, Cockerham CC (1984) Estimating F- statistics for the analysis of population structure. Evolution 38:1358–1370

Wildi W, Gurny-Masset P, Sartori M (2016) Glacial landscapes of the Val d'Hérens (Valais, Switzerland). Université de Genève, Genève

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

Williams EW, Waller DM (2012) Phylogenetic placement of species within the genus *Botrychium* ss (Ophioglossaceae) on the basis of plastid sequences, amplified fragment length polymorphisms, and flow cytometry. Int J Plant Sci 173:516–531

Winther JL, Friedman WE (2007) Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). Am J Bot 94:1248–1255

Wirsig C, Zasadni J, Ivy- Ochs S, Christl M, Kober F et al (2016) A deglaciation model of the Oberhasli, Switzerland. J Quaternary Sci 31:46–59

Wubs EJ, de Groot GA, During HJ, Vogel JC, Grundmann M, Bremer P, Schneider H (2010) Mixed mating system in the fern *Asplenium scolopendrium*: implications for colonization potential. Ann Bot 106:583–590

Supplementary Materials

Table S1 The number of variant sites retained after the successive filtering steps using the dDoccent pipeline

Step number	Stage description	Number of variant sites
1	Total raw SNPs before variants filtering	96'029
2	Genotyped in 50% of individuals, base quality > 20, minor allele count of 3, minimum coverage of 3 reads	49'597
3	Genotyped in > 95% of individuals, MAF > 0.05, minimum coverage of 10 reads	4'013
4	Genotyped in > 95% of populations	3'370
5	Biallelic only	3'066
6	Allele balance filter out loci with allele balance > 0.25 and < 0.75, and kept those are close to 0	1'793
7	Read orientation keep loci that have over 100 times more forward alternate reads than reverse alternate reads, and 100 times more forward reference reads than reverse reference reads along with the reciprocal	1'663
8	Mapping quality between reference and alternate alleles	853
9	Mapping quality remove any locus that has a quality score below 1/4 of the depth	852
10	Read quality loci with overall low read quality scores (less than 25% of read depth) were discarded	801
11	Read depth remove loci with abnormally high coverage recalculated	706
12	HWE remove erroneous variant calls based on HWE applied by population	620
13	Haplotypes identification filter paralogs and account the physical linkage	513

Conclusions and perspectives

Conclusions

Cryptic groups within Botrychium constitute a substantial reservoir of genetic diversity

The apparent morphological homogeneity among taxa misleads us on the real genetic diversity found within *Botrychium*. Based on multiple accessions per taxon throughout their known geographical distribution, our results have confirmed the monophyly of recently published cryptic taxa and provided insights on thirteen putative new taxa that also display similar morphologies to species already described. This phylogenetic study associated with observations in the field leads us to hypothesize that different ecological niches may harbor distinct cryptic taxa in close geographical ranges, for example taxa of the Lunaria species complex occurring in Alps, which show diversified habitats, from very dry to saturated soils in water or from very windy sites to those in temperate forests for instance. In total, we present a tentative number of 47 taxa, of which a high diversity of haplotype is constantly expanding with additional sampling efforts.

Botrychium have among the biggest genome sizes of vascular plants

With an average size of 26.11 Gbp for diploid genomes and up to 77.81 Gbp for the known hexaploid, *Botrychium* species have a rich story to tell about the evolution of the ancient Ophioglossaceae's lineage. The three major lineages of diploid moonwort—Lanceolatum, Lunaria, Simplex-Campestre—are characterized by different DNA content that suggest recent genomic reorganization or chromosomal structural changes. Curiously, without difference in monoploid genome size between ploidy levels, flow cytometry data show evidence for the stability of polyploid genomes after interspecific hybridization, contrasting to what is observed in flowering plants, in which large genomic or chromosomal deletions occur after allopolyploidy events. A better understanding of these underlying mechanisms is essential to finally reveal how the diploidization process takes place in natural populations to drive such a diversity of the fern karyotypes.

Rapid and recent allopolyploidization plays a major role in speciation of Botrychium

About half of *Botrychium* species are allopolyploid. A large majority are allotetraploids that were formed by the hybridization between diploid progenitors in the last two million years. Our divergence time analysis presents a first temporal scale of these speciation events within *Botrychium* that supports a rapide radiation of allopolyploids via an interesting bias of parental donors. Indeed, inter-clade allopolyploidizations occurred with non-random and non-reciprocal parental couples, where individual diploid species tend to be either the maternal or paternal contributors to polyploids, but not both. Also, no evidence of autopolyploidy was reported based on our nuclear sequences and therefore we confirm that allopolyploidy is a key speciation mechanism for generating new moonwort species.

A mixed mating system as novel speciation mechanism in early divergent vascular plants

Focusing on the Lunaria species complex, we discovered outcrossing populations of *B. lunaria* in Swiss Alps with heterozygosity at unprecedented levels and recorded among the highest within-population genetic diversity found in homosporous ferns, comparable to what is observed in seed plants. In comparison with all other taxa of the species complex, this finding provides evidence for the alternation of mating system in closely related species, in which gametophytic selfing is the predominant mode of reproduction. Our results illustrate the versatility of the life cycle of *Botrychium* and the diversification mechanisms through gametophytic selfing and sporophytic outcrossing that have driven speciation of ancestral vascular plants on terrestrial habitats.

How genetic diversity of B. lunaria was rescued in glacial refugia

Based on cpDNA, allozymes, and SNPs data, we found a substantial genetic diversity within Alps *B. lunaria* populations compared to any other geographical areas where the species has been genetically studied. Based on these empirical results, populations likely have survived in situ in central Alps refugia during the last glacial maximum, where favorably south-exposed slopes could harbor well-adapted alpine plants to extremely cold habitats, with soils free of snow and ice during the growing season. Then, post-glacial recolonization from in situ refugia has spread the allele richness maintained in those outcrossing ancestral populations along valleys and alpine meadows.

Perspectives

This thesis may be followed-up by further studies to bring a better understanding about the speciation and evolution of early vascular plants. Basically, the speciation mechanisms may be investigated at several hierarchical levels of diversity, from populations (chapters 4 and 5) to species (chapters 1 and 3) and genera (chapter 2), which require logistic resources for appropriate design samplings and new molecular tools. Future projects may address the following points:

1. Chapter 1: major expansion of sampling in Asia where few studies have been done on *Botrychium* in order to retrace the geographical origin of the Lunaria species complex.
2. Sample remaining "orphan" species and disjunct populations, notably *B. tolucaense* in Mexico, and *B. lunaria* in the Azores, Africa (Atlas Mountains of Morocco), Blue Mountains of Australia.
3. Study *Botrychium simplex* group worldwide. Do we have *B simplex* and/or *B. tenebrosum* in Switzerland?
4. Chapter 2: investigation of multiple cytotypes within Swiss Alps *B. lunaria* populations, do we have triploids, hybrids, allo-polyploids?
5. Chapter 3: expand our research and methodology to other ophioglossoid genera (see new classification PPI), especially on *Sceptridium* and *Ophioglossum*, which are the most diversified.

6. Chapter 4: more population genetic studies have to be conducted on ferns, especially to investigate their mating systems. For long time, plant biologists had in mind the low genetic variability occurring in fern taxa, resulting from frequent inbreeding populations. But it was largely underestimated and ferns may have a similar or even a higher genetic diversity at the conspecific level in comparison of seed plants. Pteridologists should perform more population genetic studies in this prospect.
7. Chapter 5: expand our study of the Val d'Herens valley to the Alps, 20 valleys targeting well-known refugia (Arosa, Gornergratt) as well as alpine localities completely covered by ice sheet during the last glacial maximum, to lead a comparative study between the phylogeographical patterns in the context of climatic oscillations of the Quaternary and the contemporary gene flow maintained among Alps *B. lunaria* populations.

Extra-contributions to the project

Ståhl, P., Ekman, J., Westerberg, S., Grant, J.R., and **B. Dauphin**. 2016. Me rom pyssling-låsbräken i Sverige. *Svensk Botanisk Tidskrift* 110 (2): 68–74.

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

Grant, J.R., and **B. Dauphin**. *Botrychium alaskense* var. *salchaketense* (Ophioglossaceae), a new variety of moonwort from the Interior of Alaska. *Phytokeys*.

Sandoz, F., Hervé, V., **Dauphin, B**., Grant, J.R., and S. Bindschedler. Environmental conditions driving the arbuscular mycorrhizal fungi communities of the *Botrychium lunaria* root system. *In preparation*.

Curriculum Vitae – BENJAMIN DAUPHIN

General-Dufour-Strasse 120, 2502 Biel, Schweiz | benjamin.dauphin@unine.ch | Tel: (+41) 77 427 66 23

EDUCATION

Ph.D. in Biology (expected by November 2017) – University of Neuchâtel (Switzerland) Since 12.2014
Subject: Evolution of moonwort ferns (*Botrychium*, Ophioglossaceae) on local to global scales

Certificate in Geomatics (with honours) – University of Geneva (Switzerland) 01.2013 – 01.2014

Master in Plant Physiology and Ecology (with honours) – University of Neuchâtel (Switzerland) 09.2011 – 12.2012

Bachelor in General Biology and Earth Sciences (with honours) – University of Cergy-Pontoise (France) and University of Montréal (Canada) 09.2008 – 05.2011

PEER-REVIEWED PUBLICATIONS

Dauphin, B. Grant, J. R. Stensvold, M. C. Maccagni, A. Nason, J. & Farrar, D. Dispersal and Diversification in Ancestral Vascular Plants. *In review.*

Sandoz, F. Hervé, V. **Dauphin, B.** Grant, J. & Bindschedler, S. Environmental conditions driving the arbuscular mycorrhizal fungi communities of the *Botrychium lunaria* root system. *In preparation.*

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

Dauphin, **B**. Farrar, D. Grant, J. & Rothfels, C. Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. *In review.*

Dauphin, **B**. Farrar, D. Maccagni, A. & Grant, J. A worldwide molecular phylogeny provides new insight on cryptic diversity within the moonworts (*Botrychium* s.s., Ophioglossaceae). *Systematic Botany*, in press.

Schuettpelz, E. Schneider, H. […] **Dauphin, B.** & al. 2016. A community-derived classification for extant lycopods and ferns. *Journal of Systematics and Evolution*, 54(6), 563–603.

Dauphin, **B**. Grant, J. & Márz, P. (2015) Ploidy level and genome size variation in the homosporous ferns *Botrychium* s.l. (Ophioglossaceae). *Plant Systematics and Evolution,* 302(5), 575–584.

Dauphin, **B**. Vieu, J. & Grant, J. (2014). Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* ss, Ophioglossaceae). *American Journal of Botany*, 101(1), 128–140.

AWARDS

Outstanding Student Award, International Botanical Congress 2017, China 07.2017

Student Award, Conference of the American Fern Society and the Botanical Society of America 07.2017

Laureate Jean-Luc Crélerot, Science Institute of the University of Neuchâtel, Switzerland 11.2013

EXPERIENCE

Ph.D. student at the University of Neuchâtel (Neuchâtel, Switzerland) — Since 12.2014
Laboratory of Evolutionary Genetics, Jason Grant's group
Focus on phylogenetics and population genetic studies

Teaching assistant at the University of Neuchâtel (Neuchâtel, Switzerland) — Since 12.2014
General biology, plant diversity and evolution

Scientific collaborator at the University of California (Berkeley, United-States) — 11.2015 – 05.2016
Laboratory of Carl Rothfels
Sequencing of non-model species using the PacBio technology

Scientific collaborator at the Forests Service and the Landscape (Sion, Switzerland) — 05.2014 – 11.2014
Manager of biodiversity and mapping of protection areas

Scientific collaborator at the University of Neuchâtel (Neuchâtel, Switzerland) — 10.2013 – 03.2014
Laboratory of Evolutionary Botany, Christian Parisod's group
Study of processes and genome dynamics of flowering plants

Internship in Ecological Genomics at the Swiss Federal Institute of Technology (Lausanne) — 02.2013 – 07.2013
Assessment of genetic diversity in relation to environmental conditions

GRANTS

Swiss Botanical Society for congress grant (750-CHF), Switzerland	07.2017 – 08.2017
Overhead Fund from the University of Neuchâtel (9,033-CHF), Switzerland	06.2016 – 12.2017
Ph.D. grant mobility from the Swiss National Science Foundation (16,400-CHF), Switzerland	11.2015 – 05.2016
Fund of Donations of the University of Neuchâtel (7,550-CHF), Switzerland	06.2012 – 12.2012
Grant Travel of the Swiss Academy of Natural Sciences (1,100-CHF), Switzerland	06.2012 – 08.2012
Fund Marguerite Wüthrich and A. Matthey-Dupraz (1,800-CHF), Switzerland	05.2012 – 07.2012
Fund Dr. Joachim of Giacomi (4,000-CHF), Swiss Academy of Natural Sciences, Switzerland	04.2012 – 12.2012
Scholarship of the Conférence des Recteurs et des Principaux des Universités du Québec, CA	08.2010 – 05.2011

LANGUAGES

English Full professional proficiency

French Mother tongue

Italian B1 European level

German A2 European level

COMPUTING

Operating system OSX, Windows and Linux

Particular skills R statistical software, GIS software, molecular biology software, experience in Unix

HOBBIES AND SPORTS

Swiss Alpine Club, climbing, hiking, running, cooking, discover new cultures and explore the world!