UNRAVELLING COMPLEX EVOLUTIONARY HISTORIES BY A MULTIDISCIPLINARY APPROACH: THE CASE OF THE COMMON MOONWORT FERNS

Thèse présentée à la Faculté des Sciences

UNIVERSITÉ DE NEUCHÂTEL

Pour l'obtention du grade de

DOCTEUR ÈS SCIENCES

Par

VINCIANE MOSSION

Acceptée par les membres du jury de thèse :

Professeur Dr. **Daniel Croll**, Université de Neuchâtel, directeur de thèse et président du jury

Docteur **Michael Kessler**, Université de Zürich, examinateur

Professeur Dr. **Sergio Rasmann**, Université de Neuchâtel, rapporteur

Professeur Dr. **Carl Rothfels**, Université de Californie, rapporteur

Docteur **Weston Testo**, Université de Göteborg, rapporteur

Défendue le **5 novembre 2021** devant les membres du jury
et le 17 février 2022 publiquement

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**

Madame Vinciane MOSSION

Titre:

"Unravelling complex evolutionary histories by a multidisciplinary approach: the case of the common moonwort ferns"

sur le rapport des membres du jury composé comme suit:

- Prof. ass. Daniel Croll, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. tit. Michael Kessler, Université de Zürich, Suisse
- Prof. Sergio Rasmann, Université de Neuchâtel, Suisse
- Prof. Carl Rothfels, University of California, Berkeley, USA
- Dr Weston Testo, University of Gothenburg, Sweden

Neuchâtel, le 9 novembre 2021

Le Doyen, Prof. A. Bangerter

Only those who dare may fly

Luis Sepúlveda, The Story of a Seagull and the Cat Who Taught Her to Fly

Acknowledgements

This PhD journey was made possible thanks to wonderful people and some institutions to which I here expressed my sincere gratitude.

Above all, I am deeply thankful to my supervisor, Prof. Daniel Croll for his inestimable support. Daniel you were very supportive in every matter, and you managed to restore my scientific curiosity and devoted enthusiasm for research while needed. Thank you for being such a great supervisor and showing me how to be a good scientist! I am furthermore grateful to Dr. Michael Kessler. I have no words to express how happy I was when you joined this PhD adventure. Thank you for getting me into the amazing world of taxonomy! Daniel and Michael, I really enjoyed working with you two! You are both amazing scientists and persons. You gave me the support, time and trust I needed to finish this thesis. I want to thank Dr. Jason Grant for giving me the opportunity to start this project. This project meant a lot to me. I wish to extend my gratitude to Dr. Benjamin Dauphin, who helped to build this project. I am obliged to Prof. Donald R. Farrar for our long discussion about *Botrychium*. You always took time for answering my thousands of questions, and never missed an opportunity to share my endless enthusiasm for these cute little ferns. All the former and current members of the evolutionary genetics laboratory were also very supportive and always willing to share scientific and non-scientific thoughts. I felt lucky to be part of this marvellous team! I enjoyed a lot teaching with Dr Ursula Oggenfuss. It was a delight to converse about nature and science with you. Besides, you turned to be a wonderful stepmother for my baby ferns. I was further fortunate to have such a great writing buddy than Dr Sabina Tralamazza. Moreover, I am grateful to the Genomic Diversity Center team who helped me to put on wheels my first chapter!

I am grateful to Clément Duckert, Cun-Feng Zhao, Amandine Pillonel, Daniel Ston, Kenan Ćatović, Kuban Uulu Zholdoshbek, Marine Ramirez, Quentin Dubois, Sandra Grünig, Siqi Liang, Zikrio and Stella Huynh who came with me in the field and made the worldwide collection a reality. It was such an amazing adventure, sometimes funny, like when a 6-year-old girl speaking Kyrgyz tried to teach us Russian; sometimes scary, like when an old man in the Tadjik mountains told me that he was not expecting me to come back alive from my trip; and sometimes extremely satisfying, like when we found one tiny specimen after 12 hours of hiking in Yunnan mountains. A special thanks to Kenančić, also called Supreme leader. I am very happy we met, and I enjoyed so much getting you crazy about the *Botrychium* of Bosnia and Herzegovina! Also, I am indebted to Sci-hub 'library' for all the papers I would probably never have access to otherwise.

Beyond work, I would like first to thank my family, especially my sister, father, mother, and grandmother who supported me during all these years. A huge thanks to all my friends! You were there to take me out while I was working too much, for listening when I was challenged with my work, even though

some of you had absolutely no idea of what I was talking about. Guys, I love you! A special thanks goes to Amandine Pillonel and Ophélie Gning who also played an important role at the University. Amandine, I don't how I would have done without you when I started laboratory work. You spent your Saturdays showing me how to properly do a PCR, prepared slides with me, measured spores, and turned my soil samples into usable data. I own you a lot of chocolate cakes! Ophélie, you helped me so much in the lab, and guided me through good practices! These genetic data presented here after are also yours! I am very happy to count you both among my friends. I am beholden to Wafa Kooli. It helped me a lot to talk with someone that could predict and understand the struggle of the last thesis months. Then, I own a large part of my personal flourishing and of my survival to Daniel Ston. I am genuinely grateful for the endless support you provided during my PhD. Thank you for not letting me eat only salad and cereals, for working with me sometimes, or simply being there when I was going through doubts and frustrations. Most of all, thank you so much for making me the naturalist I was seeking to be! No one could have showed the patience you had to teach me birdsongs and repeating over and over those insect names while I was busy thinking about *Botrychium*. I further want to thank all the people I met along the way and that were not cited above; this work finally came together also thanks to you.

Finally, and importantly, I would like to thank the Egalité des chances and Fonds des donations of the University of Neuchâtel, the Swiss Systematic Society, and the Swiss Botanical Society for contributing to the financial supports of this project.

Table of contents

Abstract

Ferns are a highly diverse group of land plant present in almost every continent and most climates. Even though ferns diverged from the seed plants about 360-430 million years ago, most extant ferns originated from recent Cenozoic radiation events. Recently diverged groups often consist of species exhibiting subtle morphological differences with frequent naturally occurring hybrids, resulting in a poorly understanding of these groups' diversity The rapid decline of biodiversity emphasizes the need to characterize the species' richness of such groups to further identify potential threats and define conservation measures. The thesis presented here focused on the understanding of the common moonwort ferns complex evolutionary history by a multidisciplinary approach. The *Botrychium lunaria* group (common moonwort) is a species complex of the Ophioglossaceae family. One of the main limitations of the studies of fern evolutionary histories, including for the *B. lunaria* group, is the availability of nuclear genomic resources. The paucity of such genome-wide datasets is mostly due to the complexity and the size of fern genomes. A recent breakthrough came with the 1KP project which led to the sequencing 1,000 plant transcriptomes. With the motivation of providing genomic tools for *Botrychium* ferns, we applied a similar transcriptomic approach as the 1KP project. We showed that transcriptome-wide dataset was a powerful tool to resolve population structure and investigate ploidy level. The newly established transcriptome is also a precious resource for defining single-copy markers for phylogenetic studies. Then, we assessed the *B. lunaria* group diversity on a global scale and investigated the role of climatic and ecological factors in the speciation process. We recovered ten well-defined monophyletic clades corresponding to eleven species-level clades. Three clades were not previously identified and need to be considered in future research. The diversification of the group coincides with the Pleistocene climatic oscillations. We also identified indications for soil preferences among species. Finally, we examined the morphological features of the uncharacterized species and subspecies-level clade. We described four new species from Eurasia, two exhibiting a large distribution range and one being restricted to the Himalayas. This thesis improves our understanding of the genetic, morphological, and ecological diversity in the *B. lunaria* group. Also, the genomic resources we developed open doors to study reticulate evolution and population differentiation in this and related groups.

Key words: Ophioglossaceae, diversity, speciation, taxonomy, transcriptomic, ferns, population genetic

Résumé

Les fougères constituent un groupe très diversifié de plantes terrestres que l'on trouve sur presque tous les continents et sous une grande variété de climats. Bien que les fougères aient divergé des plantes à graines il y a environ 360 à 430 millions d'années, la plupart des espèces actuelles sont issues d'événements de radiation récents datant du Cénozoïque. Or, les groupes ayant récemment divergé sont souvent constitués d'espèces présentant de faibles différences morphologiques et des hybrides dans les populations naturelles. Par conséquent, la diversité de ces groupes est souvent sous-estimée et rarement considérée. Cependant la biodiversité décline rapidement. Il est donc crucial de caractériser la richesse en espèces de ces groupes afin d'identifier les potentielles menaces et de définir des mesures de conservation appropriées. La thèse présentée ici a porté sur l'étude de l'histoire évolutive des Botryches lunaires en utilisant une approche multidisciplinaire. Le groupe *Botrychium lunaria* (Botryches lunaire) est un complexe d'espèces de la famille des Ophioglossaceae. L'une des principales limites pour l'étude de l'histoire de l'évolution des fougères, y compris pour le groupe *B. lunaria*, est la disponibilité des ressources génomiques. Le manque de données à l'échelle du génome est principalement dû à la complexité et à la taille des génomes de fougères. Une importante avancée a récemment été réalisée avec le projet 1KP qui a permis le séquençage de 1 000 transcriptomes de plantes. Dans le but de générer des outils génomiques pour les fougères *Botrychium*, nous avons appliqué une approche similaire à celle du projet 1KP. Nous avons montré qu'un transcriptome pouvait être un outil puissant pour résoudre la structure des populations et étudier le niveau de ploïdie. Le transcriptome nouvellement établi permettra également de développer de nouveaux marqueurs nucléaires pour des études phylogénétiques. Nous avons ensuite évalué la diversité au sein du groupe *B. lunaria* à l'échelle mondiale et étudié le rôle des facteurs climatiques et écologiques dans le processus de spéciation. Nos analyses phylogénétiques ont mis en évidence dix clades monophylétiques bien définis correspondant à onze espèces potentielles. Trois clades sont nouveaux pour la science et doivent être pris en compte dans les recherches futures. La diversification du groupe coïncide avec les oscillations climatiques du Pléistocène. Nous avons également identifié des indications de préférences édaphiques parmi les espèces. Enfin, nous avons examiné les caractéristiques morphologiques des taxa non caractérisés. Nous avons décrit quatre nouvelles espèces pour l'Eurasie : deux d'entre elles présentant une large distribution, une étant principalement connue de la péninsule des Balkans, et un autre étant limitée à l'Himalaya. Ce travail de thèse améliore notre compréhension de la diversité génétique, morphologique et écologique du groupe *B. lunaria*. De plus, les ressources génomiques développées au cours de ce projet ouvrent des nouvelles possibilités pour étudier l'évolution réticulée et la différenciation des populations dans ce groupe de fougères et les groupes apparentés.

Mots clés : Ophioglossaceae, diversité, spéciation, taxonomie, transcriptomique, fougères, génétique des populations.

General introduction

Diversity of the fern tree of life

Ferns (Polypodiopsida) are a highly diverse group of land plant which diverged from the remaining of the vascular plants long before the golden age of dinosaurs (Raubeson and Jansen 1992; Pryer et al. 2004). Intact fossils of Cladoxylopsida, thought to be the ancestor of ferns, were found in northern Scotland, dating back fern divergence from the middle Devonian (Givetian, -382.7 - 387.7 million years; (Berry and Hilton 2006). Time-scaled phylogenies involving molecular clock calibration partly agreed with the known fossil record (Wikström and Kenrick 2001; Pryer et al. 2004; Clarke et al. 2011; Rothfels et al. 2015) and partly proposed that ferns may have diverged even earlier, in mid-late Silurian (- 419.2 - 443.8 million years;(Testo and Sundue 2016). Despite of the Silurian-Devonian fern precursors extinction, ferns remain an incredibly diverse group of land plants with about 10,600 species (PPG I 2016) only outnumbered by flowering plants (about 352,000 species; (Paton et al. 2008; Joppa et al. 2011). Not only species rich, ferns are present in most climates and have colonized a multitude of habitats including deserts, grasslands, forest understory, mountainous regions, and aquatic environments (Mehltreter et al. 2010). Ferns have a unique biology characterized by two independent life stages (*i.e.* gametophyte and sporophyte Pinson et al. 2016) and the capability of versatile reproduction mode including apomixis (Grusz 2016), selfing and outcrossing (Haufler et al. 2016). Their biological features combined with their ability for long distance spore dispersal (Barrington 1993) have played a key role in the established and persistence of isolated populations from a few spores, and likely contribute to their ecological diversification.

Fern genomic resources

These past decades the accessibility of genetic data on ferns has drastically increased, especially the availability of chloroplast markers (Ravi et al. 2008; Zhong et al. 2014; Lu et al. 2015). Phylogenetic analyses resolved ferns as the sister group to seed plants (Pryer et al. 2001) and brought light into the deep divergence of fern lineages (Adams and Rohlf 2000; Des Marais et al. 2003; Magallón et al. 2013; Qi et al. 2018). Until recently fern phylogenies mostly relied on plastid data, occasionally combined with mitochondrial markers (Knie et al. 2015) or nuclear markers (Pryer et al. 2001, 2004) or both (Qiu et al. 2007). The prevalent use of uniparental inherited markers prevented analyses of reticulation events, which are known to be common in many groups of ferns (Shepherd et al. 2008; Rothfels et al. 2014; Dauphin et al. 2018). The development of the nuclear genomic resources for ferns was hampered by their huge genome sizes (Obermayer et al. 2002; Hidalgo et al. 2017), until recent studies highlighted the power of transcriptome-based approaches (Rothfels et al. 2013, 2015; Wickett et al. 2014; Shen et al. 2018; Qi et al. 2018; Leebens-Mack et al. 2019). These transcriptome-wide datasets allowed the development of highly valuable single-copy nuclear markers (Rothfels et al. 2013), gave insights on ferns position in the tree of life (Wickett et al. 2014), resolved deep evolutionary relationships among ferns (Rothfels et al. 2015; Shen et al. 2018), and uncovered features of fern genes regulation (Qi et al. 2018) and genome evolution (Leebens-Mack et al. 2019). Besides, the first insight into fern genome structures were given by the complete genome sequencing of *Azolla filiculoides* and *Salvinia cucullata* (Li et al. 2018). These two genera are known for their small genome size but most fern genomes are of enormous size (Hidalgo et al. 2017) and whole genome sequencing of those are highly complicated (Marchant et al. 2019). Enlarging transcriptome-wide datasets is an efficient way to tackle issues associated with gigantic genome sizes and the high degrees of gene duplication in ferns.

Evolution and diversification of ferns

Even though fern precursors were part of the Devonian flora, most extant ferns originated from recent Cenozoic radiation events (Schneider et al. 2004; Schuettpelz and Pryer 2009; Fiz-Palacios et al. 2011; Testo and Sundue 2016; Dauphin et al. 2018). Molecular phylogenetic studies supported an early Cenozoic burst of leptosporangiate, notably a rapid diversification of the polypods related to ecological opportunities created by angiosperm-dominated ecosystems. A similar pattern of speciation was suggested for the epiphytic *Huperzia* species (Lycopodiopsida) (Wikström and Kenrick 2001). Moreover, a late Cenozoic radiation within one of the earliest fern lineages (i.e. Ophioglossaceae) was uncovered by polyploid network and time divergence analysis (Dauphin et al. 2018). Such recent diversifications imply that some fern groups contain species complexes characterized by morphologically similar species (Paris et al. 1989), often combined with hybridization (Reichstein 1981) and polyploid networks (Lovis 1958; Rothfels et al. 2014; Williams et al. 2016; Hanušová et al. 2019). Species are often defined as a group of individuals that interbreed in nature and produce fertile offsprings (Mayr 1940; Dobzhansky 1950). This biological species concept underlies reproductive barriers which might not be complete depending from which temporal point of view one witnesses the speciation processes (Stankowski and Ravinet 2021). Indeed, speciation occurs on an evolutionary timescale along which occasional gene flow is possible until the establishment of total reproductive isolation. It is commonly the case of recently diverged groups that subtle morphological differences distinguish with frequent hybrids in natural populations. In such a context, interpreting fern species diversity and understanding the speciation processes involved is challenging and intriguing.

Evolution and taxonomy of the Ophioglossaceae fern family

The Ophioglossaceae are a fern family composed of 112 species (PPG I 2016) with a distribution centered on tropical and circumboreal regions. The Ophioglossaceae diverged from its sister family, the Psilotaceae, about 289 million years during late Paleozoic (Testo and Sundue 2016; Kumar et al. 2017; Lehtonen et al. 2017). The known fossil record is scarce with only two macrofossils of Botrychioideae, one of the four sub-families (Mankyuoideae, Helminthostachyoideae and Ophioglossoideae in Zhang et al. 2020 sense) found to date (Rothwell and Stockey 1989; Bozukov et al. 2010). The oldest fossil, *Botrychium wightonii* Rothwell and Stockey, dated of about 57 million years was found in western North America and showed resemblance to the extant species *Botrypus virginianus* (L.) Holub and *Botrypus lanuginosum* (Hook. & Grev.) Holub. The youngest fossil related to *Sceptridium underwoodianum* (Maxon) Lyon, was found in Bulgaria and dated from Oligocene era about 23 million years. These only two fossils highlight the deep time divergence between two close genera of the Botrychioideae subfamily.

The Ophioglossaceae have remarkable biological features and morphology compared to the rest of ferns (Stensvold 2008). Gametophytes are subterranean and non-photosynthetic. This characteristic shared among the Ophioglossidae (PPG I 2016) is unusual among ferns (Yatskievych 2003) and goes with endophytic fungal associations (Campbell 1908; Read et al. 2000) on which gametophytes growth relies (Farrar 1998; Winther and Friedman 2007). Distinct of most of ferns, Ophioglossaceae gametophytes have an extended life span of several years (Johnson-Groh 1998; pers. observations). The sporophyte may also last for decades (Dauphin et al. 2020). The sporophyte usually produces one leaf per year derived from single apical meristem (Farrar 2011) with a very simple morphology. The leaf is divided in two segments: a fertile segment (sporophore) and a sterile segment (trophophore). The trophophore is a photosynthetic blade and the sporophore bears globular sporangia containing thousands of spores (Stensvold 2008; pers. field observations). These two segments join into a common stalk, the base of which is enlarged to form a sheath protecting the apical meristem and the nested primordial leaves of subsequent years (Bold et al. 1987; Meza Torres 2013). Below the sheath stands an upright and unbranched subterranean stem bearing a few fleshy roots lacking hairs (Wagner and Wagner 1993). Immature leaf is a reduced image of mature leaf and does not display the circinate vernation as most of ferns do (Stensvold 2008). This peculiar morphology is overall very simple providing few traits for morphological identification and making species recognition without molecular approaches very challenging.

The Ophioglossaceae diversity assessment started in the early nineteenth century exclusively based on morphological studies despite of their simple morphology and limited diagnostic characters. This led to a long-standing disagreement between botanists on the Ophioglassaceae taxonomic treatment. Presl subdivided the Ophioglossaceae (Presl 1845) onto three subfamilies: Ophioglossoideae, Botrychioideae and Helminthostachyoideae. These subfamilies were treated by various authors at the family level (Botrychiaceae Horan, Ophioglossaceae, Helminthostachyaceae Ching) and largely used that way by Asian and Russian taxonomists till the beginning of the 21st century, meanwhile North American and European taxonomists followed the subfamily concept. The discovery on Cheju Island in the south of Korean of a new genus, Mankyua B.Y. Sun, M.H. Kim & C.H. Kim (Sun et al. 2001), morphologically intermediate between the Ophioglossoideae and Helminthostachyoideae disrupted the multiple family concept. An uniformized used of Presl's subfamilies was finally adopted with the first community classification for extant ferns and lycophytes (PPG I 2016) with the addition of the Mankyuoideae J.R. Grant & B. Dauphin. An equivalent complex taxonomical history might be told for one of the Presl's subfamily, the Botrychioideae, which has been intensively studied during the past century.

An unsettled taxonomy of the Botrychioideae subfamily

Until recently, four genera were widely accepted in the Botrychioideae (PPG I 2016): *Botrychium* Sw. (Swartz 1801), *Botrypus* Michx. (Michaux 1803), *Sceptridium* Lyon (Lyon 1905), and *JapanoBotrychium* Masam. (Masamune 1931). These genera have been considered by many authors at the section or at the subgenus level (Milde 1869; Clausen 1938; Kato 1978; Tryon and Tryon 1982; Xianchun et al. 2013). Mild (1869) in his *Botrychium* monographia lumped *Botrychium* and *Botrypus* genus into one broader '*Botrychium*' in which he defined two sections: *EuBotrychium* and *Osmundopteris*. The last corresponded to the genus *Botrypus* described by Michaux (1803). In the early 1900's Lyon segregate a new genus, *Sceptridium*, based on the gametophyte study of several *Botrychium* species and on sporophyte shape differences. Later, Clausen (1938) interpreted *Sceptridium* and Mild's sections at the sub-genus level in its monograph of the Ophioglossaceae, meanwhile Small was elevating *Osmundopteris* at the genus level (Small and Barnhart 1938). Afterwards, *Osmundopteris* was synonymised with *Botrypus* genus. The conception of Clausen was followed by Tyron and Tyron (1982) who incorporated *JaponoBotrychium* (Masamune 1931) at the subgenus level as described by Kato et Sahashi (1977) who synonymized Milde (1869) lanuginosae section of the subgenus *Osmundopteris* under the subgenus *JapanoBotrychium*. Study of Kato (1987) reassessed *Sceptridium* and *Japanotrychium* at genus level later downgraded by Wagner (1990) who considered every Botrychioideae genus as subgenus of *Botrychium*. Additionally, Wagner introduced a new section *HiemoBotrychium* within the subgenus *Sceptridium* based on idioblasts cells particularities (Arnott 1960), phenology and sporophyte morphology which was later elevated at genus level by Škoda (1997) under the name of *Holubiella*, but it was mostly ignored by taxonomists. Wagner's vision of *Botrychium* taxonomy persisted in the 1993 treatment of the Flora of North America (Wagner and Wagner 1993) and in the last treatment of the Flora of China (Xianchun et al. 2013). However, advancement in the accessibility of molecular data for ferns embraced the taxonomic vision of nineteenth and early twenty centuries' botanists.

Molecular phylogenetic studies based on plastid markers supported four genera within the Botrychioideae (Sahashi 1983; Hauk et al. 2003). The paraphyly of *Botrypus* previously suggested by Imaichi (1989) and Hauk (2003), was supported by a comprehensive plastid phylogeny of the Ophioglossaceae family recently published which led to the segregation of a fifth genus: *Sahashia* (Zhang et al. 2020). Zhang *et al.* (2020) also argued for *Holubiella* recognition as *Sceptridium* sister clade. However, phylogenetic analysis of *Sceptridium* clade including more than a hundred accessions challenged Zhang *et al.* (2020) results recovering *Sceptridium lunariodes* (*Holubiella lunariodes*) within *Sceptridium* clade (DuPasquier unpubl. data). Among the currently five recognised Botrychioideae genera, *Botrychium* exhibits a high species diversity compared to the other, of which two are monospecific (*Sahashia*, *Botrypus strictus*;

Botrypus, *B. virginianus*), one contains two species (*JapanoBotrychium*) and one is represented by about 15 species (*Sceptridium*) (Stensvold 2008). *Botrychium* is a species-rich genus containing 33 recognized species (PPG I 2016; Dauphin et al. 2017; Popovich et al. 2020), widely distributed in the northern temperate and boreal regions, with outposts in southernmost South America, Australia, and New Zealand. Molecular phylogenetic studies have recognized three well delimited clades within the genus *Botrychium*, one of which is the Lunaria group (Lunaria clade in Hauk *et al.* 2012 and Dauphin *et al.* 2017 sense) a notorious species complex.

The complexity of common moonwort ferns

The Lunaria group, also called the *Botrychium lunaria* species complex, is a group of closely related moonworts, having all a once-pinnate trophophore with basal pinnae outline from fan-shaped (flabellate) to lunate (Stensvold 2008; Farrar 2011). *B. lunaria* was long considered to be one morphologically variable and geographically widely distributed species (Milde 1869; Clausen 1938). However, detailed morphological and the genetic studies in North American showed that *B. lunaria* contained several distinct lineages (Underwood 1903; Wagner and Wagner 1981; Stensvold et al. 2002; Stensvold 2008) of which four were recognized (*B. tunux*, *B. crenulatum*, *B. neolunaria*, and B. *yaaxudakeit*) and one considered as a synonym of *B. lunaria* (*B. onondagense*). Allozyme markers became widely used for the identification of diversity within the *B. lunaria* group, and the investigation area was extended to northern Europe where additional taxa were discovered (*B. nordicum*, *B. lunaria* var. *melzeri*) (Stensvold 2008; Farrar and Stensvold 2017). Meanwhile, plastid phylogenetic analysis of western and central Europe also supported the existence of multiple lineages (Dauphin et al. 2014, 2017; Maccagni et al. 2017). These last findings questioned our understanding of the *B. lunaria* group diversity on its global distribution range.

The geographical distribution of *B. lunaria* has at least three known centers of species abundance; The Rocky Mountains and Great lakes regions in North America (Hauk et al. 2012) and the Alps in Europe (Maccagni et al. 2017). Recently, a plastid-based phylogenetic analysis (Dauphin et al. 2017) suggested that continental Europe could be the main center of diversity. However, the one past colonization route identified between Eurasia and North America was located in Beringia, which was also suggested to be a diversity center (Stensvold 2008). Beringia was a landmass connecting the eastern Siberia to Alaska. High diversity is probably to be found in Asia; a region poorly investigated so far.

Open questions remain regarding the speciation processes within the *B. lunaria* group. Polyploidization appears to be a major factor in some clades of *Botrychium*, and complex polyploids networks have been uncovered (Wagner and Lord 1956; Wagner 1993; Williams and Waller 2012; Dauphin et al. 2016), but the relevance of polyploidization in the *B. lunaria* group remains unclear. About half of the *Botrychium* species are polyploids (allotetraploids and allohexaploids), and *de novo* formation of F1 hybrids in mixed populations were found to be common in North America (Wagner et al. 1985). However, only one allopolyploid species (*B. yaaxudakeit*) is known in the *B. lunaria*, and no further evidence were found despite of extensive allozyme studies of central European populations (Dauphin et al. 2020). Nevertheless, diploid hybrids were detected by allozyme markers throughout Eurasia and North America between *B. lunaria* species (Stensvold 2008) suggesting that reticulate evolution might be at skate. Quaternary climatic variation was suggested to have played a role in the *B. lunaria* diversification. Stensvold (2008) proposed that *B. lunaria* group diversification may have been driven by successive isolation, migration, and dispersal events imposed by the Quaternary glacial-interglacial cycle. Quaternary climatic oscillations are known to have shape many alpine and boreal plants inter-specific diversity through expansion and retraction of the distribution range and adaptation to specific soil conditions (Schönswetter et al. 2005; Alvarez et al. 2009; Allen et al. 2012; Lafontaine et al. 2018; Ehlers et al. 2018). *B. lunaria* is often associated to open habitats with acid soil such as lean acid pasture (Nardion alliance) (Delarze et al. 2015) or acid alpine meadow (*Caricion curvulae)*, but the species can also be found on calcareous meadows (Seslerion alliance; (Sandoz 2015) or even in closed habitats such as forest and bush (Xianchun et al. 2013); pers. observations). In North America, *B. lunaria* species mostly occurs on basic substrate and not in acidophilic plant communities (Don Farrar, pers. comm.). Maccagni *et al.* (2017) suggested that habitats are correlated to haplotypes in the Alps. Investigating further the relationship between ecological factors and the *B. lunaria* lineages distribution would certainly bring insight into speciation processes driving this group diversification.

Major uncertainty remains regarding the actual diversity enclosed within the *Botrychium lunaria* group and the processes underlying the group's diversification. The genomic resources available for *Botrychium* are limited to a few plastids and nuclear markers. Acquisition of genome-wide dataset was hindered by the absence of information on the *Botrychium* genome size and structure. Transcriptomic resources of a closely related genus are available, but difficult to use because of the deep divergence time between genera. The work presented in this thesis investigates these gaps and improves our understanding of the *B. lunaria* group evolution, but also focuses on the development of novel genomic resources to overcome the challenges link to reticulate evolution and polyploidization events.

Outline of the chapters

The PhD thesis focused on the understanding of the common moonwort ferns evolutionary history by a multidisciplinary approach. The thesis is divided into three chapters; each embracing a different approach; and a general discussion section. **Chapter 1** contributes to implement new genomic tools for the population genetic and phylogenomic studies of *Botrychium* ferns. **Chapter 2** defines the common moonwort diversity on a global scale and investigates the potential underlying speciation processes. **Chapter 3** proposes taxonomic novelties in the *Botrychium lunaria* group with the description of newly characterized species.

One of the main limitations of the studies of fern evolutionary histories, including the common moonwort, is the availability of nuclear genomic resources. The paucity of such genome-wide datasets is mostly due to the complexity and the size of fern genomes. The recent breakthrough came with the 1KP project with the sequencing of 1,000 plant transcriptomes and their use to investigate fern evolution was the starting point of **the chapter 1**. With the motivation of providing genomic tools for *Botrychium* ferns, we applied a similar transcriptomic approach as the 1KP project. We sequenced and analyzed 12 individuals from two Swiss populations and assembled a nearly complete transcriptome based on one reference individual. We then built a transcriptome-wide single nucleotide polymorphism (SNP) dataset and tested its potential as a tool for exploring population structure. We also used the new assembly, combined with 1 KP available transcriptomes, to perform phylogenomic analyses of the fern tree of life.

The multiple lineages recently uncovered among the European populations of the *B. lunaria* group raises doubts about our true understanding of the group's global diversity. **Chapter 2** aims to assess this group's diversity on a global scale and the implications of climatic and ecological factors in the speciation process. This chapter primarily leans on a huge data collection covering all the know distribution range of the *B. lunaria* group. We built this dataset by combining published information, herbarium collections material and extensive field campaigns. We performed phylogenetic analyses based on four non-coding plastid regions to identify evolutionary units and estimate their divergence time. We then explored the climatic niches and edaphic preferences of each recovered lineage using the CHELSA database and soil pH measurements.

The phylogenetic analysis performed in chapter 2 confirmed previously found lineages and unraveled new ones, most of which lacked a proper taxonomic status. In **chapter 3**, we examined the morphological features of the uncharacterized species and subspecies clade levels as defined in chapter 2. Based on the molecular, ecological, and morphological differences, we propose the description of 3 new taxa and give insights into lineages requiring further attention before any possible taxonomical recognition.

Finally, I discuss the main results, limitations, and perspectives of this work in a general discussion section.

Literature introduction

Adams DC, Rohlf FJ (2000) Ecological character displacement in Plethodon: biomechanical differences found from a geometric morphometric study. Proc Natl Acad Sci 97:4106–4111. https://doi.org/10.1073/pnas.97.8.4106

Allen GA, Marr KL, McCormick LJ, Hebda RJ (2012) The impact of Pleistocene climate change on an ancient arctic–alpine plant: multiple lineages of disparate history in Oxyria digyna. Ecol Evol 2:649–665. https://doi.org/10.1002/ece3.213

Alvarez N, Thiel-Egenter C, Tribsch A, et al (2009) History or ecology? Substrate type as a major driver of patial genetic structure in Alpine plants. Ecol Lett 12:632–640. https://doi.org/10.1111/j.1461-0248.2009.01312.x

Arnott HJ (1960) Tracheoidal Idioblasts in *Botrychium*. Trans Am Microsc Soc 79:97–103. https://doi.org/10.2307/3223978

Barrington DS (1993) Ecological and historical factors in fern biogeography. J Biogeogr 20:275–279. https://doi.org/10.2307/2845635

Berry CM, Hilton J (2006) Givetian (Middle Devonian) cladoxylopsid 'ferns' from Orkney, northern Scotland. Earth Environ Sci Trans R Soc Edinb 97:65–73. https://doi.org/10.1017/S0263593300001401

Bold HC, Alexopoulos CJ, Delevoryas T (1987) Morphology of plants and fungi. undefined

Bozukov V, Tsenov B, Vatsev M (2010) A first find of *Botrychium* (Ophioglossaceae) in Bulgarian palaeomacroflora. Comptes Rendus Académie Bulg Sci Sci Mathématiques Nat 63:889–892

Campbell DH (1908) Symbiosis in Fern Prothallia. Am Nat 42:154–165. https://doi.org/10.1086/278916

Clarke JT, Warnock RCM, Donoghue PCJ (2011) Establishing a time-scale for plant evolution. New Phytol 192:266–301. https://doi.org/10.1111/j.1469-8137.2011.03794.x

Clausen RT (1938) A Monograph of the Ophioglossaceae. Mem Torrey Bot Club 19:1–177

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 42:620–639. https://doi.org/10.1600/036364417X696483

Dauphin B, Grant J, Mráz P (2016) Ploidy level and genome size variation in the homosporous ferns *Botrychium* s.l. (Ophioglossaceae). Plant Syst Evol 302:575–584. https://doi.org/10.1007/s00606-016-1285-7

Dauphin B, Grant JR, Farrar DR (2020) Outcrossing mating system of the early-divergent moonwort fern (*Botrychium lunaria*, Ophioglossaceae) revealed in the european Alps. Int J Plant Sci 181:926–936. https://doi.org/10.1086/710491

Dauphin B, Grant JR, Farrar DR, Rothfels CJ (2018) Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. Mol Phylogenet Evol 120:342–353. https://doi.org/10.1016/j.ympev.2017.11.025

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. https://doi.org/10.3732/ajb.1300154

Delarze R, Gonseth Y, Eggenberg S, Vust M (2015) Guide des milieux naturels de Suisse, 3rd edn. Rossolis

Des Marais DL, Smith AR, Britton DM, Pryer KM (2003) Phylogenetic relationships and evolution of extant horsetails, *Equisetum*, based on chloroplast DNA sequence data (rbcL and trnL-F). Int J Plant Sci 164:737–751. https://doi.org/10.1086/376817

Dobzhansky T (1950) Mendelian populations and their evolution. Am Nat 84:401–418. https://doi.org/10.1086/281638

Ehlers J, Gibbard PL, Hughes PD (2018) Chapter 4 - Quaternary glaciations and chronology. In: Menzies J, van der Meer JJM (eds) Past Glacial Environments (Second Edition). Elsevier, pp 77–101

Farrar DR (1998) Population genetics of moonwort *Botrychium*. In N. Berlin, P. Miller, J. Borovansky, U. S. Seal, and O. Byers [eds.], Population and habitat viability assessment for the goblin fern (*Botrychium mormo*). The Conservation Breeding Specialist Group

Farrar DR (2011) Systematics and taxonomy of genus *Botrychium*. https://www.herbarium.iastate.edu/moonwort-botrychium-systematics

Farrar DR, Stensvold MC (2017) Observations on bipolar disjunctions of moonwort ferns (*Botrychium*, Ophioglossaceae). Am J Bot. https://doi.org/10.3732/ajb.1700166

Fiz-Palacios O, Schneider H, Heinrichs J, Savolainen V (2011) Diversification of land plants: insights from a family-level phylogenetic analysis. BMC Evol Biol 11:341. https://doi.org/10.1186/1471-2148-11-341

Grusz AL (2016) A current perspective on apomixis in ferns. J Syst Evol 54:656–665. https://doi.org/10.1111/jse.12228

Hanušová K, Čertner M, Urfus T, et al (2019) Widespread co-occurrence of multiple ploidy levels in fragile ferns (*Cystopteris fragilis* complex; Cystopteridaceae) probably stems from similar ecology of cytotypes, their efficient dispersal and inter-ploidy hybridization. Ann Bot 123:845–855. https://doi.org/10.1093/aob/mcy219

Haufler CH, Pryer KM, Schuettpelz E, et al (2016) Sex and the single gametophyte: revising the homosporous vascular plant life cycle in light of contemporary research. BioScience 66:928–937. https://doi.org/10.1093/biosci/biw108

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. https://doi.org/10.1600/036364412X635377

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

Hidalgo O, Pellicer J, Christenhusz MJM, et al (2017) Genomic gigantism in the whisk-fern family (Psilotaceae): *Tmesipteris obliqua* challenges record holder Paris japonica. Bot J Linn Soc 183:509–514. https://doi.org/10.1093/botlinnean/box003

Imaichi R (1989) Early leaf development and leaf sheath formation of *Botrychium strictum* and *B. virginianum* (Ophioglossaceae). Ann Bot 63:249–256

Johnson-Groh CL (1998) Population demographics, underground ecology and phenology of *Botrychium mormo*. in N. Berlin, P. Miller, J. Borovansky, U. S. Seal, and O. Byers, eds. Population and habitat viability assessment (PHVA) for the goblin fern (*Botrychium mormo*) Final Report. Conservation Biology Specialist Group, Apple Valley, MN., 103-108.

Joppa LN, Roberts DL, Pimm SL (2011) How many species of flowering plants are there? Proc R Soc B Biol Sci 278:554–559. https://doi.org/10.1098/rspb.2010.1004

Kato M (1978) On the sporangia of the Ophioglossaceae. Acta Phytotaxa Geobot 29:25–32

Knie N, Fischer S, Grewe F, et al (2015) Horsetails are the sister group to all other monilophytes and Marattiales are sister to leptosporangiate ferns. Mol Phylogenet Evol 90:140–149. https://doi.org/10.1016/j.ympev.2015.05.008

Kumar S, Stecher G, Suleski M, Hedges SB (2017) TimeTree: A resource for timelines, timetrees, and tivergence times. Mol Biol Evol 34:1812–1819. https://doi.org/10.1093/molbev/msx116

Lafontaine G de, Napier JD, Petit RJ, Hu FS (2018) Invoking adaptation to decipher the genetic legacy of past climate change. Ecology 99:1530–1546. https://doi.org/10.1002/ecy.2382

Leebens-Mack JH, Barker MS, Carpenter EJ, et al (2019) One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574:679–685. https://doi.org/10.1038/s41586-019-1693-2

Lehtonen S, Silvestro D, Karger DN, et al (2017) Environmentally driven extinction and opportunistic origination explain fern diversification patterns. Sci Rep 7:4831. https://doi.org/10.1038/s41598-017-05263-7

Li F-W, Brouwer P, Carretero-Paulet L, et al (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants 4:460–472. https://doi.org/10.1038/s41477-018-0188-8

Lovis JD (1958) An evolutionary study of the fern *Asplenium trichomanes*. PhD thesis, University of Leeds

Lu J-M, Zhang N, Du X-Y, et al (2015) Chloroplast phylogenomics resolves key relationships in ferns. J Syst Evol 53:448–457. https://doi.org/10.1111/jse.12180

Lyon HL (1905) A new genus of Ophioglossaceae. Bot Gaz 40:455–458

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. Alp Bot 127:185–196. https://doi.org/10.1007/s00035-017-0188-8

Magallón S, Hilu KW, Quandt D (2013) Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. Am J Bot 100:556–573. https://doi.org/10.3732/ajb.1200416

Marchant DB, Sessa EB, Wolf PG, et al (2019) The C-Fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. Sci Rep 9:1–14. https://doi.org/10.1038/s41598-019-53968-8

Masamune (1931). J Soc Trop Agric 3:

Mayr E (1940) Speciation phenomena in birds. Am Nat 74:249–278. https://doi.org/10.1086/280892

Mehltreter K, Walker LR, Sharpe JM, et al (2010) Fern ecology. Cambridge University Press

Meza Torres MSF (2013) Ophioglossaceae: an exomorphological approach with emphasis in South American species. Indian Fern J 30:192–205

Michaux A (1803) Flora Boreali-Americana, sistens caracteres plantarum, Parisiis et Argentorati. Levrault

Milde J (1869) *Botrychium* Monographia. Zool-Bot Ges Wien Verhandlungen 19:55–190

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. https://doi.org/10.1093/aob/mcf167

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

Paton AJ, Brummitt N, Govaerts R, et al (2008) Towards target 1 of the global strategy for plant conservation: a working list of all known plant species—progress and prospects. TAXON 57:602–611. https://doi.org/10.2307/25066027

Pinson JB, Chambers SM, Nitta JH, et al (2016) The separation of generations: biology and biogeography of long-Lived sporophyteless fern gametophytes. Int J Plant Sci 178:1–18. https://doi.org/10.1086/688773

Popovich SJ, Farrar DR, Gilman AV (2020) *Botrychium furculatum* (Ophioglossaceae), a new moonwort species from the Rocky Mountains of North America. Am Fern J 110:165–182. https://doi.org/10.1640/0002-8444-110.4.165

PPG I (2016) A community-derived classification for extant lycophytes and ferns. J Syst Evol 54:563–603. https://doi.org/10.1111/jse.12229

Presl CB (1845) Supplementum tentaminis pteridographiae, continens genera et species ordinum dic- torum Marattiaceae, Ophioglossaceae, Osmundaceae, Schizaeaceae et Lygodiaceae, e Typographia Caes. reg. aulica filiorum Amadei Haase. Prague

Pryer KM, Schneider H, Smith AR, et al (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409:618–622. https://doi.org/10.1038/35054555

Pryer KM, Schuettpelz E, Wolf PG, et al (2004) Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. Am J Bot 91:1582–1598. https://doi.org/10.3732/ajb.91.10.1582

Qi X, Kuo L-Y, Guo C, et al (2018) A well-resolved fern nuclear phylogeny reveals the evolution history of numerous transcription factor families. Mol Phylogenet Evol 127:961–977. https://doi.org/10.1016/j.ympev.2018.06.043

Qiu Y, Li L, Wang B, et al (2007) A nonflowering land plant phylogeny inferred from nucleotide sequences of seven chloroplast, mitochondrial, and nuclear genes. Int J Plant Sci 168:691–708. https://doi.org/10.1086/513474

Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science 255:1697–1699. https://doi.org/10.1126/science.255.5052.1697

Ravi V, Khurana JP, Tyagi AK, Khurana P (2008) An update on chloroplast genomes. Plant Syst Evol 271:101–122. https://doi.org/10.1007/s00606-007-0608-0

Read DJ, Duckett JG, Francis R, et al (2000) Symbiotic fungal associations in 'lower' land plants. Philos Trans R Soc Lond B Biol Sci 355:815–831. https://doi.org/10.1098/rstb.2000.0617

Reichstein T (1981) Hybrids in European Aspleniaceae (Pteridophyta): significance, recognition, genome analysis, and fertility; checklist of species and hybrids; description of some new hybrids and cytology of several already known hybrids. Bot Helvetica 89–139

Rothfels CJ, Johnson AK, Windham MD, Pryer KM (2014) Low-copy nuclear data confirm rampant allopolyploidy in the Cystopteridaceae (Polypodiales). TAXON 63:1026–1036. https://doi.org/10.12705/635.32

Rothfels CJ, Larsson A, Li F-W, et al (2013) Transcriptome-mining for single-copy nuclear markers in ferns. PLOS ONE 8:e76957. https://doi.org/10.1371/journal.pone.0076957

Rothfels CJ, Li F-W, Sigel EM, et al (2015) The evolutionary history of ferns inferred from 25 low-copy nuclear genes. Am J Bot 102:1089–1107. https://doi.org/10.3732/ajb.1500089

Rothwell GW, Stockey RA (1989) Fossil Ophioglossales in the Paleocene of western North America. Am J Bot 76:637–644. https://doi.org/10.2307/2444411

Sahashi N (1983) Morphological and taxonomical studies on Ophioglossales in Japan and the adjacent regions. New taxa of *Sceptridium* in Isl. Oshima, the Izu Islands. J Jpn Bot 58:16–23

Sandoz A (2015) Influence of ecological factors on arbuscular mycorrhizal fungi (AMF) of *Botrychium lunaria*. Master, Neuchâtel

Schneider H, Schuettpelz E, Pryer KM, et al (2004) Ferns diversified in the shadow of angiosperms. Nature 428:553–557. https://doi.org/10.1038/nature02361

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. https://doi.org/10.1111/j.1365-294X.2005.02683.x

Schuettpelz E, Pryer KM (2009) Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. Proc Natl Acad Sci 106:11200–11205. https://doi.org/10.1073/pnas.0811136106

Shen H, Jin D, Shu J-P, et al (2018) Large-scale phylogenomic analysis resolves a backbone phylogeny in ferns. GigaScience 7:1–11. https://doi.org/10.1093/gigascience/gix116

Shepherd LD, Perrie LR, Brownsey PJ (2008) Low-copy nuclear DNA sequences reveal a predominance of allopolyploids in a New Zealand *Asplenium* fern complex. Mol Phylogenet Evol 49:240–248. https://doi.org/10.1016/j.ympev.2008.06.015

Škoda B (1997) Taxonomic comments on the "Flora of NorthAmerica north of Mexico", vol. 2, with some new nomenclatural combinations for Pteridophyta. Preslia 68:341–359

Small JK, Barnhart JH (1938) Ferns of the southeastern states. Descriptions of the fern-plants growing naturally in the states south of the Virginia-Kentucky state line and east of the Mississippi river., The Science Press Printing Co. Lancaster, Pa.

Stankowski S, Ravinet M (2021) Defining the speciation continuum. Evolution 75:1256–1273. https://doi.org/10.1111/evo.14215

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

Stensvold MC, Farrar DR, Johnson-Groh C (2002) Two new species of Moonworts (*Botrychium* subg. *Botrychium*) from Alaska. Am Fern J 92:150–160. https://doi.org/10.1640/0002-8444(2002)092[0150:TNSOMB]2.0.CO;2

Sun B-Y, Kim MH, Kim CH, Park C-W (2001) *Mankyua* (Ophioglossaceae): a new fern genus from Cheju Island, Korea. TAXON 50:1019–1024. https://doi.org/10.2307/1224718

Swartz O (1801) Filicum, ordine systematico redactarum. J Für Bot 1800:4–120

Testo W, Sundue M (2016) A 4000-species dataset provides new insight into the evolution of ferns. Mol Phylogenet Evol 105:200–211. https://doi.org/10.1016/j.ympev.2016.09.003

Tryon RM, Tryon AF (1982) Ophioglossaceae. In: Tryon RM, Tryon AF (eds) Ferns and Allied Plants: With Special Reference to Tropical America. Springer, New York, NY, pp 25–39

Underwood LM (1903) An index to the described species of *Botrychium*. Bull Torrey Bot Club 30:47–48

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

Wagner WH (1990) Ophioglossaceae. In: The families and genera of vascular plants: pteridophytes and gymnosperms, vol. 1., Springer-Verlag. Kramer, K. U. & P. S. Green, New-York, pp 193–197

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

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. https://doi.org/10.2307/2482592

Wagner WH, Wagner FS (1993) Ophioglossaceae Agardh. In: Flora of North America - Pteridophytes and gymnosperms, Oxford University Press. The Editorial Committee of the Flora of North America, Oxford, pp 85–106

Wagner WH, Wagner FS, Beitel JM (1985) Evidence for interspecific hybridisation in pteridophytes with subterranean mycoparasitic gametophytes. Proc R Soc Edinb Sect B Biol Sci 86:273–281. https://doi.org/10.1017/S026972700000823X

Wickett NJ, Mirarab S, Nguyen N, et al (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc Natl Acad Sci 111:E4859–E4868. https://doi.org/10.1073/pnas.1323926111

Wikström N, Kenrick P (2001) Evolution of Lycopodiaceae (Lycopsida): Estimating Divergence Times from rbcL Gene Sequences by Use of Nonparametric Rate Smoothing. Mol Phylogenet Evol 19:177–186. https://doi.org/10.1006/mpev.2001.0936

Williams EW, Farrar DR, Henson D (2016) Cryptic speciation in allotetraploids: Lessons from the *Botrychium matricariifolium* complex. Am J Bot 103:740–753. https://doi.org/10.3732/ajb.1500281

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. https://doi.org/10.1086/664711

Winther JL, Friedman WE (2007) Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). Am J Bot 94:1248–1255. https://doi.org/10.3732/ajb.94.7.1248

Woodhams MD, Lockhart PeterJ, Holland BR (2016) Simulating and summarizing sources of gene tree incongruence. Genome Biol Evol 8:1299–1315. https://doi.org/10.1093/gbe/evw065

Xianchun Z, Lui Q, Sahashi N (2013) Ophioglossacae. In: Flora of China, Beijing: Science Press&St. Louis: Missouri Botanical Garden Press. Zhengyi, W., P. H. Raven & H. Deyuan, St. Louis

Yatskievych G (2003) Pteridophytes (Ferns). In: eLS. American Cancer Society

Zhang L, Fan X-P, Petchsri S, et al (2020) Evolutionary relationships of the ancient fern lineage the adder's tongues (Ophioglossaceae) with description of *Sahashia* gen. nov. Cladistics 36:380–393. https://doi.org/10.1111/cla.12408

Zhong B, Fong R, Collins LJ, et al (2014) Two new fern chloroplasts and decelerated evolution linked to the long generation time in tree ferns. Genome Biol Evol 6:1166–1173. https://doi.org/10.1093/gbe/evu087

CHAPTER 1: Transcriptome-wide SNPs for *Botrychium lunaria* ferns enable fine-grained analysis of ploidy and population structure

Vinciane Mossion[1], Benjamin Dauphin[1,2], Jason Grant[1], Michael Kessler[3], Niklaus Zemp[4], Daniel Croll[1 *]

[1] Laboratory of Evolutionary Genetics, University of Neuchâtel, Neuchâtel, Switzerland

[2] Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

[3] Department of Systematic and Evolutionary Botany, University of Zürich, Zurich, Switzerland

[4] Genetic Diversity Centre (GDC), ETH Zurich, Zurich, Switzerland

This chapter is published in *Molecular Ecology Resources*, 22(1), 254-271, 2021. https://doi.org/10.1111/1755-0998.13478

Abstract

Ferns are the second most diverse group of land plants after angiosperms. Extant species occupy a wide range of habitats and contribute significantly to ecosystem functioning. Despite the importance of ferns, most taxa are poorly covered by genomic resources and within-species studies based on high-resolution markers are entirely lacking. The genus *Botrychium* belongs to the family Ophioglossaceae, which includes species with very large genomes and chromosome numbers (*e.g. Ophioglossum reticulatum* 2n = 1,520). The genus has a cosmopolitan distribution with 35 species, half of which are polyploids. Here, we establish a transcriptome for *Botrychium lunaria* (L.) Sw., a diploid species with an extremely large genome of about ~19.0–23.7 Gb. We assembled 25,677 high-quality transcripts with an average length of 1,333 bp based on deep RNA-sequencing of a single individual. We sequenced eleven additional transcriptomes of individuals from two populations in Switzerland, including the population of the reference individual. Based on read mapping to reference transcript sequences, we identified 374,463 single nucleotide polymorphisms (SNPs) segregating among individuals for an average density of 14 SNPs per kilobase. We found that all 12 transcriptomes were most likely from diploid individuals. The transcriptome-wide markers provided unprecedented resolution of the population genetic structure, revealing substantial variation in heterozygosity among individuals. We also constructed a phylogenomic tree of 92 taxa representing all fern orders to ascertain the placement of the genus *Botrychium*. High-quality transcriptomic resources and SNP sets constitute powerful population genomic resources to investigate the ecology, and evolution of fern populations.

Keywords: Ferns, Ophioglossaceae, population genomics, transcriptome assembly, SNP development, ploidy

Introduction

Ferns (Polypodiopsida) are a highly diverse group of cosmopolitan vascular plants and present in most climates. Around 85% of the species richness is found in the tropics (Page, 2002; Ranker & Haufler, 2008), where ferns have diversified into a multitude of habitats including deserts, grasslands, forest understory, mountainous regions, and aquatic environments (Mehltreter et al., 2010). Ferns play key roles in ecosystem functioning, including serving as a habitat for invertebrates (Ellwood & Foster, 2004), shaping plant recolonization of disturbed habitats (Walker, 1994) and influencing the composition of tree species communities (George & Bazzaz, 1999a, 1999b). Most ferns are characterized by two independent life stages (*i.e.* gametophytic and sporophytic; Pinson, Chambers, Nitta, Kuo, & Sessa, 2016). Ferns have complex and idiosyncratic life cycles. For example, many fern species are capable of versatile reproductive modes (Sessa, Zimmer, & Givnish, 2012) including apomixis (Grusz, 2016), sporophytic and gametophytic selfing, and outcrossing (Barker & Wolf, 2010; Haufler et al., 2016). Flexibility in the reproductive mode plays a crucial role in the foundation and persistence of isolated populations from a few spores, and thus likely contributed to the evolution of fern lifestyles. In addition, the reproductive mode likely impacts the tempo of evolution and lineage diversification.

Phylogenetic analyses have resolved the position of ferns as the sister group to seed plants (Pryer et al., 2001) and as the second earliest diverging lineage of vascular land plants (Pryer et al., 2004; Raubeson & Jansen, 1992). The root age of ferns has been estimated to be ca. 360–431 million years, underlining the deep divergence among ferns lineages (Des Marais, Smith, Britton, & Pryer, 2003; Lehtonen et al., 2017; Magallón, Hilu, & Quandt, 2013; Pryer et al., 2004; Qi et al., 2018; Rothfels et al., 2015; Testo & Sundue, 2016; Wikström & Kenrick, 2001; Zhong, Fong, Collins, McLenachan, & Penny, 2014). Fern phylogenies have been mostly established based on chloroplast markers (Grewe, Guo, Gubbels, Hansen, & Mower, 2013; Kuo, Li, Chiou, & Wang, 2011; Lu, Zhang, Du, Wen, & Li, 2015; Rai & Graham, 2010; Schuettpelz & Pryer, 2007; Testo & Sundue, 2016), sometimes combined either with mitochondrial markers (Knie, Fischer, Grewe, Polsakiewicz, & Knoop 2015) or nuclear markers (Pryer et al., 2001, 2004) or both (Qiu et al., 2007). The heavy reliance on plastid and mitochondrial markers largely prevents analyses of reticulation events (*i.e.* hybridization and introgression). Yet polyploidizations occur at one third of the fern speciation rate (Wood et al., 2009). In contrast, a series of recent phylogenomic studies have highlighted the power of transcriptome-based approaches (Leebens-Mack et al., 2019; Qi et al., 2018; Rothfels et al., 2013, 2015; Shen et al., 2018; Wickett et al., 2014). Expanding genome- or transcriptome-wide datasets are essential to overcome challenges imposed by high degrees of paralogy in many fern genomes.

The first insights into the structure of fern genomes were provided by the complete genome sequences of *Azolla filiculoides* and *Salvinia cucullata* (Li et al., 2018). These two genera are known for their small genome sizes (*A. filiculoides* 753 Mb and *S. cucullata* 255 Mb; Li et al., 2018). For example, the genome sizes of *A. microphylla* and *S. molesta* were estimated at 1C = 0.77 pg or ~0.75 Gb and 1C = 2.25 pg, respectively (Obermayer et al., 2002; Clark et al., 2016). Conversely, many fern genomes are of enormous size (*e.g., Tmesipteris obliqua 1C = 150.61 pg;* Hidalgo, Pellicer, Christenhusz, Schneider, & Leitch, 2017). Major progress in the establishment of genomic resources was made with the sequencing of 73 fern transcriptomes (Carpenter et al., 2019; Leebens-Mack et al., 2019). Such datasets have been successfully used to develop single copy nuclear markers to resolve deep evolutionary relationships among ferns (Rothfels et al., 2013, 2015). Transcriptome assemblies are also an important tool to develop genotyping approaches and overcome challenges associated with extremely large fern genomes (Bennett & Leitch, 2001; Hanson & Leitch, 2002; Obermayer et al., 2002). These approaches typically reduce genome complexity but still provide sufficient polymorphic markers to conduct population genomics analyses (Seeb et al., 2011). Establishing transcriptomic datasets for understudied fern clades will bring new insights into fern diversification.

An important genus lacking transcriptomic resources is *Botrychium* belonging to subclass Ophioglossidae (PPG I, 2016). This subclass is characterized by a subterranean gametophytic stage (Field et al., 2015; Jeffrey, 1898; Winther & Friedman, 2007) and extremely large and complex genomes (*e.g.*, *Ophioglossum petiolatum* 1C = 65.55pg; Obermayer *et al.*, 2002). *Botrychium* occurs in open habitats on nearly every continent across a broad temperate and boreal distribution. The genus is divided into three monophyletic clades as estimated by non-coding plastid markers (Simplex-Campestre, Lanceolatum and Lunaria; Dauphin, Farrar, Maccagni, & Grant, 2017), containing 35 recognized taxa (PPG I, 2016). The challenge of identifying *Botrychium* taxa based on morphology is underlined by claims of cryptic species (Clausen, 1938; Hauk, 1995). Ambiguous morphologies are sometimes caused by polyploidization which is a major driver of speciation as half of the known *Botrychium* species are allopolyploids (Dauphin, Grant, Farrar, & Rothfels, 2018). Nuclear markers have resolved the parental origins of these allopolyploid taxa and have provided insights into the genus radiation (Dauphin et al., 2018). Additionally, the reconstruction of maternal lineages of *Botrychium* has revealed high genetic diversity within the Lunaria clade, highlighting the uncertainty of taxonomic assignments (Dauphin et al., 2017; Dauphin, Vieu, & Grant, 2014; Maccagni, Parisod, & Grant, 2017). Previous population genetic studies based on isozymes have shown a lack of genetic differentiation among morphologically recognized types (Williams, Farrar, & Henson, Don, 2016), and the low amount of genetic variation detected within *Botrychium* populations suggests pervasive self-fertilization (Farrar, 1998; Hauk & Haufler, 1999; Williams, 2021). Furthermore, genetic differentiation among populations and regions was found to be low, suggesting that gene flow may occur (Birkeland, Borgenkjetne, Brysting, Elven, & Alsos, 2017; Camacho & Liston, 2001; Swartz & Brunsfeld, 2002). These studies highlight the need for powerful, genome-wide marker systems to resolve population structures, life histories, and taxonomy of eusporangiate ferns.

In this study, we developed transcriptome-wide SNPs for the fern species *Botrychium lunaria* (L.) Sw. which has a large genome size of about 24.1 pg (2C; ~23.7 Gb; Veselý, Bureš, Šmarda, & Pavlíček, 2012) establishing the first such resource for ferns. By analyzing the transcriptome of 12 individuals mapped against the newly established reference transcriptome, we show the power of the discovered SNPs to resolve population-level variation in heterozygosity and to test for evidence of ploidy variation. We further demonstrate the power of transcriptome-wide markers to resolve phylogenetic relationships at the genus level and among deeply divergent fern lineages.

Material and Methods

Sampling, library preparation and sequencing

Leaf material of *B. lunaria* was obtained from three locations in Switzerland: two in the Val d'Hérens in the Pennine Alps, Mase and Forclaz, within approximately 30 km, and one at the Chasseral in the Jura Mountains (Table 1). The two regions are separated by ~120 km and the Alpine population was sampled on meadows over an altitudinal range of 1,500 to 2,400 meters (Table 1). Leaves of six individuals from Val d'Hérens and from Chasseral were collected in July 2015 and June 2017, respectively. Plant material was wrapped in aluminum foil and frozen immediately in liquid nitrogen. Total RNA was extracted from trophophores (i.e., the sterile part of the leaves) using the RNAeasy Plant Mini Kit (Qiagen) and DNA was eliminated using DNase I digestion. Total RNA was quantified using a Qubit fluorometer (Invitrogen, Thermo Fisher Scientific) with the RNA Broad-Range assay kit (Invitrogen, Thermo Fisher Scientific) and quality-checked using an Agilent 2200 Tape Station (Agilent Technologies, Inc.). Samples were diluted to 100 ng/µl in RNase free ultra-pure water before library preparation. The RNA-sequencing libraries were prepared following a TruSeq RNA library preparation protocol (Illumina, Inc.) enriching for polyadenylated RNAs. After quality assessment on an Agilent 2200 Tape Station, libraries were pooled and sequenced in 150 bp single-end mode on one lane of an Illumina HiSeq 4000 sequencer.

Absolute genome sizes were estimated by flow cytometry to assess the ploidy level of the six specimens for which frozen material was available, following the one-step methodology of Doležel et al. (1998). Frozen material was silica-dried before being processed.

De novo assembly, filtering and quality assessment

Sequencing reads were quality-checked using FastQC v. 0.11.7 (Andrews, 2010) and trimmed using Trimmomatic v. 0.38 (Bolger, Lohse, & Usadel, 2014). Reads were retained if the leading and trailing base qualities were > 5, had a 4-bp sliding window > 15, and a minimum read length of 36 bp. Trimmed reads from a single reference individual were then *de novo* assembled using Trinity v. 2.8.3 (Haas et al., 2013). The other individual transcriptomes were used for SNP calling (see below). The reference individual choice was made based on the highest trimmed transcript sequence number. We used the pseudo-alignment percentage calculated by Kallisto v. 0.45.0 (Bray, Pimentel, Melsted, & Pachter, 2016) to assess the representativeness of the raw assembly across the twelve sequenced individuals in total. Candidate coding regions were identified using TransDecoder v. 5.3.0 (Haas et al., 2013). Only transcripts with an open reading frame (ORF) of at least 100 amino acids were kept. We also retained only the longest isoform per transcript using the Trinity v. 2.8.3 toolkit. We used Diamond v. 0.9.24 (Buchfink, Xie, & Huson, 2015) to screen the transcript assembly against the NCBI non-redundant protein (nr) and UniVec databases to identify potential foreign RNA contaminants. The best hit for each transcript was assigned at the phylum-level using the R package *taxise* v. 0.9.7 (Chamberlain & Szöcs, 2013) in RStudio v. 1.2.1335 (R Development Core Team, 2020; RStudio Team, 2015). We excluded all transcripts with a best hit outside of the plant kingdom (*i.e.*, Viridiplantae). We did not filter for hits within the plant kingdom due to the scarce representation of seedless plant sequences in the NCBI nr and UniVec databases. Transcript redundancies were detected using CD-HIT-EST v.4.8.1 (Weizhong & Godzik, 2006) with a global sequence identity of 0.95% and a word size of 10. Redundant transcripts were pruned from the final assembly retaining the longest transcript per cluster. The completeness of the transcriptome assemblies was assessed using BUSCO v. 5.0.0 with the viridiplantae_odb10 database (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015). Data were visualized using the R package *ggplot2* v. 3.3.2 (Wickham, 2016).

Variant calling

We generated alignment BAM files for each individual against the transcriptome using the short read aligner Bowtie2 v. 2.4.2 (Langmead, 2010) and SAMtools v. 1.11 (Li et al., 2009). Depth coverage of the reference individual was extracted using SAMtools idxstats. Alignments were processed with HaplotypeCaller implemented in the Genome Analysis Toolkit (GATK) v. 4.1.8.1 (DePristo et al., 2011; McKenna et al., 2010; Van der Auwera et al., 2013) for single nucleotide polymorphism (SNP) calling. The resulting gvcf files were combined and genotyped using the GATK CombineGVCF and GenotypeGVCF tools, respectively. We excluded monomorphic sites from further analysis. We filtered SNPs for the number of genotyped chromosomes (AN $\geq$ 20) out of a maximum of 24 (12 diploid individuals). Quality criteria QUAL > 100, QualByDepth > 5.0, RMSMappingQuality > 20.0, MappingQualityRankSumTest retained values > -2.0 and < 2.0, and ReadPosRankSumTest and BaseQualityRankSumTest retained values > -2.0 and < 2.0 were defined following the best practices and were applied to flag low-quality loci (Suppl. Figure S1). We removed SNPs failing the above filters using VCFtools v. 0.1.16 (Danecek et al., 2011) and added a filter to retain only bi-allelic SNPs. Analyses were performed using the R packages *vcfR* v. 1.12.0 (Knaus & Grünwald, 2017) and the SNP statistics among transcripts were visualized using *ggplot2* v. 3.3.2.

Population genetics analyses

Intra-individual allele frequencies were calculated for each individual and SNP locus using the mapped read depth per allele (AD). The frequency distributions were plotted per individual. We subsampled the number of SNPs by selecting one SNP every 1,000 bp of transcriptomic sequence using VCFtools v. 0.1.16. The subsampling was performed to obtain a representative set of markers across all polymorphic transcripts. Furthermore, the subsampling reduced linkage disequilibrium among markers to reduce biases in the representation of the population genetic structure. We performed principal component analyses (PCA) and calculated the pairwise Nei's F_{ST} and the mean heterozygosity (H_e) per location and per individual (Nei, 1987). These analyses were performed using the R packages *vcfR* v. 1.12.0, *adegenet* v. 2.1.3 (Jombart & Ahmed, 2011) and *hierfstat* v. 0.5-7 (Goudet, 2005), and data were visualized using *ggplot2* v. 3.3.2.

Functional annotation

We functionally characterized encoded protein sequences based on gene ontology (GO) terms. We summarized GO terms by selecting the least redundant annotations among the 30 most frequent terms per ontology (cellular component CC, molecular function MF, and biological process BP). Analyses were performed using the Bioconductor packages *AnnotationDbi* v. 1.46.0 (Pagès, Carlson, Falcon, & Li, 2020), *GO.db* v. 3.8.2 (Carlson, 2020), *GSEABase* v. 1.34.0 (Morgan, Falcon, & Gentleman, 2020), *annotate* v. 1.62.0 (Gentleman, 2020), and data were visualized using the R package *ggplot2* v. 3.3.2.

Genus-level phylogenetic analyses

To estimate the phylogenetic placement of the twelve individuals included in our study, we retrieved sequences of four previously analyzed nuclear regions (i.e. ApPEFP_C, CRY2cA, CRY2cB, and transducin) of both diploid and polyploid *Botrychium* taxa (Dauphin et al., 2018). We searched homologous sequences in the transcriptome assembly using BLAST v. 2.9.0 (Altschul, Gish, Miller, Myers, & Lipman, 1990). If the associated transcript was found in the assembly, we used BCFtools v. 1.9 (Li, 2011) to retrieve the corresponding transcript from the 11 remaining individuals using the VCF file information. Sequence alignments were performed with MAFFT v. 7.470 under the G-INS-i strategy and default parameters (Katoh, Misawa, Kuma, & Miyata, 2015; Katoh & Standley, 2013). Multiple alignments were visually inspected and manually adjusted using Geneious v. 8.1.9 (Kearse et al., 2012). Phylogenetic trees were inferred using maximum likelihood (ML) in RAxML-NG v. 0.9.0 (Kozlov, Darriba, Flouri, Morel, & Stamatakis, 2019). We ran tree inferences with a fixed random seed of 2 under the HKY+GAMMA model based on model settings by Dauphin et al. (2018) to ensure reproducibility. The tree search was performed using 25 random- and 25 parsimony-based starting trees. The branch support was estimated using 1,000 bootstrap replicates and calculated according to the transfer bootstrap expectation matrix (Lemoine et al., 2018). The support values were depicted on the best-scoring ML tree. The tree was visualized using the R packages ape v. 5.4-1 (Paradis & Schliep, 2019) and ggtree v.2.4.0 (Yu, Smith, Zhu, Guan, & Lam, 2017).

Phylogenomic analyses

We performed phylogenomic analyses across ferns by including the newly established *B. lunaria* transcriptome for a total of 95 transcriptomes including 85 fern species (18 eusporangiates and 67 leptosporangiates), six Spermatophyta and two Lycopodiopsida (Suppl. Table S1; Leebens-Mack et al., 2019; Qi et al., 2018; Shen et al., 2018). The Spermatophyta and Lycopodiopsida species represent the outgroup in the analysis. We first performed an ortholog search using OrthoFinder v. 2.3.12 (Emms & Kelly, 2015) including the newly established *B. lunaria* transcriptome and protein sequences of other transcriptomes. We retained orthogroups shared by every member of the taxa set regardless of the gene copy number. We used ASTRAL-Pro v. 1.1.5, which was specifically developed to summarize multi-copy gene trees into a species tree (Zhang, Scornavacca, Molloy, & Mirarab, 2020). To reduce the computational load, we processed only orthogroups for which at least half of the species had a single-copy gene and excluded two orthogroups with very high paralog counts ($n \geq 700$). Sequences of the orthogroups subset were subsequently aligned with MAFFT v. 7.475 under the G-INS-i strategy and default parameters. The optimal substitution model was assessed for each orthogroup alignment using Modeltest-NG v. 0.1.6 (Darriba et al., 2020). Finally, unrooted gene trees were built using maximum likelihood (ML) in RAxML-NG v. 1.0.1. We ran tree inferences under the best model according to the Akaike information criterion (AICc) criterion with a fixed random seed of 12345. The tree search was performed on 25 random and 25 parsimony-based starting trees. The inferred gene trees (*i.e.* best ML trees) were used to estimate a species tree with ASTRAL-Pro. Branch support was calculated using local posterior probabilities (Sayyari & Mirarab, 2016). In order to display internal branch lengths on our final trees, we first added arbitrary terminal branch lengths using the python script (add-bl.py) provided with ASTRAL. Species trees were edited using the R packages *treeio* v.1.14.0 (Wang et al., 2020) and *ape* v. 5.4-1. The arbitrary terminal branch lengths were removed for the final tree visualizations.

Results

Sample collection and transcriptome assembly

We successfully analyzed transcriptomes of twelve *B. lunaria* individuals. The transcriptome sequencing produced 14.6–50.1 million reads per individual. After quality trimming, we retained 97.0–99.2% of the reads (Figure 1A, Suppl. Table S2). The highest number of high-quality reads (49.5 million) was obtained for the Chasseral individual CHA_I_1. We selected this individual to assemble a reference transcriptome for the species. The raw assembly for CHA_I_1 contained 167,306 transcripts for a total of 87,537 candidate genes before redundancy filtering. Mapping reads from all twelve individuals to the raw transcriptome assembly showed a pseudo-alignment rate (*i.e.* percentage of mapped reads) between 74.2–82.5% regardless of the population of origin (Figure 1A, Suppl. Table S2). We analyzed all assembled transcripts for the presence of high-confidence open reading frames (ORF; ≥100 amino acids). We retained 69,280 transcripts (41.4%) covering 26,139 predicted genes (Figure 1B). Next, we selected the longest transcript for each gene (Figure 1D). We screened each gene against the complete non-redundant protein and the UniVec database of NCBI and found evidence for contamination in 438 transcripts. Most contaminant sequences were associated with viruses, fungi, or bacteria (Figure 1C). We performed an additional screen for redundant transcripts using CD-HIT-EST. We identified 16 clusters and retained only a unique representative transcript per cluster for further analyses. The final assembly consisted of 25,677 unique transcripts spanning a total of 34,23 Mb. The average and median transcript lengths were 1,333 and 967 bp, respectively (Figure 1D). The N50 of the final transcriptome was 1,995 bp with an average GC content of 44.3% (Table 2). GO terms were assigned to 11,137 transcripts (43.4%; Figure 2D, Suppl. Table S3).

A highly complete B. lunaria *transcriptome*

The completeness of the assembled *B. lunaria* transcriptome assessed using BUSCO matched the completeness of complete genomes assemblies of other ferns. Importantly, the absence of fern species among the 30 species constituting the BUSCO viridiplantae_odb10 database might lead to an underestimation of the assembly completeness. We found 90.1% complete single-copy, 0.9% complete duplicates, 5.9% fragmented and 3.1% missing genes for the *B. lunaria* transcriptome (Figure 1E, Table 3). This is comparable to the only two complete genome assemblies of ferns: *A. filiculoides* with 79.3% and *S. cucullata* with 90.4% complete single-copy genes (Figure 1E, Table 3). The two Salviniaceae genomes and the *B. lunaria* transcriptome exhibited a comparable number of missing BUSCO genes (2.8–1.6% and 3.1%, respectively; Figure 1E, Table 3) and of fragmented BUSCO genes (9.9–4.7% and 5.9% respectively; Figure 1E, Table 3). The mapped reads coverage depth of the reference individual to the assembled transcripts is on average 1,649× with a range of 4 to 514,622×. Transcript coverage is strongly skewed towards low coverage, whereas ~7% show coverage >4000 reads (Figure 2A). The read coverage shows no obvious association with transcript length (Figure 2B).

Phylogenetic assignment of all analyzed transcriptomes

The analysis of nuclear barcoding loci confirmed the twelve individuals included in the transcriptome analyses as *B. lunaria*. Among the four nuclear loci previously sequenced in a broad sample of *Botrychium* species, three loci displayed sequence variation nearly exclusively in intronic sequences (Dauphin et al., 2018). Therefore, no comparison with our transcriptomic sequences was possible. We focused on the locus CRY2cA carrying sufficient informative sites in the coding regions to produce a well-supported phylogeny.

The combined dataset for CRY2cA included 67 individuals, representing 38 *Botrychium* taxa and an outgroup constituted by *Sceptridium multifidum* and *Botrypus virginianum* (Dauphin et al., 2018; Suppl. Table S4). The multiple sequence alignment contained a total of 3,579 sites and 153 patterns. The main clades Lanceolatum, Lunaria, and Simplex-Campestre were resolved as monophyletic (Suppl. Figure S2). Lanceolatum was resolved as a sister group to the Simplex-Campestre and Lunaria clade. All individuals from the Chasseral and Val d'Hérens grouped with *B. lunaria* var. *lunaria* and formed a well-supported clade.

Transcriptome-wide phylogenomic trees of ferns

Phylogenomic analyses based on the transcriptomes of 85 fern species (Leebens-Mack et al., 2019; Qi et al., 2018; Shen et al., 2018) including the new transcriptome for *B. lunaria* were congruent with previous studies. We identified 41,186 orthogroups in total (Suppl. Table S5) of which we retained 761 orthogroups and 93 species to construct a phylogenomic tree (Suppl. Table S6). The species tree was well-resolved and branch support values from the local posterior probability (LPP) were high (> 0.95; Figure 5). The species tree topology was consistent with the most recent fern phylogenies (Knie et al., 2015; Kuo et al., 2011; Leebens-Mack et al., 2019; Lu et al., 2015; Qi et al., 2018; Rai & Graham, 2010; Rothfels et al., 2015; Shen et al., 2018; Testo & Sundue, 2016) and with the current consensus classification (PPG I, 2016). The *B. lunaria* transcriptome clustered with the sister genus *Sceptridium* and the closely related genus *Botrypus*. Among the earliest divergent ferns (*i.e.* eusporangiate and early leptosporangiate), we identified Equisetales as the sister clade to all other ferns, and Marattiales as the sister clade to all leptosporangiates. Consistent with recent work, the Gleicheniales were recovered as paraphyletic and the deep eupolypod relationships remained largely unresolved in our phylogeny (Suppl. Figure S3).

Ploidy assessments and analyses of within-species transcriptome-wide polymorphism

The mapping rate of reads from each individual against the assembled transcriptome sequences varied between 82.1–86.7% (Figure 2C). The highest mapping rate was found for the individual IIT1_H5 (86.7%), which was slightly higher than the mapping rate of the reference individual used to establish the transcriptome (CHA_I_1, 85.5%; Suppl. Table S4). We found no meaningful difference in mapping rates among populations. Based on reads aligned against reference transcripts, we assessed intra-individual allele frequency distributions based on read counts. Diploid individuals should show a singular, dominant peak at 0.5 corresponding to heterozygous SNPs (Chen et al., 2018). We indeed found a major peak around a frequency of 0.5 without any secondary peaks at 0.25 or 0.75, suggesting that all individuals are likely diploid (Figure 3A). However, very recent autopolyploids would also share this signature. The broad peak shape is likely due to stochasticity in read coverage at each SNP among individuals. Low read coverage can produce individual allele frequency estimates far from 0.5. For independent evidence of ploidy levels, we used flow cytometry. The estimated absolute genome sizes varied between 19.38–20.58 pg among the six analyzed individuals (2C-values; Table 1). Hence, there is no indication for ploidy variation and the genome size estimates match previous estimates for the diploid genome of *B. lunaria* (Veselý et al., 2012). Given this evidence for diploidy, we proceeded with calling SNPs assuming a diploid state in all individuals. We recovered a total of 376,463 high-quality bi-allelic SNPs after filtering. The average number of SNPs per transcript was 17 and the maximum number was 257 (Figure 3B). The SNP density per transcript had a mean of 14, a median of 10, and a maximum of 153 SNPs per kb (Figure 3C). The median SNP density decreased slightly with transcript length (Figure 3D).

Population structure and heterozygosity

The SNP genotyping data from the twelve individuals revealed a clear population differentiation between the two main sampling locations (Figure 4A). The first principal component (PC1, 17% of total variance explained) of the principal component analysis (PCA) identified a divergent genotype in the Chasseral population (CHA_I_7, Figure 4B). The second principal component (PC2, 12%) separated the two populations Chasseral and Val d'Hérens (Figure 4B). We performed a second PCA excluding the CHA_I_7 individual and found the Chasseral population to be more diverse than Val d'Hérens (Figure 4C). We found no apparent differentiation between the two locations sampled in Val d'Hérens, but the sampling coverage was limited (*n* = 3 per location). The pairwise F_{ST} between populations was low (0.040). Mean heterozygosity was slightly higher in Val d'Hérens (H_e = 0.20) than in the Chasseral population (0.17; Figure 4D). We found similar levels of variation in individual heterozygosity among populations ranging from 0.16 to 0.21, except for CHA_I_7, which was an outlier in the PCA (Figure 4E). CHA_I_7 showed less than half the heterozygosity (H_e = 0.05) compared to other members of the same population.

Discussion

We established a high-quality transcriptome for the genus *Botrychium* filling an important gap in the coverage of eusporangiate ferns. The completeness of the transcriptomic gene space was comparable to well-assembled fern genomes. Using twelve individuals of the same species sampled in two regions, we were able to generate the first intra-species transcriptome-wide SNP dataset for ferns in general. Our analyses revealed a clear genetic distinction between the two *B. lunaria* populations and among the twelve individuals challenging the general assumption of gametophytic selfing being the dominant reproductive mode in *Botrychium* populations. A phylogenomic tree based on 761 orthologous genes confirmed the phylogenetic position of the genus among other fern lineages.

Establishment of a transcriptome for B. lunaria

Generating a representative transcriptome assembly is challenging because not all genes are expressed in all tissues and life cycle stages. Across the life cycle of ferns, gene expression patterns are largely overlapping (Sigel, Schuettpelz, Pryer, & Der, 2018), but the covered gene space is usually increased by including multiple target tissues. For *Botrychium*, only the trophophore and the sporophore were adequate tissues for the extraction of RNA since underground tissues are colonized by arbuscular mycorrhizal fungi (AMF; Winther & Friedman, 2007) leading to numerous contaminants. Because we only included trophophore tissue, the assembled transcriptome potentially underrepresents sporogenesis- and root-specific genes. Despite these challenges, our *B. lunaria* transcriptome has a fairly complete gene space in comparison to a wide range of assembled transcriptomes (Der, Barker, Wickett, dePamphilis, & Wolf, 2011; Leebens-Mack et al., 2019; Qi et al., 2018; Shen et al., 2018; Figure 1E). It is important to note that database-dependent tools such as BUSCO consistently underestimate transcriptome completeness if the database was compiled without closely related species. The challenge in using BUSCO is exemplified by the absence of ferns species in the viridiplantae dataset. The gene space of assembled fern genomes tends to show less fragmented BUSCO genes compared to the *B. lunaria* transcriptome (Li et al., 2018). However, the *B. lunaria* transcriptome is consistent with other high-quality fern transcriptomes (Leebens-Mack et al., 2019; Qi et al., 2018; Shen et al., 2018). Missing gene segments in assembled transcriptomes are often caused by uneven read depth among genes or alternative splicing complicating gene recovery. The completeness of the *B. lunaria* transcriptome compared to other fern genomes and transcriptomes provides a powerful tool for phylogenomic and population genetic analyses.

The *B. lunaria* transcriptome enables strong phylogenomic inference at different taxonomic levels, overcoming challenges associated with the small number of nuclear and chloroplast markers available for most ferns. *Botrychium* taxa have subtle morphological characteristics, hence taxonomy relies largely on molecular data (Dauphin et al., 2017, 2014; Maccagni et al., 2017; Stensvold, Farrar, & Johnson-Groh, 2002; M.C. Stensvold & Farrar, 2017). We have placed the individual *B. lunaria* transcriptomes among other closely related taxa by retrieving orthologous genes, which were previously used for phylogenetic analyses. The newly established transcriptome will enable investigations of the gene content of *Botrychium*. The tens of thousands of markers across the coding sequences allow powerful genome-wide studies of population and species differentiation. Highly dense marker sets are critical for ecological genomics investigations to for example, identify loci underlying climatic adaptation, recent gene flow events and consequences of variation in the reproductive mode. Beyond this, refined analyses of transcriptomic markers can help to retrace the evolution of the extensive ploidy variation among *Botrychium*.

As an expansion of the phylogenetic analyses within *Botrychium*, we analyzed orthologous genes across all ferns. The genus *Botrychium* was placed within the Ophioglossales with strong support (Figure 5). Furthermore, the phylogenetic placements of most fern clades were congruent with previous studies and highlighted unresolved discordances. For example, the paraphyly observed for the Gleicheniales in our species tree (Figure 5, Suppl. Figure S3) corroborates recent findings (Qi et al., 2018; Shen et al., 2018). Sparse sampling can strongly influence tree topologies. For instance, the Matoniaceae, constituting one of the three Gleicheniales families (PPG I, 2016), are not represented in phylogenomic studies. Phylogenies based only on few barcoding loci, but with a more representative sampling, identified Gleicheniales as being a monophyletic clade (Pryer et al., 2004; Schuettpelz, Korall, & Pryer, 2006; Schuettpelz & Pryer, 2007). Phylogenomic studies including the Matoniaceae will be needed to ascertain the placement of the Dipteridaceae outside of the order Gleicheniales. The resolution of our fern-wide phylogenomic analyses revealed some limitations in resolution at the within-genus level. *Equisetum hyemale* and *E. arvense* were recovered as sister species distant from *E. diffusum* even though *E. diffusum* and *E. arvense* belong to the same subgenus *Equisetum* which does not include *E. hyemale* (subgenus. *Hippochaete*; Christenhusz et al., 2021). The reduced power to resolve some phylogenetic relationships with genera might be explained by the scarce species representation within some genera and the degree of conservation of the included orthologs.

Fine-grained resolution of population structure

The transcriptome-wide SNPs revealed clear geographical structuring between the two *B. lunaria* populations sampled from Switzerland (Figure 4B, 4C). The differentiation was apparent even when subsampling SNPs containing a maximum of 1 SNP per kb to avoid biases by highly polymorphic transcripts and high linkage disequilibrium. The full set of SNPs will provide a powerful tool for ecological genomics investigations to unravel loci contribution to adaptation. It has been generally assumed *Botrychium* species show no meaningful genetic differentiation within populations (Farrar, 1998; Hauk & Haufler, 1999; Williams, 2021) or low genetic differentiation among populations (Birkeland et al., 2017; Camacho & Liston, 2001; Swartz & Brunsfeld, 2002; Dauphin, Grant, & Farrar, 2020). However, the absence of genetic differentiation reported by previous studies may well stem from low marker resolution. The transcriptome-wide markers used here showed every individual was clearly distinct, and populations showed marked differentiation. The Chasseral and Val d'Hérens populations were collected in the Jura Mountains and Pennine Alps, respectively. The two sites are 120 km apart and separated by habitats unsuitable for *B. lunaria.* Therefore, reduced gene flow and genetic differentiation among populations is expected. We found no indication of genetic substructure among the two locations Mase and Forclaz within the Val d'Hérens valley. This suggests sufficient gene flow at the local scale or recent recolonization at the upper front of the valley, which is consistent with restriction-site associated DNA sequencing-based analyses of the same field sites (Dauphin, 2017). We found no evidence for higher levels of ploidy based on mapped read depths per individual (Figure 3A). The intra-individual allele frequency distributions showed some spread around the singular peak at frequency 0.5 (expected for diploids). The spread in the estimated frequencies likely reflects noise caused by low coverage SNP positions and was observed in other systems as well (Chen et al., 2018). The individual CHA_1_7 has a less pronounced and a shifted major peak. The individual does not cluster as closely as the other genotypes of the same population and is less heterozygous. We find no difference in sequencing quality or quantity that could explain the observed peak pattern. Unequal mapping of reads matching the two alleles may be a possible reason. Such patterns may be expected for introgressed sequences of higher divergence. Absolute genome size measurements corroborate the absence of higher level ploidy. This is in accordance with a broader sampling of genome sizes of 57 individuals across the *B. lunaria* group, which likewise recovered only diploid conditions (V. Mossion & M. Kessler, unpubl. data), except for the allotetraploid species *B. yaaxudakeit*, which is only known from northwestern North America (Stensvold et al., 2002). Also, consistent with our findings, a recent study of Swiss populations based on allozyme markers found no evidence for fixed heterozygosity, a typical indicator of polyploidy (Dauphin *et al.*, 2020). However, we cannot rule out that very recent polyploidization or autopolyploidization events may remain undetected.

Population-level genetic diversity is indicative of the reproductive mode of *Botrychium* populations. Self-fertilization is considered common in homosporous ferns and includes sporophytic and gametophytic selfing (Klekowski & Baker, 1966; Gastony & Gottlieb, 1985; McCauley, Whittier, & Reilly, 1985; Soltis & Soltis, 1986; Soltis & Soltis, 1990; Soltis & Soltis, 1992; Sessa, Testo, & Watkins Jr, 2016). In sporophytic selfing, zygotes are produced by gametes from two distinct gametophytes that originate from a single sporophyte. In contrast, in gametophytic selfing, zygotes are produced from gametes of the same gametophyte. Gametophytic selfing is thought to be the main reproductive mode for the genus *Botrychium* (Hauk & Haufler, 1999) and does lead to completely homozygous plants within one generation (Klekowski & Lloyd, 1968). In a population undergoing largely gametophytic selfing, low genetic diversity would also be expected among individuals due to genetic drift. The genetic diversity of *B. lunaria* populations (Figure 4D, E) and the clear structure among sites (Figure 4B, C) suggest that sporophytic selfing or outcrossing are predominant. Only one individual in the Chasseral population showed signature of a recent gametophytic selfing event (Figure 4E). These findings contrast with the general assumption that gametophytic selfing is the dominant reproductive mode in the genus but corroborate evidence of outcrossing in Swiss populations (Dauphin et al., 2020).

Our study establishes a high-quality set of transcriptome-wide SNPs in a species of the early diverging fern genus *Botrychium.* With an estimated genome size of 19.4–24.6 (2C; ~19.0-23.7 Gb) for *B. lunaria*, the assembly of transcriptomes provides the only currently feasible approach to generate extensive genome-wide markers information. *Our* transcriptome-wide SNPs enable fine-grained demographic history analyses and investigate the co*nsequences of the complex* mating systems and putative polyploidization events. Furthermore, ecological genomics studies based on our transcriptomic markers *will provide the first opportunities* to dissect adaptation to the local environment within fern species.

Acknowledgements

We thank Frederic Sandoz for assistance in the field work. We thank Aria Minder and Sylvia Kobel for input on laboratory methods. We acknowledge the 1KP sequencing consortium for advance access to transcriptomic sequences. Emilie Chanclud, Ursula Oggenfuss and Erik Koenen provided advice on data analysis. Ursula Oggenfuss, Leen N. Abraham, Simone Fouché, Pierre-Emmanuel Du Pasquier, Rosa-Angela Ston, Erik Koenen and Giacomo Zilio provided helpful comments on a previous manuscript version. Rosa-Angela Ston mounted herbarium vouchers. Elke Kessler conducted the flow cytometry analyses. Data produced and analysed in this paper were generated in collaboration with the Genetic Diversity Centre (GDC), ETH Zurich and the Functional Genomics Center (FGC), Zurich. This work was supported by the Overhead fund of the University of Neuchâtel.

Literature chapter 1

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. doi: 10.1016/S0022-2836(05)80360-2

Andrews, S. (2010). FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved August 30, 2019, from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Barker, M. S., & Wolf, P. G. (2010). Unfurling Fern Biology in the Genomics Age. *BioScience*, *60*(3), 177–185. doi: 10.1525/bio.2010.60.3.4

Bennett, M. D., & Leitch, I. J. (2001). Nuclear DNA Amounts in Pteridophytes. *Annals of Botany*, *87*(3), 335–345. doi: 10.1006/anbo.2000.1339

Birkeland, S., Borgenkjetne, I., Brysting, A. K., Elven, R., & Alsos, I. G. (2017). Living on the edge: Conservation Genetics of Seven Thermophilous Plant Species in a High Arctic Archipelago. *AoB PLANTS*, *9*(1),1–21. doi: 10.1093/aobpla/plx001

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527. doi: 10.1038/nbt.3519

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59–60. doi: 10.1038/nmeth.3176

Camacho, F. J., & Liston, A. (2001). Population structure and genetic diversity of *Botrychium pumicola* (Ophioglossaceae) based on inter-simple sequence repeats (ISSR). *American Journal of Botany*, *88*(6), 1065–1070. doi: 10.2307/2657089

Carlson, M. (2020). GO.db: A set of annotation maps describing the entire Gene Ontology (Version 3.8.2). Bioconductor version: Release (3.10). doi: 10.18129/B9.bioc.GO.db

Carpenter, E. J., Matasci, N., Ayyampalayam, S., Wu, S., Sun, J., Yu, J., … Wong, G. K.-S. (2019). Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience*, *8*(10). doi: 10.1093/gigascience/giz126

Chamberlain, S. A., & Szöcs, E. (2013). taxize: Taxonomic search and retrieval in R. *F1000Research*, *2*, 191. doi: 10.12688/f1000research.2-191.v2

Chen, E. C., Mathieu, S., Hoffrichter, A., Sedzielewska-Toro, K., Peart, M., Pelin, A., … Corradi, N. (2018). Single nucleus sequencing reveals evidence of inter-nucleus recombination in arbuscular mycorrhizal fungi. *ELife*, *7*, e39813. doi: 10.7554/eLife.39813

Christenhusz, M. J. M., Chase, M. W., Fay, M. F., Hidalgo, O., Leitch, I. J., Pellicer, J., & Viruel, J. (2021). Biogeography and genome size evolution of the oldest extant vascular plant genus, Equisetum (Equisetaceae). *Annals of Botany*, *127*(5), 681–695. doi: 10.1093/aob/mcab005

Clark, J., Hidalgo, O., Pellicer, J., Liu, H., Marquardt, J., Robert, Y., … Schneider, H. (2016). Genome evolution of ferns: Evidence for relative stasis of genome size across the fern phylogeny. *The New Phytologist*, *210*(3), 1072–1082. doi: 10.1111/nph.13833

Clausen, R. T. (1938). A Monograph of the Ophioglossaceae. *Memoirs of the Torrey Botanical Club*, *19*(2), 1–177.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. doi: 10.1093/bioinformatics/btr330

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (n.d.). ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular Biology and Evolution*. doi: 10.1093/molbev/msz189

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). *Systematic Botany*, *42*(4). doi: 10.1600/036364417X696483

Dauphin, B., Grant, J. R., & Farrar, D. R. (2020). Outcrossing Mating System of the Early-Divergent Moonwort Fern (*Botrychium lunaria*, Ophioglossaceae) Revealed in the European Alps. *International Journal of Plant Sciences*, *181*(9), 926–936. doi: 10.1086/710491

Dauphin, B., Vieu, J., & Grant, J. R. (2014). Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* s.s., Ophioglossaceae). *American Journal of Botany*, *101*(1), 128–140. doi: 10.3732/ajb.1300154

Dauphin, B. (2017). *Evolution of moonwort ferns (Botrychium, Ophioglossaceae) on local to global scales*. University of Neuchâtel.

Dauphin, B., Grant, J. R., Farrar, D. R., & Rothfels, C. J. (2018). Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. *Molecular Phylogenetics and Evolution*, *120*, 342–353. doi: 10.1016/j.ympev.2017.11.025

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., … Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. doi: 10.1038/ng.806

Der, J. P., Barker, M. S., Wickett, N. J., dePamphilis, C. W., & Wolf, P. G. (2011). De novo characterization of the gametophyte transcriptome in bracken fern, *Pteridium aquilinum*. *BMC Genomics*, *12*, 99. doi: 10.1186/1471-2164-12-99

Des Marais, D. L., Smith, A. R., Britton, D. M., & Pryer, K. M. (2003). Phylogenetic Relationships and Evolution of Extant Horsetails, Equisetum, Based on Chloroplast DNA Sequence Data (rbcL and trnL-F). *International Journal of Plant Sciences*, *164*(5), 737–751. doi: 10.1086/376817

Doležel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysák, M. A., Nardi, L., & Obermayer, R. (1998). Plant Genome Size Estimation by Flow Cytometry: Inter-laboratory Comparison. *Annals of Botany*, *82*, 17–26. doi: 10.1006/anbo.1998.0730

Ellwood, M. D. F., & Foster, W. A. (2004). Doubling the estimate of invertebrate biomass in a rainforest canopy. *Nature*, *429*(6991), 549–551. doi: 10.1038/nature02560

Emms, D. M., & Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157. doi: 10.1186/s13059-015-0721-2

Farrar, D. R. (1998). *Population genetics of moonwort Botrychium. In N. Berlin, P. Miller, J. Borovansky, U. S. Seal, and O. Byers [eds.], Population and habitat viability assessment for the goblin fern (Botrychium mormo)* (pp. 109–113). The Conservation Breeding Specialist Group. Retrieved from The Conservation Breeding Specialist Group website: http://www.cpsg.org/content/goblin-fern-phva-1998

Field, K. J., Leake, J. R., Tille, S., Allinson, K. E., Rimington, W. R., Bidartondo, M. I., … Cameron, D. D. (2015). From mycoheterotrophy to mutualism: Mycorrhizal specificity and functioning in *Ophioglossum vulgatum* sporophytes. *New Phytologist*, *205*(4), 1492–1502. doi: 10.1111/nph.13263

Gastony, G. J., & Gottlieb, L. D. (1985). Genetic Variation in the Homosporous Fern *Pellaea andromedifolia*. *American Journal of Botany*, *72*(2), 257–267. doi: https://doi.org/10.1002/j.1537-2197.1985.tb08290.x

Gentleman, R. (2020). annotate: Annotation for microarrays (Version 1.64.0). Bioconductor version: Release (3.10). doi: 10.18129/B9.bioc.annotate

George, L. O., & Bazzaz, F. A. (1999a). The Fern Understory as an Ecological Filter: Emergence and Establishment of Canopy-Tree Seedlings. *Ecology*, *80*(3), 833–845. doi: 10.1890/0012-9658(1999)080[0833:TFUAAE]2.0.CO;2

George, L. O., & Bazzaz, F. A. (1999b). The Fern Understory as an Ecological Filter: Growth and Survival of Canopy-Tree Seedlings. *Ecology*, *80*(3), 846–856. doi: 10.1890/0012-9658(1999)080[0846:TFUAAE]2.0.CO;2

Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184–186. doi: 10.1111/j.1471-8286.2004.00828.x

Grewe, F., Guo, W., Gubbels, E. A., Hansen, A. K., & Mower, J. P. (2013). Complete plastid genomes from *Ophioglossum californicum*, *Psilotum nudum*, and *Equisetum hyemale* reveal an ancestral land plant genome structure and resolve the position of Equisetales among monilophytes. *BMC Evolutionary Biology*, *13*, 8. doi: 10.1186/1471-2148-13-8

Grusz, A. L. (2016). A current perspective on apomixis in ferns. *Journal of Systematics and Evolution*, *54*(6), 656–665. doi: https://doi.org/10.1111/jse.12228

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*, 1494–1512. doi: 10.1038/nprot.2013.084

Hanson, L., & Leitch, I. J. (2002). DNA amounts for five pteridophyte species fill phylogenetic gaps in C-value data. *Botanical Journal of the Linnean Society*, *140*(2), 169–173. doi: 10.1046/j.1095-8339.2002.00083.x

Haufler, C. H., Pryer, K. M., Schuettpelz, E., Sessa, E. B., Farrar, D. R., Moran, R., … Windham, M. D. (2016). Sex and the Single Gametophyte: Revising the Homosporous Vascular Plant Life Cycle in Light of Contemporary Research. *BioScience*, *66*(11), 928–937. doi: 10.1093/biosci/biw108

Hauk, W. D. (1995). A Molecular Assessment of Relationships among Cryptic Species of *Botrychium* Subgenus *Botrychium* (Ophioglossaceae). *American Fern Journal*, *85*(4), 375–394. JSTOR. doi: 10.2307/1547816

Hauk, W. D., & Haufler, C. H. (1999). Isozyme variability among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). *American Journal of Botany*, *86*(5), 614–633. doi: 10.2307/2656570

Hidalgo, O., Pellicer, J., Christenhusz, M. J. M., Schneider, H., & Leitch, I. J. (2017). Genomic gigantism in the whisk-fern family (Psilotaceae): *Tmesipteris obliqua* challenges record holder *Paris japonica*. *Botanical Journal of the Linnean Society*, *183*(4), 509–514. doi: 10.1093/botlinnean/box003

Jeffrey, E. C. (1898). The gametophyte of *Botrychium virginianum*. *Univ. of Toronto Studies. Biological Series.*, *1*, 3-32 pl.i-iv.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. doi: 10.1093/bioinformatics/btr521

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2015). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. doi: 10.1093/nar/gkf436

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)*, *28*(12), 1647–1649. doi: 10.1093/bioinformatics/bts199

Klekowski, E. J., & Baker, H. G. (1966). Evolutionary Significance of Polyploidy in the Pteridophyta. *Science*, *153*(3733), 305–307. doi: 10.1126/science.153.3733.305

Klekowski, E. J., & Lloyd, R. M. (1968). Reproductive biology of the Pteridophyta: 1. General considerations and a study of Onoclea sensibilis L. *Botanical Journal of the Linnean Society*, *60*(383), 315–324. doi: 10.1111/j.1095-8339.1968.tb00093.x

Knaus, B. J., & Grünwald, N. J. (2017). vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, *17*(1), 44–53. doi: 10.1111/1755-0998.12549

Knie, N., Fischer, S., Grewe, F., Polsakiewicz, M., & Knoop, V. (2015). Horsetails are the sister group to all other monilophytes and Marattiales are sister to leptosporangiate ferns. *Molecular Phylogenetics and Evolution*, *90*, 140–149. doi: 10.1016/j.ympev.2015.05.008

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 1–3. doi: 10.1093/bioinformatics/btz305

Kuo, L.-Y., Li, F.-W., Chiou, W.-L., & Wang, C.-N. (2011). First insights into fern matK phylogeny. *Molecular Phylogenetics and Evolution*, *59*(3), 556–566. doi: 10.1016/j.ympev.2011.03.010

Langmead, B. (2010). Aligning Short Sequencing Reads with Bowtie. *Current Protocols in Bioinformatics*, *32*(1), 11.7.1-11.7.14. doi: 10.1002/0471250953.bi1107s32

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., … One Thousand Plant Transcriptomes Initiative. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, *574*(7780), 679–685. doi: 10.1038/s41586-019-1693-2

Lehtonen, S., Silvestro, D., Karger, D. N., Scotese, C., Tuomisto, H., Kessler, M., … Antonelli, A. (2017). Environmentally driven extinction and opportunistic origination explain fern diversification patterns. *Scientific Reports*, *7*(1), 4831. doi: 10.1038/s41598-017-05263-7

Lemoine, F., Entfellner, J.-B. D., Wilkinson, E., Correia, D., Felipe, M. D., Oliveira, T. D., & Gascuel, O. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, *556*(7702), 452. doi: 10.1038/s41586-018-0043-0

Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., Vries, J. de, Delaux, P.-M., … Pryer, K. M. (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants*, *4*(7), 460–472. doi: 10.1038/s41477-018-0188-8

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Lu, J.-M., Zhang, N., Du, X.-Y., Wen, J., & Li, D.-Z. (2015). Chloroplast phylogenomics resolves key relationships in ferns. *Journal of Systematics and Evolution*, *53*(5), 448–457. doi: 10.1111/jse.12180

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*(2), 185–196. doi: 10.1007/s00035-017-0188-8

Magallón, S., Hilu, K. W., & Quandt, D. (2013). Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *American Journal of Botany*, *100*(3), 556–573. doi: 10.3732/ajb.1200416

McCauley, D. E., Whittier, D. P., & Reilly, L. M. (1985). Inbreeding and the rate of self-fertilization in a grape fern, *Botrychium dissectum*. *American Journal of Botany*, *72*(12), 1978–1981. (McCauley1985). doi: 10.1002/j.1537-2197.1985.tb08471.x

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. doi: 10.1101/gr.107524.110

Mehltreter, K., Walker, L. R., Sharpe, J. M., Kessler, M., Richardson, S. J., Hietz, P., … Sheffield, E. (2010). *Fern ecology*. Cambridge University Press. Retrieved from https://www.cambridge.org/ch/academic/subjects/life-sciences/plant-science/fern-ecology

Morgan, M., Falcon, S., & Gentleman, R. (2020). GSEABase: Gene set enrichment data structures and methods (Version 1.48.0). Bioconductor version: Release (3.10). doi: 10.18129/B9.bioc.GSEABase

Nei, M. (1987). Molecular Evolutionary Genetics. In *Molecular Evolutionary Genetics*. Columbia University Press.

Obermayer, R., Leitch, I. J., Hanson, L., & Bennett, M. D. (2002). Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. *Annals of Botany*, *90*(2), 209–217. doi: 10.1093/aob/mcf167

Page, C. N. (2002). Ecological strategies in fern evolution: A neopteridological overview. *Review of Palaeobotany and Palynology*, *119*(1), 1–33. doi: 10.1016/S0034-6667(01)00127-0

Pagès, H., Carlson, M., Falcon, S., & Li, N. (2020). *AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor*. Bioconductor version: Release (3.10).

Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. doi: 10.1093/bioinformatics/bty633

Pinson, J. B., Chambers, S. M., Nitta, J. H., Kuo, L.-Y., & Sessa, E. B. (2016). The Separation of Generations: Biology and Biogeography of Long-Lived Sporophyteless Fern Gametophytes. *International Journal of Plant Sciences*, *178*(1), 1–18. doi: 10.1086/688773

PPG I. (2016). A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution*, *54*(6), 563–603. doi: 10.1111/jse.12229

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

Pryer, K. M., Schneider, H., Smith, A. R., Cranfill, R., Wolf, P. G., Hunt, J. S., & Sipes, S. D. (2001). Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature*, *409*(6820), 618–622. doi: 10.1038/35054555

Pryer, K. M., Schuettpelz, E., Wolf, P. G., Schneider, H., Smith, A. R., & Cranfill, R. (2004). Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. *American Journal of Botany*, *91*(10), 1582–1598. doi: 10.3732/ajb.91.10.1582

Qi, X., Kuo, L.-Y., Guo, C., Li, H., Li, Z., Qi, J., … Ma, H. (2018). A well-resolved fern nuclear phylogeny reveals the evolution history of numerous transcription factor families. *Molecular Phylogenetics and Evolution*, *127*, 961–977. doi: 10.1016/j.ympev.2018.06.043

Qiu, Y., Li, L., Wang, B., Chen, Z., Dombrovska, O., Lee, J., … Ambros, M. (2007). A Nonflowering Land Plant Phylogeny Inferred from Nucleotide Sequences of Seven Chloroplast, Mitochondrial, and Nuclear Genes. *International Journal of Plant Sciences*, *168*(5), 691–708. doi: 10.1086/513474

R Development Core Team. (2020). R: The R Project for Statistical Computing. Retrieved February 9, 2020, from https://www.r-project.org/

Rai, H. S., & Graham, S. W. (2010). Utility of a large, multigene plastid data set in inferring higher-order relationships in ferns and relatives (monilophytes). *American Journal of Botany*, *97*(9), 1444–1456. doi: 10.3732/ajb.0900305

Ranker, T. A., & Haufler, C. H. (2008). *Biology and Evolution of Ferns and Lycophytes*. Cambridge University Press. doi: 10.1017/CBO9780511541827

Raubeson, L. A., & Jansen, R. K. (1992). Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science (New York, N.Y.)*, *255*(5052), 1697–1699. doi: 10.1126/science.255.5052.1697

Rothfels, C. J., Larsson, A., Li, F.-W., Sigel, E. M., Huiet, L., Burge, D. O., … Pryer, K. M. (2013). Transcriptome-Mining for Single-Copy Nuclear Markers in Ferns. *PLOS ONE*, *8*(10), e76957. doi: 10.1371/journal.pone.0076957

Rothfels, C. J., Li, F.-W., Sigel, E. M., Huiet, L., Larsson, A., Burge, D. O., … Pryer, K. M. (2015). The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *American Journal of Botany*, *102*(7), 1089–1107. doi: 10.3732/ajb.1500089

RStudio Team. (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston. Retrieved August 30, 2019, from https://www.rstudio.com/

Sayyari, E., & Mirarab, S. (2016). Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Molecular Biology and Evolution*, *33*(7), 1654–1668. doi: 10.1093/molbev/msw079

Schuettpelz, E., Korall, P., & Pryer, K. M. (2006). Plastid atpA data provide improved support for deep relationships among ferns. *TAXON*, *55*(4), 897–906. doi: 10.2307/25065684

Schuettpelz, E., & Pryer, K. M. (2007). Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon*, *56*(4).

Seeb, J. E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., & Seeb, L. W. (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in non-model organisms. *Molecular Ecology Resources*, *11*(s1), 1–8. doi: 10.1111/j.1755-0998.2010.02979.x

Sessa, E. B., Testo, W. L., & Watkins Jr, J. E. (2016). On the widespread capacity for, and functional significance of, extreme inbreeding in ferns. *New Phytologist*, *211*(3), 1108–1119. doi: 10.1111/nph.13985

Sessa, E. B., Zimmer, E. A., & Givnish, T. J. (2012). Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology*, *12*(1), 104. doi: 10.1186/1471-2148-12-104

Shen, H., Jin, D., Shu, J.-P., Zhou, X.-L., Lei, M., Wei, R., … Yan, Y.-H. (2018). Large-scale phylogenomic analysis resolves a backbone phylogeny in ferns. *GigaScience*, *7*(2), 1–11. doi: 10.1093/gigascience/gix116

Sigel, E. M., Schuettpelz, E., Pryer, K. M., & Der, J. P. (2018). Overlapping Patterns of Gene Expression Between Gametophyte and Sporophyte Phases in the Fern *Polypodium amorphum* (Polypodiales). *Frontiers in Plant Science*, *9*, 1–17. doi: 10.3389/fpls.2018.01450

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. doi: 10.1093/bioinformatics/btv351

Soltis, D. E., & Soltis, P. S. (1986). Electrophoretic Evidence for Inbreeding in the Fern *Botrychium virginianum* (ophioglossaceae). *American Journal of Botany*, *73*(4), 588–592. doi: https://doi.org/10.1002/j.1537-2197.1986.tb12078.x

Soltis, D. E., & Soltis, P. S. (1992). The Distribution of Selfing Rates in Homosporous Ferns. *American Journal of Botany*, *79*(1), 97–100. doi: https://doi.org/10.1002/j.1537-2197.1992.tb12628.x

Soltis, P. S., & Soltis, D. E. (1990). Evolution of Inbreeding and Outcrossing in Ferns and Fern-Allies. *Plant Species Biology*, *5*(1), 1–11. doi: 10.1111/j.1442-1984.1990.tb00187.x

Stensvold, Mary Clay, Farrar, D. R., & Johnson-Groh, C. (2002). Two New Species of Moonworts (*Botrychium* subg. *Botrychium*) from Alaska. *American Fern Journal*, *92*(2), 150–160. doi: 10.1640/0002-8444(2002)092[0150:TNSOMB]2.0.CO;2

Stensvold, M.C., & Farrar, D. R. (2017). Genetic Diversity in the Worldwide *Botrychium* lunaria (Ophioglossaceae) Complex, with New Species and New Combinations. *Brittonia*, *69*(2), 148–175. Scopus. doi: 10.1007/s12228-016-9455-z

Swartz, L. M., & Brunsfeld, S. J. (2002). The morphological and genetic distinctness of *Botrychium minganense* and *B. crenulatum* as assessed by morphometric analysis and RAPD markers. *American Fern Journal*, *92*(4), 249–269. doi: 10.1640/0002-8444(2002)092[0249:TMAGDO]2.0.CO;2

Testo, W., & Sundue, M. (2016). A 4000-species dataset provides new insight into the evolution of ferns. *Molecular Phylogenetics and Evolution*, *105*, 200–211. doi: 10.1016/j.ympev.2016.09.003

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., … DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*(1), 11.10.1-33. doi: 10.1002/0471250953.bi1110s43

Veselý, P., Bureš, P., Šmarda, P., & Pavlíček, T. (2012). Genome size and DNA base composition of geophytes: The mirror of phenology and ecology? *Annals of Botany*, *109*(1), 65–75. doi: 10.1093/aob/mcr267

Walker, L. R. (1994). Effects of fern thickets on woodland development on landslides in Puerto Rico. *Journal of Vegetation Science*, *5*(4), 525–532. doi: 10.2307/3235979

Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., … Yu, G. (2020). Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Molecular Biology and Evolution*, *37*(2), 599–603. doi: 10.1093/molbev/msz240

Weizhong, L., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, *22*(13), 1658–1659. doi: 10.1093/bioinformatics/btl158

Wickett, N. J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., … Leebens-Mack, J. (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences*, *111*(45), E4859–E4868. doi: 10.1073/pnas.1323926111

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). Springer International Publishing. doi: 10.1007/978-3-319-24277-4

Wikström, N., & Kenrick, P. (2001). Evolution of Lycopodiaceae (Lycopsida): Estimating Divergence Times from rbcL Gene Sequences by Use of Nonparametric Rate Smoothing. *Molecular Phylogenetics and Evolution*, *19*(2), 177–186. doi: 10.1006/mpev.2001.0936

Williams, E. W. (2021). Population Genetics of Species in the Genera *Botrychium* and *Botrypus* (Ophioglossaceae). *American Fern Journal*, *111*(2), 129–146. doi: 10.1640/0002-8444-111.2.129

Williams, E. W., Farrar, D. R., & Henson, Don. (2016). Cryptic speciation in allotetraploids: Lessons from the *Botrychium matricariifolium* (Ophioglossaceae) complex. *American Journal of Botany*, *103*(4), 1–14.

Winther, J. L., & Friedman, W. E. (2007). Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). *American Journal of Botany*, *94*(7), 1248–1255. doi: 10.3732/ajb.94.7.1248

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, *106*(33), 13875–13879. doi: 10.1073/pnas.0811575106

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. doi: https://doi.org/10.1111/2041-210X.12628

Zhang, C., Scornavacca, C., Molloy, E. K., & Mirarab, S. (2020). ASTRAL-Pro: Quartet-Based Species-Tree Inference despite Paralogy. *Molecular Biology and Evolution*, *37*(11), 3292–3307. doi: 10.1093/molbev/msaa139

Zhong, B., Fong, R., Collins, L. J., McLenachan, P. A., & Penny, D. (2014). Two New Fern Chloroplasts and Decelerated Evolution Linked to the Long Generation Time in Tree Ferns. *Genome Biology and Evolution*, *6*(5), 1166–1173. doi: 10.1093/gbe/evu087

Data accessibility statement: Raw sequencing reads and the assembled transcriptome are available at the NCBI Sequence Read Archive under the Bioproject accession PRJNA605155 (Vinciane Mossion, Daniel Croll; 2020; TSA: RNA sequencing of two Swiss *Botrychium* lunaria populations; NCBI SRA; BioProject PRJNA605155). The assembly was deposited under the accession number GIJZ00000000.1 (Vinciane Mossion, Daniel Croll; 2020; TSA: *Botrychium* lunaria isolate CHA_I_1, transcriptome shotgun assembly; NCBI GenBank; GIJZ00000000.1). Phylogenetic trees, alignments and protein sequences are available as Supplementary Files S1–S7 on Zenodo (https://doi.org/ 10.5281/zenodo.5061267).

Author contributions: VM, BD and DC designed the study. VM, MK and DC performed analyses, NZ contributed to analyses, BD and JG acquired funding. VM and DC wrote the manuscript with input from co-authors.

Figures

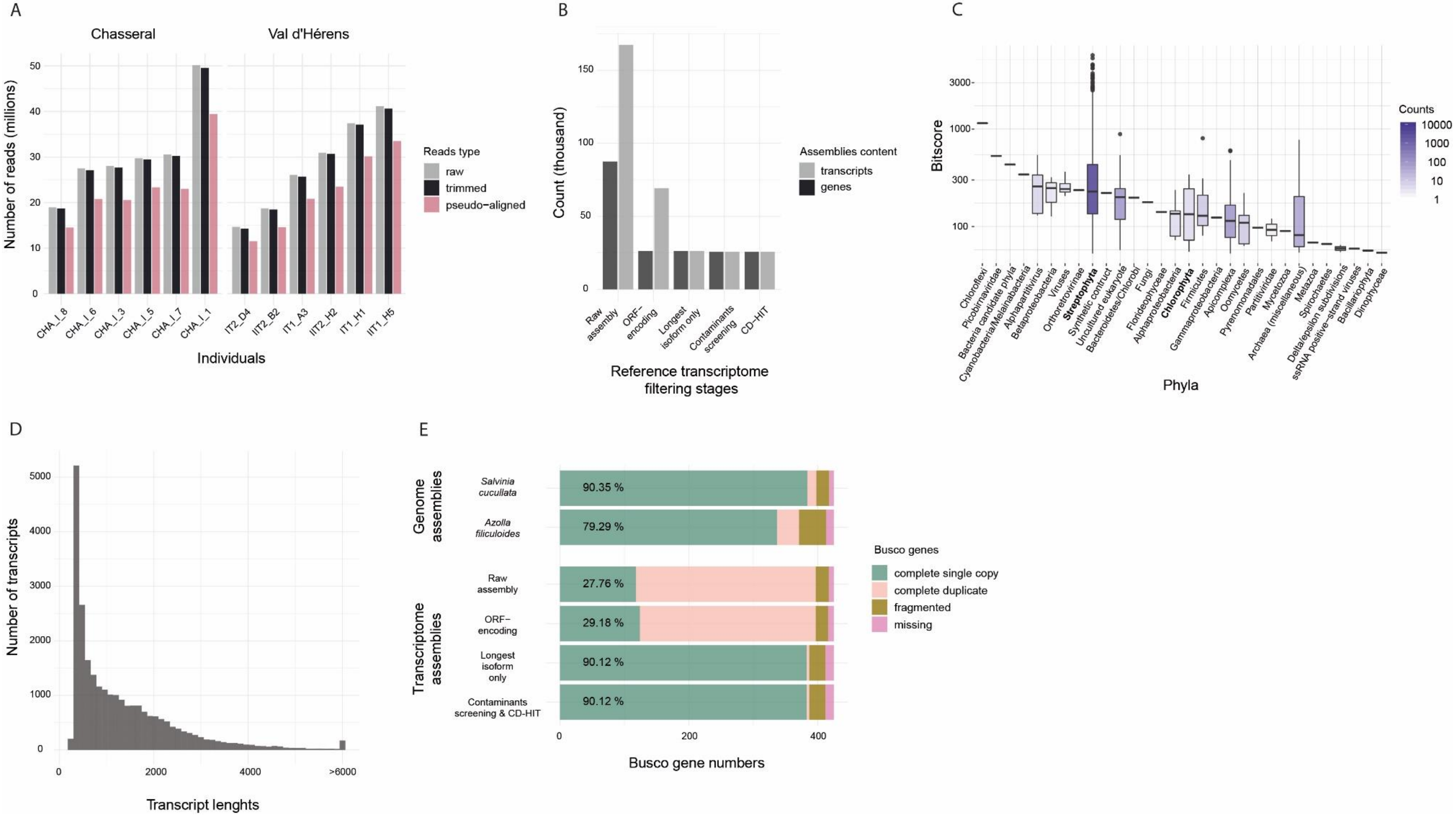


FIGURE 1: De novo assembly of the *Botrychium lunaria* transcriptome. (A) Distribution of the raw, trimmed, and pseudo-aligned read numbers per individual. (B) Transcripts and gene assembly content at each filtering stage of the transcriptome. (C) Contaminant transcript sequences detected by phylum. Retained phyla are indicated in bold. See methods for further details. (D) Distribution of the assembled transcript lengths given in bp. (E) BUSCO genes detected at each filtering stage of the *B. lunaria* transcriptome and for two genome assemblies of ferns (*Azolla filliculoides* and *Salvinia cucullata*; Li *et al.*, 2018).

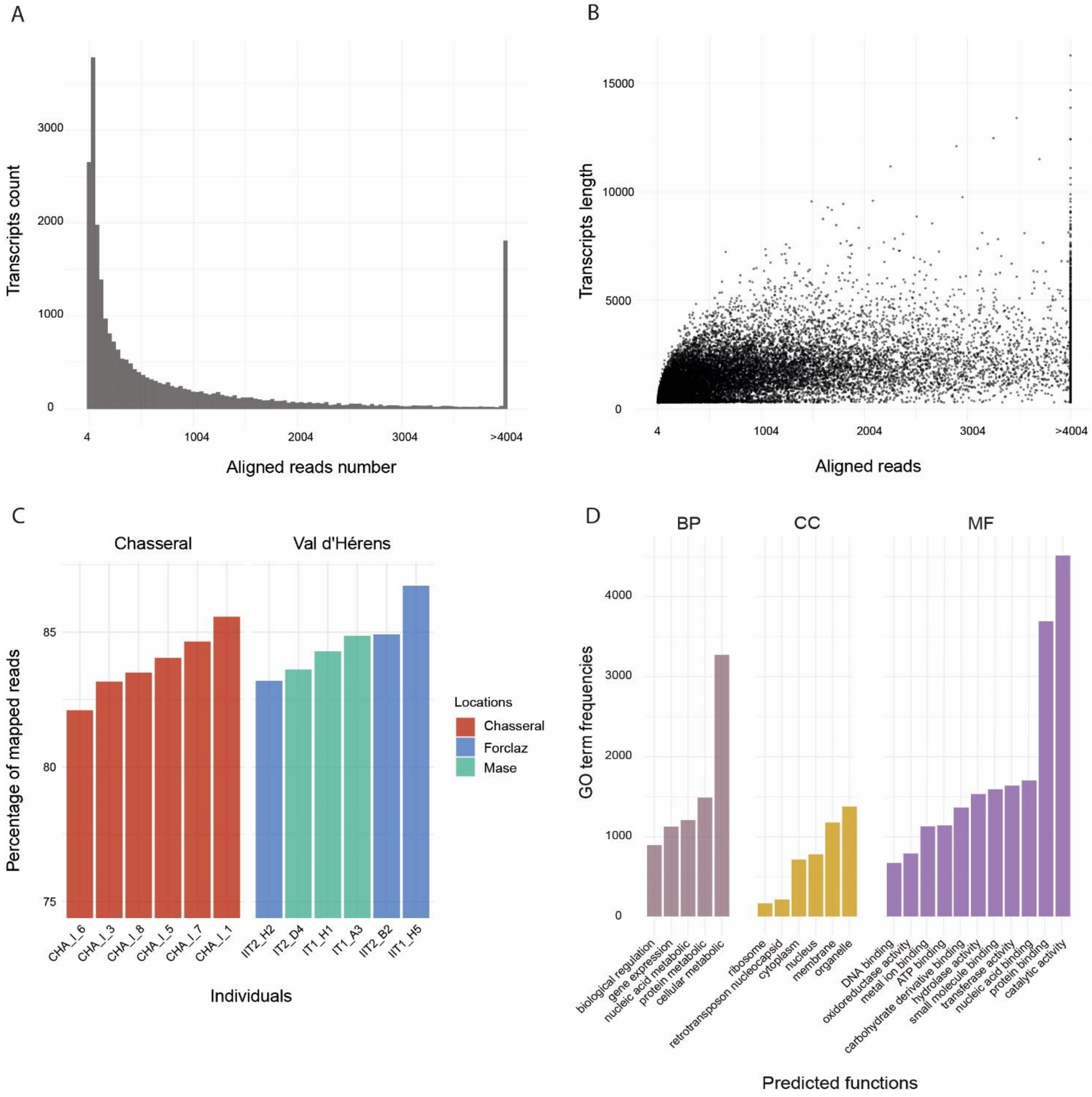


FIGURE 2: Analyses of the assembly coverage. (A) Number of aligned reads per assembled transcript for the reference individual. (B) Aligned reads of the reference individual according to the assembled transcript length. (C) Mapping rate of all 12 individuals from the Chasseral and Val d'Hérens populations (including the subpopulations Forclaz and Mase). The reference individual used to assemble the transcriptome was CHA_I_1. (D) Characterizations of predicted functions encoded by the transcriptome. Gene ontology (GO) term annotations are shown for the 30 most frequent terms per ontology (BP = biological process, CC = cellular component, and MF = molecular function). GO terms with highly similar functions are excluded from the representation.

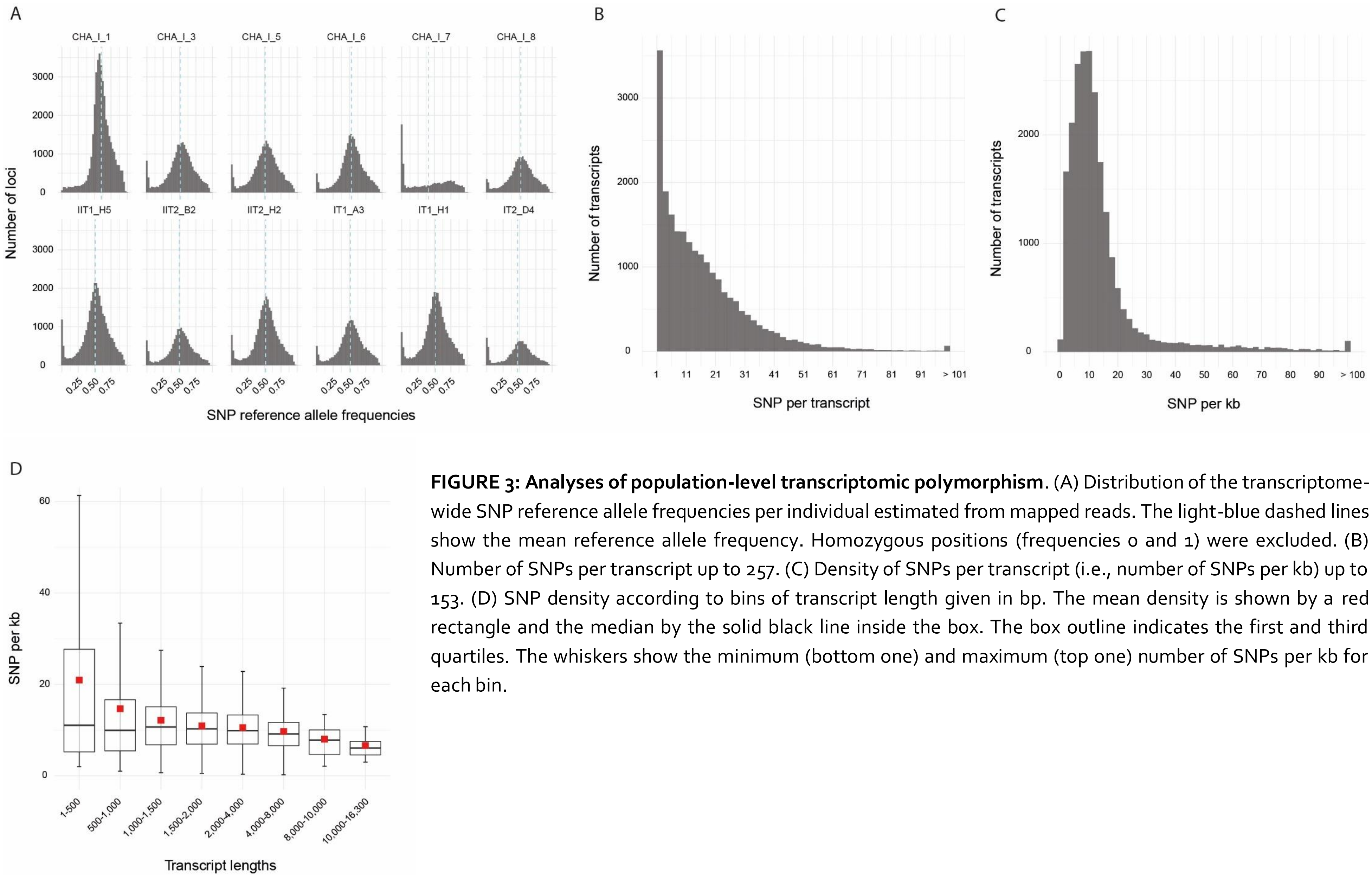


FIGURE 3: Analyses of population-level transcriptomic polymorphism. (A) Distribution of the transcriptome-wide SNP reference allele frequencies per individual estimated from mapped reads. The light-blue dashed lines show the mean reference allele frequency. Homozygous positions (frequencies 0 and 1) were excluded. (B) Number of SNPs per transcript up to 257. (C) Density of SNPs per transcript (i.e., number of SNPs per kb) up to 153. (D) SNP density according to bins of transcript length given in bp. The mean density is shown by a red rectangle and the median by the solid black line inside the box. The box outline indicates the first and third quartiles. The whiskers show the minimum (bottom one) and maximum (top one) number of SNPs per kb for each bin.

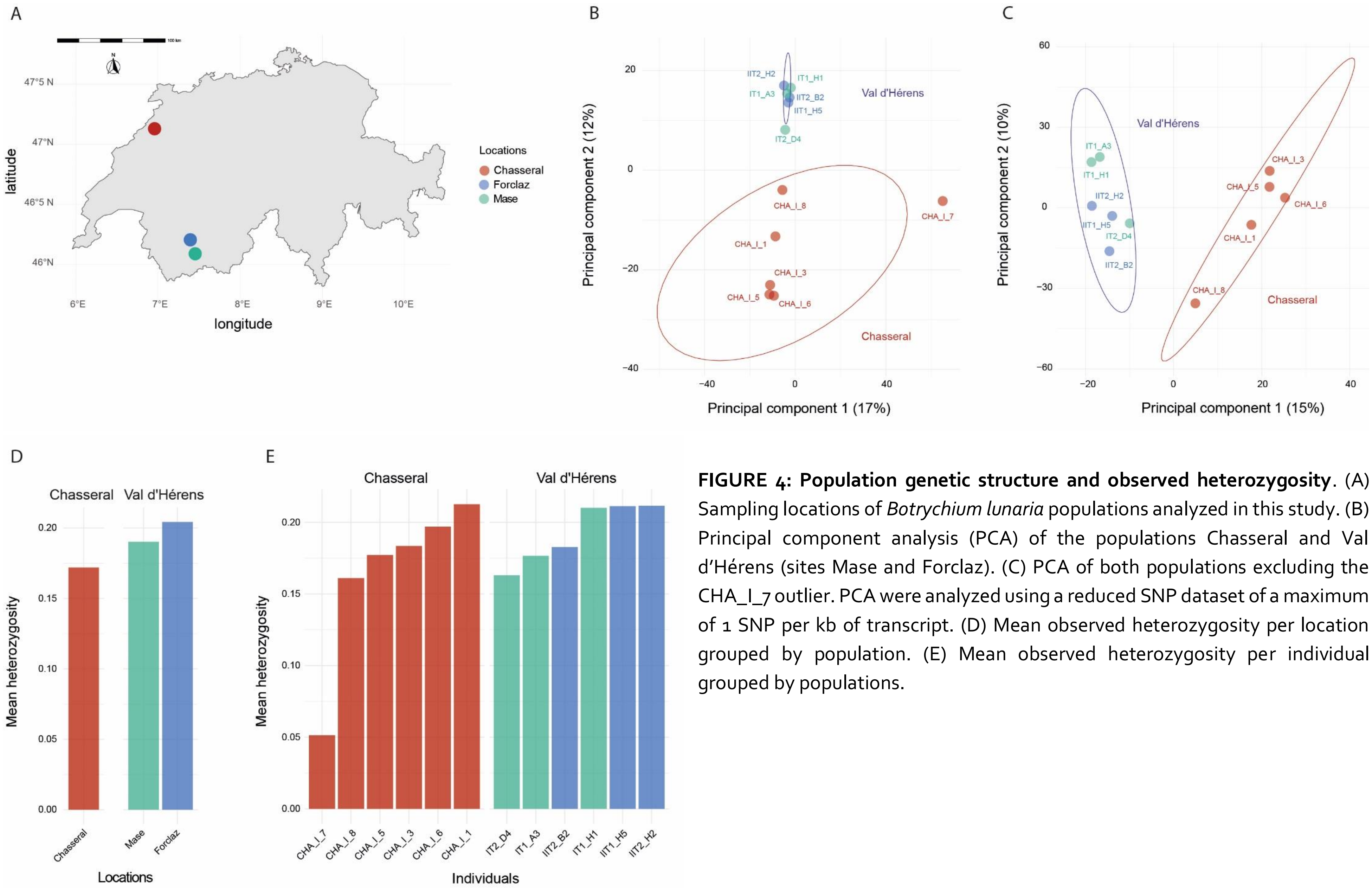


FIGURE 4: Population genetic structure and observed heterozygosity. (A) Sampling locations of *Botrychium lunaria* populations analyzed in this study. (B) Principal component analysis (PCA) of the populations Chasseral and Val d'Hérens (sites Mase and Forclaz). (C) PCA of both populations excluding the CHA_I_7 outlier. PCA were analyzed using a reduced SNP dataset of a maximum of 1 SNP per kb of transcript. (D) Mean observed heterozygosity per location grouped by population. (E) Mean observed heterozygosity per individual grouped by populations.

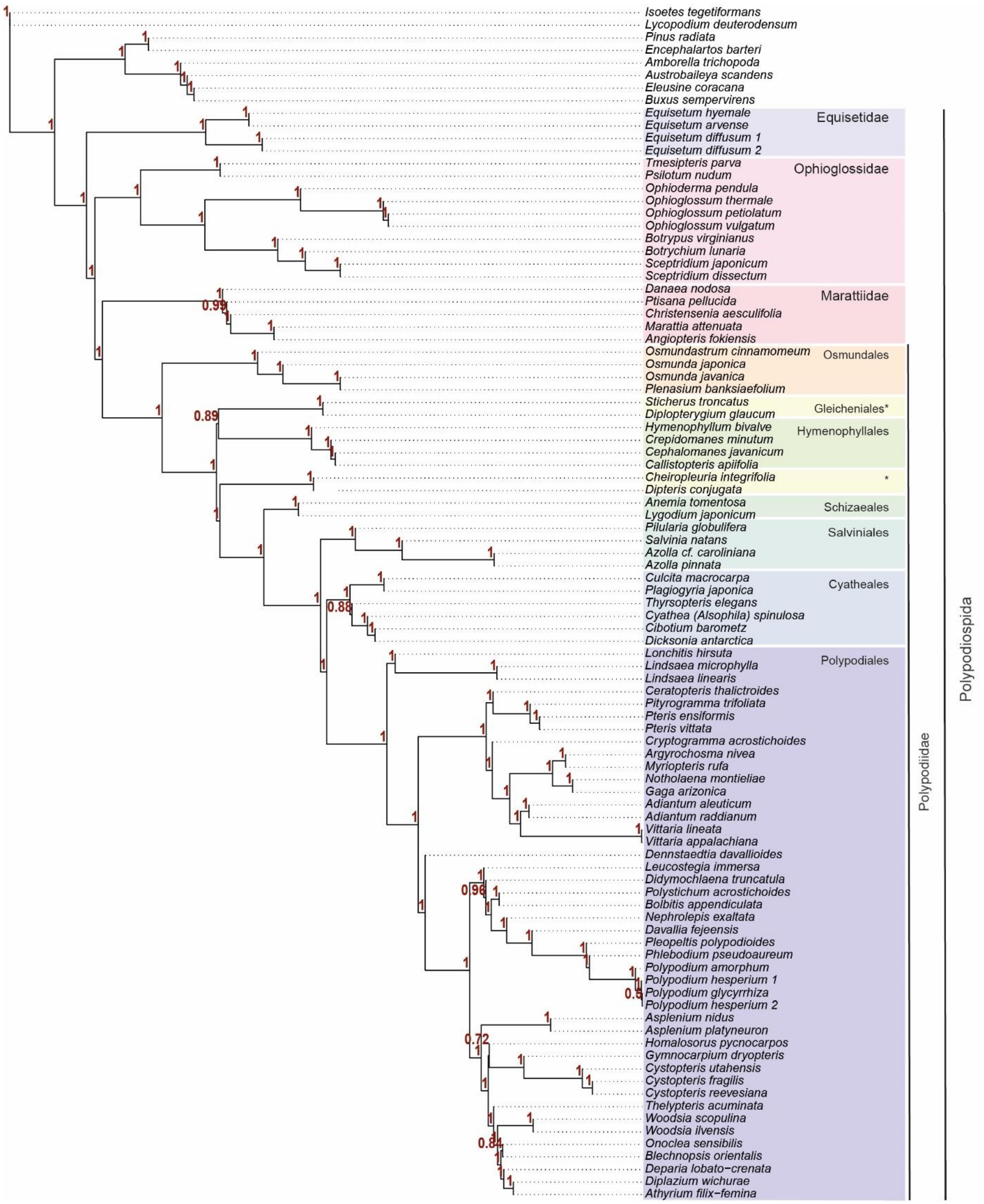


FIGURE 5: Phylogenomic relationships among ferns including *Botrychium lunaria*. A species tree of 761 orthologous genes including 93 taxa inferred by coalescence-based method implemented in ASTRAL-Pro. The branch support is indicated by local posterior probability (LPP) values. LPP values of the main topology are displayed in dark red above branches on left node sides. Fern subclasses (Equisetidae, Ophioglossidae, Marattiidae and Polypodiidae) are denoted by colored rectangles or a vertical dark grey line on clade sides. Polypodiidae orders (Osmundales, Hymenophyllales, Gleicheniales, Schizaeales Salviniales, Cyatheales and Polypodiales) are designated by colored rectangles. The internal branch lengths are in coalescent units. The tip branch lengths are not estimated. The tree with the full ASTRAL annotations is available as supplementary file S7.

Tables

TABLE 1: Populations, accessions and voucher information. Coordinates are given in WGS84.

Individual identifier	Population	Location	Latitude	Longitude	Altitude (m)	Date	Voucher number	Absolute genome size[1]	Deposit[2] institute
CHA_I_1	Chasseral	Chasseral	47.12974	7.04934	1549.71	07.06.2017	NE000101258	-	NEU
CHA_I_3	Chasseral	Chasseral	47.12974	7.04934	1549.71	07.06.2017	NE000101257	-	NEU
CHA_I_5	Chasseral	Chasseral	47.12974	7.04934	1549.71	07.06.2017	NE000101256	-	NEU
CHA_I_6	Chasseral	Chasseral	47.12974	7.04934	1549.71	07.06.2017	NE000101255	-	NEU
CHA_I_7	Chasseral	Chasseral	47.12974	7.04934	1549.71	07.06.2017	CHA_I_7	20.58	UNINE
CHA_I_8	Chasseral	Chasseral	47.12974	7.04934	1549.71	07.06.2017	NE000101254	-	NEU
IIT2_H2	Val d'Hérens	Forclaz	46.08741	7.54379	2346.15	2015	IIT2_H2	19.38	UNINE
IIT1_H5	Val d'Hérens	Forclaz	46.08851	7.53906	2346.15	2015	IIT1_H5	19.97	UNINE
IIT2_B2	Val d'Hérens	Forclaz	46.08741	7.54379	2406.10	2015	IIT2_B2	19.84	UNINE
IT1_A3	Val d'Hérens	Mase	46.20432	7.48350	2406.62	2015	IT1_A3	19.40	UNINE
IT1_H1	Val d'Hérens	Mase	46.20432	7.48350	2406.62	2015	IT1_H1	19.84	UNINE
IT2_D4	Val d'Hérens	Mase	46.19642	7.48502	2424.07	2015	IT2_D4	-	UNINE

[1] Expressed as 2C-values and picograms

[2] NEU: Herbarium of the University of Neuchâtel; UNINE: University of Neuchâtel, Institute of Biology Herbarium. The specimens deposited at UNINE are frozen samples stored at -80°C.

TABLE 2: Overview of assembly statistics over the different transcript filtering stages.

Filtering stage	Assembled bases	Transcripts	Genes	N50-longest isoform	N50-all	GC%
Raw assembly	56,273,802	167,306	87,537	1,689	1,089	43.68
ORF-encoding	34,588,465	69,280	26,139	2,152	1,988	44.00
Longest isoform only	34,588,465	26,139	26,139	1,988	1,988	44.31
Contaminant screening	34,245,455	25,701	25,701	1,995	1,995	44.30
CD-HIT	34,230,067	25,677	25,677	1,995	1,995	44.30

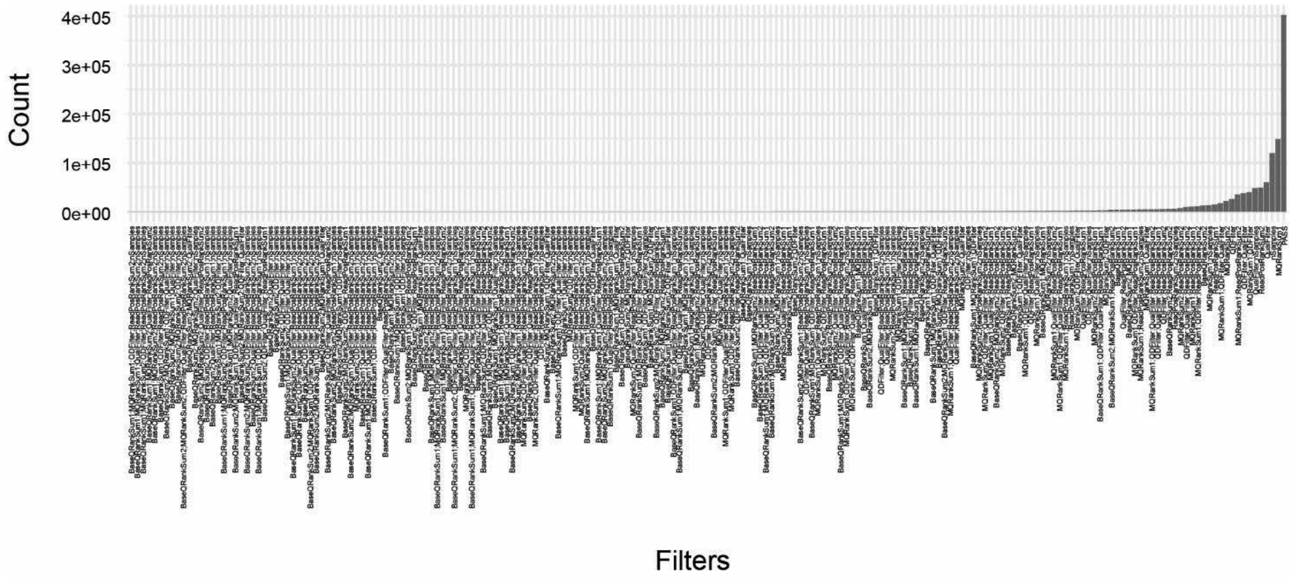
TABLE 3: Analyses of assembly completeness using BUSCO genes. The *B. lunaria* transcriptome is compared to two genome assemblies of ferns (*Azolla filliculoides* and *Salvinia cucullata*).

Species name	Filtering stage	C[1]	CS[2]	CD[3]	F[4]	M[5]	n[6]	Dataset
Azolla filliculoides	-	371	337	34	42	12	425	viridiplantae_odb10
Salvinia cucullata	-	398	384	14	20	7	425	viridiplantae_odb10
B. lunaria	Raw assembly	397	118	279	20	8	425	viridiplantae_odb10
B. lunaria	ORF-encoding	397	124	273	19	9	425	viridiplantae_odb10
B. lunaria	Longest isoform only	387	383	4	25	13	425	viridiplantae_odb10
B. lunaria	Contaminant screening	387	383	4	25	13	425	viridiplantae_odb10
B. lunaria	CD-HIT	387	383	4	25	13	425	viridiplantae_odb10

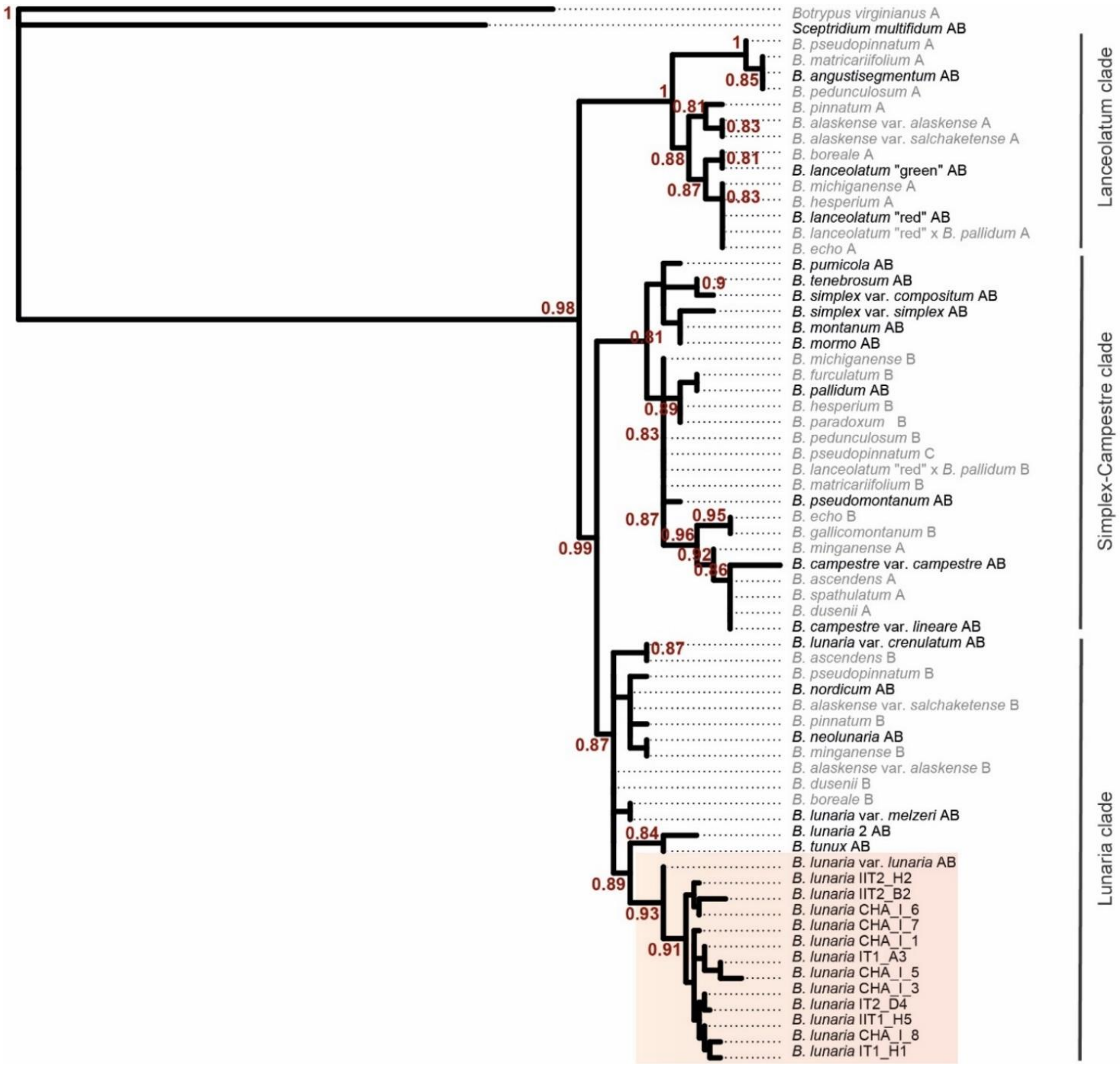
Notes: [1]Complete genes, [2]Complete and single copy genes, [3]Complete and duplicated genes, [4]Fragmented genes, [5]Missing genes, [6]Total number of BUSCO genes in the dataset.

Supplementary material

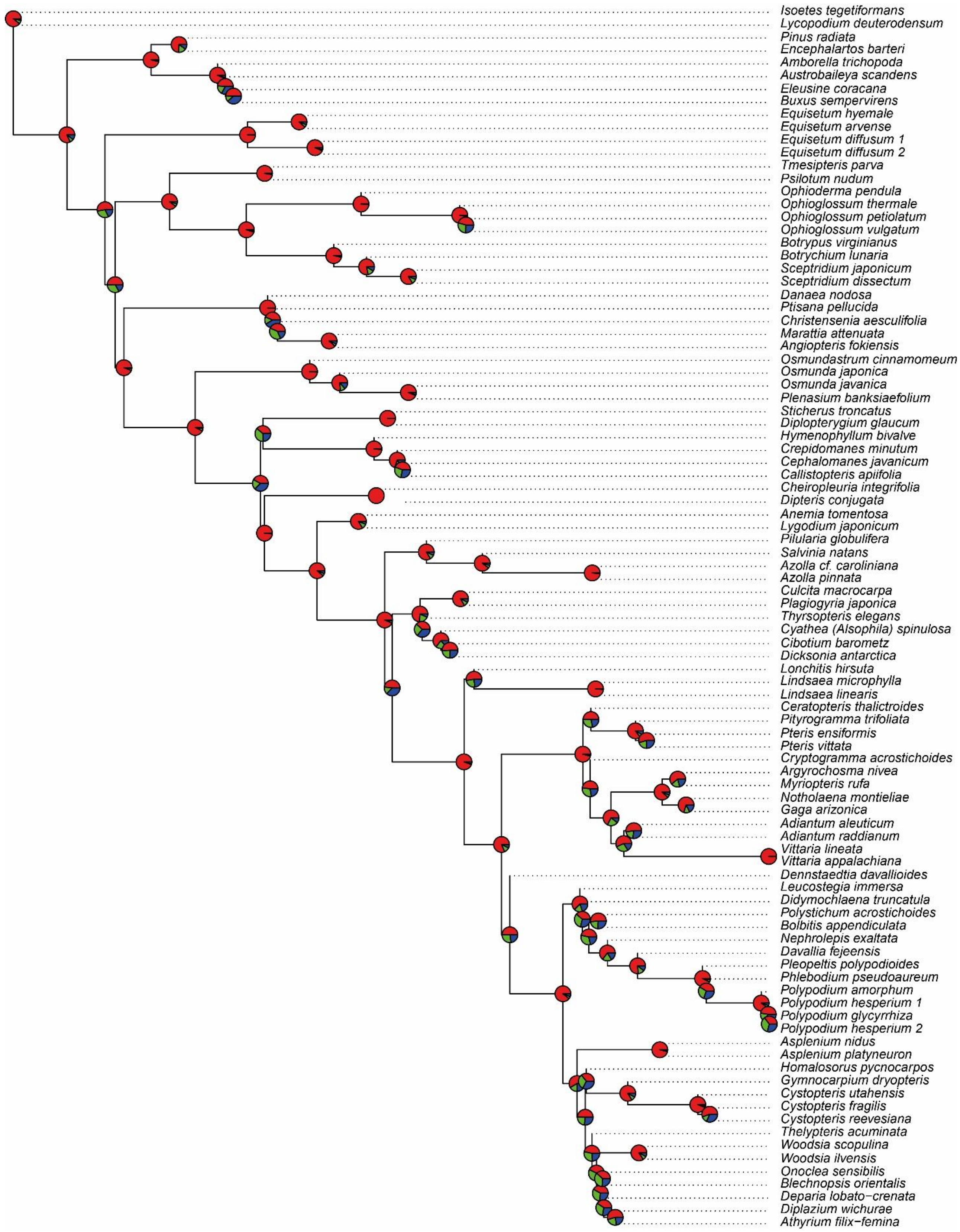
Supplementary Figures



SUPPLEMENTARY FIGURE S1: SNP filtering outcome. The number of filtered flagged loci per individual (or combination of quality criteria). PASS identifies the number of SNPs passing all filtering criteria.



SUPPLEMENTARY FIGURE S2: Phylogenetic positions of the *Botrychium lunaria* samples from Chasseral and Val d'Hérens within *Botrychium* genus*.* Maximum likelihood (ML) tree of the CRY2cA locus. The bootstrap support values above 0.80 are indicated next to nodes in dark red. Diploid taxa are shown in black and polyploids are shown in grey. For polyploids, the subgenome is shown individually on the tree and specified by a letter after the species name: "A" for maternal subgenome, "B" for paternal subgenome, and "C" for additional subgenome for the hexaploid (modified from Dauphin *et al.*, 2018). The three *Botrychium* main clades are delimited by vertical dark grey lines. Individuals from Chasseral and Val d'Hérens populations are highlighted by a red rectangle.



SUPPLEMENTARY FIGURE S3: Astral species tree on which the branches quartet support values were depicted by pie charts. Inside each pie chart, the red, the green and the blue slices correspond to the normalized quartet scores for the main topology, the first and the second alternative, respectively. The internal branch lengths are in coalescent units.

Supplementary Tables

[see https://doi.org/10.1111/1755-0998.13478]

SUPPLEMENTARY TABLE S1: Information about the transcriptomes used in the phylogenomic analyses. The identifiers refer to those originally published by Leebens-Mack et al. (2019), Qi et al. (2018) and Shen et al. (2017) and used in this study.
SUPPLEMENTARY TABLE S2: **Raw and trimmed read distribution for the *Botrychium lunaria* transcriptomes.** Percentage of retained reads after trimming, pseudo-aligned reads by Kallisto and reads aligned by Bowtie.
SUPPLEMENTARY TABLE S3: **Interproscan results.**
SUPPLEMENTARY TABLE S4: Orthogroups table given by OrthoFinder.
SUPPLEMENTARY TABLE S5: Filtered orthogroups table.
SUPPLEMENTARY TABLE S6: Information regarding the CRY2cA sequence dataset used for the phylogenetic inference of *Botrychium* species. The individual identifiers refer to those used by Dauphin et al. (2018) and in this study.

Supplementary Files

[Available from Zenodo: https://doi.org/10.5281/zenodo.5061267]

FILE S1: Alignment of *Botrychium* CRY2cA sequences used to infer the genus-level phylogeny.
File name: File_S1_CRY2_alignment.zip
FILE S2: Peptide sequences used to infer the orthogroups named according to species names or identifiers (see Table S1).
File name: File_S2_peptide.sequences.7z
FILE S3: Alignments of the orthogroup sequences subset used to infer the phylogeny.
File name: File_S3_orthogroups.alignments.zip
FILE S4: Output files from Modeltest-ng named by orthogroup names.
File name: File_S4_modeltestng.files.zip
FILE S5: Output files from RAxML-NG named by orthogroup names.
File name: File_S5_raxmlng.trees.zip
FILE S6: Astral consensus tree with full support and branch length annotations.
File name::File_S6_astral.tree.full.annotation.zip
FILE S7: VCF file containing the full SNP dataset.
File name: File_S7_vcf.filtered.file.zip

CHAPTER 2: Global diversification of the common moonwort ferns (*Botrychium lunaria* group, Ophioglossaceae) was mainly driven by Pleistocene climatic shifts

Vinciane Mossion[1], Erik Koenen[2], Jason Grant[1], Daniel Croll[1], Donald R. Farrar[3], Michael Kessler[2]

[1] Laboratory of Evolutionary Genetics, University of Neuchâtel, Neuchâtel, Switzerland

[2] Department of Systematic and Evolutionary Botany, University of Zürich, Zurich, Switzerland

[3] Department of Ecology, Evolution and Organismal Biology, Iowa State University of Science and Technology, United-States of America

This chapter is in preparation for its submission to *Molecular phylogenetics and Evolution*

Abstract

The cosmopolitan *Botrychium lunaria* group belong to the most species rich genus of the family Ophioglossaceae and was considered to consist of two species until molecular studies in North America and northern Europe led to the recognition of multiple new taxa. Recently, additional genetic lineages were found scattered in Europe, emphasizing our poor understanding of the global diversity of the *B. lunaria* group, while the processes involved in the diversification of the group remain unexplored. We conducted the first global phylogenetic study of the group including 513 ingroup accessions sequenced for four non-coding plastid loci. We recovered ten well-supported clades, although relationships between clades were inconsistent between Bayesian and Maximum Likelihood analyses. We treated each clade at the species level, except for one clade including two ploidy levels, resulting in the recognition of 11 species, 5 of which are unnamed. In contrast to previous studies, we found species diversity to be equally distributed across the northern hemisphere, with 7-8 species per continent. We estimated the stem age of the *B. lunaria* group at 2.4-5.1 million years, with most species 1.8-2.6 million years old, and subspecies 0.5-1.0 million years old. Diversification thus coincided with Pleistocene climatic fluctuations that strongly affected the areas inhabited by the group, suggesting that diversification was driven by climatically induced cycles of extinction, dispersal, and migration. Furthermore, ecological differentiation between species suggests these complex population dynamics were associated with adaptations to specific environmental conditions. We found limited evidence that speciation is driven by polyploidization and hybridization. We show that the *B. lunaria* group has greater species level diversity than previously assumed and suspect that further cryptic species may await discovery, especially in the *B. neolunaria* clade.

Keywords: Ophioglossaceae, ferns, speciation, diversity, phylogenetics, climate, ploidy

Abbreviations: MARK: Cadi Ayyad University herbarium, MW: Moscow State University herbarium, ISC: Iowa State University herbarium, AIX: Muséum d'Histoire Naturelle d'Aix-en-Provence herbarium, PE: Institute of Botany, Chinese Academy of Sciences herbarium, NEU: University of Neuchâtel herbarium, Z/ZT: herbarium of the University of Zurich and the Federal Institute of Technology in Zurich.

Introduction

Ferns are an ancient group of vascular plants, with a fossil records dating back to the middle Devonian (387.7 – 382.7 million years) (Berry and Hilton, 2006; Taylor et al., 2009) and divergence from the seed plants estimated to have taken place about 360-430 million years (Des Marais et al., 2003; Lehtonen et al., 2017; Magallón et al., 2013; Pryer et al., 2004; Qi et al., 2018; Rothfels et al., 2015; Testo and Sundue, 2016; Wikström and Kenrick, 2001; Zhong et al., 2014). Devonian fern lineages such as Cladoxylopsida or Rhacophytales (Taylor et al., 2009) have long gone extinct, while most of the about 11,000 extant fern species arose from a Cenozoic radiation (Schneider et al., 2004; Schuettpelz and Pryer, 2009; Testo and Sundue, 2016). The recent nature of most extant fern species and their overall slow (Smith, 1972), but heterogeneous (Rothfels et al., 2012; Testo and Sundue, 2016) rates of diversification, imply that many groups comprise species complexes characterized by morphologically poorly-defined species (Paris et al., 1989; Vasques et al., 2019; Williams et al., 2016), often in combination with polyploid networks in which diploid species give rise to allopolyploid hybrids which then in turn diversify (Chang et al., 2013; Hanušová et al., 2019; Lovis, 1958; Rothfels et al., 2014; Williams et al., 2016). In such a context, interpreting species diversity and understanding the speciation processes involved in diversification is challenging and intriguing.

The Ophioglossaceae is an example of an early diverging family of ferns estimated at 262-322 million years old (Kumar et al., 2017; Lehtonen et al., 2017; Testo and Sundue, 2016), composed of recently diverged species (Dauphin et al., 2018; Lehtonen et al., 2017; Rothfels et al., 2015) many of which are genetically distinct but morphologically hardly distinguishable (Dauphin et al., 2017; Hauk, 1995; Zhang et al., 2020). Ophioglossaceae are characterized by subterranean gametophytes that are colonized by mycorrhizal fungi (Winther and Friedman, 2007), and an unusual sporophyte leaf morphology: the single leaf produced each year is composed of a sterile photosynthetic blade, the trophophore, and of a fertile portion, the sporophore (Figure 1A). This simple morphology provides only a limited number of distinctive features, making morphological differentiation of species challenging (Farrar, 2011; Stensvold, 2008; Williams, 2015; Williams et al., 2016). Accordingly, the number of species recognized in the family has historically been quite low. Clausen (1938) only recognized six species in *Botrychium sensu stricto* (subgenus *EuBotrychium*). Later Wagner (Wagner and Grant, 2002; Wagner and Wagner, 1981, 1994, 1990a, 1990b, 1986, 1983a, 1982) and other taxonomists (Farrar and Johnson-Groh, 1991, 1991; Gilman et al., 2015; Mickel and Smith, 2004; Popovich et al., 2020; Stensvold et al., 2002; Stensvold and Farrar, 2017), utilizing cytology, allozyme data, and detailed morphological analyses gradually increased the number of taxa to currently 36 (Dauphin et al., 2017; PPG I, 2016), of which 30 have been described in the last four decades. The newly defined species exhibit subtle morphological differences often along with geographical and ecological distinctness. More recently, phylogenetic analysis based on non-coding plastid markers (rbcL, trnL L-F, rpl16 intron, psbA-trnH^{GUG}, matK intron) (Dauphin et al., 2017, 2014; Hauk, 1995; Hauk et al., 2012, 2003) uncovered further lineage diversity within *Botrychium*, especially among the European populations treated as *Botrychium lunaria* (L.) Sw. (Dauphin et al., 2017, 2014; Maccagni et al., 2017). These findings suggested that deeper investigations on *B. lunaria* populations might shed further light on the diversity within the group.

The *Botrychium lunaria* group is one of three monophyletic clades within *Botrychium*, called the Lunaria clade by (Dauphin et al., 2014). It includes what was long considered to be a single, morphologically variable species that occurs across the boreal and alpine regions of North America, Europe, and Asia with outposts in New Zealand and Australia (Clausen, 1938; Kato, 1987; Milde, 1869). As early as 1903, Underwood suggested that at least two distinct taxa occurred in North America, prompting his description of *B. onondagense*. However, subsequent studies treated *B. onondagense* as a variety or form of a *B. lunaria* (Butters and Abbe, 1953; Clarke and House, 1923; Clausen, 1938; Clute, 1905). Later, based on extensive morphological surveys, Wagner (1981) described *B. crenulatum*, a North American species, which has been widely accepted (Dauphin et al., 2017; Farrar et al., 2017; Stensvold and Farrar, 2017). When allozyme markers were used more extensively and the investigation area was extended to northern Europe, five additional taxa were recognized (*B. tunux*, *B. nordicum*, *B. neolunaria*, *B. yaaxudakeit*, and *B. lunaria* var. *melzeri*) (Stensvold et al., 2002; Stensvold and Farrar, 2017). The recognition of *B. neolunaria* as the common *B*. "*lunaria*" in North America re-evaluated Underwood's *B*. "*onondagense*" as the North American representative of European *B.* "*lunaria*". Until a few years ago, most taxa in the *B. lunaria* group were from in North America. More recently, however, it has been found that the genetic diversity among *B.* "*lunaria"* populations in western and central Europe may exceed the diversity found in North America (Dauphin et al., 2014, 2014; Maccagni et al., 2017). Some European lineages could be assigned to species previously described from North America, such as *B. tunux* (Stensvold et al., 2002). Yet, there remain several poorly understood lineages within *Botrychium* possibly representing additional species (Dauphin et al., 2017). These hitherto undescribed lineages are widespread throughout Europe, highlighting the uncertainty of taxonomic assignments of European *B.* "*lunaria"* and the limited understanding of the global diversity of the *B. lunaria* group.

Significant advances have been made to understand speciation processes within *Botrychium.* Polyploidization appears to be a major factor in some clades of *Botrychium*, and complex polyploid networks have been uncovered (Dauphin et al., 2016; Wagner, 1993; Wagner and Lord, 1956; Williams and Waller, 2012), but the relevance of polyploidization in the *B. lunaria* group remains unclear. Among the 33 accepted *Botrychium* species, nearly half are allopolyploids. Of the new taxa in the *B. lunaria* group, *B. yaaxudakeit* is noteworthy as the first and only allopolyploid species discovered in the group. Using allozyme analysis, Stensvold *et al.* (2002) found that *B. yaaxudakeit* carries a combination of alleles matching *B.* "*lunaria"* and *B. neolunaria* and shows morphological intermediacy. These criteria support the hypothesis that *B.*

yaaxudakeit was derived by hybridization between the two diploid species followed by chromosome doubling. Scattered sampling in North America and central Europe (Dauphin et al., 2016; Veselý et al., 2012; Williams and Waller, 2012) found no further evidence of polyploids in the *B. lunaria* group, suggesting that the speciation processes involved in the diversification of this group may differ from other *Botrychium* clades. However, diploid hybrids between *B. neolunaria* and *B. "lunaria"* have been detected by allozyme markers in North America, in western and northern Europe, and on the east coastline of Asia and Oceania (Stensvold et al., 2002; Stensvold and Farrar, 2017). These hybrids, commonly called introgressed hybrids, originate from the germination of spores produced by F1 hybrids. The term reflects the unequal contributions by the two diploid parents during the meiosis. Introgressed hybrids are diploid and fertile plants which are supposedly as isolated from their parents as allotetraploids are. The success of the hybrids suggests that introgression may favor rapid speciation by promoting adaptive divergence (Abbott et al., 2013). Phylogenetic analysis of plastid data showed that whereas the maternal donor of these hybrids was *B. neolunaria,* hybrids formed a distinct subclade (Dauphin et al., 2017), supporting this idea. Hence, both polyploidization and hybridization events have contributed to the diversification within the *B. lunaria* group, but their specific contributions remain poorly understood.

Diversification within the *B. lunaria* group may have been driven by successive isolation, migration, and dispersal events imposed by the Quaternary glacial-interglacial cycles (Stensvold, 2008). The *B. lunaria* group mainly occurs in circumboreal-temperate regions strongly affected by Quaternary glacial cycles (Ehlers et al., 2011; Stensvold, 2008). As in many other boreal and alpine plant groups, this may have led to complex population dynamics (Schönswetter et al., 2005), possibly associated with adaptations to different climatic or soil conditions (Allen et al., 2012; Alvarez et al., 2009; Lafontaine et al., 2018). For some species of the genus *Botrychium*, local ecological, climatic, and edaphic specialization has been reported. For example, populations of *B. campestre* from northeastern Iowa (USA) have slope preferences within habitats (Nekola and Schlicht, 1996), whereas populations of *B. matricariifolium* from Vosges (France) occur on sandy, acid soils, in areas with high annual mean precipitation and cool summers (Muller, 1986). The species *B. simplex* and *B. tenebrosum*, which form a closely related species pair that occurs in North America, Sweden, and Switzerland, also tend to have different preferences for wetter or dryer habitats, respectively (Ståhl et al., 2016; A. Maccagni pers. comm.). However, these species belong to other clades of *Botrychium*, and little is known on the climatic and edaphic niches of species from the *B. lunaria* group. Ecological niche segregation among two genetic lineages of *B.* "*lunaria*" has been suggested based on habitat preferences in the Alps (Maccagni et al., 2017). In North America, *B. tunux* and *B. yaaxudakeit* favor well drained soils whereas *B. neolunaria* grows on poorly to moderately drained substrates (Stensvold et al., 2002; Stensvold and Farrar, 2017). In the absence of further characterization of the ecological niches of the lineages within the *B. lunaria* group, it is currently impossible to address to which degree climatic and edaphic adaptation contributed to the diversification of the group.

In this study, we analyzed the worldwide diversity and diversification time of the *B. lunaria* group and associated the occurrence of individual lineages with climatic and edaphic factors. We performed a plastid-based phylogenetic reconstruction including a large set of newly collected samples from eastern Europe and Asia. We also investigated genome size variation to identify potential polyploidization events throughout the diversification of the *B. lunaria* group. Specifically, we asked the following questions: 1) How many species-level taxa can be recognized in the *B. lunaria* group? 2) What is the species diversity of the group in eastern Europe and Asia? 3) What is the role of polyploidization in the diversification of the group? 4) Can the timing of species divergences can be related to Pleistocene climatic fluctuations? 5) Is there ecological differentiation between the species that may be linked to the diversification of the group?

Material and Methods

Sample collection

We collected 172 fresh sporophytes at 63 locations within Eurasian Mountain ranges, including 41 sites in central and eastern Europe (Alps, Carpathians, Dinaric Alps, Rhodopes, Central Balkans, and Caucasus) and 14 in central and eastern Asia (Pamir-Alay, Tien-San, Eastern Himalayas, Qinling, and Yann) (Figure 1B). At each location, we sampled 3-15 plants at least 20 cm apart, collected soil samples around the roots of one or two individuals, measured the soil pH using a Hellige pH-meter and recorded the geographical coordinates at the center of the sampling site using a Garmin eTrex 30X GPS. From each plant, a few pinnae were silica-dried on site and the remaining plant kept as herbarium voucher. Vouchers were deposited in the herbaria of the University of Zurich (Z/ZT) and University of Neuchâtel (NEU), Switzerland, and in herbaria in the countries of origin (supplementary Table S1). In addition, we also incorporated 203 samples from China, Russia, Europe, North Africa, and North America provided by collaborators (see acknowledgments) and herbaria (PE, MW, AIX, ISC, MARK) (supplementary Table S1), which included type material of five recognized taxa (*B. tunux*, *B. neolunaria*, *B. crenulatum*, *B. yaaxudakeit*, *B. lunaria* var. *melzeri*). The sampling was designed to cover the whole distribution range of the *B. lunaria* group as described by Stensvold (2008) (Figure 1B).

DNA extraction, amplification, and sequencing

We extracted total genomic DNA from silica-dried and herbarium material using the DNeasy plant mini kit (Qiagen, Hilden, Germany). We modified the manufacturer's protocol as follows: the mixture was incubated in buffer AP1 and RNase A for 20 minutes, a centrifugation step of 1 minutes at 8000 rpm was added before the centrifugation of the 17th step, and DNA was eluted in ultra-pure water. Total genomic DNA was quantified and quality-checked by spectrophotometry (NanoDrop 2000). The four non-coding plastid loci targeted (trnH^{GUG}-psbA intergenic spacer, trnL^{UAA}-trnF^{GAA} intergenic spacer, matK intron, and rpl16 intron) were amplified by polymerase chain reaction (PCR) following protocols and thermocycling conditions published by Dauphin *et al.* (2014; 2017) and Small *et al.* (2005). Besides, we optimized thermocycling conditions for touchdown protocol and the use of a high-fidelity DNA polymerase to amplify DNA degraded and extracted from herbarium material (supplementary Table S2). Briefly, the PCR reactions were performed in a volume of 50 µl containing 10 ng of DNA template, 0.2 µM of each primer (Microsynth, Balgach, Switzerland), 0.2 mM of DNTPs mix (Promega, Madison, USA), 1 unit of GoTaq DNA polymerase (Promega), 1X GoTaq buffer (Promega), and an additional 1 mM of $MgCl_2$ (Promega) for trnL^{UAA}-trnF^{GAA}. Herbarium material DNAs were amplified using Q5 high-fidelity DNA polymerase with the following PCR product concentrations: 0.3 µM of each primer (Microsynth), 0.2 mM of DNTPs mix (Promega), 0.5 unit of Q5 (New England Biolabs Inc., Ipswich, England), 1X Q5 buffer (New England Biolabs Inc.), and an additional 1 mM of Q5 enhancer (New England Biolabs Inc.) for trnL^{UAA}-trnF^{GAA}. Each PCR product was quality-checked by electrophoresis on agarose gel (1.5 %) stained with GelRed® (Biotium, Fremont, USA) and visualized under UV light on a transilluminator (UVI Tec Gel Documentation System). Template purifications and sequencing were outsourced to Macrogen Europe (Amsterdam, Netherlands). Succinctly, sequencing reactions were performed in the Master Cycler pro 384 (Eppendorf) using the ABI BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), following the protocols supplied by the manufacturer. Single-pass sequencing was performed on each template using the same primers combination than PCR. The fluorescent-labeled fragments were purified from the unincorporated terminators with the BigDye XTerminator® Purification Kit (Applied Biosystems). The samples were injected to electrophoresis in an ABI 3730xl DNA Analyzer (Applied Biosystems). The four regions were paired end sequenced. The remaining silica material and DNA extractions were deposited at the Botanical Institute of Zurich, Switzerland.

Sequence cleaning and alignment

Sequences were trimmed and ambiguous bases annotated using the default settings of Geneious v.8.1.9 (https://www.geneious.com) *Trim ends* and *Find heterozygotes* plug-ins. Consensus sequences were obtained from paired end contigs, and uncertain position were manually cleaned according to the base quality of the chromatograms. In the case of ambiguous bases presenting equal chromatogram quality on both sequences, ambiguity was kept, and the base coded according to the IUB code (Cornish-Bowden, 1985). Consensus sequences were aligned per locus using MAFFT v. 7.017 (Katoh and Standley, 2013) under the G-INS-i strategy and default parameters. Alignments were visually inspected and manually curated using Geneious. For instance, gaps in low quality 3' and 5' end regions were annotated as missing data. A minority of individuals presented partial sequencing data (6 %, supplementary Table S1, Table 1). Additional published sequences of *Botrychium* (*e.g.,* outgroups, specimens of the synonymized *B. onondagense* and supplementary type material) and of the Botrychioideae sub-family were incorporated to our dataset (Dauphin et al., 2017, 2014; Maccagni et al., 2017; Zhang et al., 2020) (supplementary Table S1, S3). A third of these supplementary accessions had information at the trnH^{GUG}-psbA and trnL^{UAA}-trnF$^{GAA.}$ loci only. The uniparentally inheritance (Gastony and Yatskievych, 1992; Guillon and Raquin, 2000; Kuo et al., 2018; Vogel et al., 1998) and the absence of recombination (Ravi et al., 2008) in fern chloroplasts allowed to analyze the four targeted loci as a concatenated single locus. Alignments per locus were concatenated using the R v.4.0.3 (R Core Team, 2018) seqinr v.4.2-4 (Charif and Lobry, 2007) package and the *concat* function (Khang, 2016). The ingroup was represented by 521 individuals with type material of each named taxon within the *B. lunaria* group except for *B. lunaria* and *B. onondagense*. The outgroup contained nine diploid species of the two other *Botrychium* clades (Lanceolatum and Simplex-Campestre) (supplementary Table S1).

Phylogenetic reconstruction

We assessed the phylogenetic relationships within the *B. lunaria* group based on two datasets (*i.e.*, multiple alignment): one containing all the 530 accessions (full dataset - A2; supplementary File S1) and one containing a subset of 323 individuals that excluded 82 samples with partial sequencing data and 116 accessions carrying low phylogenetic signal (rogue individuals) (reduced dataset - A1; supplementary File S2). Rogues were identified using RogueNaRok (Aberer et al., 2013). The calculations were based on one hundred bootstrap trees estimated with the best tree inferenced from the multiple alignment exempt of partial data under a TIM1+I+G substitution model using RAxML-NG v. 1.0.1 (Kozlov et al., 2019) (Table 1). For each dataset, the best scheme and the optimal substitution model were assessed using PartitionFinder 2 (Frandsen et al., 2015) with *branch lengths* set as *unlinked*, and the *substitution models* set as *all* or *MrBayes*. Optimal substitution models were chosen according to the Akaike information criterion (AICc). Indels were scored for the A1 dataset using the *2-matrix* script (Salinas and Little, 2014) which implemented the simple indel coding algorithm described by (Simmons and Ochoterena, 2000). In total, 36 indels were scored (supplementary File S2). Unrooted species trees were built with maximum likelihood (ML) using RAxML-NG and Bayesian inferences (BI) using Bayesphylogenies 2 (Pagel et al., 2004).

The reduced dataset was analyzed both with and without indels. We ran ML inferences under a TIM1+I+G substitution model and BIN model for the indels with a fixed random seed of 2. The tree search was performed on 25 random and 25 parsimony-based starting trees and branch support was estimated over 1,000 bootstrap replicates. The minimum number of bootstrap replicates was also determined using the autoMRE bootstrap convergence test (Pattengale et al., 2010). Bootstrap values were depicted on the best tree using the ***Transfer Bootstrap Expectation*** metric (Lemoine et al., 2018). BI were conducted under GTR+I+G substitution model and M1P models for indels using the reversible jump method. We set two runs of 50 million MCMCMC generations, each including one cold and 2 heated chains, with a seed number of 946432 and tree sampled every 5,000 generations. Parameter files were inspected with tracer v. 1.7 (Rambaut et al., 2018) and a burn-in of 10 and 14% was applied to the analyses with and without indels, respectively. Majority-rule consensus (MRC) trees of 50% were built based on posterior probabilities (PP) using SumTrees v. 4.4.0 and DendroPy library v. 4.4.0 (Sukumaran and Holder, 2010). The full dataset was analyzed with a constraint backbone to minimize the effect of incomplete data and rogue individual on tree inference. We used the best ML tree of the reduced dataset as the constraint backbone. This phylogenetic inference was performed by ML only under a TIM1+I+G substitution model with a fixed random seed of 2. The tree search was performed on 25 random and 25 parsimony-based starting trees and branch support was estimated over 100 bootstrap replicates. Trees were depicted using R ggtree v. 2.4.0 (Yu et al., 2018, 2017), ape v. 5.4-1 (Paradis and Schliep, 2019), treeio v. 1.14.0 (Wang et al., 2020)and ggplot2 v. 3.3.3 (Wickham, 2016), tidyverse v.1.3.0 (Wickham et al., 2019), viridis v.0.5.1 (Garnier, 2018), scales v.1.1.1 (Wickham and Seidel, 2020), ggtreeExtra v.0.4.5 (Xu and Yu, 2021), ggstar v.1.0.1 (Xu, 2021), RColorBrewer v.1.1-2 (Neuwirth, 2014), and ggnewscale v.0.4.5 (Campitelli, 2021)packages and Figtree v.1.4.4 (Rambaut, 2009).

Ploidy level assessment

We assessed the ploidy level of 57 individuals covering the whole phylogeny using flow cytometry following the one-step methodology of Dolezel and Bartos (2005). Briefly, approximately 0.125 cm^2 of fresh or silica-dried internal standard leaf (*Pisum sativum* ´Ctirad`; 2C= 9.09 pg, Doležel et al., 1998) and 0.125-0.250 cm^2 of silica-dried fern tissue were co-chopped using a fresh razor blade in a Petri dish containing 1 ml of ice-cold Otto I buffer (0.1 M citric acid, 0,5% Tween 20). The resulting suspension was incubated for 15 min at room temperature (20°C) with occasional shaking and filtered through a 42µm nylon mesh. The filtrate was stained with 1 ml of Otto II buffer (0.4 M Na_2HPO_4*12 $H_2$0; fluorochrome DAPI: 4 µg/ml) for 1-2 min at room temperature, after which the relative fluorescens was recorded on a CytoFLEX S (Beckman Coulter, Indianapolis, USA). The excitation beam was provided by a laser tuned to a wavelength of 351 nm. Fluorescence emission was detected by a filter permitting passage of light of wavelength of 450 nm. The 2C-values were calculated by comparing the mean *Botrychium* peak to the mean standard peak. Ploidy level were also derived from spore sizes (Popovich et al., 2020) for 14 individuals with genetic proximity to the tetraploid *B. yaaxudakeit* and its maternal donor *B. neolunaria* for which silica-dried material were not available. Spore size of three *B. neolunari*a and five *B. yaaxudakeit* specimens identified by allozymes (Stensvold 2008, D. Farrar unpubl. data) were used as reference. Dried spores were placed in a 1.5 ml Eppendorf tube and suspended in 300 ml of Euparal (Roth, Karlsruhe, Germany) mounting medium. Nine drops of spore solution were equally distributed on a cover slip on the top of which a slide was gently placed. The slides were dried cover slip faced down during a minimum of two weeks before spore measurement. Spores were photographed using inverted microscope (Leica DMI3000B). The spore's longest length was measured on a minimum of ten spores per individual using LAS v.4.13. Leica Software and ImageJ v.1.52a (Rasband, 1997). Average and standard deviation of spore's lengths were calculated in R.

Species delimitation

We considered taxa as independently evolving metapopulations lineages following the concept of de De Queiroz (2007) by applying an integrative approach (Padial et al., 2010). We selected well-supported monophyletic clades to be candidate species (Solís-Lemus et al., 2015; Sukumaran and Knowles, 2017). Well-supported monophyletic sub-clades nested within species were treated as subspecies when the remainder of the species-level clade did not form coherent units, so that a separation into two or more species was impractical. Species status of clades was further supported by coherent geographical distribution, or distinct ecological features, in combination with previously published information on reproductive isolation between taxa, when available (Wagner and Wagner, 1983b).

Divergence time

We estimated divergence times within the *B. lunaria* group in a two-step time calibration. First, we estimated the divergence time of the *Botrychium* crown, and second, we inferred the divergence times of the *B. lunaria* group with the *Botrychium* crown age as a secondary derived calibration. In the first step, we analyzed the concatenated alignment of the subfamily Botrychioideae (supplementary File S3, table S3) under strict and uncorrelated lognormal relaxed clock models using BEAST v.1.10.4 (Drummond et al., 2012). The alignment was partition according PartitionFinder 2 output (Table 1) and the site models were unlinked. A birth-death *tree prior* was applied with a given starting tree as *tree model.* The starting tree was built using RAxML-NG and transformed into an ultra-metric tree using the penalized likelihood method (Sanderson, 2002) implemented in the chronopl function of the R ape v. 5.4-1 package. The *clock rate* or the *ucldMean* priors were set with an exponential distribution and defaults parameters. We used two macrofossils to constrain the stem ages of the *Botrypus virginianus* and *Sceptridium* (Bozukov et al., 2010; Rothwell and Stockey, 1989). Morphological analysis related the youngest fossil to *Sceptridium underwoodianum* (Bozukov et al., 2010), whereas the oldest fossil was described as an extinct species, *Botrychium wightonii* Rothwell & Stockey, related to *Botrypus virginianus* (Rothwell and Stockey, 1989). We applied to the *tmrca priors* a uniform distribution with the minimum age of fossils as lower boundaries (56.8 and 20.44 million years, respectively) and the Ophioglossaceae family age estimated by Lehtonen et al. (2017) as upper boundaries (148.9 million years). We set a normal distribution for the *treemodel-rootheight* prior with a mean age of 148.9 million years (Lehtonen et al., 2017) and a standard deviation of 1.0, and we truncated it with an upper boundary in the middle Jurassic (175.6 million years). We constrained the calibrated clades as monophyletic and set *Helminthostachys zeylanica* as outgroup of the subfamily Botrychioideae to facilitate the convergence of the analysis. We ran one analysis per clock model of 20 and 50 million MCMC generations for the strict and the relaxed clock respectively, with parameters sampled every 1,000 generations. Besides, those two analyses were performed without sequence data and without fossil calibrations to control for the influence of the priors on the time-divergence estimates.

Second, we analyzed the A1 alignment containing only unique sequences created by RAxML-NG (supplementary File S4) under strict and uncorrelated lognormal relaxed clock models using BEAST. A yule *tree prior* was applied with a given starting tree as *tree model*. The starting tree was built with the same method than the Botrychioideae starting tree. The *clock rate* or the *ucldMean* priors were set with an exponential distribution and defaults parameters.

We set the distributions parameters using the time-divergence estimates of the genus *Botrychium* given by the analysis of the subfamily Botrychioideae. The *treemodel-rootheight* priors were defined by a normal and a gamma distribution for the strict and the relaxed clock, respectively. The normal distribution had a mean of 14.1 million years, a standard deviation of 1.6 and we truncated it with an upper boundary of 20.7 million years (upper range boundary of *Botrychium* mean). The gamma distribution had a *shape* of 19.5 million years (mean age), a *scale* of 0.7 and we truncated it with an upper boundary of 73.1 million years. We constrained the calibrated clade and *Lunaria* and *Simplex-Campestre* taxon sets as monophyletic to facilitate the convergence of the analysis. We ran one analysis per clock model of 50 million MCMC generations with parameters sampled every 1,000 generations.

The convergence of all analysis was inspected in Tracer v.1.7.1. The effective sample size of every parameter was above 200 and we defined a burn-in of 10 %. We summarized our post burn-in trees using TreeAnnotator v.1.10.4 (Drummond et al., 2012) to generate a maximum clade credibility tree. Besides, we sampled priors for 100 million generations and compared the time divergence estimates with those obtained with the run including the data to ensure the estimates were not driven by priors. Trees were depicted using R ape, phytools v.0.7-70 (Revell, 2012), treeio, ggtree, ggplot2, deeptime v.0.0.5.3. (Gearty, 2021) and strap v.1.4 (Bell and Graeme, 2014) packages. The densitrees were drawn using subsamples of posterior trees obtained with Logcombiner v. 1.10.4 (Drummond et al., 2012).

Distribution ranges and climatic analysis

We used 511 genotyped individuals with geographical coordinates to define species distribution ranges. The species distribution ranges were visualized using R ggplot2, sf v.0.9-7 (Pebesma, 2018), ggspatial v.1.1.5 (Dunnington, 2021), rnaturalearth v.0.1.0 (South, 2017), and RColorBrewer packages.

We extracted the 19 bioclimatic variables from CHELSA (Karger et al., 2018, 2017) for the collecting localities of a subset of 465 individuals which excluded 46 individuals: 30 individuals with inaccurate geographic coordinates, one individual with uncertain ploidy level, and 15 individuals presumably being introgressed hybrids (Stensvold, 2008; D. Farrar unpubl. data) (supplementary Table S1). The 39 individuals of the *B.* aff. *neolunaria* group were only used to compare the climatic niches of the *B. neolunaria* s.str. and *B.* aff. *neolunaria* groups. The subclade 2 of the *B.* aff. *neolunaria* group had an insufficient number of geo-referenced individuals to be considered in the analysis. Despite of the removal of some accessions, the subset of individuals covered all species and subspecies we recognized. The variable scales varied widely and because variables such as the minimum temperature of the coldest month (BIO 6) had negatives values, log transformation could not be uniformly applied. Thus, we normalized the data using the Z-score standardization. The overall differences between species were visualized for each bioclimatic variable with boxplot representations (supplementary Figure S1). The variables showing differences between taxa were identified both at species and at subspecies level using Analysis of Variance (ANOVA) (supplementary Files S11 to S14). Then, to see which species differed from each other, we performed Tukey post-hoc tests (supplementary File S15). We retained all variables for the analysis at the species level, and 4-16 variables for the analysis at the subspecies level (supplementary Files S16-S19). To visualize the climatic niches of both species and subspecies, we performed principal component analyses (PCA) to account for covariance between climatic variables (supplementary Files S16-S23). The data normalization was done using R *scale* function and analysis were conducted using the R rstatix v.0.6.0 (Kassambara, 2021) and FactoMineR v.2.4 (Lê et al., 2008) packages. The data and results were visualized using the R ggplot2 tidyverse and factoextra v.1.0.7 (Kassambara and Mundt, 2020), cowplot v.1.1.1 (Wilke, 2020), and ggpubr v.0.4.0 (Kassambara, 2020) packages.

Soil pH measurements obtained through field expeditions (see sampling collection section) was refined by measurements realized in the soil biology laboratory of the University of Neuchâtel. Briefly, soil samples were dried at 40°c. Then, one volume of dried soil was mixed with 2.5 volumes of water and stirred 3 times for 5 seconds with 20 minutes between each. The pH was measured 20 minutes after the last agitation using a 914 pH/Conductometer (Metrohm SA, Zofingen, Switzerland). In total, we measured soil pH for 50 populations represented by 54 samples covering eight of the ten clades (supplementary Table S1). The association patterns of soil pH and species were visualized with boxplots using R ggplot2 and tidyverse packages.

Results

Locus diversity and plastid-based phylogenetic analysis

The concatenated plastid DNA regions consisted of 2,374 aligned nucleotides of which 4-10 % were parsimony-informative sites (Table 1). A large proportion of individuals (> 80 %) were analyzed at all four target loci (supplementary Table S1). The loci were not equally informative (Table 1). The trnH^{GUG}-psbA intergenic spacer and the rpl16 intron were the most variable (114 and 142 parsimony informative sites, respectively), whereas the trnL^{UAA}-trnF^{GAA} intergenic spacer was the least informative (81 sites). However, the rpl16 intron presented the lowest sampling completeness due to repeated PCR amplification failure and limited availability of published data (supplementary Table S1, Table 1).

The phylogenetic analysis recovered ten well-supported monophyletic clades inside the *B. lunaria* group (Figure 2, see supplementary Figure S2, S3 and File S5 for tip labels and support values of all branches), consistently arranged into five main groups of clades (Figure 2B, 2C). A first group included three clades (clades 3, 4, and 5), two included each a unique clade (clade 8 and clade 9), another included two clades (clades 2 and 7), and the last included three big clades (clades 1, 6, and 10). The trees resulting from the ML and BI analysis based on the reduced dataset retrieved the same main groups of clades and clades (Figure 3; supplementary Files S6 and S7), except for clades 7 and 2 that were not clustered in the BI analysis, and clade 8 which was only recovered in the ML analysis. However, the relationships between clades were not concordant between the ML and BI analyses (Figure 3).

The addition of indels slightly increased the branch lengths and the deeper node support in both ML and BI trees (supplementary Figure S4 and Files S8 and S9). The topologies recovered with indels were largely concordant with the analyses based on the reduced trees, and individuals were assigned to the same clades, with few exceptions (supplementary Figure S4). The use of a constrained topology (supplementary File S10) for the ML analysis based on the full dataset allowed the assignment of most individuals with partial sequencing data and low phylogenetic signal. Of the 521 accessions in our study, we were able to reliably assign 509 (97.7 %) to one of the 10 clades. The remaining 12 specimens had uncertain placement, mostly due to incomplete or unreliable sequence data. Eight of these individuals appeared on exceedingly long branches and their placement remained debatable (Figure 2, supplementary Figure S2). Moreover, the assignment of one rogue individual from the Alps (AM-lun33-3-CH) to a clade otherwise restricted to Eastern Asia was suspicious enough to prune it from the final tree.

Some of the ten main clades recovered corresponded to previously recognized species (e.g., *B. tunux* and *B. crenulatum* to clades 5 and 10, respectively), whereas in other cases they combined recognized species, such as *B. neolunaria* and its polyploid derivate *B. yaaxudakeit* (clade 10; Figure 2B, 2C). Some clades showed clear phylogenetic structure within them. We identified six such subclades (Figure 2C) in the tree resulting from the ML analysis based on the full dataset. Three were within clade 1, namely subclades 4 (*B. lunaria* var. *melzeri*), 5, and 6 (*B. nordicum*), two were in clade 10 (subclades 2 and 3), and one was found in clade 6 (subclade 1). Of these subclades, three are geographically restricted (central Alps, Carpathians, and Baikal; subclades 5, 3 and 2, respectively) and three were also recovered in the BI trees based on the reduced dataset (subclades 4, 2 and 3).

Ploidy level assessments

The assessment of the ploidy level within the *B. lunaria* group indicated a dominant diploid condition. Absolute genome size showed no evidence of ploidy level variation with 2C-values of 55 specimens between 19.38 and 22.91 pg which attest to a diploid condition (Dauphin et al., 2016; Veselý et al., 2012; Williams and

Waller, 2012), except for two individuals from clade 10 subclade 3 with 2C-values of 17.49 and 17.91 (supplementary Table S4). Spore measurements made on 14 additional individuals showed a similar pattern with one exception. Most had length averages between 31 µm and 39 µm (supplementary Table S5) which corresponds to the range found for the diploid *B. neolunaria* (Stensvold and Farrar, 2017). However, one specimen of Caucasus Mountains had a length average of 43 µm which enters the range estimated for the tetraploid *B. yaaxudakeit* (Stensvold et al., 2002).

Species-clade level classification and candidate species

Well-supported monophyletic clades containing type material or independently identified specimens by allozyme markers (Stensvold, 2008; Stensvold and Farrar, 2017; D. Farrar unpubl. data) of previously recognized species were assigned to the corresponding species, with clades 5, 1, 9 and 2 corresponding to *B. tunux*, *B. lunaria*, *B. crenulatum*, and *B. onondagense*, respectively. Among these, *B. onondagense* is currently synonymized with *B. lunaria*. Our species criterion of well-supported monophyly was not applicable to clade 10 because it contained both diploids and polyploids described as distinct species (*i.e.*, *B. neolunaria* and *B. yaaxudakeit;* Stensvold et al., 2002; Stensvold and Farrar, 2017). In this case, all the specimens of clade 10 were grouped under the name of *B. neolunaria* group, including tetraploids previously identified by allozymes (Stensvold 2002, 2008; D. Farrar unpubl. data) recognized as *B. yaaxudakeit,* specimens previously identified as *B. neolunaria* called *B. neolunaria* s.str. and all remaining specimens known or assumed to be diploids recognized as *B.* aff *neolunaria* group. We did not treat all diploids as the species *B. neolunaria* because they are known to include introgressed hybrids with *B. neolunaria* as maternal parent (Stensvold, 2008; Stensvold and Farrar, 2017). Thus, while they cluster in the same cade, they have different genetic backgrounds and very variable morphology (V. Mossion unpubl. data). The remaining five clades are treated as unnamed candidate species and named by their clade numbers (species 3, 4, 6, 7 and 8). We further recognized four monophyletic clades nested within species that are treated here at subspecies level for consistency, including the taxa recognized as species level as *B. nordicum* and as variety level as *B. lunaria* var. *melzeri* (Stensvold and Farrar, 2017). We thus use the names *B. lunaria* "subsp. *nordicum*" and *B. lunaria* "subsp. *melzeri*", but without formal recombination. The subclades nested within the *B. neolunaria* group (clade 10) are not treated at the subspecies level as their species assignment is unclear.

Divergence time estimates

Our age estimates suggest that the radiation of the *B. lunaria* group started at the end of the Miocene and extended to the Pliocene, about 45 million years (myr) after the genus *Botrychium* diverged from its sister genera *JapanoBotrychium* and *Sceptridium* (Figure 4A, supplementary Figure S5). The results are comparable between the two clock models even though the uncorrelated lognormal relaxed clock model recovered higher divergence time estimates for both genus- and species-level analyses (Figure 4B, supplementary Figures S5, S6, and S7). For example, the mean estimates of the crown ages of the genus *Botrychium* were 19.5 (95 % HDP 7.1-22.6 myr) and 14.1 (95 % HDP 8.7-20.8 myr) under the uncorrelated lognormal relaxed clock model and the strict clock model, respectively. The *B. lunaria* group arose in the Pliocene epoch (mean 3.6 myr, 95 % HDP 2.38-5.053 myr) while the species within mostly diverged from one another during the Gelasian, the first Pleistocene age at 2.58-1.8 myr (Figure 4C, supplementary Figure S6, and S7). The divergence times of the subspecies were estimated within the 95 % interval confidence only for subspecies "a" of species 6 and for *B. lunaria* "subsp. *nordicum*" due to missing data, rogue individual, and a lack of a sufficient number of unique sequences. They diverged from the remaining of the clade at the end of the Calabrian and Chibanian ages, the second and third quarter of the Pleistocene epoch at ~1 myr and ~0.5 myr, respectively (Figure 4C, supplementary Figure S6 and S7).

Distribution ranges of species in the B. lunaria *group*

Our classification into species results in the recognition of seven species in North America (USA, Canada, Greenland), seven in Asia, one each in northwestern Africa and Oceania, and eight to nine (considering the potential tetraploid from Georgia) species in Europe (Figure 2A, Figure 5A-C). Of the 11 candidate species found among our samples, five (species 4, 7, 8, *B. crenulatum*, and *B. yaaxudakeit*) had narrow geographical ranges. These species were restricted to northern and eastern Europe, the Himalayas, as well as western and northern North America. The remaining species had wider distributions. For instance, *B. lunaria* (Figure 5A) was recovered from northwestern Africa to Scandinavia and Central Asia. *Botrychium tunux* and *B. onondagense* both had a disjunct distribution in Europe and North America with a disjunction between Eastern European and Central Asian populations, and no sample from western North America was assigned to *B. onondagense*. Species 3 and 6 were both found from Central and Eastern Europe to eastern Asia with a marked disjunction in Central Asia for species 3 and a few records in Alaska for species 6. The most widespread species-level group was the *B. neolunaria* group (Figure 5B), which we recorded across almost the entire range of the *B. lunaria* group.

Climatic and soil analyses

The climatic niches largely overlapped between taxa, albeit with slight local differences. The principal component analysis (PCA) clustered the species into two groups. The first group consisted of species 1, 3, 8, and *B. crenulatum*, and the second group of species 2, 4, 6, 7, *B. neolunaria* s.str., and *B. tunux* (Figure 6). The first component (PC1, 40.0 %) separated the species by temperature-related variables (BIO 1, 11, and 6) showing a gradient of strength of winter from mild winters (first group) to strong winters (second group) (Figure 6A-B). The second component (PC2, 23.8 %) was correlated with precipitation-related variables (BIO 12, 13, 14, 16, 17, 19) and temperature seasonality (BIO 7 and 4), corresponding to a gradient of climatic continentality from less seasonal climates with high annual precipitations in areas close to oceans to strongly seasonal climates with low annual precipitations in continental regions. Within both groups, species were arranged along this continentality gradient. For example, in the second group species were arranged from species 7 (oceanic) to species 4 (continental) (Figure 6C-D). The third component (PC3, 12.6 %) was related to temperature-related variables (BIO 10, 5 and 1) expressing a gradient of summer temperature, with the first group under cooler summer climates than the second group. PC3 was also correlated to diurnal temperature fluctuations and their constancy through the year (BIO 2 and 3; supplementary Figure S8). Thus, the PCA recovered that *B. crenulatum* and species 7 occur in habitats with a large range of diurnal and annual temperature fluctuations, with the former subjected to even stronger temperature variations.

We also observed differences in the climatic niches between subspecies (supplementary Figure S9) and between the *B. neolunaria* s.str. and *B.* aff. *neolunaria* groups. The PCA within species 1 showed substantial climatic niche differentiation explained by seasonality (PC1, 40.3 %), high temperatures (PC2, 26.8%), and the strength of the winter (PC3, 15.7 %). In brief, *B. lunaria* "subsp. *melzeri*" occurred in dry climates with a strong seasonality and cold winters, *B. lunaria* "subsp. *nordicum"* in wet climates with low seasonality and low annual temperatures, and the two others under more temperate climates. The PCA within species 6 revealed similar importance of the seasonality (PC1, 62.3 %) and precipitation (PC2, 16.9 %), with the nominal subspecies in dry habitats characterized by medium seasonality and cold winters, and subspecies 1 in wet habitats characterized by low seasonality and mild winters. The PCA of the B. *neolunaria* showed differences between *B. neolunaria* s.str., and *B.* aff. *neolunaria* group related to summer temperature (PC1, 50.78 %), and seasonality (PC2, 26.2 %) with *B. neolunaria* s.str. and the *B.* aff. *neolunaria* group in climate with hot and dry summer, and higher summer temperatures for the former, and except subclade 3 individuals in climates with cold and wet summers.

Soil pH showed differences between species even though the sample size was limited. *Botrychium lunaria*, species 6, and species 8 were only found on acidic soils, whereas *B. tunux*, species 3 and species 7 were recovered exclusively from neutral soils (Figure 6E-F). *Botrychium* aff. *neolunaria* group was mostly found on both acidic and neutral soils. Interestingly *B. onondagense* was found across the whole range of pH values recorded in the *B. lunaria* group with a trend towards the acidic soils.

Discussion

Based on over 500 sampled individuals, we provide the first global molecular phylogenetic analysis of the *Botrychium lunaria* group. The lineage diversity recovered is congruent with previous, geographically more limited studies and uncovered three novel monophyletic clades, resulting in the recognition of 11 candidate species and 4 candidate subspecies. These are evenly distributed among whole northern hemisphere, with seven or eight species per continent. We found little evidence that polyploidization played a role in the diversification of the group, although hybridization appears to be important in one clade. The time-calibrated phylogeny shows a concordance of Pleistocene climatic oscillations with the radiation of the group. Moreover, climate and soil pH showed slight differences among candidate species and subspecies, suggesting that ecological drivers played a role in the diversification of the *B. lunaria* group.

We freely acknowledge that one limitation of our study is that it only includes non-coding plastid regions, so that hybrids and polyploid individuals may not have been recognized in the absence of nuclear markers. However, as discussed below, the clades recovered by us make up consistent geographical and ecological units, lending support to the notion that they are real evolutionary entities.

Taxonomic diversity in the B. lunaria *group*

Our plastid-based phylogeny recovered ten well-supported clades, of which three are novel (clades 6, 7, and 8) compared to earlier studies based on the same non-coding plastid regions (Dauphin et al., 2017, 2014; Maccagni et al., 2017). The remaining seven clades are consistent with previous studies even though their relationships differ. Clades are arranged in five main groups in our analysis against two in the previous studies. For example, clades 9 (*B. crenulatum*) and 2 were recovered separated from each other, whereas they were previously retrieved as sister clades (Dauphin et al., 2017).

The ten clades found in the *B. lunaria* group include 17 candidate taxa. Eight of these correspond to currently accepted taxa, namely *B. lunaria* (Swartz, 1801), *B. crenulatum* (Wagner and Wagner, 1981), *B. tunux* (Stensvold et al., 2002), *B. yaaxudakeit* (Stensvold et al., 2002), *B. nordicum* (Stensvold and Farrar, 2017); our *B. lunaria* "subsp. *nordicum*"), *B. neolunaria* (Stensvold and Farrar, 2017), and *B. lunaria* var. *melzeri* (Stensvold and Farrar, 2017); our *B. lunaria* "subsp. *melzeri*"). Four other clades have previously been recovered in phylogenetic analyses but have not been described or were synonymized, namely *B. onondagense* (Underwood, 1903) (Clade 2), synonymized with *B. lunaria* by Clute (1905) as well as species 4, species 3, and *B. lunaria* subsp. "a" (Dauphin et al., 2017, 2014; Maccagni et al., 2017). Finally, four clades are novel candidate taxa (species 6, 7, 8, species 6 subsp. "a"). This increased number of taxa is the result of our geographically more extensive sampling and overall increased sample size.

The consistency of the ten species-level clades in the ML and Bayesian analyses, despite different arrangements among the clades in the two analyses, in combination with their ecological, geographical, and morphological distinctness (Stensvold et al., 2002; Stensvold and Farrar, 2017; Wagner and Wagner, 1981; V. Mossion unpublish. data), supports the notion that these clades are best treated at species level. Although hybrids have been documented in the *B. lunaria* group (Stensvold, 2008), the fact that co-occurring species such as *B. neolunaria* and *B. tunux* in North America (Figure 1Ag) conserve their morphological distinctiveness suggest these species are reproductively largely isolated (Wagner and Wagner, 1983b). Unfortunately, directly testing for reproductive isolation between lineages of the *B. lunaria* group by breeding experiments is technically restrained by the difficulty of cultivating the subterranean gametophytes (Campbell, 1911; Whittier, 1972, 1981; Mossion unpubl. data). We are nevertheless confident that our treatment of the ten major clades recovered in our analyses as species is sound, along with the recognition of allotetraploid *B. yaaxudakeit* as an eleventh species.

The taxonomic treatment of clades nested within the species-level clades is less straightforward, as exemplified by the case of *B. nordicum*, here referred as *B. lunaria* "subsp. *nordicum"*, which was found to be deeply nested within *B. lunaria*, corroborating previous findings (Dauphin et al., 2017). This taxon was described at species level based on morphological distinctness and unique alleles separating it from *B. "lunaria"* (Stensvold and Farrar, 2017). In our analysis, this taxon was recovered as a monophyletic clade nested within *B. lunaria* so that treating it as specifically distinct would render the remainder of *B. lunaria* paraphyletic. Our divergence time estimates revealed that *B. lunaria* "subsp. *nordicum"* has only recently derived from the rest of *B. lunaria*, and its nested placement is thus likely the result of incomplete lineage sorting (Pamilo and Nei, 1988; Wendel and Doyle, 1998). For consistency, we treat *B. lunaria* "subsp. *nordicum"* at subspecies level but considering its morphological distinctness and the allozyme patterns (Stensvold and Farrar, 2017), further investigations involving nuclear sequencing data are necessary to decide on any taxonomic rank modification. The same is true for the other subspecies recognized here.

Finally, the *B. neolunaria* group presents special taxonomic problems. Not only does it include allotetraploid *B. yaaxudakeit*, which shares the plastid genome with it maternal parent *B. neolunaria* and thus cannot be distinguished with our data, but there also appears to be considerable genome size variation within this group. Thus, our genome size measurements for Eurasian individuals of the *B. neolunaria* group showed much lower 2C-values (17.49 – 21.88 pg) than reported for "true" *B. neolunaria* from North America (27.26-27.75 pg; (Dauphin et al., 2016). To date, no record of "true" *B. neolunaria* identified by allozyme markers is known from Eurasia, even though hundreds of plants have been analyzed (Stensvold, 2008). However, a considerable amount of fertile introgressed hybrids between *B. neolunaria* and B. "*lunaria"* has been recorded from Eurasia. We propose that the differences in genome size in the *B. neolunaria* group reflect the presence of several taxa originating from multiple hybridization events between *B. neolunaria* and one or several other species from the *B. lunaria* group. We would not be surprised if this group contains several species-level taxa, but our plastid markers only indicate the maternal donor, and do not allow to test this hypothesis. Further studies involving nuclear markers are needed to address the evolutionary history and taxonomic diversity of *B. neolunaria* group.

Global geographical distribution of the B. lunaria *group*

Based on intensive sampling in Western and Central Europe, Maccagni *et al.* (2017) proposed that continental Europe is the center of diversity of the *B. lunaria* group. However, our geographically more exhaustive sampling shows that species numbers are quite evenly distributed between continents, with seven species in North America, eight in Europe, and seven or eight in Asia. Importantly, our sampling is less dense in Asia (94 versus 346 specimens in Europe), with large areas yet to be sampled. It is thus conceivable that further species diversity might occur in Asia.

Dauphin et al. (2017) hypothesized that the *B. lunaria* group may have originated in Asia, since a clade mainly composed of Asian specimens was sister to the remainder of the group. Testing this hypothesis requires to firmly resolve the relationships between clades, which is not the case in our analysis. Indeed, the relationships between clades were inconsistent between the maximum likelihood (ML) and Bayesian inference (BI) methods. For example, clade 7, exclusively composed of Asian specimens, was sister to all the others in our BI trees, whereas the group of clades 4, 5 (*B. tunux*), and 3, which contain mostly European and North American specimens, was recovered as such in the ML trees. It appears that our increased sampling size is not sufficient to resolve the deeper nodes and that additional sequencing data, preferably including nuclear data, are needed to address question on the geographical origin of the *B. lunaria* group.

Considering the intensive previous studies in North America (Stensvold, 2008; Stensvold et al., 2002; Stensvold and Farrar, 2017; Wagner and Wagner, 1981), unsurprisingly most of the newly found candidate species have ranges restricted to Europe or Asia. Our study extends the known geographic distribution of most widespread species but confirms the narrow distributions of *B. crenulatum* and *B. lunaria* subspecies "a" (Dauphin et al., 2017; Stensvold, 2008; Wagner and Wagner, 1981). Perhaps the most surprising range extension concerns the *B. neolunaria* group, which was previously only known from Eastern Asia, Oceania, and North America (Dauphin et al., 2017; Stensvold and Farrar, 2017) but was recorded by us across all Eurasia. However, disentangling the distribution of introgressed hybrids and "true" *B. neolunaria* is not possible here as both share the same plastid genome. The same statement applies for the distribution of *B. yaaxudakeit*, the derivative polyploid of *B. neolunaria*. However, genome size estimates showed no evidence of tetraploid plants that might be referable to *B. yaaxudakeit* in Eurasia, except for a single specimen with large spores from Georgia, whose taxonomic identity remains unresolved.

Limited roles of polyploidization and hybridization in the diversification of the B. lunaria *group*

A previous study of other clades in the genus *Botrychium* has uncovered complex polyploid networks, revealing an important role of allopolyploidization in species formation in the genus (Dauphin et al., 2018). In contrast, we found little evidence of ploidy variation within the *B. lunaria* group despite extensive sampling and additionally selecting morphologically aberrant specimens that might be hybrids. In combination with previous studies (Dauphin et al., 2016; Veselý et al., 2012; Williams and Waller, 2012), we thus show that the diploid condition is predominant in the *B. lunaria* group. This finding indicates that ploidy variation is not a major driver of diversification in the *B. lunaria* group, with the exception of allotetraploid

B. yaaxudakeit (Stensvold et al., 2002). Nevertheless, as detailed above, hybridization may have led to still poorly understood diversification in the *B. neolunaria* group.

Pleistocene climatic oscillations influenced the radiation of the B. lunaria *group*

Our study shows that the current species diversity of the *B. lunaria* group originated during the last 3 million years, suggesting a possible link with the massive Pleistocene climatic fluctuations. Large parts of the current ranges of the species in the *B. lunaria* group were covered by ice during glacial times, including the entire known range of species 4 in boreal Sweden, and most of the habitat of the species occurring in the Alps (Ehlers et al., 2018; Ivy-Ochs, 2015). This is a well-known situation for alpine and boreal plant species which has frequently been invoked as driving the divergence of specific and intraspecific genetic differentiation by repeated cycles of extinction, migration, allopatry, and sympatry (Tribsch and Schönswetter, 2003; Westergaard et al., 2011). In the *B. lunaria* group, this has not resulted in the formation of primarily allopatric species, as for example in the European gentians (Hungerer and Kadereit, 1998), but rather in the formation of numerous widespread species that mostly broadly overlap geographically and that often co-occur. For example, in the Swiss Alps, *B. tunux* has mostly been found in mixed populations with species 3, while *B. lunaria* subsp. "a" and species 3 were also recovered in sympatry (Maccagni et al., 2017). In our study, we found similar patterns, with for example *B. onondagense* and species 3 co-occurring at a site in Romania, or species 6 and *B. lunaria* co-occurring at several sites in Bosnia and Herzegovina. This co-occurrence of species is certainly a result of the high dispersal ability of ferns by their spores (Barrington, 1993), especially long-distance dispersal by wind, and attests to a highly dynamic history of population movements which cannot be reconstruct based on plastid data. Although hybrids are known to occur between species of the *B. lunaria* group (Stensvold, 2008), it is evident that species identity is not primarily maintained by limited gene flow between geographically remote populations, but rather by pre- or postzygotic mating barriers which further supports the treatment of clades as species (Wagner et al., 1985). However, whether original species divergence took place in sympatry or parapatry, or whether it was kick-started by allopatry, remains open to speculation.

Climatic and edaphic differentiation between species

To our knowledge, no study has directly investigated the importance of ecological factors as drivers of the diversification in the *B. lunaria* group. However, Maccagni et al. (2017) observed possible habitat differences between two species (*B. lunaria* and our species 3) suggesting a role of ecological segregation. Also, the distribution of arbuscular mycorrhizal fungi (AMF) on which the growth of *Botrychium* gametophytes relies, is driven by abiotic factors (Davison et al., 2021; Sandoz et al., 2020) which implies potential indirect edaphic preferences.

We found that the climatic niches of both species and subspecies within the *B. lunaria* group showed a certain degree of differentiation, but with marked overlap between many species. The main differentiation occurred along gradients of elevation (as reflected by various temperature-related variables) and climatic continentality. The two main groups of species recovered in the PCA on climatic factors did not correspond to the main group of clades recovered in the phylogenetic analyses, showing that low- and high-elevation species occur in several clades and that their differentiation was not driven by elevational specialization. Unfortunately, the high uncertainty regarding species-level relationships within the main groups precludes any conclusions about evolutionary tendencies related to species diversification within these clades. In any case, the conclusion from these analyses would be that while there is some climatic niche differentiation between species, this is relatively weak and unlikely to suggest that climatic divergence would be a major driver of diversification in the *B. lunaria* group.

Interestingly, however, in the analyses of climatic niches within variable species, we found that subspecies were mostly better segregated in climatic niche space than the species. This likely reflects the fact that analyses within single species encompass a narrower range of climatic conditions as well as fewer taxa, allowing for clearer segregation between taxa. Yet, considering that differentiation at subspecies level might be a precursor to species-level differentiation, and contrary to our conclusions above, this would suggest that climatic niche differentiation might indeed play an important role in the diversification of the *B. lunaria* group.

Our soil pH data from a subsample of sites also reveal patterns of edaphic preferences between species, with some species recorded only on neutral soils (*B. tunux*, species 3, species 7) and others on acidic soils (*B. lunaria*, species 6, species 8), although *B. onondagense* and *B.* aff. *neolunaria* group occurred across the entire pH range. This suggests that there might be some degree of edaphic differentiation between species. In North America, *Botrychium* species have been found associated with calcareous bedrock, basalt bedrock and coastal soils influenced by ocean spray (D. Farrar pers. obs.) with a higher species abundance and richness on neutral or basic soils (Farrar et al., 2017). The soil pH data from North America were extrapolated from bedrock substrates, but our data showed partial dissociation between bedrocks and soil pH. Indeed, two third of the populations with acidic soil pH were found on limestone, dolomite, marble, sandstone, or conglomerate bedrocks, suggesting small scale soil heterogeneity (Hartmann and Moosdorf, 2012).

Soil pH and temperature have been shown to explain the realized niche of arbuscular mycorrhizal fungi (AMF) (Davison et al., 2021), with clear differences among *Glomus* species, which are known to be the main group of AMF associated with *Botrychium* species (Sandoz et al., 2020; Winther and Friedman, 2007). In the Alps, AMF communities exhibit distinct composition in acidic and neutral soils (Sandoz et al., 2020). As the growth of *Botrychium* is dependent on their AMF symbiosis, it might well be that pH preferences between species reflect association with distinct Glomeraceae species and that the co-evolution between AMF and *Botrychium* played a role in the diversification of the *B. lunaria* group.

In conclusion, we find that there is some ecological differentiation between species of the *B. lunaria* group, but evidence of this is rather weak and at present we are unable to identify it as a prime driver of speciation in the group. Even though limited, the soil data available to us evokes potential small-scale ecological differentiation that may not be recovered by our climatic data (at 1 km^2 resolution). We suggest that in-depth studies of ecological conditions at sites where several species co-occur might help to further understand the role of ecological specialization in this group.

Conclusions

Based on a global sample of over 500 specimens, our study reshapes the understanding of the species-level diversity within the *B. lunaria* group and the geographical distribution of its diversity. We find that the timing of diversification corresponds to that of Pleistocene climatic oscillations, suggesting that diversification was mainly driven by repeated cycles of dispersal and extinction. This may then have been combined with some cases of polyploidization (*B. yaaxudakeit*, perhaps a further one in Georgia) and hybridization (in the *B. neolunaria* group), as well as slight climatic and edaphic differentiation between many species and subspecies. Although most of species-level taxa we recognize are 2-3 million years old, the fact that we also found several much younger, genetically, and climatically distinct sub-clades, which we treat at subspecies level, suggests that diversification in the group is ongoing. Future studies should focus on clarifying the clade relationships within the *B. lunaria* group, on understanding the evolution and taxonomic implications of variation within the *B. neolunaria* group, and on documenting the role of small-scale ecological differentiation between species.

Acknowledgements

For their help organising and carrying out field work, we are in debt to Alexandru Coltoiu and the Bucegi Natural Park, Calimani National Park administration (Romania), Claudia Danau and the Retezat National Park (Romania), Clément Duckert, Cun-Feng Zhao, Daniel Ston, Kenan Ćatović, Kuban Uulu Zholdoshbek and the administration of *Sarychat-Ertash State Nature Reserve* (Kyrgyzstan), Luca Lässer, Marine Ramirez, Quentin Dubois, Sabin Neatu and the Ceahlau National Park (Romania), Sandra Grünig, Sergei Aleksandrov and the Central Balkan National Park (Bulgaria), Shamil Shetekauri, Siqi Liang, Tolkha Shetekauri and the team of Tbilisi Botanical Garden (Georgia), Xianchun Zhang, and Zikriyo. We are also thankful to Ahmed Ouhammou (MARK), Alessio Maccagni, Alexey P. Seregin (MW), Andreas Beiser, Dauphin Benjamin, Iowa State University (ISC), Filippo Prosser, Hedwig Meindl, Isabelle Chanaron (AIX), Markus Grabher, Natalia Gamova, Robin Walls, Sigitas Juzėnas, Siri Birkeland, Teddy Dolstra, Vincenzo

Ferri, Wittmann Helmunt, Xian-Yun Mu, and Xianchun Zhang (PE) for providing specimens. Many thanks to Amandine Pillonel, Ophélie Gning, Pierre-Emmanuel DuPasquier, Luyinda Lukau, Bondo Mateus, and Benite Abayo who helped with the molecular laboratory work. Amandine Pillonel prepared the soil samples and made the ex-situ pH measurements. Elke Kessler conducted the flow cytometry analyses. We thank Pierre-Emmanuel DuPasquier, Libing Zhang, and Liang Zhang for the advanced access to Botrychioideae sequencing data. We also acknowledge Sabina Moser Tralamazza and Ursula Oggenfuss for their thoughtful revisions of the manuscript.

Funding: This work was supported by grants from the 'Fond des donations' and the 'Subvention égalité' of the University of Neuchâtel.

Literature chapter 2

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Aberer, A.J., Krompass, D., Stamatakis, A., 2013. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. Syst. Biol. 62, 162–166. https://doi.org/10.1093/sysbio/sys078

Allen, G.A., Marr, K.L., McCormick, L.J., Hebda, R.J., 2012. The impact of Pleistocene climate change on an ancient arctic–alpine plant: multiple lineages of disparate history in *Oxyria digyna*. Ecol. Evol. 2, 649–665. https://doi.org/10.1002/ece3.213

Alvarez, N., Thiel-Egenter, C., Tribsch, A., Holderegger, R., Manel, S., Schönswetter, P., Taberlet, P., Brodbeck, S., Gaudeul, M., Gielly, L., Küpfer, P., Mansion, G., Negrini, R., Paun, O., Pellecchia, M., Rioux, D., Schüpfer, F., Loo, M.V., Winkler, M., Gugerli, F., 2009. History or ecology? Substrate type as a major driver of patial genetic structure in Alpine plants. Ecol. Lett. 12, 632–640. https://doi.org/10.1111/j.1461-0248.2009.01312.x

Barrington, D.S., 1993. Ecological and historical factors in fern biogeography. J. Biogeogr. 20, 275–279. https://doi.org/10.2307/2845635

Bell, M.A., Graeme, T.L., 2014. strap: Stratigraphic tree analysis for palaeontology.

Berry, C.M., Hilton, J., 2006. Givetian (Middle Devonian) cladoxylopsid 'ferns' from Orkney, northern Scotland. Earth Environ. Sci. Trans. R. Soc. Edinb. 97, 65–73. https://doi.org/10.1017/S0263593300001401

Bozukov, V., Tsenov, B., Vatsev, M., 2010. A first find of *Botrychium* (Ophioglossaceae) in Bulgarian palaeomacroflora. Comptes Rendus Académie Bulg. Sci. Sci. Mathématiques Nat. 63, 889–892.

Butters, F.K., Abbe, E.C., 1953. A floristic study of Cook county, Northeastern Minnesota. Rhodora 55, 94.

Campbell, D.H., 1911. The eusporangiate: the comparative morphology of the Ophioglossaceae and the Marattiaceae., Carnegie Institution of Washington. ed. Washington, D.C., USA.

Campitelli, E., 2021. ggnewscale: Multiple Fill and Colour Scales in "ggplot2."

Chang, Y., Li, J., Lu, S., Schneider, H., 2013. Species diversity and reticulate evolution in the *Asplenium normale* complex (Aspleniaceae) in China and adjacent areas. TAXON 62, 673–687. https://doi.org/10.12705/624.6

Charif, D., Lobry, J.R., 2007. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis., in: Structural approaches to sequence evolution: molecules, networks, populations, biological and medical physics, biomedical engineering. Springer Verlag, pp. 207–232.

Clarke, J.M., House, H.D., 1923. Report of the state botanist for 1921. N. Y. State Mus. Bull.

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

Clute, W.N., 1905. A check list of the North American fernworts. Fern Bull. 13, 109–123.

Cornish-Bowden, A., 1985. Nomenclature for incompletely specified bases in nucleic-acid sequences. Recommendations 1984. Nucleic Acids Res., the IUBMB Nomenclature Committee 13, 3021–3030.

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, 620–639. https://doi.org/10.1600/036364417X696483

Dauphin, B., Grant, J., Mráz, P., 2016. Ploidy level and genome size variation in the homosporous ferns *Botrychium* s.l. (Ophioglossaceae). Plant Syst. Evol. 302, 575–584. https://doi.org/10.1007/s00606-016-1285-7

Dauphin, B., Grant, J.R., Farrar, D.R., Rothfels, C.J., 2018. Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. Mol. Phylogenet. Evol. 120, 342–353. https://doi.org/10.1016/j.ympev.2017.11.025

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. https://doi.org/10.3732/ajb.1300154

Davison, J., Moora, M., Semchenko, M., Adenan, S.B., Ahmed, T., Akhmetzhanova, A.A., Alatalo, J.M., Al-Quraishy, S., Andriyanova, E., Anslan, S., Bahram, M., Batbaatar, A., Brown, C., Bueno, C.G., Cahill, J., Cantero, J.J., Casper, B.B., Cherosov, M., Chideh, S., Coelho, A.P., Coghill, M., Decocq, G., Dudov, S., Fabiano, E.C., Fedosov, V.E., Fraser, L., Glassman, S.I., Helm, A., Henry, H.A.L., Hérault, B., Hiiesalu, Indrek, Hiiesalu, Inga, Hozzein, W.N., Kohout, P., Kõljalg, U., Koorem, K., Laanisto, L., Mander, Ü., Mucina, L., Munyampundu, J.-P., Neuenkamp, L., Niinemets, Ü., Nyamukondiwa, C., Oja, J., Onipchenko, V., Pärtel, M., Phosri, C., Põlme, S., Püssa, K., Ronk, A., Saitta, A., Semboli, O., Sepp, S.-K., Seregin, A., Sudheer, S., Peña-Venegas, C.P., Paz, C., Vahter, T., Vasar, M., Veraart, A.J., Tedersoo, L., Zobel, M., Öpik, M., 2021. Temperature and pH define the realised niche space of arbuscular mycorrhizal fungi. New Phytol. 231, 763–776. https://doi.org/10.1111/nph.17240

De Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879–886. https://doi.org/10.1080/10635150701701083

Des Marais, D.L., Smith, A.R., Britton, D.M., Pryer, K.M., 2003. Phylogenetic relationships and evolution of extant horsetails, *Equisetum*, based on chloroplast DNA sequence data (rbcL and trnL-F). Int. J. Plant Sci. 164, 737–751. https://doi.org/10.1086/376817

Dolezel, J., Bartos, J., 2005. Plant DNA flow cytometry and estimation of nuclear genome size. Ann. Bot. 95, 99–110. https://doi.org/10.1093/aob/mci005

Doležel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysák, M.A., Nardi, L., Obermayer, R., 1998. Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann. Bot. 82, 17–26. https://doi.org/10.1006/anbo.1998.0730

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973. https://doi.org/10.1093/molbev/mss075

Dunnington, D., 2021. ggspatial: spatial data framework for ggplot2.

Ehlers, J., Gibbard, P.L., Hughes, P.D., 2018. Chapter 4 - Quaternary glaciations and chronology, in: Menzies, J., van der Meer, J.J.M. (Eds.), Past Glacial Environments (Second Edition). Elsevier, pp. 77–101. https://doi.org/10.1016/B978-0-08-100524-8.00003-8

Ehlers, J., Gibbard, P.L., Hughes, P.D., 2011. Quaternary glaciations - extent and chronology, first. ed. Elsevier.

Farrar, D.R., 2011. Systematics and taxonomy of genus *Botrychium*.

Farrar, D.R., Gilman, A.V., Moran, R.C., 2017. Ophioglossales. Online edition. R.F.C. Naczi, J. R. Abbott, and Collaborators, New manual of vascular plants of northeastern United States and adjacent Canada, NYBG Press. ed. New York.

Farrar, D.R., Johnson-Groh, C., 1991. A new prairie moonwort (*Botrychium* subgenus *Botrychium*) from northwestern Minnesota. Am. Fern J. 81, 1–6.

Frandsen, P.B., Calcott, B., Mayer, C., Lanfear, R., 2015. Automatic selection of partitioning schemes for phylogenetic analyses using iterative k-means clustering of site rates. BMC Evol. Biol. 15, 13. https://doi.org/10.1186/s12862-015-0283-7

Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27, 480–488. https://doi.org/10.1016/j.tree.2012.04.012

Garnier, S., 2018. viridis: default color maps from "matplotlib."

Gastony, G.J., Yatskievych, G., 1992. Maternal inheritance of the chloroplast and mitochondrial genomes in Cheilanthoid ferns. Am. J. Bot. 79, 716–722. https://doi.org/10.1002/j.1537-2197.1992.tb14613.x

Gearty, W., 2021. deeptime: plotting tools for anyone working in deep time.

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

Guillon, J.-M., Raquin, C., 2000. Maternal inheritance of chloroplasts in the horsetail *Equisetum variegatum* (Schleich.). Curr. Genet. 37, 53–56. https://doi.org/10.1007/s002940050008

Hanušová, K., Čertner, M., Urfus, T., Koutecký, P., Košnar, J., Rothfels, C.J., Jarolímová, V., Ptáček, J., Ekrt, L., 2019. Widespread co-occurrence of multiple ploidy levels in fragile ferns (*Cystopteris fragilis* complex; Cystopteridaceae) probably stems from similar ecology of cytotypes, their efficient dispersal and inter-ploidy hybridization. Ann. Bot. 123, 845–855. https://doi.org/10.1093/aob/mcy219

Hartmann, J., Moosdorf, N., 2012. Global lithological map database v1.0 (gridded to 0.5° spatial resolution). Suppl. Hartmann Jens Moosdorf Nils 2012 New Glob. Lithol. Map Database GLiM Represent. Rock Prop. Earth Surf. Geochem. Geophys. Geosystems 13 Q12004 Httpsdoiorg1010292012GC004370. https://doi.org/10.1594/PANGAEA.788537

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

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. https://doi.org/10.1600/036364412X635377

Hauk, W.D., Parks, C.R., Chase, M.W., 2003. Phylogenetic studies of Ophioglossaceae: evidence from rbcL and trnL-F plastid DNA sequences and morphology. Mol. Phylogenet. Evol. 28, 131–151.

Hungerer, K.B., Kadereit, J.W., 1998. The phylogeny and biogeography of *Gentiana* L. sect. Ciminalis (Adans.) Dumort.: A historical interpretation of distribution ranges in the European high mountains. Perspect. Plant Ecol. Evol. Syst. 1, 121–135. https://doi.org/10.1078/1433-8319-00055

Ivy-Ochs, S., 2015. Glacier variations in the European Alps at the end of the last glaciation. Cuad. Investig. Geográfica 41, 295–315. https://doi.org/10.18172/cig.2750

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P., Kessler, M., 2018. Data from: Climatologies at high resolution for the earth's land surface areas. https://doi.org/10.5061/DRYAD.KD1D4

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P., Kessler, M., 2017. Climatologies at high resolution for the earth's land surface areas. Sci. Data 4, 170122. https://doi.org/10.1038/sdata.2017.122

Kassambara, A., 2021. rstatix: pipe-friendly framework for basic statistical tests.

Kassambara, A., 2020. ggpubr: "ggplot2" based publication ready plots.

Kassambara, A., Mundt, F., 2020. factoextra: extract and visualize the results of multivariate data analyses.

Kato, M., 1987. A phylogenetic classification of Ophioglossaceae. Gard. Bull. Singap. 40, 1–14.

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010

Khang, T. fei, 2016. Concatenation R. figshare.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., Stamatakis, A., 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 1–3. https://doi.org/10.1093/bioinformatics/btz305

Kumar, S., Stecher, G., Suleski, M., Hedges, S.B., 2017. TimeTree: A resource for timelines, timetrees, and tivergence times. Mol. Biol. Evol. 34, 1812–1819. https://doi.org/10.1093/molbev/msx116

Kuo, L.-Y., Tang, T.-Y., Li, F.-W., Su, H.-J., Chiou, W.-L., Huang, Y.-M., Wang, C.-N., 2018. Organelle genome inheritance in *Deparia* ferns (Athyriaceae, Aspleniineae, Polypodiales). Front. Plant Sci. 0. https://doi.org/10.3389/fpls.2018.00486

Lafontaine, G. de, Napier, J.D., Petit, R.J., Hu, F.S., 2018. Invoking adaptation to decipher the genetic legacy of past climate change. Ecology 99, 1530–1546. https://doi.org/10.1002/ecy.2382

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Softw. 25, 1–18. https://doi.org/10.18637/jss.v025.i01

Lehtonen, S., Silvestro, D., Karger, D.N., Scotese, C., Tuomisto, H., Kessler, M., Peña, C., Wahlberg, N., Antonelli, A., 2017. Environmentally driven extinction and opportunistic origination explain fern diversification patterns. Sci. Rep. 7, 4831. https://doi.org/10.1038/s41598-017-05263-7

Lemoine, F., Entfellner, J.-B.D., Wilkinson, E., Correia, D., Felipe, M.D., Oliveira, T.D., Gascuel, O., 2018. Renewing Felsenstein's phylogenetic bootstrap in the era of big data. Nature 556, 452. https://doi.org/10.1038/s41586-018-0043-0

Lovis, J.D., 1958. An evolutionary study of the fern *Asplenium trichomanes* (PhD thesis). University of Leeds.

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. Alp. Bot. 127, 185–196.

Magallón, S., Hilu, K.W., Quandt, D., 2013. Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. Am. J. Bot. 100, 556–573. https://doi.org/10.3732/ajb.1200416

Mickel, J.T., Smith, A.R., 2004. The Pteridophytes of Mexico. Mem. N. Y. Bot. Gard. 88, 1–1054.

Milde, J., 1869. *Botrychium* Monographia. Zool.-Bot. Ges. Wien Verhandlungen 19, 55–190.

Muller, S., 1986. *Botrychium matricariifolium* (Retz.) A. Braun ex Koch dans les pelouses sableuses du Pays de Bitche (Vosges du Nord). Bull. Société Bot. Fr. Lett. Bot. 133, 189–197. https://doi.org/10.1080/01811797.1986.10824699

Nekola, J.C., Schlicht, D.W., 1996. Distribution of *Botrychium campestre* in northeastern Iowa. Am. Fern J. 86, 119–123. https://doi.org/10.2307/1547151

Neuwirth, E., 2014. RColorBrewer: ColorBrewer Palettes.

Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. Front. Zool. 7, 16. https://doi.org/10.1186/1742-9994-7-16

Pagel, M., Meade, A., Barker, D., 2004. Bayesian estimation of ancestral character states on phylogenies. Syst. Biol. 53, 673–684. https://doi.org/10.1080/10635150490522232

Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583. https://doi.org/10.1093/oxfordjournals.molbev.a040517

Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Paris, C.A., Wagner, F.S., Wagner, W.H., Jr., 1989. Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. Am. Fern J. 79, 46–54. https://doi.org/10.2307/1547159

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E., Stamatakis, A., 2010. How many Bootstrap replicates are necessary? J. Comput. Biol. 17, 337–354. https://doi.org/10.1089/cmb.2009.0179

Pebesma, E., 2018. Simple features for R: standardized support for spatial vector data. R J. 10, 439–446.

Popovich, S.J., Farrar, D.R., Gilman, A.V., 2020. *Botrychium furculatum* (Ophioglossaceae), a new moonwort species from the Rocky Mountains of North America. Am. Fern J. 110, 165–182. https://doi.org/10.1640/0002-8444-110.4.165

PPG I, 2016. A community-derived classification for extant lycophytes and ferns. J. Syst. Evol. 54, 563–603. https://doi.org/10.1111/jse.12229

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

Qi, X., Kuo, L.-Y., Guo, C., Li, H., Li, Z., Qi, J., Wang, L., Hu, Y., Xiang, J., Zhang, C., Guo, J., Huang, C.-H., Ma, H., 2018. A well-resolved fern nuclear phylogeny reveals the evolution history of numerous transcription factor families. Mol. Phylogenet. Evol. 127, 961–977. https://doi.org/10.1016/j.ympev.2018.06.043

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A., 2009. FigTree.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032

Rasband, W.S., 1997. ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA.

Ravi, V., Khurana, J.P., Tyagi, A.K., Khurana, P., 2008. An update on chloroplast genomes. Plant Syst. Evol. 271, 101–122. https://doi.org/10.1007/s00606-007-0608-0

Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

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. https://doi.org/10.12705/635.32

Rothfels, C.J., Larsson, A., Kuo, L.-Y., Korall, P., Chiou, W.-L., Pryer, K.M., 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of Eupolypod II ferns. Syst. Biol. 61, 490–490. https://doi.org/10.1093/sysbio/sys001

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., Pokorny, L., Chen, T., dePamphilis, C., DeGironimo, L., Chen, L., Wei, X., Sun, X., Korall, P., Stevenson, D.W., Graham, S.W., Wong, G.K.-S., Pryer, K.M., 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. Am. J. Bot. 102, 1089–1107. https://doi.org/10.3732/ajb.1500089

Rothwell, G.W., Stockey, R.A., 1989. Fossil Ophioglossales in the Paleocene of western North America. Am. J. Bot. 76, 637–644. https://doi.org/10.2307/2444411

Salinas, N.R., Little, D.P., 2014. 2matrix: A utility for indel coding and phylogenetic matrix concatenation1. Appl. Plant Sci. 2. https://doi.org/10.3732/apps.1300083

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109. https://doi.org/10.1093/oxfordjournals.molbev.a003974

Sandoz, F.A., Bindschedler, S., Dauphin, B., Farinelli, L., Grant, J.R., Hervé, V., 2020. Biotic and abiotic factors shape arbuscular mycorrhizal fungal communities associated with the roots of the widespread fern *Botrychium lunaria* (Ophioglossaceae). Environ. Microbiol. Rep. 12, 342–354. https://doi.org/10.1111/1758-2229.12840

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

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. https://doi.org/10.1111/j.1365-294X.2005.02683.x

Schuettpelz, E., Pryer, K.M., 2009. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. Proc. Natl. Acad. Sci. 106, 11200–11205. https://doi.org/10.1073/pnas.0811136106

Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.

Smith, A.R., 1972. Comparison of fern and flowering plant distributions with some evolutionary interpretations for ferns. Biotropica 4, 4–9. https://doi.org/10.2307/2989639

Solís-Lemus, C., Knowles, L.L., Ané, C., 2015. Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution 69, 492–507. https://doi.org/10.1111/evo.12582

South, A., 2017. rnaturalearth: world map data from natural earth.

Ståhl, P., Ekman, J., Westerberg, S., Grant, J.R., Dauphin, B., 2016. Mer om pysslinglåsbräken i Sverige. Sven. Bot. Tidskr. 110, 68–74.

Stensvold, M.C., 2008. A taxonomic and phylogeographic study of the *Botrychium lunaria* complex (Doctor of Philosophy). Iowa State University.

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. https://doi.org/10.1007/s12228-016-9455-z

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. https://doi.org/10.1640/0002-8444(2002)092[0150:TNSOMB]2.0.CO;2

Sukumaran, J., Holder, M.T., 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26, 1569–1571. https://doi.org/10.1093/bioinformatics/btq228

Sukumaran, J., Knowles, L.L., 2017. Multispecies coalescent delimits structure, not species. Proc. Natl. Acad. Sci. 114, 1607–1612. https://doi.org/10.1073/pnas.1607921114

Swartz, O., 1801. Filicum, ordine systematico redactarum. J. Für Bot. 1800, 4–120.

Taylor, T.N., Taylor, E.L., Krings, M., 2009. 11 - Ferns and early fernlike plants, in: Taylor, T.N., Taylor, E.L., Krings, M. (Eds.), Paleobotany (Second Edition). Academic Press, London, pp. 383–478. https://doi.org/10.1016/B978-0-12-373972-8.00011-5

Testo, W., Sundue, M., 2016. A 4000-species dataset provides new insight into the evolution of ferns. Mol. Phylogenet. Evol. 105, 200–211. https://doi.org/10.1016/j.ympev.2016.09.003

Tribsch, A., Schönswetter, P., 2003. Patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. TAXON 52, 477–497. https://doi.org/10.2307/3647447

Underwood, L.M., 1903. An index to the described species of *Botrychium*. Bull. Torrey Bot. Club 30, 47–48.

Vasques, D.T., Ebihara, A., Hirai, R.Y., Prado, J., Motomi, I., 2019. Phylogeny of *Hymenophyllum* subg. *Mecodium* (Hymenophyllaceae), with special focus on the diversity of the *Hymenophyllum polyanthos* species complex. Plant Syst. Evol. 305, 811–825. https://doi.org/10.1007/s00606-019-01609-y

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. https://doi.org/10.1093/aob/mcr267

Vogel, J.C., Russell, S.J., Rumsey, F.J., Barrett, J.A., Gibby, M., 1998. Evidence for maternal transmission of chloroplast DNA in the genus *Asplenium* (Aspleniaceae, Pteridophyta). Bot. Acta 111, 247–249. https://doi.org/10.1111/j.1438-8677.1998.tb00704.x

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

Wagner, W.H., Grant, J.R., 2002. *Botrychium alaskense*, a new moonwort from the interior of Alaska. Am. Fern J. 92, 164–170.

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

Wagner, W.H., Lord, L.P., 1956. The morphological and cytological distinctness of *Botrychium minganense* and *B. lunaria* in Michigan. Bull. Torrey Bot. Club 83, 261–280. https://doi.org/10.2307/2482592

Wagner, W.H., Wagner, F.S., 1994. Another widely disjunct, rare and local North American moonwort (Ophioglossaceae: *Botrychium* subg. *Botrychium*). Am. Fern J. 84, 5–10. https://doi.org/10.2307/1547552

Wagner, W.H., Wagner, F.S., 1990a. Moonworts (*Botrychium* subg. *Botrychium*) of the upper great lakes region, U.S.A. and Canada, with descriptions of two new species. Contrib. Univ. Mich. Herb. 17, 312–325.

Wagner, W.H., Wagner, F.S., 1990b. Notes on the fan-leaflet group of moonworts in North America with descriptions of two new members. Am. Fern J. 80, 73–81. https://doi.org/10.2307/1547172

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. https://doi.org/10.2307/1547557

Wagner, W.H., Wagner, F.S., 1983a. Two moonworts of the Rocky Mountains; *Botrychium hesperium* and a new species formerly confused with It. Am. Fern J. 73, 53–62. https://doi.org/10.2307/1546534

Wagner, W.H., Wagner, F.S., 1983b. Genus communities as a systematic tool in the study of New World *Botrychium* (Ophioglossaceae). Taxon 32, 51–63.

Wagner, W.H., Wagner, F.S., 1982. *Botrychium rugulosum* (Ophioglossaceae), a newly recognized species of evergreen grapefern in the Great Lakes area of North America [Pteridophyta; North Central States (USA)]. Contrib. Univ. Mich. Herb. 15, 315.

Wagner, W.H., Wagner, F.S., Beitel, J.M., 1985. Evidence for interspecific hybridisation in pteridophytes with subterranean mycoparasitic gametophytes. Proc. R. Soc. Edinb. Sect. B Biol. Sci. 86, 273–281. https://doi.org/10.1017/S026972700000823X

Wang, L.-G., Lam, T.T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C.W., Jones, B.R., Bradley, T., Zhu, H., Guan, Y., Jiang, Y., Yu, G., 2020. Treeio: an R package for phylogenetic Tree input and output with richly annotated and associated data. Mol. Biol. Evol. 37, 599–603. https://doi.org/10.1093/molbev/msz240

Wendel, J.F., Doyle, J.J., 1998. Phylogenetic incongruence: window into genome history and molecular evolution, in: Soltis, D.E., Soltis, P.S., Doyle, J.J. (Eds.), Molecular Systematics of Plants II: DNA Sequencing. Springer US, Boston, MA, pp. 265–296. https://doi.org/10.1007/978-1-4615-5419-6_10

Westergaard, K.B., Alsos, I.G., Popp, M., Engelskjøn, T., Flatberg, K.I., Brochmann, C., 2011. Glacial survival may matter after all: nunatak signatures in the rare European populations of two west-arctic species. Mol. Ecol. 20, 376–393. https://doi.org/10.1111/j.1365-294X.2010.04928.x

Whittier, D.P., 1972. Gametophytes of *Botrychium dissectum* as grown in sterile culture. Bot. Gaz. 133, 336–339.

Whittier, P., 1981. Spore germination and young gametophyte development of *Botrychium* and *Ophioglossum* in axenic culture. Am. Fern J. 71, 13–19. https://doi.org/10.2307/1546671

Wickham, H., 2016. ggplot2: elegant graphics for data analysis, 2nd ed, Use R! Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. J. Open Source Softw. 4, 1686. https://doi.org/10.21105/joss.01686

Wickham, H., Seidel, D., 2020. scales: scale functions for visualization.

Wikström, N., Kenrick, P., 2001. Evolution of Lycopodiaceae (Lycopsida): estimating divergence times from rbcL gene sequences by use of nonparametric rate smoothing. Mol. Phylogenet. Evol. 19, 177–186. https://doi.org/10.1006/mpev.2001.0936

Wilke, C.O., 2020. cowplot: streamlined plot theme and plot annotations for "ggplot2."

Williams, E.W., 2015. Tracking morphological change and demographic dynamics in ephemeral *Botrychium s.s.* (Ophioglossaceae) populations1. J. Torrey Bot. Soc. 142, 152–165. https://doi.org/10.3159/TORREY-D-14-00036.1

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. https://doi.org/10.3732/ajb.1500281

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. https://doi.org/10.1086/664711

Winther, J.L., Friedman, W.E., 2007. Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). Am. J. Bot. 94, 1248–1255. https://doi.org/10.3732/ajb.94.7.1248

Xu, S., 2021. ggstar: Star Layer for "ggplot2."

Xu, S., Yu, G., 2021. ggtreeExtra: an R package to add geom layers on circular or other layout tree of "ggtree."

Yu, G., Lam, T.T.-Y., Zhu, H., Guan, Y., 2018. Two methods for mapping and visualizing associated data on phylogeny using Ggtree. Mol. Biol. Evol. 35, 3041–3043. https://doi.org/10.1093/molbev/msy194

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8, 28–36. https://doi.org/10.1111/2041-210X.12628

Zhang, L., Fan, X.-P., Petchsri, S., Zhou, L., Pollawatn, R., Zhang, X., Zhou, X.-M., Lu, N.T., Knapp, R., Chantanaorrapint, S., Limpanasittichai, P., Sun, H., Gao, X.-F., Zhang, L.-B., 2020. Evolutionary relationships of the ancient fern lineage the adder's tongues (Ophioglossaceae) with description of *Sahashia* gen. nov. Cladistics 36, 380–393. https://doi.org/10.1111/cla.12408

Zhong, B., Fong, R., Collins, L.J., McLenachan, P.A., Penny, D., 2014. Two new fern chloroplasts and decelerated evolution linked to the long generation time in tree ferns. Genome Biol. Evol. 6, 1166–1173. https://doi.org/10.1093/gbe/evu087

Competing interests: none

Author contributions: VM, JG and MK designed the study. JG, MK and VM acquired the funding. VM organized and performed the fieldwork. DF provided essential type and identified reference plant material. EK, VM, DC and MK performed the analysis. MK, VM and DF interpretated the results. VM and MK wrote the first draft with substantial inputs from DC and DF. All co-authors reviewed the manuscript and approved the final version.

Data accessibility: Sequences will be deposited on Genbank. Alignments, and phylogenetic trees will be available as supplementary Files S1-S10 (https://doi.org/10.5281/zenodo.6531504).

Figures

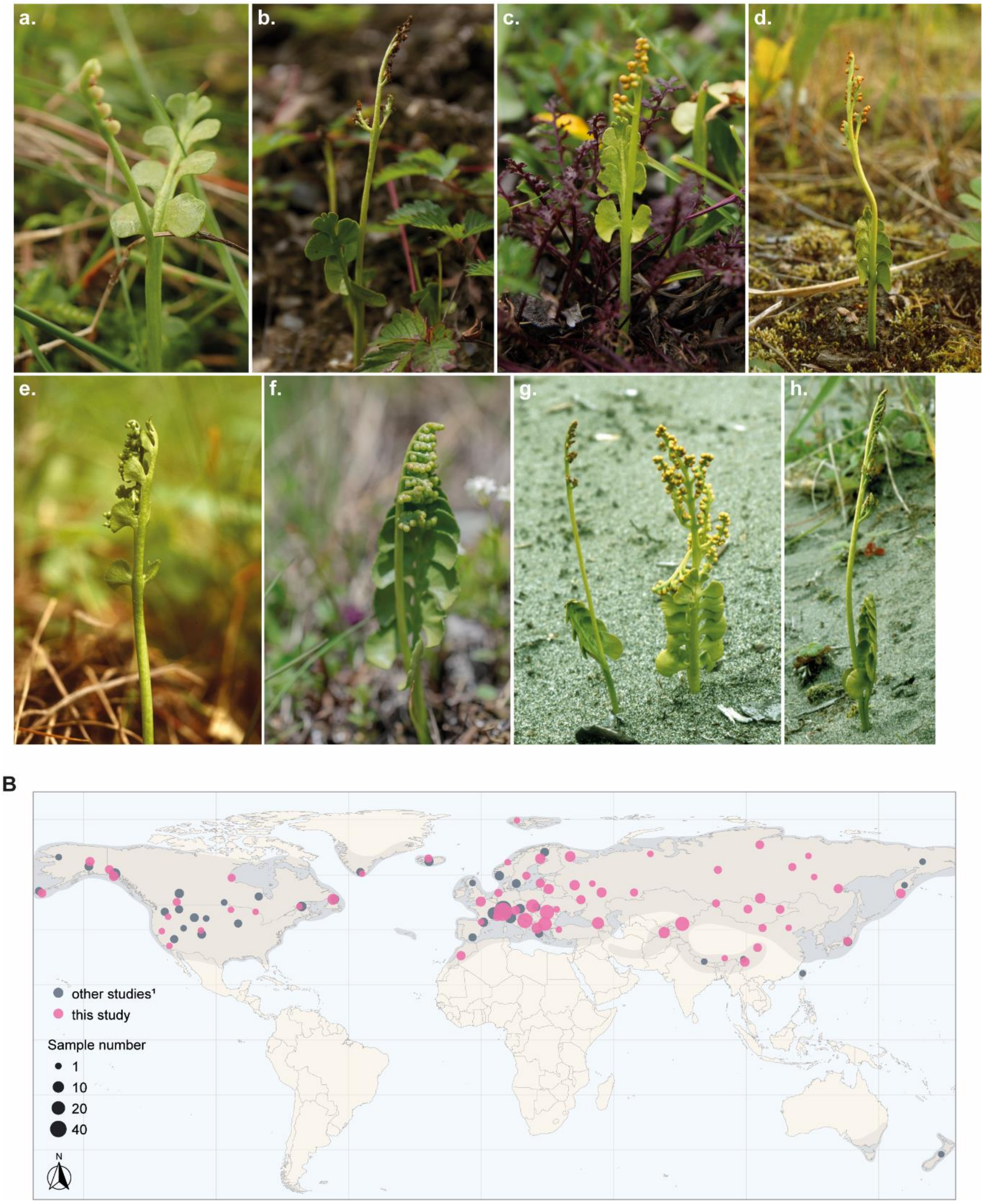


FIGURE 1: **Morphological diversity of the *B. lunaria* group and geographical distribution of the analyzed individuals**. (A) **a**. VM-lun26-5-RO (photo by Q. Dubois, Romania). **b.** VM-lun55-1-CN (photo by V. Mossion, China). **c.** VM-lun54-CN (photo by V. Mossion, China). **d**. *B. neolunaria* (photo by D. Farrar, USA). **e.** *B. crenulatum* (photo by D. Farrar, USA). **f.** *B.* lunaria, VM-lun89-CH (photo by V. Mossion, Switzerland). **g.** *B. neolunaria s.str.* on the left, *B. tunux* on the right (photo by D. Farrar, USA). **h.** *B. yaaxudakeit* (photo by D. Farrar, USA). (B) Sampling map, differentiating between sequences from previous studies (gray) and this study (pink). The distribution range of the *B. lunaria* group (Stensvold, 2018) is showed by the grey layer. Individuals less than 500 km apart were clustered and are indicated by a unique circle. The circle sizes are proportional to individual number. [1]Dauphin *et al.* (2014, 2017), Maccagni *et al.* (2017)

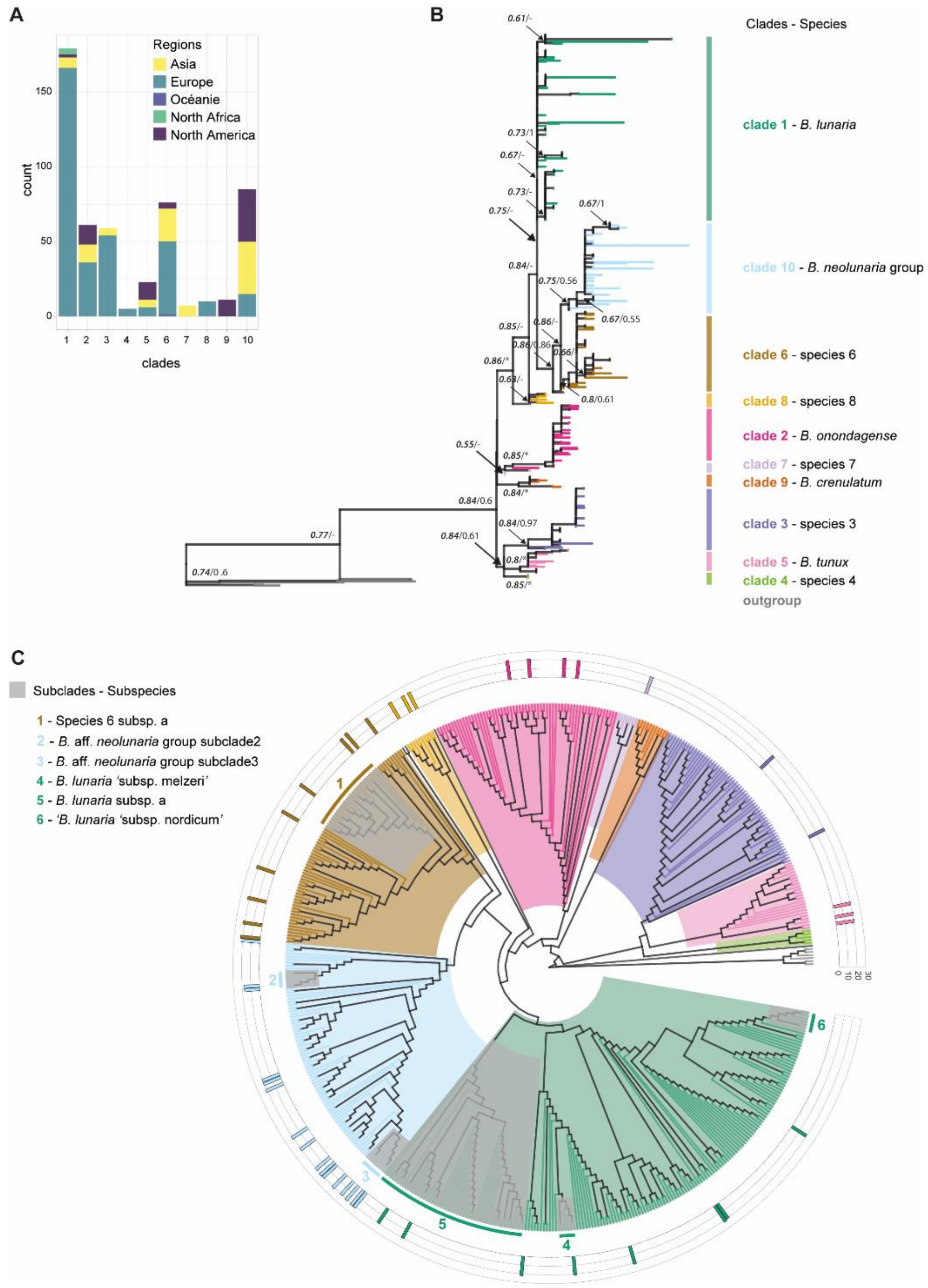


FIGURE 2: **Maximum likelihood (ML) phylogenetic trees depicting the lineages of the *B. lunaria* group with branch support values and genome size distribution**. Trees are based on the complete dataset (*i.e.*, including accessions with partial sequence data). Part of the outgroup was pruned of the trees to reduce the root length. (A) Distribution of specimens grouped per clade by large geographic units. (B) Phylogram of the ML tree. ML bootstrap supports (MLBS), and Bayesian inference posterior probabilities (BIPP) are on the left and right, respectively, along the main clade and subclade branches (*i.e.*, * fully supported, - not applicable; see **supplementary Figure S3, Files S6 and S5** for full BIPP and MLBS values, respectively). Each color represents a distinct clade. (C) Cladogram of the ML tree in fan layout. Branch and background are colored according to clades (tip labels are depicted on **supplementary Figure S2 and S3**). Grey layers on the top of the tree and colored side bars indicate the subclade positions. The layer around the tree shows the absolute genome size distribution given in picograms for the 2C-Values (see supplementary **Table S4** for the values).

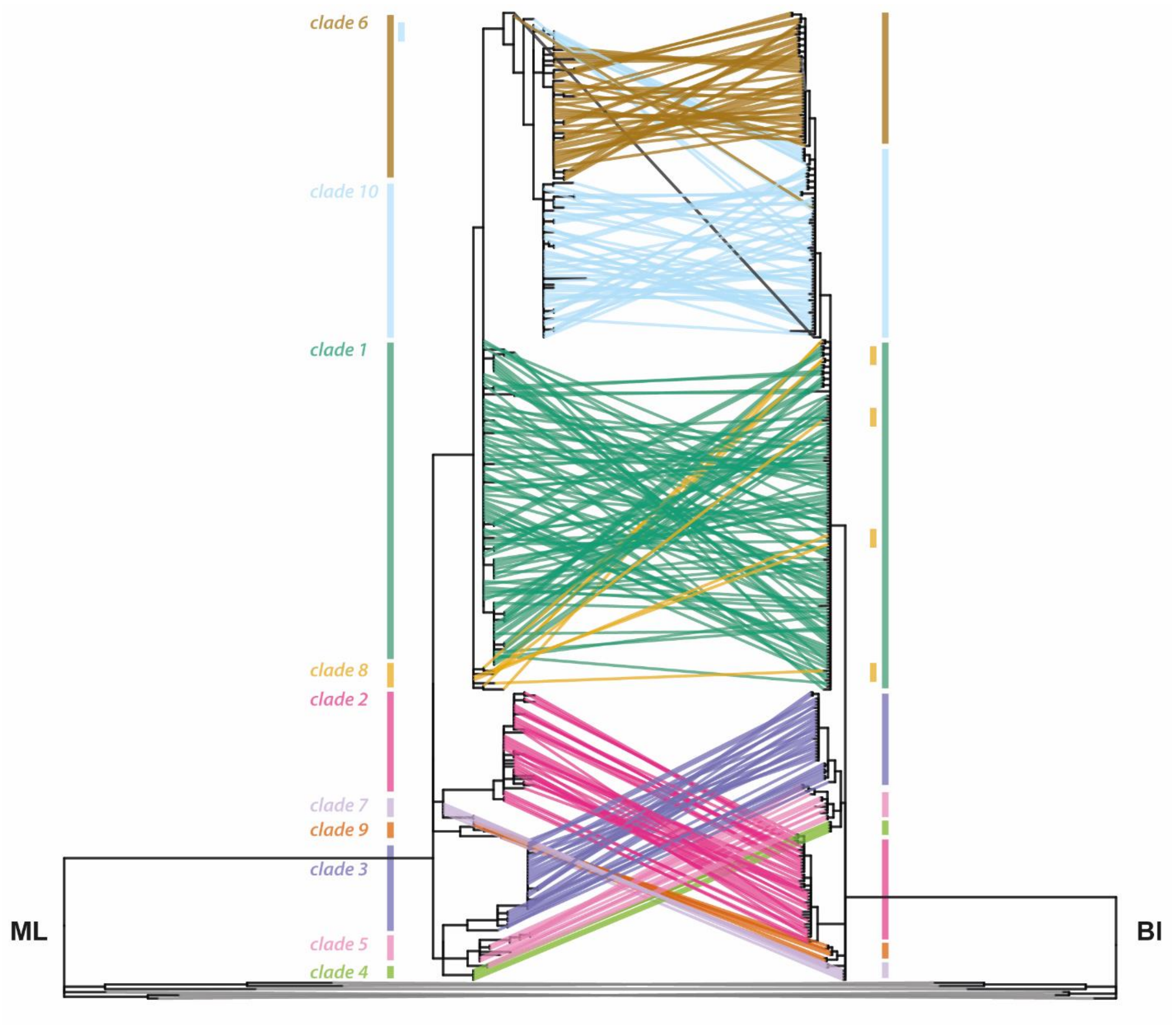


FIGURE 3: **Topological congruence of phylogenetic trees inferred from the reduced dataset using Maximum likelihood (ML) and Bayesian inference (BI) methods**. The tree on the right side corresponds to the BI tree and on the left side to the ML tree. Corresponding tips are linked by lines colored according to clade assignment.

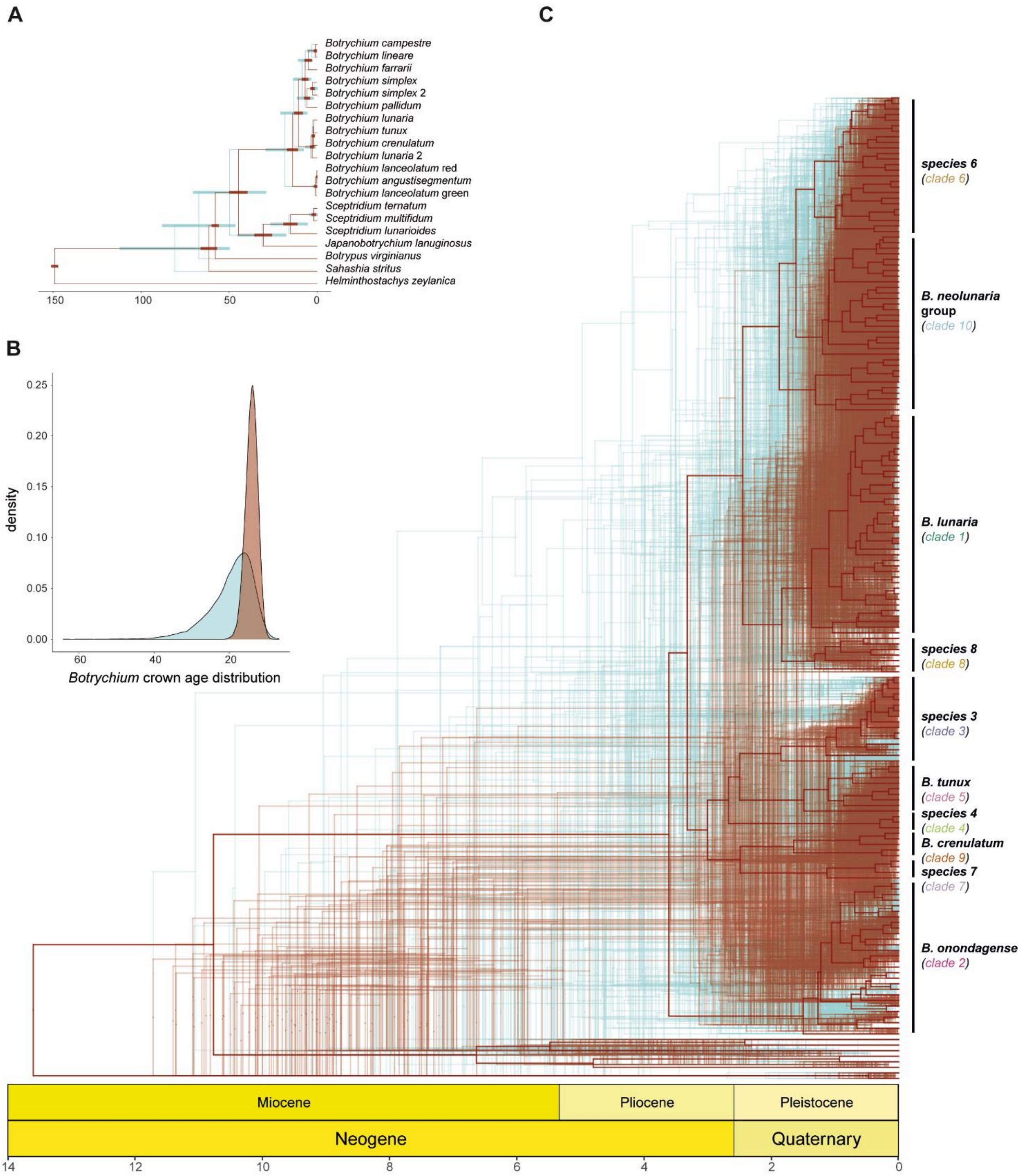


FIGURE 4: **Divergence time estimates of the subfamily Botrychioideae and of the species of *B. lunaria* complex**. Time scales are in million-year unit. The red color refers to the divergence time analysis ran under the strict clock model, and the blue color to the relaxed clock model. (A) Time calibrated phylogenies of the subfamily Botrychioideae. The 0.95 Highest Posterior Density (HPD) is represented by bars at the calibrated nodes. (B) Distribution of the *Botrychium* crown age density estimated by the Botrychioideae time-divergence analysis. (C) Time calibrated phylogeny of the *B. lunaria* complex. The uncertainty around the node ages is displayed by density trees. The density trees were each drawn on subset of 90 posterior trees. The geological time scale shows the periods and the epochs and follows the International Chronostratigraphic Chart (2018; https://stratigraphy.org).

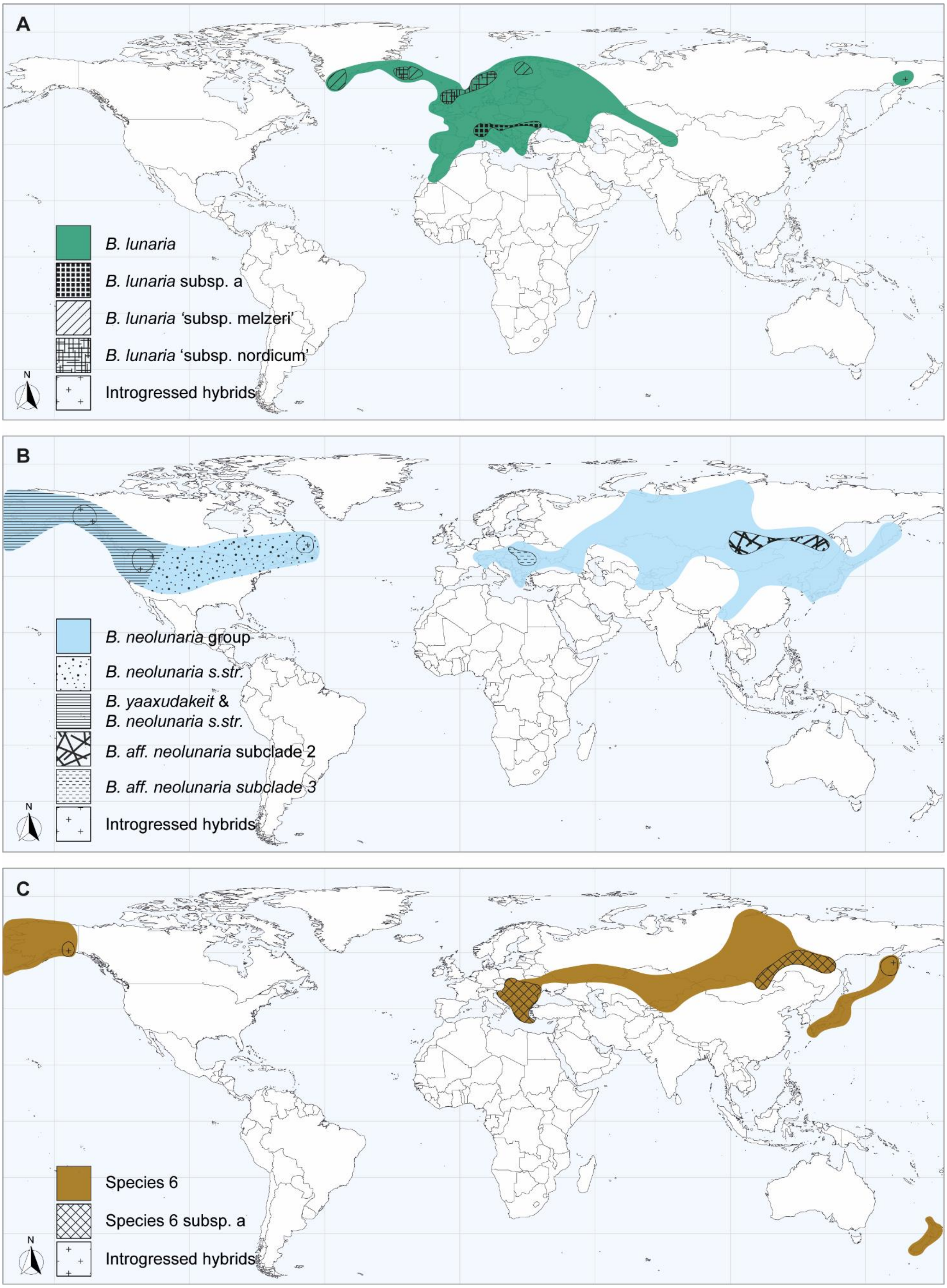


FIGURE 5: **Geographic distributions of the species of the *B. lunaria* group**. (A) Global distribution of *B. lunaria* (clade 1). (B) Global distribution of the *B. neolunaria* group (clade 10). (C) Global distribution of the species 6 (clade 6).

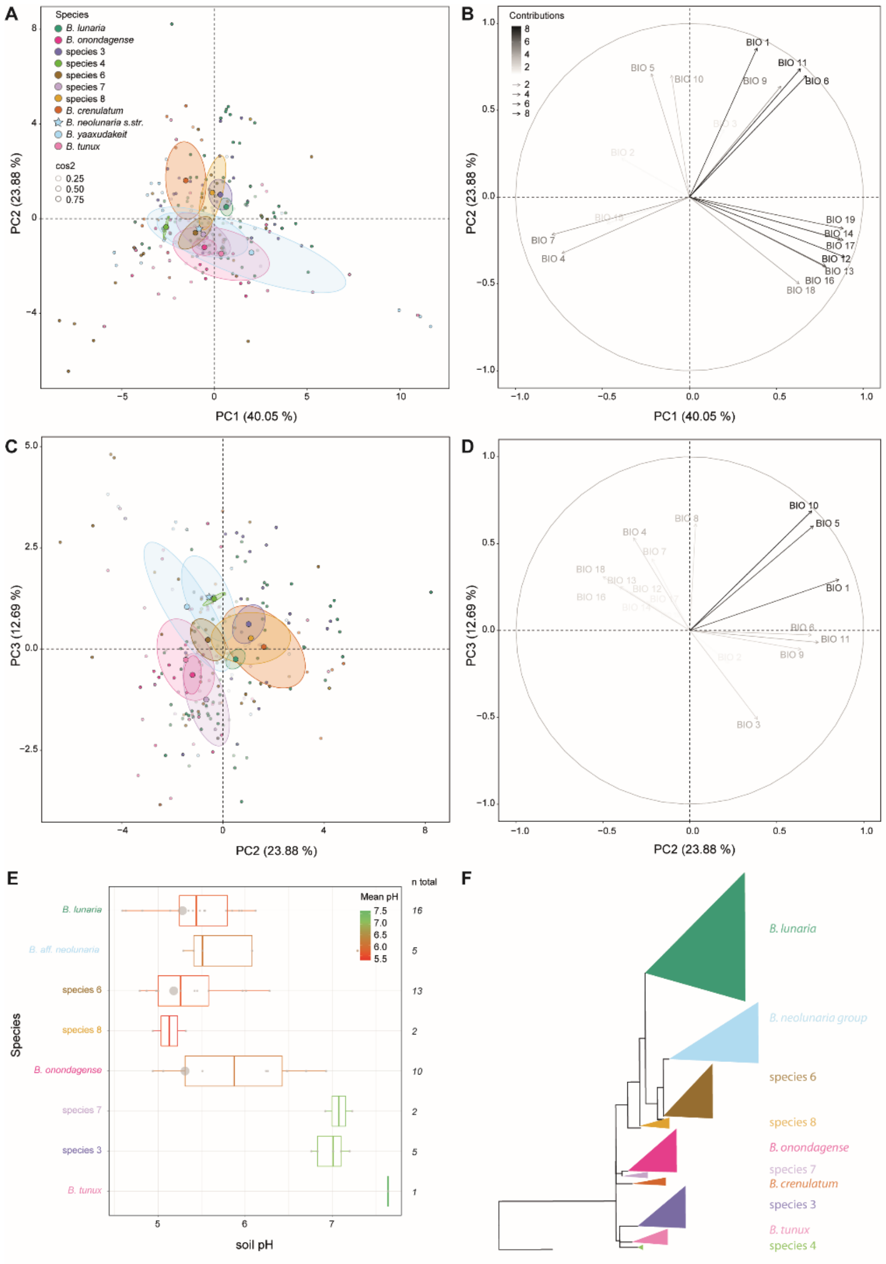


FIGURE 6: **Ecological differentiation between species of the *B. lunaria* group**. (A-D) Principal Component Analyses (PCA) of climatic factors comparing the species occurrences. Point transparencies increase with lower cos2 values. Points and 95 % confidence ellipses are colored according to clade assignment, with centroids indicated by larger circles. Arrow transparency represents the variable contributions to the PC. The definitions of the CHELSA bioclimatic variables are given in the **supplementary Table S6**. (A) Individual factor map showing the PC 1 and 2. (B) Variable factor map showing the PC1 and PC2. (C) Individual factor map showing the PC 2 and 3. (D) Variable factor map showing the PC2 an PC3. (E) Boxplots of soil pH. The boxplot colors represent the pH mean per species. The grey dots show the distribution of the individual pH values, and their size is relative to the number of individuals with similar pH values. The sample size per species is given by ***n***. (F) ML phylogenetic tree depicted with collapsed clades. Triangle height relates the clade sizes, and the triangle width the maximum branch length.

Tables

TABLE 1: **Characteristics of the multiple alignments including the outgroup.**

Alignments	*Individuals*	*Alignment sites (bp)*	*Parsimony-informative sites* [1]	*Unique sequences* [2]	*Unique patterns* [3]	*Gaps %* [4]	*Variable sites %* [5]	*Indels* [6]	*Indels gaps %* [7]	*Substitution model* [8]	*Partition* [9]	*Auto MRE*[10]
trnH^{GUG}-psbA	524	582	139 (114)	-	-	-	-	-	-	-	-	-
trnL^{UAA}-trnF^{GAA}	514	414	100 (81)	-	-	-	-	-	-	-	-	-
matk	478	652	116 (99)	-	-	-	-	-	-	-	-	-
rpl16	458	726	157 (142)	-	-	-	-	-	-	-	-	-
A1 (reduced)	323	2374	341 (259)	177	317	3.49	8.93	37	19.31	**ML**: TIM1+I+G, BIN **BI**: GTR+I+G, M1P	1	**ML**: 450
A2 (full)	530	2374	99.71 (99.54)	303	440	10.93	9.86	-	-	**ML**: TIM1+I+G	1	-
Botrychioideae	20	1588	17.19	20	368	2.58	35.9	-	-	**Part 1**: GTR+I **Part 2**: GTR+I+G	**Part 1**: trnH^{GUG}-psb^{A}, trnL^{UAA}-trnF^{GAA} **Part 2**: matK	-

Notes: [1] Number of parsimony-informative sites given by the pis function of R phyloch package. The value in parentheses correspond to the number of parsimony-informative sites while outgroup was excluded. [2, 3, 4, 5, 7, 10] These values were reported on Raxml-ng outputs. [6]. Number of indels scored by the simple indel coding algorithm, [8,9] Optimal substitution models and best scheme selected by PartitionFinder2 for the ML and the BI analysis, respectively. The indels models used (BIN and M1P) were the simplest model available for indels.

Supplementary material

Supplementary Figures

[Available from Zenodo: https://doi.org/10.5281/zenodo.6531504]

SUPPLEMENTARY FIGURE S1: Boxplots of the differences between species for each bioclimatic variable.

SUPPLEMENTARY FIGURE S2: Phylogram of the Maximum likelihood (ML) tree including all accessions with tip labels. The tree is depicted in the rectangular layout. Tip labels correspond to the sequence code, type information (*i.e.*, taxa name and type status) when applicable, the region of origin and the three letters ISO code of the country. For further information on accessions refer to the supplementary Table 1. The tip colors indicate clade assignment, and the Bootstrap supports of the main clades and subclades are specified on left node sides.

SUPPLEMENTARY FIGURE S3: Posterior probabilities (PP) of the Bayesian (BI) tree based on reduced dataset. Phylogram of the BI tree in rectangular layout. The PP are depicted on the right side of nodes. Tip labels correspond to the sequence code, type information (*i.e.*, taxa name and type status) when applicable, the region of origin and the three letters ISO code of the country.

SUPPLEMENTARY FIGURE S4: Effect of indels on the trees' topology inferred from the reduced dataset. (A) Tanglegram of the trees inferred using Maximum likelihood (ML) method. The tree on the right side corresponds to the ML tree inferred with indels and on the left side to the ML tree inferred without indels. Corresponding tips are link by lines colored according to clade assignments. (B) Tanglegram of the trees inferred using Bayesian (BI) method. The tree on the right side corresponds to the BI tree inferred with indels and on the left side to the BI tree inferred without indels. Corresponding tips are link by lines colored according to clade assignments.

SUPPLEMENTARY FIGURE S5: Time calibrated phylogenies of the Botrychioidae sub-family genus. On the left-side time divergence analysis ran under strict clock model and on the right-side time divergence analysis ran under relaxed clock model. The dark red and cadet blue bars represent the height of 0.95 HPD. The height medians are specified above the bars. The black dots indicate a posterior probability above 0.95. The time scales are in million-year unit.

SUPPLEMENTARY FIGURE S6: Time calibrated phylogeny of the *B. lunaria* complex inferred under the strict clock model. The dark red bars represent the height of 0.95 Highest Posterior Density (HPD) of the calibrated nodes. The height medians are specified above the bars. The black dots indicate a posterior probability equal or above 0.95 and the grey dots show a posterior probability comprises between 0.90 and 0.95. The time scale is in million-year unit.

SUPPLEMENTARY FIGURE S7: Time calibrated phylogeny of the *B. lunaria* complex inferred under the relaxed clock model. The cadet blue bars represent the height of 0.95 HPD of the calibrated nodes. The height medians are specified above the bars. The black dots indicate a posterior probability equal or above 0.95 and the grey dots show a posterior probability comprises between 0.90 and 0.95. The time scale is in million-year unit.

SUPPLEMENTARY FIGURE S8: PC 3 and 4 of the Principal Component Analysis (PCA) at the species level. (A) Individual factor map. (B) Variable factor map**.** The definitions of the CHELSA variables are given in the **supplementary Table S6**.

SUPPLEMENTARY FIGURE S9: Principal Component Analysis (PCA) at the sub-species level. (A) Biplot for *B. lunaria* subspecies showing PC 1 and 2. (B) Biplot for *B. lunaria* subspecies showing PC 2 and 3. (C) Biplot for *species 6* subspecies showing PC 1 and 2. (D) Biplot for *species 6* subspecies showing PC 2 and 3. (E) Biplot for *B. neolunaria* subspecies showing PC 1 and 2. (F) Biplot for *B. neolunaria* group subspecies showing PC 2 and 3. The definitions of the CHELSA variables are given in the **supplementary Table S6**.

Supplementary Tables

[See separate file, available from Zenodo: https://doi.org/10.5281/zenodo.6531504]

SUPPLEMENTARY TABLE S1: Information about the specimens used in this study.
SUPPLEMENTARY TABLE S2: PCR protocols and thermocycling conditions developed in this study.
SUPPLEMENTARY TABLE S3: Information about the sequences used in the divergence time analyses at the Botrychioideae subfamily level.
SUPPLEMENTARY TABLE S4: Flow cytometry results.
SUPPLEMENTARY TABLE S5: Spore length measurements.
SUPPLEMENTARY TABLE S6: CHELSA climatic variable meanings.

Supplementary Files

[Available from Zenodo: https://doi.org/10.5281/zenodo.6531504]

FILE S1: Multiple alignment of the full dataset (A2) used for the phylogenetic reconstruction.
File name: File_S1_A2_full_alignment.zip
FILE S2: Multiple alignment of the reduced dataset (A1) used for the phylogenetic reconstruction.
File name: File_S2_A1_reduced_alignment.zip
FILE S3: Multiple alignment of the subfamily Botrychioideae used for the time divergence analyses.
File name: File_S3_time_calibration_Botrychioideae_alignment.zip

FILE S4: Multiple alignment A1 used for the time divergence analyses. This alignment is the alignment containing no duplicated sequenced produced by RAxML-NG.
File name: File_S4_time_calibration_B.lunaria_group_alignment.zip

FILE S5: Maximum likelihood tree inferred from A2 alignment with bootstrap support values.
File name: File_S5_ML_A2_12102020.zip

FILE S6: Bayesian tree inferred from A1 alignment with posterior probabilities.
File name: File_S6_BP_A1_15102020.zip

FILE S7: Maximum likelihood tree inferred from A1 alignment with bootstrap support values.
File name: File_S7_ML_A1_15102020.zip

FILE S8: Bayesian tree inferred from A1 alignment with indels scored. Contains the posterior probabilities.
File name: File_S8_BP_A1_indels_15102020.zip

FILE S9: Maximum likelihood tree inferred from A1 alignment with indels scored. Contains the bootstrap support values.
File name: File_S9_ML_A1_indels_15102020.zip

FILE S10: RAxML-NG bestTree of the phylogenetic tree inferred from A1 alignment (File S7) used to constraint the backbone of Maximum likelihood tree inferred from A2 alignment (File S5)
File name: File_S10_ML_A1_bestTree_15102020.zip

FILES S11-S23: Summary of the Analysis of Variance, Tukey post-hoc tests results, Principal component analysis eigenvalues and variable contributions. Contained in Folder S1.
Folder name: Folder_S1_climatic_analyses.zip

CHAPTER 3: Taxonomical novelties in the *Botrychium lunaria* group (Ophioglossaceae)

Vinciane Mossion[1], Donald Farrar[2], Jason Grant[1], Zhang Xian-Chun[3], Alessio Maccagni[4] & Michael Kessler[5]

[1] Laboratory of Evolutionary Genetics, University of Neuchâtel, Neuchâtel, Switzerland

[2] Department of Ecology, Evolution and Organismal Biology, Iowa State University of Science and Technology, United-States of America

[3] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China

4 Giardino botanico del Cantone Ticino, Isole di Brissago, CH-6614 Brissago

[5] Department of Systematic and Evolutionary Botany, University of Zürich, Zurich, Switzerland

This chapter is in preparation for its submission to *Phytotaxa.*

Herbarium collected during this thesis project have not yet been deposited in herbaria. By consequences, I here do not provide information on the herbaria for those when listing the specimens examined. This will be added before manuscript submission.

Abstract

Recent molecular studies have revealed that the Botrychium lunaria group (Ophioglossaceae) consists of at least 11 species, with only 6 of these named so far. Based on previously published molecular and ecological evidence, and on the morphological differences presented here, in the present study we describe three new species (Botrychium himalayense, B. orientale, B. rotundum), one new subspecies (B. orientale subsp. nemus) and redescribe a formerly synonymized species (B. onondagense). We further provide notes on the remaining species and further undescribed species that are too poorly understood at present to be described.

Keywords: *Botrychium*, taxonomy, Ophioglossaceae, diversity, morphology, identification, habitats

Introduction

The fern genus *Botrychium* L. (Swartz 1801) is widely distributed in the northern temperate and boreal regions, with outposts in southernmost South America, Australia, and New Zealand. It has a relatively simple morphology, with each plant usually producing a single leaf per year that is subdivided into a photosynthetic, pinnate segment (the trophophore) and a fertile segment of variable degree of division (the sporophore). The genus has been defined differently over the years. Mild (1869) consider it to contain two subgenera (*EuBotrychium* and *Osmundopteris*), while later Clausen (1938) and Kato & Sahashi (1977) made it even broader by lumping *Sceptridium* (Lyon 1905) and *JapanoBotrychium* (Masamune 1931), respectively, into *Botrychium*. This broad circumscription of *Botrychium* was conserved in the last treatment of the Flora of China (Xianchun et al. 2013). Meanwhile, phylogenetic studies based on molecular data (Hauk et al. 2003; Shinohara et al. 2013; Zhang et al. 2020) support the segregation into six genera, defining *Botrychium* in a fairly narrow way, and *Botrypus* Michaux (1803), *Sceptridium* Lyon (1905), *JapanoBotrychium* Masamune (1931), *Holubiella* (Škoda 1997), and *Sahashia* Zhang and Zhang (2020) being now considered to be distinct (PPG I 2016; Zhang et al. 2020). Within this narrowly defined genus *Botrychium*, molecular and morphological studies have recognized three well delimited clades (Hauk et al. 2012; Dauphin et al. 2014, 2017), of which the *Botrychium lunaria* group (Lunaria clade in Hauk *et al.* 2012 and Dauphin *et al.* 2017 sense) is the focus of the current study.

The *Botrychium lunaria* group contains what was long considered to be one morphologically variable and geographically widespread species (Milde 1869; Clausen 1938). This species was described by Fuchs (1542) as *Lunaria minor*, defined by Linnaeus (1753) as *Osmunda lunaria*, which was transferred into *Botrychium* by Swartz (1801) when he described the genus. Later, three further species-level names were proposed, namely *B. lunatum* (Salisbury) Gray (1821), *B. racemosum* Bubani (1901), and *B. onondagense* Underwood (1903) of which the first two are currently regarded as synonyms of *B. lunaria* and the last was recently used to name one of the genetic lineages in the *B. lunaria* group (Mossion *et al.* in prep). In the last two decades of the twentieth century, detailed studies in North America by W.H. Wagner and colleagues started to show that *B.* "*lunaria*" contains several genetically and morphologically distinct forms. The first of these was described as *B. crenulatum* by Wagner & Wagner (1981). The distinctness of this species was later confirmed based on isozyme data by Hauk & Haufler (1999). Following these findings, Stensvold *et al.* (2002) described *B. tunux* and the allotetraploid *B. yaaxudakeit* from mixed populations with *B.* "*lunaria*" in Alaska. Finally, Stensvold & Farrar (2017) described two further species, namely *B. neolunaria* from western North America and *B. nordicum* from Iceland and Norway. Thus, currently six species-level taxa are recognized in the *B. lunaria* group.

Most of the above research was centered on North America, with Europe and Asia poorly explored until recently. In the last decade, however, several studies have extensively explored the genetic diversity of the *B. lunaria* group in western and central Europe using chloroplast (Dauphin et al. 2014, 2017; Maccagni et al. 2017) and allozyme markers (Dauphin et al. 2020). Even more recently, this sampling has been extended to cover eastern Europe and Asia (Mossion *et al.* in prep.). These studies have revealed further unexpected cryptic diversity in the *B. lunaria* group. In the most comprehensive sampling to date, Mossion *et al.* (in prep.) recovered ten well-supported clades, several of which correspond to the previously described species whereas others did not have any applicable names (Figure 1).

In the present study, we describe some of the previously unnamed clades at species and subspecies level, while acknowledging that several taxonomically unresolved groups remain. For this reason, our study should not be viewed as a comprehensive taxonomic treatment of *B. lunaria* group, but rather as a step towards a better understanding of a particularly taxonomically challenging fern group. We list all species of the group in alphabetical order, with short notes on some previously described, well-known species, and full descriptions on new and re-circumscribed taxa.

Material and Methods

Species delimitation

We followed the delimitation of species-level taxa proposed by the phylogenetic study of Mossion *et al.* (in prep.). Their delimitation relied on several lines of evidence. First, the species correspond to genetically well-defined, monophyletic lineages, typically with an estimated age of 2.6-1.8 million years (Mossion *et al.* in prep.). Second, for some well-studied lineages it has been shown that they conserve their morphological distinctness even in sympatry, suggesting they are reproductively isolated (Wagner and Wagner 1983), as previously documented in North America (Stensvold *et al.* 2002) and Europe (Maccagni et al. 2017), as well as by field observations by Mossion *et al.* (in prep.). Third, species in sympatry conserve their genetic characteristics, as shown for *B. lunaria* and *B. nordicum* (Stensvold and Farrar 2017), even though hybrid individuals are regularly found in mixed species populations. These typically have intermediate or unusual morphology and often malformed spores. However, fertile introgressed hybrids have also been documented (Stensvold 2008; Stensvold and Farrar 2017) and an allopolyploid species has evolved from past hybridization events (Stensvold et al. 2002). Fourth, the species have coherent geographical ranges that are often coupled with slight differences in climatic and edaphic niches between species (Mossion *et al.* in prep.). Finally, as described in the present study and is typical in other species groups in the genus *Botrychium* (Farrar and Johnson-Groh 1991; Wagner and Grant 2002; Gilman et al. 2015; Farrar et al. 2017; Popovich et al. 2020), the species are morphologically distinguishable by subtle traits, even though there is considerable morphological overlap between species.

Subspecies-level taxa are described based on less well-defined, younger (0.5-1 my) clades nested within species. Typically, the remainder of the species is paraphyletic relative to the subspecies. Nevertheless, we are aware that recently evolved species may be nested within their parent species due to incomplete lineage sorting (Rieseberg and Brouillet 1994; Syring et al. 2007).

Morphological analysis

We examined the vouchers of the individual included in the phylogenetic analysis of Mossion *et al.* (in prep) when available, and additional individuals collected between 2016 and 2020 in the same populations, for a total of 277 studied specimens. We defined a set of qualitative and quantitative traits based on diagnostic characters previously used for the study of *Botrychium* morphology (Farrar 2011; Williams et al. 2016), *B. lunaria* group species descriptions (Wagner and Wagner 1981; Stensvold et al. 2002; Stensvold and Farrar 2017), and identification keys of the North American Ophioglossaceae (Farrar et al. 2017). We measured 19 quantitative traits (Table 1) on the sporophyte from high-resolution scans of dried pressed specimens using the software ImageJ v.1.52a (Rasband 1997; Schneider et al. 2012). We used these measurements to derive ten ratios (partly given in Table 1). We also considered qualitative traits such as outline, color, and the habit of living plants which were recorded on herbarium specimens, high-resolution scans, and photographs of living plants. The terminology used for the qualitative traits was based on Lellinger (2002) and previous *Botrychium* species descriptions (Wagner and Wagner 1981; Stensvold et al. 2002; Stensvold and Farrar 2017; Popovich et al. 2020). Spore size average was estimated by measuring the longest diameter of a minimum of 10 spores mounted in Euparal medium for several specimens per species selected from different populations when possible (n=4 plants for *B. rotundum*, n=3 for *B. himalayense*, n=5 for *B. onondagense*, n=4 for *B. orientale* subsp. *orientale*, n=3 for *B. orientale* subsp*. nemus*, n=8 for *B. lunaria* including n=2 for *B. lunaria* 'helveticum'). The spores were photographed and measured using an inverted microscope (Leica DMI3000B – Olympus IX81) and measured with the softwares LAS v4.13 and Olympus CellSens Dimension 2.2. Known or suspected hybrids were measured but not included in the description, and immature individual were not used to describe minimum sizes and ratios.

To assess the genome size of new taxa, we used a subset of examined specimens that were previously studied by flow cytometry (Mossion *et al.* 2021; Mossion *et al.* in prep.). We completed these data by estimating the genome size of species 4 of which, unfortunately, no silica-dried material was available from the sequenced specimens. Thus, we used five specimens collected at the same time, at the same localities, and with a comparable morphology to the sequenced ones. Their genome size were estimated by flow cytometry following the one-step method of (Dolezel and Bartos 2005), as described in Mossion *et al.* (in prep).

Distribution range

Global distribution range were defined by combining geo-referenced data from the phylogenetic studies of Dauphin *et al.* (2017), Maccagni *et al.* (2017) and Mossion *et al.* (in prep), type localities (Underwood 1903; Wagner and Wagner 1981; Stensvold et al. 2002; Stensvold and Farrar 2017), herbarium specimens of confirmed identities, and living plants that could be confidentially identified from photographs.

Species accounts

Botrychium crenulatum Wagner in Wagner & Wagner (1981)

This species was described on morphological grounds, being characterized by its herbaceous texture, its yellowish color, a long common stalk compared to *B. 'lunaria'*, with a stalked trophophore, an average of 3 pinna pairs, much more remote, smaller, and narrower segments, with fewer veinlets terminating along the distal margin, and the prevalence of crenulate, as opposed to nearly entire, outer margin (Wagner and Wagner 1981). The genetic distinctness of the species was later confirmed by isozyme (Hauk and Haufler 1999) and more recently by plastid markers (Dauphin *et al.* 2017; Mossion *et al*. in prep.). Its known range is limited to North America.

Botrychium himalayense **V. Mossion, X. Zhang, M. Kessler sp. nov.**

Type:-- China, YUNNAN, Shangri-La, slope on the roadside, 4334 m, 6.8.2017, 28.38212804, 99.00178296, *VM_lun54_3_CN* (Figure 2).

DIAGNOSIS:-- *Botrychium himalayense* is characterized by its yellow-green sporophyte; flabellate to lunate pinnae with a rounded appearance, opposite and slightly asymmetrical with larger upper sides and straight basiscopic margins; trophophores long-stalked, with the junction between the sporophores and trophophores conspicuously below the first pinna pair; sporophore stalks about as long as or longer than the mature trophophores, with the fertile sporophore portions shorter at maturity than the sporophore stalk, representing 35-45% of the total sporophores length; sporangia-bearing branches only 3-4 pairs, with sporangia arranged by pairs, mostly imparipinnate on lateral branches but paripinnate at fertile portion apices; pinnae folded towards the sporophore; fertile portions of the sporophores tend to be pointed towards the trophophore in immature living plants and partly outward in mature plants.

Description:-- *Underground stems* (rhizomes) erect, unbranched, bearing 8 fleshy roots. *Aboveground leaves* erect, glabrous, with yellow-green pinnae and rachises turning reddish when dry. Total leaf height from the ground 62-100 mm, with common stalks 18-31(50) mm long, 0.4–1 time the length of the mature sporophores. Junction between the sporophore and trophophore stalk is conspicuously below the first pinna pair. *Trophophores* narrowly ovate with apices round, once pinnate, 15-28 mm long.

Trophophore stalks (1)2–5(7) mm long, (0.05)0.2–0.25 times the entire length of the mature trophophores which represent (5)20-30 % of the trophophore blade length. Trophophore blades 12–22 mm long, basally 5–8 mm wide. *Pinna outlines* fan-shaped to lunate, rounded, the stalk short but thick, somewhat opposite, mostly asymmetrical with an oversized upper part but longer lower part. Lateral margins straight, rarely concave, outer pinna margins entire to crenulate. Junction between outer and lateral margins sharp-cornered. Angle between basiscopic margin of middle pinna and rachis 50-90°**.** Basal pinnae spanning an arc of 90-140°. *Pinna pairs* 3-5, basal pairs more or less parallel to the rachises, those above the basal pairs slightly oriented upward, remote or overlapping. Lower pinnae mostly remote, distance between basal and 2nd pairs 2–5 mm, distance between 2nd and 3rd pairs 1.5–2 mm, the former distance 1–3 times the latter. Basal pinna pairs slightly larger in size than the adjacent pairs, 9–13 mm long, 5–8 mm wide, stalked. The subsequent pairs gradually but slightly reduced in size, short-stalked with a large stalk, basiscopic margins tend to be extended, and outer margins not overlapping the rachises. Pinnae tend to fold towards the sporophore in living plants. *Pinna veins* dichotomous, 2 major veins entering the pinna base, 8-27+ ending near the margins. *Pinnae* with mostly moderately thin texture, veins partly visible on herbarium specimens. Rachises of a thinner texture than the pinnae. *Sporophores* erect, 31–47 mm long, mature sporophores 1.4–2 times the entire length of the trophophores. Sporophore stalks 20–29 mm long, mature sporophore stalk 0.9–1.2 times the entire length of the trophophores (less than 0.8 times while immature), fertile portion of sporophores 11–21 mm long and 0.5-0.8 times the length of sporophore stalks, mostly branched, narrowly ovate to triangulate in outline. The fertile portion represents at maturity 35-45% of the total sporophore length. Sporangia-bearing branches 3-4 pairs, lowermost branches the longest, 4-7 mm long, once pinnate, stalked, slightly ascending, opposite branches straight. Distance between basal branch pairs and 2nd pairs 1.7-3 times the distance between 2nd and 3rd branch pairs. The basal branch pairs conspicuously distant from the other branch pairs, distance between subsequent pairs gradually reduced. Fertile portion of sporophores tend to be pointed towards the trophophore in immature living plants and partly outwards in mature living plants (Figure 3). *Sporangia* (8)30-80, arranged by pairs, opposite, with usually one terminal spore on lateral branches but not at the fertile portion apices. *Sporangia* 0.9-1.5 mm in diameter. *Sporophores* longer at maturity than the entire length of the trophophores, with a stalk equally or slightly longer than the trophophore. *Spores* 34–36 µm in longest diameter (Figure 4). 2 C-values 20.00 pg; diploid.

Etymology:--We name this species for its distribution centered on the Himalayan mountain chain.

Distribution:-- *Botrychium himalayense* occurs at 3960-4334 m in the Himalayas where it is known from a few localities in Nepal (Manasly province) and China (Yunnan province and Tibet autonomous region). Additional localities may be expected to be found in suitable habitats throughout the mountain range.

Ecology:-- *Botrychium himalayense* grows on neutral to basic soil, limestone bedrocks, in steep, gravely slope dominated by dwarf shrubs of *Rhododendron* spp., in *Kobresia* meadows with shrubs of *Potentilla* spp., *Sibiraea* spp., *Rhododendron* spp., and *Salix* spp., and in alpine grass meadows dominated by *Rhododendron* shrubs. Sporangia open in August.

Additional specimens examined:**-- China**, YUNNAN, Shangri-La, slope on the roadside, 4334 m, 6.8.2017, 28.38212804, 99.00178296, *VM-lun54-7-CN, VM-lun54-8-CN, VM-lun54-9-CN, VM-lun54-10-CN* (paratypes, PE!), *VM-lun54-1-CN, VM-lun54-2-CN, VM-lun54-4-CN, VM-lun54-5-CN, VM-lun54-11-CN* (paratypes); Haba Xueshan, 4099 m, 4.8.2021, 27.34709803, 100.095795, *VM-lun52-1-CN* ; TIBET AUTONOMOUS REGION, 4200 m, 8.8.2013, 29.387909, 92.04351, Jin X.H., Wang L.S., Wang Q., Wei R., Fu Z.X., Feng Y.L., *01314569* (PE)

Notes:-- This species was referred to as Species 7 in Mossion *et al.* (in prep.) and to the Asian group at the base of Lunaria 2 in Dauphin *et al.* (2017). It is one of the more distinct species in the *B. lunaria* group and should be reliably identified based on herbarium specimens.

Botrychium lunaria (Linnaeus 1753) Swartz 1801 s.str.

= *Botrychium lunatum* Gray (1821) nom. illeg. This name was presented by (Gray 1821) to refer to *B. lunaria* of Swartz without any justification for the change

The nomenclature of *B. lunaria* in the strict sense is challenging to resolve, since the type specimen of Linnaeus (Jonsell & Jarvis 1994; Linnaeus 1753) has no locality information. The type locality has been suggested to be somewhere in Sweden (Dauphin *et al.* 2017; Jonsell & Jarvis 1994), even though the type comes from Clifford herbarium which is mainly composed of cultivated plants and plants from various collectors around the world (Jarvis 2016). Thus, it is difficult to trace back the original collection site of the Linnaeus type. Moreover, even if the type is from Sweden, four taxa of the *B. lunaria* groups have so far been found to occur in Sweden (*B. lunaria*, *B. lunaria* var. *melzeri*, *B. rotundum*, and species 4; Dauphin *et al.* 2017, Mossion *et al.* in prep). Morphologically, the Linneaus type is close to *B. lunaria*, but its immature condition makes reliable name assignment difficult without molecular study. The poor condition of the type prevents the necessary molecular analysis, and the absence of type locality avert further material collection for morphological or molecular analysis. For these reasons, epitypification of *B. lunaria* should be considered by future studies. We refrain from this because of uncertainty regarding the identity of species 4 and because even as narrowly defined as here, *B. lunaria* contains a number of well-defined genetic clades, including *B. nordicum* (Figure 1). Thus, it is conceivable that further study may break up the species into even finer species-level taxa. In such a situation, premature epitypification may result in the name being restricted to a clade that does not correspond to the intention of the authors during epitypification. Since the botanical code only allows a single epitypification, this should only be conducted when the situation is clearer than it is now.

Botrychium racemosum (Bubani 1902) is based on *Lunaria minor* (Fuchs 1542) which was considered by Linnaeus (1753) to be a synonym of *Osmunda lunaria*. However, we are unable to assign this name to any species of the *B. lunaria* complex recognized here because of the uncertainty surrounding the typification of *B. lunaria* s.str., and because the original description is very sketchy. Further, numerous varieties and forms have described under *B. lunaria* by Milde (1869) and (Clausen 1938), but the inadequate descriptions and the absence of many types make the assignment to current species almost impossible.

For the present moment, we keep the name *B. lunaria* for the large clade (clade 1 Mossion *et al.* in prep.) containing *B. lunaria* var. *melzeri*, *B. nordicum*, a multitude of smaller nested clades, and a well-supported sister clade that we refer to as *B. lunaria* 'helveticum'. As defined here, *B. lunaria* occurs from Greenland across Europe to Morocco and Kyrgyzstan. Our circumscription of *B. lunaria* is much narrower than that of previous authors, because North American '*B. lunaria*' are now assigned to *B. onondagense* (which see for further details) which also narrows the morphological definition of *B. lunaria*. However, morphological variation among the specimens we examined is still considerable. This variation could be the result of phenotypic plasticity related to habitat characteristics like the height of the surrounding vegetation or the amount of shade as well as hybridization events or recent evolutionary units which would not have been detected by the more slowly evolving plastid genome.

B. lunaria var. ***lunaria***

As defined here, "typical" B. lunaria can be recognized by a number of traits. Plants have ovate to narrowly ovate trophophores, with 5-7 pinna pairs having slightly asymmetrical outline. The middle and upper pinnae are ascending. Usually, the stalk of mature sporophores is slightly shorter or as long as the trophophores but tends to be longer while plants are growing in the shade. For most of the measured plants, the fertile portion of the sporophore fertile is about 30-50 % of the total length of mature sporophores, with a few exceptions (up to 60%, see *SJ-lun59-3-LT* and *SJ-lun59-4-LT* for example). Also, the distances between sporangia bearing branches are gradually reduced toward the apices. Live plants have some characters that are not visible on the herbarium specimens. For example, live plants are green to dark green and the difference in color with *B. tunux* or *B. rotundum* is noteworthy. Also, *B. lunaria* has a thinner pinna texture than the species cited above, and the sporangia are often turned outward (not facing the trophophore) even in immature plants.

Within *B. lunaria* var. *lunaria*, two groups can be separated morphologically based on trophophore stalk, sporangia diameter and pinna outlines. The first group is characterized by the absence of trophophore stalks, larger sporangia (0.8-1.2 mm), more lunate pinnae, whereas the second group is defined by long trophophore stalks, smaller sporangia (0.6-0.8 mm), and narrower fan-shaped to almost rhombic pinnae. Interestingly, specimens of the first group occurr mostly in the western part of the distribution range and specimens of the second group in the eastern part.

Moreover, some of the examined specimens are here considered to represent possible hybrids (0209486*; 0209483; VM-lun79-1-GR; VM-lun79-5-GR; VM-lun74-8-BA; VM-lun69-1-HR; VM-lun72-5-BA*). *VM-lun74-8-BA* is very robust with twice pinnate trophophore, deeply incised pinnae, several unusually robust sporophores, and a junction between the trophophore and the sporophore almost at the bottom of the pinna petioles of the second pinna pair. This plant was collected in a mixed population of *B. lunaria* and *B. orientale*. *VM-lun69-1-HR* is morphologically intermediate between *B. lunaria* and *B. onondagense*, and *0209483*[2] is intermediate between *B. lunaria* and *B. rotundum* or *B. tunux*. We also identified two specimens with a very singular morphology (*VM-lun79-1-GR*, *VM-lun79-5-GR*) from a mixed population *B. lunaria* and *B. orientale* in Greece. These specimens have pinna outlines reminiscent of specimens from Vogar Iceland, called '*B. lunaria f. struckii*', with pinnae outlines between lunate and flabellate, long stalked, a thin texture, and outer margins crenulate. The most remarkable traits are the sporophore rachises and sporophore branch width which appear excessively large compared to similar-sized specimens, especially on *VM-lun79-5-GR*. Average spore sizes of *VM-lun79-5-GR* and *VM-lun72-5-BA* (44-46 µm) enter the range of the allopolyploid species of *Botrychium* (Farrar pers. obs.). However, genome sizes measured for two specimens reveal a diploid condition (*VM-lun79-1-GR*, 2C-value 21.98 pg; *VM-lun69-1-HR*, 2C-value 22.26 pg).

Botrychium lunaria var. *lunaria* has been recorded from a large range of habitats (Figure 5). It grows on acidic soils, mostly in natural or semi-natural meadows like alpine pastures where it often occurs on steep, exposed parts of meadows, dominated by *Nardus* spp. or *Festuca* spp. (Alps) and associated with *Alchemilla vulgaris* and *Hieracium pilosella*. It has also been found in high alpine meadows dominated by small grasses, lichens, and mosses, often associated with *Thymus* spp. And *Gentiana campestris*, in sparse birch pine forest, on the border of forest and mountain steppe among bushes like *Juniperus communis*, in ditches inside dry meadows associated with *Hieracium* spp. and *Dianthus deltoides* (Sweden), in wet meadows, on rocky slope between *Pinus mugo* patches where tiny mountain streams are running (Carpathians), on pozzine (Atlas), in *Calluna vulgaris* fields on dry sand (Netherlands), and in mixed boreal forest among dense *Calamagrostis arundinacea* tussocks (Sweden). The phenology of *B. lunaria* differs depending on elevation and latitudes. Sporangia open in early June at low elevations (below 1000 meters) to mid-August at higher elevations (up to 2700+ meters) in the boreal-temperate regions, whereas in the south of the distribution range sporangia open at high elevation (above 3000 meters) early to mid-July (*e.g.*, Moroco, south Spain).

Examined specimens:-- Herb. Clifford 472 (lectotype, BM!)

Austria, TYROL, Grossglockner, 2070 m, 13.7.2014, 42,542, 12,827, *000148305* (NEU); VORARLBERG, Zwemähder, 2030 m, 24.8.2016, 46,9241405360121, 10,0517144476628, *AB-lun30-1-AT*

Andorra, ENCAMP, Envalira, 2320 m, 7.8.2014, 42,5417082, 1,7117429, *000148304* (NEU)

Belarus, MINSK, ~ 0.5 km from the bridge over the river Berezina, upstream of the river, 170 m, 28.5.1998, 54,138649, 26,407082, 0209486 (MW)

Bosnia & Herzegovina, HERZEGOVINA-NERETVA, At the edge of Veliky vran peak, 1993,8 m, 12.7.2018, 43,6660809628665, 17,5043629948049, *VM-lun72-1-BA*, *VM-lun72-5-BA*; On the side of the path going to Veliki Vran peak, 1630,06 m, 12.7.2018, 43,65986100398, 17,5123269762843, *VM-lun72-19-BA*, *VM-lun72-20-BA*; SARAJEVO, Treskavica mountains, Ilijaš, 1730,3 m, 14.7.2018, 43,6210370343178, 18,3682409953325, *VM-lun73-1-BA*;

Croatia, LIKA-SENJ, Bunovac valley, 1218,3 m, 07.7.2018, 44,3547539692372, 15,5414149817079, *VM-lun70-1-HR*; PRIMORJE-GORSKI-KOTAR, Close to Risjnak peak, on the border of the path, 1351,7 m, 02.7.2018, 45,4286999814212, 14,624719005078, *VM-lun69-1-HR*; ŠIBENIK-KNIN, Dinara mountain, 1355,7 m, 09.7.2018, 44,0748470183461, 16,3684759754687, *VM-lun71-2-HR*, *VM-lun71-7-HR*, *VM-lun71-9-HR*

France, ALSACE, Grand Ballon, 1270 m, 23.7.2014, 47,8974632, 7,1031091, *000148321* (NEU); RHÔNE-ALPES, Béal pass, 1600 m, 9.8.2014, 46,427, 3,807, *000148325* (NEU); SAVOIE, Galibier pass, 2300 m, 9.8.2014, 45,080021, 6,4232254, *000148326* (NEU)

Georgia, KAKHETI, path between "Meteo" to "Black rock lake". Population located in a slope upper the path, 2580,5 m, 29.6.2016, 41,8742339964956, 46,3848690222948, *VM-lun2-1-GE*; MTSKHETA-MTIANETI, close to Gergeti Trinity church, on the left side of the path in the direction of the top, 2457,4 m, 31.6.2016, 42,6663909852504, 44,601653991267, *VM-lun3-1-GE*; RACHA-LECHKHUMI-AND-LOWER-SVANETI, Forest road from Tsanashi village to Lebarde crosspath, near Lentekhi village, 1921,4 m, 05.7.2016, 42,7624419797211, 42,6096080336719, *VM-lun6-1-GE*, *VM-lun6-2-GE*

Germany, BADEN-WÜRTTEMBERG, Feldberg, 1450 m, 0.0.2014, 47,869927, 8,012723, *JG-lun35-1-DE*, *JG-lun35-2-DE*

Great Britain, ENGLAND, Appleslade Bottom, 55 m, 11.6.2016, 50,876294, -1,631549, *RMW-lun32-1-GB*

Greece, EASTERN-MACEDONIA-AND-THRACE, Paggaio ski center, mount Paggeo, 1894 m, 845825 m, 26.7.2018, 40,9143849648535, 24,0944249834865, *VM-lun79-1-GR*, *VM-lun79-5-GR*; KALLII, Vardousia, 1900 m, 19.6.2015, 38,6965, 22,1315, *000147846* (NEU)

Iceland, NOROURLAND-EYSTRA, Myvatn Lake, 0 m, 13.7.2005, 65,60386, -16,996105, *DF12942* (NEU)

Kyrgyzstan, ISSYK-KUL, Barskoon mountain, 2946 m, 15.7.2017, 42,0888699777424, 77,6415430102497, *VM-lun51-1-KG*; NARYN, On the plateau above Kalek's yurt Camp, 2615 m, 09.7.2017, 41,5842000301927, 76,4296020381152, *VM-lun47-3-KG*

Lithuania, TELŠIAI, 0.5 ha dray hill, 180 m to the west of the place where the Uošna2 river flows through the gravel road., 140 m, 23.6.2017, 55,994122, 21,930136, *SJ-lun58-3-LT*, *SJ-lun58-5-LT*; VILNIUS, 0.04 ha forest opening 70 m east of the junction with the road leading to Zelvė Recreation Center., 127 m, 22.7.2017, 54,822308, 24,8063, *SJ-lun59-3-LT*, *SJ-lun59-4-LT*

Norway, TELEMARK, Holtan, 6.2 m, 59,392008, 9,21743, *DF17167* (NEU)

Romania, HARGHITA, Campsite next to the road (unpaved), from pensiunea (meteo station), 1730,7 m, 07.8.2016, 47,0957679860293, 25,2665400132536, *VM-lun17-2-RO*; MARAMUREȘ, between Pasul Prislop and Isorul Bristita, 30-40 minutes from the lake, 1588,4 m, 04.8.2016, 47,5839030370116, 24,8078459780663, *VM-lun15-2-RO*

Russia, IVANOVO, near village Pashki, on the south-eastern shore of Ryabo lake, 0 m, 12.6.2014, 56,695871, 41,9586773, *0209575* (MW); MURMANSK, Porya lip, Medvezhiy, eastern bay, 0 m, 06.7.2010, 66,72414, 33,699, *0209436* (MW); REPUBLIC-OF-BASHKORTOSTAN, Bashkir State Nature reserve, southern Kraka mountain range, suqre 104, section 26, upper part of the slope of the Northern exposure, 0 m, 27.7.1999, 53,3720479127158, 57,7696427002768, *0209631* (MW), *0209631* (MW); SMOLENSK, Smolensk, 204 m, 22.6.2001, 55,4833, 32,133, *0209483* (MW); TVER, 2 km north of the village Greshnevo, 151 m, 25.6.2014, 58,20319, 34,78221, 0209455 (MW); Russia, TVER, right Doybitsa river bank, 131 m, 16.6.1987, 56,504425, 36,540844, *0209461* (MW)

Slovakia, BANSKÁ-BYSTRICA, side of the hiking trail fom Trangoska (green marks), 50m from Stefanika refugia, 1697,9 m, 17.7.2016, 48,926213029772, 19,6471289824694, *VM-lun7-1-SK*, *VM-lun7-2-SK*, *VM-lun7-7-SK*; PREŠOV, Belianske Tatry, south part of the mountain range. Dolina Bielych plies, on the red path to Kopské sedlo, 10 minutes from Biele pleso, 1566 m, 22.7.2016, 49,2250220105051, 20,2241569664329, *VM-lun12-1-SK*; ŽILINA, Western Tatras, between Malé Závraty and Ziarske seldo on both side of the path, 1691,1 m, 25.7.2016, 49,1904179938137, 19,7396060172468, *VM-lun13-1-SK*, *VM-lun13-2-SK*

Spain, ANDALUSIA, Sierra Nevada, 3100 m, 0.8.2014, 37,0488751, -3,3684683, *000148313* (NEU)

Sweden, NORBOTTEN, In Archipelago near to Lulea, 7 m, 6.7.2012, 65,588078, 22,621655, *000147808* (NEU)

Switzerland, SCHWYZ, Rigi, 1750 m, 28.6.2014, 47,897, 8,485, *000100747* (NEU); TICINO, Gola di Lago, 900 m, 23.6.2014, 46,487, 8,965, *000100763* (NEU); VALAIS, Chemin de Sotier, 150 m de loc 1, 2215 m, 20.7.2019, 46,159807, 7,554648, *VM-lun94-3-CH*; Sex de la Brinta, sur la pente en amont du Sentier, 2606 m, 14.7.2019, 46,210898, 7,531453, *VM-lun89-1-CH*; Dans les pentes au dessus du lac de Moiry en direction du Sex de Marinda, 2305 m, 16.7.2019, 46,135475, 7,566306, *VM-lun91-1-CH*; Val d'Anniviers, on the left side of the educational path toward the black lake., 2498 m, 17.8.2018, 46,24905, 7,620656, *VM-lun86-36-CH*, *VM-lun86-41-CH*

Botrychium lunaria* var. *melzeri Stensvold & Farrar (2017)

This variety was described from Greenland and Iceland based on allozyme and morphological evidence. Stensvold & Farrar (2017) did not consider it to be sufficiently distinct to warrant species status, a result supported by plastid phylogenies (Dauphin *et al.* 2017, Mossion *et al*. in prep.) who recovered it as a monophyletic, well supported clade deeply nested within *B. lunaria*. Mossion *et al.* (in prep.) named this taxon *B. lunaria* "*subsp. melzeri*" for consistency with other infraspecific taxa recognized in that study (all which were treated as subspecies), but without proposing a formal recombination.

Botrychium lunaria var. *melzeri* is characterized by a long and narrow almost glaucous trophophore and remote pinnae. Pinnae are often rhombic or lobed in outline, with the basal pinnae often bearing or even replaced by small sporophore. This variety tends to be mature earlier than *B. lunaria* var. *lunaria* (Stensvold and Farrar 2017). It has been recently recorded from Sweden and Romania (Mossion *et al.* in prep.) through plastid markers.

Examined specimens:-- Greenland, NARSARSUAQ, 25 M north of airport runway, 52 m, 04.7.2005, 61,15916667, -45,42805556, *DF12773*; KUJALLEQ, Narsarsuag, 0 m, 5.7.2005, 61,174664, -45,389705, *DF12806* (isotype, NEU!)

Iceland, NORÐURLAND-EYSTRA, 0, 279 m, 13.7.2005, 65,61833, -17,023333, *DF12913* (paratype)

Romania, NEAMȚ, On the slope of Toaca peak near the path., 1850,166138 m, 08.8.2016, 46,9768700189888, 25,94929497689, *VM-lun18-2-RO*

***Botrychium lunaria* 'helveticum'**

Botrychium lunaria 'helveticum' forms a genetically well-supported clade, sister to the remainder of the *B. lunaria* clade (Mossion *et al.* in prep.). It was referred to *B. lunaria* subspecies 5 in Mossion *et al.* (in prep.), Lunaria 5 in Dauphin *et al.* (2017), and clade 4.3 in Maccagni *et al.* (2017).

Morphologically, it is very similar to *B. lunaria*, but we noticed a few differences in trophophore outline and sporophore ratios (Table 1, Figure 6). *B. lunaria* 'helveticum' has terminal pinnae of the same length as the adjacent pinnae, but narrower, giving the terminal pinnae a more elongated aspect, and trophophore apices a pointy outline. More striking, basiscopic pinna margins of pinnae above the basal ones are straight and angled at 90° with the rachises, giving a ladder-like aspect to the trophophores. Also, sporophores tend to be shorter, and at maturity sporophore stalks are mainly conspicuously shorter in length than the trophophores.

A few individuals from this clade could possibly be hybrids (*VM-lun90-1-CH; VM-lun71-1-HR*). *VM-lun71-1-HR* has a long and thin sporophore stalk and a marked trophophore stalk. It was collected from a mixed population of *B. lunaria* and *B. orientale*. *VM-lun90-1-CH* is a very robust plant with deeply incised pinnae and a junction between the trophophore and the sporophore almost at the bottom of the pinna petioles of the second pinna pair. Such high insertion of the sporophore was also observed in other potential hybrids identified by plastid data as *B. lunaria*. These plants were found at one location in Switzerland and one in Croatia.

Specimens of *B. lunaria* 'helveticum' have been found mainly in the Alps and the Jura mountains with most of the localities known from Switzerland and a few in Austria, France, and Italy. However, the specimens outside of Switzerland were phylogenetically assigned only based on partial sequencing data, except for one individual from Austria. Thus, further molecular studies including additional markers should be performed to conclude on the extend of the geographic distribution of this clade.

Examined specimens: -- Austria, STYRIA, Sölk pass, 1570 m, 14.7.2014, 47,2573882965362, 14,0824556012667, *000148309* (NEU); TYROL, Timmelsjoch, 2150 m, 14.7.2014, 44,121, 11,053, *000148308* (NEU)

Croatia, ŠIBENIK-KNIN, Dinara mountain, 1323,680542 m, 09.7.2018, 44,0741830039769, 16,3667020294815, *VM-lun71-1-HR*

France, PROVENCE-ALPES-CÔTE-D'AZUR, St Véran, près de l'Est de la chapelle des Clausis, 2359 m, 27.8.1978, 44,673462, 6,922021, *014790* (AIX)

Italy, LIGURIA, Maggiorasca, 1750 m, 4.6.2014, 44,5504632351699, 9,48864161968231, *000148336* (NEU)

Republic of Moldova, OCNITA, Gîrbova, 0 m, 17.7.1886, 47,305083, 28,693868, *77734* (MARS)

Switzerland, BERN, Chasseral, 1550 m, 29.6.2014, 47,1280437977979, 7,04639106988906, *000100769* (NEU); 1549,71 m, 7.6.2017, 47,12974, 7,04934, *000101256* (NEU), *000101257* (NEU); GRAUBÜNDEN, Bernina pass, 2050 m, 11.7.2014, 46,4110852, 10,0494969, *000100756* (NEU); VALAIS, Tsirouc, en descendant de Sorbois, à côtè d'une cabane en bois, 2197 m, 15.7.2019, 46,164413, 7,589963, *VM-lun90-1-CH*; Val d'Anniviers, on the left side of the educational path toward the black lake., 2498 m, 17.8.2018, 46,24905, 7,620656, *VM-lun86-8-CH*, *VM-lun86-17-CH*, *VM-lun86-38-CH*, *VM-lun86-40-CH*; Sex de la Brinta, sur la pente en amont du Sentier, 2606 m, 14.7.2019, 46,210898, 7,531453, *VM-lun89-3-CH*; Chemin de Sotier, 150 m de loc 1, 2215 m, 20.7.2019, 46,159807, 7,554648, *VM-lun94-1-CH*, *VM-lun94-2-CH*; Entre la cabane du Grand Mountet et le début de la montée, 2865 m, 18.7.2019, 46,062462, 7,650156, *VM-lun92-2-CH*; Descente du grand Mountet vers Zinal, bord du chemin, 2463 m, 18.7.2019, 46,082099, 7,640803, *VM-lun92-6-CH*; Le Biolec, 1954 m, 19.7.2019, 46,166823, 7,594218, *VM-lun93-1-CH*, *VM-lun93-2-CH*, *VM-lun93-5-CH*.

Botrychium neolunaria Stensvold & Farrar (2017)

This species was described from North America and Greenland based on a combination of molecular and morphological data. It is morphologically characterized by an oblong to narrowly ovate trophophore outline, pinnae mostly remote to somewhat overlapping with a junction between the basiscopic side margins and the outer margins sharply angled. The sporophore stalk largely exceeds the trophophore in length and the fertile portion is about a third of the entire sporophore length with strongly ascending branches (Stensvold and Farrar 2017). It is one of the parent taxa of *B. yaaxudakeit* (Dauphin *et al.* 2014; Stensvold *et al.* 2002).

Already Stensvold & Farrar (2017) noted that the genotype of *B. neolunaria* is found in specimens from both North America and Eurasia that do not correspond morphologically to typical *B. neolunaria*. They interpreted these as fertile, introgressed hybrids between *B. neolunaria* and *B*. "*lunaria*" and did not award them taxonomic rank. Mossion *et al.* (in prep.) documented that such individuals are very widespread in Eurasia. The taxonomic status of these plants is difficult to resolve. The plastid phylogeny of Mossion *et al.* (in prep.) recovered these plants mixed with "true" *B. neolunaria* and *B. yaaxudakeit*, suggesting that *B. neolunaria* is the maternal parent of these forms. It is thus not possible to reach a better understanding of the evolutionary independency of these plants without nuclear data. The presumed introgressed specimens are further characterized by extreme morphological variability, with some closely approaching *B. neolunaria* and others *B. onondagense* or *B. lunaria*. Furthermore, genome size varies considerably among individuals, ranging from 17.5 pg (2C-values) in two specimens from Siberia (Baikal Natural Reserve) to more typically 20-22 pg (2C-values) (Mossion *et al.* in prep.). We interpret all this evidence to indicate that the introgressed hybrids are the result of various hybridization events, possible including more than one paternal taxon, and that the genetic contribution of the parental taxa varies between individuals and populations. Much more detailed studies using both nuclear markers and morphological analyses will be necessary to understand the evolution and taxonomy of the populations.

Based on these complications, for the time being, we propose to recognize "true" *B. neolunaria* only from North America and to place all introgressed specimens in a "*B. aff. neolunaria* group".

Botrychium nordicum Stensvold & Farrar (2017)

This species was described from Iceland and Norway based on a combination of molecular and morphological data. It is morphologically distinct from *B. lunaria* primarily by its deeply incised pinna margins, shorter common stalk, and thinner pinna texture with highly visible veins (Stensvold and Farrar 2017). More recently, it has been confirmed from Scotland based on plastid markers (Dauphin et al. 2017).

In the phylogenetic study of Mossion *et al.* (in prep.), *B. nordicum* was recovered as a monophyletic group deeply nested within *B. lunaria*. For consistency with other nested clades and to avoid rendering *B. lunaria* as a non-monophyletic species, Mossion *et al.* (in prep.) treated this taxon at subspecies level, but without making a formal taxonomic recombination. We believe that *B. nordicum* is a recently evolved taxon (most of its range was covered by glaciers during the ice ages) that may be an evolutionary independent species but that remains nested in *B. lunaria* due to incomplete lineage sorting. For the present time, we propose no change of its taxonomic status.

Botrychium onondagense Underwood (1903)

This species was described by Underwood from Onondaga county, New York. It was treated as a synonym of *B. lunaria* by subsequent authors (Clute 1905; Clarke and House 1923; Clausen 1938; Butters and Abbe 1953), but the molecular phylogenetic analysis (Dauphin et al. 2017) recovered North American specimens of '*B. lunaria*' and European specimen of '*B. lunaria*' in two distinct clades. This segregation of North American and European genotypes was simultaneously suggested by allozyme markers (Stensvold and Farrar 2017). Since the morphology of the type of *B. onondagense* corresponds to that of sequenced specimens of the North American clade, we here use this name for this clade. Underwood's description included four specimens, of which three are immature, including the type. Immature specimens are inappropriate for morphological characterization because the trophophores and sporophores grow until sporangia ripen. However, the overall pinna shape and arrangement are consistent between Underwood's specimens and our specimens. Because of the much larger sample size available to us and the broader geographical distribution resulting in more morphological variability (Figure 7), we here provide a refined description of *B. onondagense*.

Description:-- *Underground stems* (rhizomes) erect, unbranched, bearing 6-16+ fleshy roots. *Aboveground leaves* erect, glabrous, mostly with bud green pinnae to light green rachises which turn red-green when dry. Total leaves height from the ground to the higher part of the rachises of mature plants (31) 41-125 mm, with common stalks 11-50 mm long, 0.2-1 times the length of the mature sporophores. Mostly, the mature sporophores largely exceed the common stalks in length. Junction between the sporophore and trophophore stalk mostly below the petioles of the first pinna pair. Sometimes at the base of the petioles of the first pinna pairs, rarely above. *Trophophores* ovate to narrowly oblong with apices round, once pinnate, 12-55 mm long. Trophophore stalks sometimes absent; when present, (0.2)0.6–4(8) mm long, mostly less than 0.1 (0.2, see specimen DF16494 and DF16506 for examples) times the entire length of the mature trophophores, which represent 1 to 12 (30) % of the trophophore blades' length, making stalks appearing extremely small and often absent even if present. Trophophore blades 13–58 mm long, basally 3–20 mm wide. *Pinna outlines* cuneate to flabellate, stalked even in small plants with sometimes exceedingly long and thin stalk, basal pinnae somewhat opposite with the subsequent mostly alternate, almost symmetrical while cuneate to slightly asymmetrical while fan-shaped with a wider upper side and extended lower side. Lateral margins straight to moderately concave, basiscopic margins extended on upper pinnae but not spreading, outer margins entire to slightly crenulate mostly in unusually large plants. Junction of outer and lateral margins rounded. Angles between basiscopic margin of the middle pinna and rachis about 40-90°. Basal pinnae spanning an arc of 70° to nearly 180°. Occasionally pinnae bear few sporangia on their margins. *Pinna pairs* (2) 3-5 (8), basal pairs more or less parallel to the rachises, upper pairs slightly ascending, remote to overlapping. Lower pinnae remote with upper pinnae remote to overlapping; distance between basal and 2nd pairs 2–10 mm, distance between 2nd and 3rd pairs 1–5 mm, the former distance (0.9) 1.1–3.3 times the latter. Basal pinna pairs equal to slightly larger in size than the adjacent pairs, 11–28 mm long, 6–13 mm wide, stalked. The subsequent pairs gradually reduced in size, stalked; acroscopic margins of upper pinna pairs do not overlap the rachises, except in noticeably large or immature plants. Pinnae sometimes fold toward sporophores in living plants. *Pinna veins* dichotomous, 2 major veins entering the pinna base, 17-30+ ending near the margins. *Pinnae* with thick to thin texture, veins poorly to nicely visible, especially on herbarium. Rachises always have a thick texture. *Sporophores* erect, 22–100 mm long, mature sporophores 0.9–2.3 times the entire length of the mature trophophores. Sporophore stalks 11–46 mm long, 0.5–1.5 times the entire length of the trophophores, fertile portion of sporophores 10–60 mm long and 0.3-1.6 times the length of sporophore stalks, branched, broadly to narrowly lanceate in outline. The fertile portion represents at maturity 25-60 % of the total mature sporophore length. Sporangia-bearing branches (2) 4–6 pairs, lowermost branches usually the longest, 2-17 mm long, 1-2 pinnate, moderate- to short-stalked, rarely subsessile, mostly opposite sometimes alternate. Distance between basal branch pairs and 2nd pairs 1-1.9 (2.5) times the distance between 2nd and 3rd branch pairs. The basal branch pairs sometimes conspicuously distant from the other branch pairs. The basal, 2sd and 3rd branch pairs often appear more distant from each other than the remain of the pairs among them. After the third branch pairs the distance between the subsequent branch pairs tend to be abruptly reduced toward the fertile portion apices. Sporangia-bearing branches not ascending, basal branches almost perpendicular to the rachises, oriented outward or inward of the trophophore in mature living plants. *Sporangia* 15 to 100+ in large plants, arranged by pairs, opposite, imparipinnate. Sporangia 0.7-1.2 mm in diameter. Sporophores longer at maturity than the entire length of the trophophores, with a stalk mostly shorter and rarely exceeding the trophophore. *Spores* 33–39 µm in longest diameter. 2C-values 20.66-22.1 pg; diploid.

Distribution:-- *B. onondagense* occurs in North America from north-USA across Canada to Alaska and Greenland. It further occurs in Iceland, Svalbard, in central and eastern Europe from France to Romania and Italy to Norway, in northwestern Russia, in Tajikistan, and in Kyrgyzstan.

Ecology:-- *Botrychium onondagense* grows from sea level to 4200 m on acidic, neutral and basic soils, in a large range of habitats including alpine meadows, seaside rocky-sandy meadows, beaches protected from waves action, damp meadows, disturbed soil areas like roadsides, unstable rocky slopes, in grassy swales, and on flat limestone outcrops.

Sporangia open in late June at lower elevations (sea level) to mid of August at higher elevations (above 3000 m).

Examined specimens:-- USA, NEW YORK, Onondaga county, Split Rock, near Salvay Process Quarry, 20.06.1890, Underwood L.M., *00144293* (syntypes, NY!); Jamesville Road near Syracuze, Rust M. O., *0014492* (NY!)

Austria, SALZBURG, NNW der Sticklerhütte, Riedingscharte, 2270 m, 19.9.2014, 47,15883, 13,37175, *MH-lun39-1-AT*, *MH-lun39-2-AT*; STEIERMARK, Stuhleck, 1770 m, 14.7.2014, 47,5739490201965, 15,7906150817871, *000100752* (NEU)

Canada, MANITOBA, "Tank Farm" south of Cape Henry , 7 m, 14.7.2005, 58,779004, -94,184629, *DF15597*; NEWFOUNDLAND, Fogo Island, 0 m, 5.7.2001, 49,669017, -54,140424, *DF5060* (NEU); coastal sand dunes, 25 m, 06.8.2019, 50,71416667, -57,39111111, *DF20196*; along roadside across cape, 4 m, 25.7.2019, 51,56361111, -55,74305556, *DF20300*; blufftops above ocean, 58 m, 25.7.2019, 51,58194444, -55,74305556, *DF20315;* ONTARIO, James Bay, 0 m, 6.2012, 52,642995, -82,020802, *DF19250* (NEU); QUEBEC, Bic NP, 0 m, 27.6.2008, 48,33619, -68,819094, *DF16494* (NEU); Mingan Islands, 0 m, 15.8.2002, 50,219187, -63,902083, *DF7874* (NEU); 10 miles SE of Rimouski, 10 m, 27.6.2008, 48,35944444, -68,77916667, *DF16506*; WININPEG, Tank Farm, 0 m, 14.7.2007, 49,875393, -97,118415, *DF15589* (NEU)

France, SAVOIE, Galibier pass, 2300 m, 9.8.2014, 45,080021, 6,4232254, *000148326* (NEU)

Greenland, KUJALLEQ, Narsarsuag, 0 m, 4.7.2005, 61,211243, -45,275267, *DF12833* (NEU)

Italy, BELLUNO, Giau pass, 2120 m, 13.7.2014, 46,4867069557722, 12,0423243888027, *000148337* (NEU)

Kazakhstan, ALMATY, On the top of Mokhnataya Sopka near Medko tract, 2600 m, 4.8.1936, 43,1357456, 77,0506852, *0800423* (MW)

Kyrgysztan, CHUY, Valley of high altitude close tothe road in the south direction, 2730 m, 01.7.2017, 42.2018870338797, 73.2397759985178, *VM-lun44-1-KG*; NARYN, On the plateau above Kalek's yurt Camp, 2615 m, 07.7.2017, 41,5842000301927, 76,4296020381152, *VM-lun47-1-KG*; NARYN, Dolon Pass. On the slope near the parking 500m before the pass, to the south, 2987 m, 08.7.2017, 41,839481033847, 75,7483200263231, *VM-lun48-2-KG*; Montenegro, DURMITOR, 0, 1900 m, 0.0.0, 43,1117, 18,9949, *000147847* (NEU)

Romania, DÂMBOVIȚA, gully close to the refugia, red triangle hiking trail between Pestera and Omu, 2105,2 m, 14.8.2016, 45,4226690344512, 25,4121400415897, *VM-lun21-1-RO*, *VM-lun21-2-RO*; DÂMBOVIȚA, Mecetul Turcese, between red strip path to Omu and blue strip path to Pestera., 2320,5 m, 14.8.2016, 45,4380609747022, 25,4492530133575, *VM-lun22-1-RO*, *VM-lun22-2-RO*; HUNEDOARA, near Stâna Scorota, 1818,08 m, 18.8.2016, 45,2960430365055, 22,8538060002028, *VM-lun26-1-RO*, *VM-lun26-2-RO*; PRAHOVA, between Cabana Valea Dorului and Varful cu Dor, 1830 m, 13.8.2016, 45,35514, 25,48783, *VM-lun20-1-RO*; more or less 100-200 m below Babele telecabin, 2157,3 m, 14.8.2016, 45,4056940041482, 25,4748749826103, *VM-lun23-1-RO*, *VM-lun23-2-RO*, *VM-lun23-3-RO*, *VM-lun23-4-RO*

Russia, MURMANSK, Turiy cape, by the road, 3 m, 27.6.2008, 66,53777, 34,51086, *0209424* (MW); Island Big Velichanka on the White Sea, south-eastern cape of the island, 0 m, 29.7.1963, 67,099354, 32,38457, *0209432* (MW); Porya lip, lonely luda, the central part of the island, 0 m, 23.7.2011, 66,77817, 33,61504, *0209434* (MW); Medvezhiy, eastern bay, 0 m, 06.7.2010, 66,72414, 33,699, *0209436* (MW); Porya Guba, Tatianina Severnaya Luda, 2 m, 14.7.2008, 66,7021, 33,86397, *0209435* (MW)

Taiwan, XIULIN, Nanhu Mt, 0 m, 19.6.2002, 24,28584, 121,427991, *DF7268* (NEU);

Tajikistan, DISTRICTS-UNDER-CENTRAL-GOVERNMENT-JURISDICTION, Peter I mountain, Fortambek glacier, Suloev (Сулоева) glade, on the southern slope of Beliy Kamen mountain, 4200 m, 20.7.1981, 38,976582, 71,895179, *0800417* (MW); SUGHD, Between Maslihatteppe and Kalsit, from Saratog village follow the river Arch then follow the river Biob then go up to Kalsit mountain, 3324 m, 25.6.2017, 39,1266359947621, 68,2712559588253, *VM-lun42-1-TJ*, *VM-lun42-2-TJ*, *VM-lun42-7-TJ*; along a stream (spring water), on the side of Kasnok river, close to the path to Biruzovoe lake, just after the path up to Alaudin lake., 3344 m, 26.6.2017, 39.1647419705986, 68.2283169776201, *VM-lun43-1-TJ*, *VM-lun43-5-TJ*, *VM-lun43-6-TJ*

Notes:-- This species forms a genetically well-supported clade which was refered to Species 2 in Mossion *et al.* (in prep.), Lunaria 2 in Dauphin *et al.* (2017) and clade 2 in Maccagni *et al.* (2017). Morphologically, it is characterized by its bright green pinnae and light green rachises colors, short-stalked trophophores, and pinnae cuneate to flabellate and slightly asymmetrical, with the basal pinna pairs opposite and the subsequent ones often alternate. The basiscopic pinna margins are mostly straight, and the upper pinna pairs are slightly ascending, whereas the basal pairs are somewhat parallel to the rachises. The mature sporophore stalks are shorter than, rarely exceeding, trophophores in length, with the fertile portion of the sporophore about a third, rarely more than half, of the total length of mature sporophores. The three basal sporangia-bearing branch pairs typically are more distant from each other than the more distal pairs. Thus, after the third branch pairs the distances between the branch pairs tend to be abruptly reduced. Sporangia-bearing branches are not ascending, and the basal branches are almost perpendicular to the rachises, with the sporangia organized by imparipinnate pairs.

Among the examined specimens, numerous are confirmed or putative hybrids between *B. onondagense* and other species. Introgressed hybrids between *B. onondagense* and *B. neolunaria* (*DF7268, DF15597, DF7874, DF5060, DF15589*, possibly *DF20196*) confirmed by allozymes (Stensvold 2008) tend to have intermediate pinna morphology varying from almost flabellate to broadly lunate, sometimes with incised margins, with pinnae spreading and ascending. Sporangia-bearing branches tend to be strongly ascending, with sporophore stalks and sporophore fertile portions length and ratios towards the extreme of the range size, often exceeding the upper side of the size range (*e.g.*, sporophore length up to 160 mm). An additional number of specimens (*0209432; 0209434 [MW24_2]; 0209435; 000148337* [AM48O]*; DF12833; VM-lun22-2-RO*) are here considered to represent possible hybrids between this species and *B. lunaria.* Their pinnae are broadly lunate, asymmetrical, sometimes with a long and thin stalk and with deeply incised margins (*DF12833*). Plants are bigger and can be very robust, such as *DF12833*, of which the fertile portion of the sporophore is reminiscent of *B. lunaria* var. *melzeri*. *VM-lun22-2-RO* shows no morphological peculiarity, except being a very tiny plants and alternate basal pinna pair, but its average spore size (45 µm) (Figure 4, 7) enters the range of the allotetraploid *B. yaaxudakeit* (Stensvold et al. 2002). Putative hybrids were found as scattered individuals at various sites in North America, Russia, and Romania.

Botrychium orientale **V. Mossion, & M. Kessler**

Type:-- Bulgaria, BLAGOEVGRAD, on the side of hiking trail to Xaha Pametaka, 2134 m, 29.07.2018, 41.77291397, 23.41994103, *VM-lun81-3-BG* (Figure 8).

DIAGNOSIS:-- *Botrychium orientale* is characterized by being a well-defined monophyletic clade based on plastic markers (Mossion *et al*. in prep.), and by strongly rippled spore coats, making the coat ornamentation peculiarly protruding like big bubbles appearing on the surface (Figure 4).

Description:-- *Underground stems* (rhizomes) erect, unbranched, bearing 6-16+ fleshy roots. *Aboveground leaves* erect, glabrous, mostly with light yellow-green pinnae and rachises, and red-green rachises when dry. Total leaf height from the (13)46-258 mm, with common stalks (4)8-126 mm long, 0.2-1.2 times the length of the mature sporophores. In two thirds of the cases, the mature sporophores largely exceed the common stalks in length. Junction between the sporophore and trophophore stalk between or above the petioles of the first pinna pair, less often below. *Trophophores* narrowly ovate, with apices round to oblong, once pinnate, (8)20-98 mm long. Trophophore stalks lacking in 76% of the examined specimens; when present, 3–7(34) mm long, 0.05–0. (0.45) times the entire length of the mature trophophores, which represent 5-30(80) % of the trophophore blade length. Trophophore blades (8)21–103 mm long, basally 3.3–30 mm wide. *Pinna outlines* broadly lunate to flabellate, the latter mostly in small plants and in plants from the eastern part of the distribution range, stalked even in small plants, mostly alternate, occasionally opposite, sometimes only the basal pinnae, asymmetrical with a disproportion toward the lower side. Lateral margins straight to concave, basiscopic margins rarely slightly to strongly recurved and often spreading on upper pinnae, outer margins entire to deeply incised, so that the pinnae appear divided onto several lobes. Junction of outer and lateral margins sharp-cornered. Angles between basiscopic margin of the middle pinna and rachis 37-120°. Basal pinnae spanning an arc of 90-180°. Occasionally pinnae bear few sporangia on their margins. *Pinna pairs* (2)4-7(9), more or less perpendicular to the rachises, remote to completely overlapping especially in plants with deeply incised pinnae, those above the basal pair spreading. Lower pinnae usually remote to approximate with upper pinnae approximate to overlapping; distance between basal and 2nd pairs 3–18 mm, distance between 2nd and 3rd pairs 2–13 (17) mm, the former distance 0.9–2.3 times the latter. Basal pinna pairs slightly larger to approximately equal in size to the adjacent pairs, (7)10–43 mm long, (3)7–22 mm wide, stalked. The subsequent pairs gradually reduced in size, stalked; sometimes acroscopic margins of upper pinna pairs overlapping the rachises. *Pinna veins* dichotomous, 2 major veins entering the pinna base, 25-40+ ending near the margins. *Pinnae* with very fleshy texture, veins poorly visible. Rachises of thicker texture than the pinnae. *Sporophores* erect, (16)30–173 mm long, mature sporophores 1.2–2.2(3.6) times the entire length of the mature trophophores. Sporophore stalks 17–88 mm long, 0.5–1.5(2) times the entire length of the trophophores, fertile portion of sporophores (9)13–82(92) mm long and 0.4-1.2(1.6) times the length of sporophore stalks, branched, broadly triangular to narrowly ovate in outline. The fertile portion represents at maturity (13)30-60% of the total sporophore length. Sporangia-bearing branches (4)6–9 pairs, lowermost branches usually the longest, 3-26 mm long, 1–2(3) pinnate, long- to short-stalked, opposite, or alternate. Distance between basal branch pairs and 2nd pairs is 0.9-1.9 times the distance between 2nd and 3rd branch pairs. The distance between branch pairs is gradually reduced toward the fertile portion apices. Sporangia-bearing branches strongly ascending, partly or completely twisted in outward of the trophophore in mature living plants. *Sporangia* 26 to 450+ in large plants, arranged by pairs, opposite, paripinnate. Sporangia 0.6-1(1.15) mm in diameter. Sporophores longer at maturity than the entire length of the trophophores, with a stalk often shorter or equally and rarely exceeding the trophophore. *Spores* 36–39 µm in longest diameter. 2C-values 21.37-22.91 pg; diploid.

Etymology:--The name refers to the eastern distribution of the species, centered in the Balkans and extending to China.

Distribution:-- *Botrychium orientale* is distributed from eastern Europe to eastern Asia, with an isolated population in New Zealand. In Europe it occurs throughout the Balkans (Bosnia & Herzegovina, Macedonia, Greece, Bulgaria) and the Carpathians (Slovakia, Ukraine, Romania), with a few outposts in the central-western part of Russia. In Asia, it has been recorded in Tadjikistan, Kyrgysztan, Mongolia, Japan, and easternmost Russia (Kamchatka Krai, Khaborovsk Krai and the Republic of Sakha). The identity of some plants from Alaska is uncertain, as detailed under subspecies *nemus*.

Ecology:-- *Botrychium orientale* grows from sea level to 4,200 m on acidic to semi-acidic soil, in alpine moors, alpine heath in alpine meadows, in open areas (scree, sliding snow corridors) inside alpine pine forest (*Pinus mugo*), on stony ridges, often associated with acidophilic communities including *Juniperus communis* var. *saxatilis* and Ericaceae such as *Vaccinium uliginosum* and *V. vitis-idea*. In the eastern part of its distribution range, *B. orientale* grows in glades in spruce forest, in larch forest, in ravine forest, in mixed coniferous forest of larch with spruce and Siberian cedar, and in damp meadows.

Sporangia open from mid-June at lower elevations (below 300 m) to mid-August at higher elevations (above 2,400 m).

Notes:-- This species was referred to as Species 6 in Mossion *et al.* (in prep.). *Botrychium orientale* is morphologically very similar to *B. lunaria*. The distinction based on morphology is easier when considering the two subspecies of *B. orientale* separately. Subspecies *orientale* differs from *B. lunaria* by its somewhat higher number of pinna pairs (6-9 vs. 5-7); pinnae broadly lunate and not flabellate, and less oriented upwards than *B. lunaria* (Figure 8, 9); trophophores light yellow green in living plants (vs. bright to dark green); junction between the sporophores and the trophophores usually above or between the petioles of the first pinna pair (vs. at the base of the first pinna pair); and fertile portion of the sporophyte with a higher number of sporangia-bearing branches (6-9 vs. 5-7), with up to twice as many sporangia, with branches bi- to tripinnate. The differentiation between *B. lunaria* and *B. orientale* subsp. *nemus* based on the sporophyte morphology is very challenging. Subspecies *nemus* tends to have fewer pinna pairs; pinnae more remote, less oriented upwards; and longer trophophore stalks, but there is large overlap. However, the distinction can be done based on spore coat ornamentations (Figure 4).

A number of specimens (*VM-lun71-4-HR, VM-lun72-22-BA, VM-lun79-3-GR*) are here considered to represent possible hybrids between this species and *B. lunaria*. In these plants, less than 20% of the spores were well developed, suggesting hybridogenic spore malformation and sterility. These specimens are very robust plants, with sometimes a basal pinna replaced by a second trophophore and with additional small sporophores appearing at the base of a regular, somewhat robust sporophore. They have pinnae somewhat spathulate (basal pinnae spanning an arc of ~40°) or lunate to fan-shaped and deeply incised. Putative hybrids were found as scattered individuals at various sites where *B. orientale* occurs in sympatry with *B.*

lunaria. Surprisingly, average size of the well-developed spores of *VM-lun72-22-BA* and *VM-lun71-4-HR* was 43-44 µm, corresponding to the range of the allotetraploid *B. yaaxudakeit* (Stensvold et al. 2002). However, the absolute genome size of *VM-lun71-4-HR* that we measured by flow cytometry (2C-value 21.87 pg; Mossion *et al.* in prep.) fell within the range of the diploid species of the *B. lunaria* group. The inconsistence between genome size and spore size could be explained by the formation of diplo-spores as observed in *Asplenium* hybrids (Reichstein 1981). Germination experiments would be required to test this hypothesis.

B. orientale subspecies *orientale*

DIAGNOSIS:--The nominal subspecies can be distinguished from subspecies *nemus* by its longer trophophores (up to 258 mm) lacking trophophore stalks, and a higher number of pinna pairs (6-8(9) vs. (2)4-5(6)), mostly broadly lunate in outline. The junction between the sporophore and trophophore stalk is mostly above the petioles of the first pinna pairs, and less often in between. The fertile portion of the sporophores is in general shorter than, rarely about, the length of the sporophore stalks, representing 30-50% of the total mature sporophore length (Figure 9).

Distribution:-- Known from the Balkans (Croatia, Bosnia-Herzegovina, Macedonia, Greece, Bulgaria), Slovakia, Romania, and Ukraine, and with few possible records from Central Asia and the eastern part of Russia.

Ecology:-- *Botrychium orientale* subsp. *orientale* grows at (400)1,000-4,200 m, on acidic to semi-acidic soils, in alpine moors, alpine heath in alpine meadows, in open areas (scree, sliding snow corridors) inside alpine pine forest (*Pinus mugo*), on stony ridges, often associated with acidophilic communities including *Juniperus communis* var. *saxatilis* and Ericaceae such as *Vaccinium uliginosum* and *V. vitis-idea*, more rarely in damp meadows or glades in a spruce or larch forest.

Additional specimens examined:-- Bosnia & Herzegovina, HERZEGOVINA-NERETVA, At the edge of Veliky vran peak, 1903,3 m, 12.7.2018, 43,6638620216399, 17,501989994198, *VM-lun72-10-BA*, *VM-lun72-21-BA*, VM-lun72-8-BA, VM-lun72-9-BA; SARAJEVO, Treskavica mountains, Ilijaš, 1730,3 m, 14.7.2018, 43,6210370343178, 18,3682409953325, *VM-lun73-2-BA*; on the path toward the highest peak, 1777,03 m, 15.7.2018, 43,6030300240963, 18,3628009725362, *VM-lun74-1-BA*, *VM-lun74-2-BA*; on the ridge, on the way back from highest peak. Pašina Planina., 2039 m, 15.7.2018, 46,5928333333, 18,35725, *VM-lun75-1-BA*, *VM-lun75-2-BA*, *VM-lun75-4-BA*

Bulgaria, BLAGOEVGRAD, Mountain above Popovi livadi, on the way to Orelek peak, out of the path following the ridge. Close to the red-yellow stick with yellow sign., 1959,2 m, 28.7.2018, 41,5624130330979, 23,6117880418896, *VM-lun80-1-BG*; On the side of the hiking path to Xaha Pemataka, 2133,6 m, 29.7.2018, 41,7729139700531, 23,4199410304427, *VM-lun81-1-BG, VM-lun81-2-BG, VM-lun81-4-BG, VM-lun81-5-BG,* *VM-lun81-8-BG*, *VM-lun81-10-BG*, *VM-lun81-13-BG* (paratypes, Figure 8); KYUSTENDIL, 7 lakes area. Population located about 100 m from Trilistnika lake, mostly found on southern oriented slope, close to the large rocks, 2223,04 m, 31.7.2018, 42,20523099415, 23,3199910260736, *VM-lun82-1-BG*, *VM-lun82-3-BG*, *VM-lun82-4-BG*; LOVECH, Papa (Pupa) peak, western part of the main ridge, close to the winter path, 1970,6 m, 07.8.2018, 42,7487769629806, 24,3762389849871, *VM-lun85-1-BG*, *VM-lun85-2-BG*; SOFIA, At the door of Rila National Park, on the way to Musala hut from the lift, 2410,7 m, 01.8.2018, 42,2062260098755, 23,5811070073395, *VM-lun83-2-BG*, *VM-lun83-7-BG*; STARA-ZAGORA, Eastern slope of Mazalat peak, "Evil peak"., 1953,04 m, 06.8.2018, 42,7197729796171, 25,1051079947501, *VM-lun84-3-BG*

Greece, EASTERN-MACEDONIA-AND-THRACE, Paggaio ski center, mount Paggeo, 1894,8 m, 26.7.2018, 40,9143849648535, 24,0944249834865, *VM-lun79-2-GR*, *VM-lun79-7-GR*; WESTERN-MACEDONIA, On the mount in front of the ski resort Vigla, up to the dirt road below the line of wind power generator, 1949,4 m, 22.7.2018, 40,797294974327, 21,2644089851528, *VM-lun78-1-GR*, *VM-lun78-3-GR*

Kazakhstan, ALMATY, On the top of Mokhnataya Sopka near Medko tract, 2600 m, 4.8.1936, 43,1357456, 77,0506852, *0800423* (MW)

Kyrgyzstan, ISSYK-KUL, Taarylga-Tor, 2996 m, 14.7.2017, 42,0757490303367, 77,6482200343161, *VM-lun50-2-KG*

Macedonia, POLOG, Popova Šapka ski resort, above the path to Jelac, 2059 m, 18.7.2018, 42,022193, 20,869077, *VM-lun76-1-MK*, *VM-lun76-11-MK*

Mongolia, KHENTII, Eren-Daba ridge, about 20 km NW from the village Norovlin on the road to the river crossing the river Onon, 0 m, 15.8.1989, 48,8364388, 111,6509524, *0168258* (MW)

Romania, BISTRIȚA-NĂSĂUD, colossus mountain, SW of Rodnei Muntii Natural Park, 1706,8 m, 05.8.2016, 47,4733530264347, 24,6829400304704, *VM-lun16-1-RO*, *VM-lun16-2-RO*; MARAMUREȘ, between Pasul Prislop and Isorul Bristita, 30-40 minutes from the lake, 1588,4 m, 04.8.2016, 47,5839030370116, 24,8078459780663, *VM-lun15-1-RO*; NEAMȚ, On the slope of Toaca peak near the path., 1850,1 m, 08.8.2016, 46,9768700189888, 25,94929497689, *VM-lun18-1-RO*

Russia, KHABAROVSK-KRAI, Dzhugdzhur, river valley Dzhangin, 853 m, 3.8.2010, 55,54771, 134,99221, 0003991 (MW); REPUBLIC-OF-SAKHA, Surroundings of the village, Yakokit, 369 m, 14.7.1945, 58,90856, 125,816597, 0003995 (MW)

Slovakia, KOŠICE-HIGHER-TERRITORIAL-UNIT, Slovakian Paradise, meadow on each side of the yellow hiking trail from Havrania skala to Statenský. 5 minutes from Havrania skala., 1103,6 m, 19.7.2016, 48,8903239835053, 20,3416760265827, *VM-lun9-1-SK*, *VM-lun9-5-SK*; PREŠOV, follow the yellow path to

Panská hol'a, pasture on both side near Smrečiny peak along the border of Low Tatras National Park, 1349,5 m, 20.7.2016, 48,9489869959652, 20,0542199984192, *VM-lun9-1-SK*, *VM-lun9-2-SK*

Ukraine, ZAKARPATSKA, Above the ski resort, near the path to Landarn. Blyznytsia mountain, 1748,1 m, 30.7.2016, 48,2323650363832, 24,2311059962958, *VM-lun14-1-UA*, *VM-lun14-2-UA*

USA, ALASKA, Solo Strip, 1021 m, 26.6.2007, 62,065833, -142,04111111, *DF15465*; 63,619999, -146,716112, *DF15460* (NEU)

***B. orientale* subspecies *nemus* V. Mossion & M. Kessler**

Type:-- Mongolia, KHENTII, Norovlin, Eren-Daba ridge about 20 km NW from the village Norovlin (Норовлин), on the road to the river crossing the river Onon, 15.08.1989, Diriymaa S., Grubov V. I., Gubanov I. A., *0168258* (MW) (Figure 10).

DIAGNOSIS:--Distinguished from the nominal subspecies by its smaller trophophores (rarely exceeding 70 mm), stalked trophophores, and a lower number of pinna pairs ((2)3-5(6) vs. 6-8(9)), lunate to flabellate in outline. The junction between the sporophore and the trophophore is mostly conspicuously below the petioles of the first pinna pairs. The fertile portion of sporophores is about the length of sporophore stalks or longer, representing (40)45-65% of the total mature sporophore length.

Distribution:-- Distributed throughout Russia, in the Central Asian mountain ranges, Mongolia, Japan, and with an outpost in New Zealand. Specimens from Alaska may belong to this species or to *B. neolunaria* (see under Notes).

Ecology:-- *Botrychium orientale* subsp. *nemus* grows from sea level to 2,000(3,000) on semi-acidic soils, in glades in spruce forest, in birch forest, in ravine forest, on the edge of willow forest, and more rarely in dry alpine pasture.

Additional specimens examined:-- Bosnia-&-Herzegovina, SARAJEVO, Treskavica mountains, on the ridge, on the way back from highest peak. Pašina Planina., 2039 m, 15.7.2018, 46,592833, 18,35725, *VM-lun75-5-BA*

Japan, NAGANO, Matsumoto, 0 m, 0.0.2014, 36,326106, 137,675719, *WS20140803001*; SHIZUOKA, Mt Fuji, 0 m, 0.0.2014, 35,360555, 138,727778, *WS20140903001*

Kyrgyzstan, NARYN, Son Kul, 3124 m, 03.7.2017, 41.7141040228307, 75.030067032203, *VM-lun45-2-KG*, *VM-lun45-3-KG*; Dolon Pass. On the slope near the parking 500m before the pass, to the south, 2987 m, 08.7.2017, 41,839481033847, 75,7483200263231, VM-lun48-3-KG

Mongolia, BAYAN-ÖLGII, Yolt river basin (upper reaches of the Black Irtysh) near the border with **China**, 30 km south of the village of Altai, 0 m, 23.7.1988, 48,302222, 89,515278, *0168257* (MW); KHENTII, Eren-Daba ridge, about 20 km NW from the village Norovlin on the road to the river crossing the river Onon, 0 m, 15.8.1989, 48,8364388, 111,6509524, *0168258* (MW)

Russia, KRASNOYARSK-KRAI, northwest of the Anabar plateau, the confluence of the river Ilya in r. Kotuikan, 164 m, 1.8.2007, 70,6164055, 105,2163405, *0004005* (MW); KAMCHATKA-KRAI, Yelizovo, river valley Avatcha, 18 m, 29.7.1966, 53,209954, 158,381313, 0003973 (MW); BELGOROD, near village Gorozhenoye (Горожeноe), Gorozhenoe forest tract, 0 m, 26.5.2007, 50,7474141, 37,3894184, *0209595* (MW); SARATOV, 2.5 km south-west of the village Orkino (between villages of Orkino and Ozerki), 244 m, 4.6.2008, 52,024083, 45,547055, *0209622* (MW); REPUBLIC-OF-SAKHA, Right bank of the river Viluy. The middle course of the river. Elgyay, 105 m, 14.7.1957, 62,473755, 117,512579, *0003998* (MW)

United-States of America, ALASKA, Eielson Farm Road, 150 m, 13.7.2017, 64,754176, -147,431142, *JG-neo1-1-US*; Selawik, 0 m, 0.0.2012, 66,567983, -159,4975, *MD12060* (NEU)

Notes:--This subspecies could be of hybridogenic origin as it contains individuals identified by allozymes as introgressed hybrids between *B. neolunaria* and '*B. lunaria*' (Stensvold 2008; Stensvold and Farrar 2017). Morphologically the introgressed hybrids display the three main characteristics of *B.* subspecies *nemus*, namely a junction between the sporophopre and the trophophore below the first pinna pairs, a trophophore stalk, and a fertile portion of sporophores about the length of sporophore stalks or longer, even though the examined specimens were not completely mature. More detailed molecular studies are needed to clarify the taxonomic relationships of the Alaskan plants and of the evolution of subspecies *orientale* in general.

Botrychium rotundum **V. Mossion, J. Grant, M. Kessler & D. Farrar**

Type:-- Switzerland, VALAIS, Fully, on the trail side going from the parking at Fully entrance to the 'Rocher des Follatères', 544 m, 13.05.2017, 46.1235, 7.0759, *VM-lun1-1-CH* (Figure 11).

DIAGNOSIS:-- Distinguished from the remainder of the taxa of the *B. lunaria* group by the following combination of traits: opposite, flabellate, rounded, and symmetrical pinnae in outline, except the basal pinna pairs that sometimes have a slightly extended lower side; pinnae remote to approximate and somewhat parallel to the rachises; trophophores ovate to triangulate with very thick texture and a light bright green color in living plants; junction between the sporophore and the trophophore is mostly between or at the base of the petioles of the proximal pinna pairs; sporophore stalk shorter or about the length of the trophophore, with the fertile part of the sporophore fertile about half to two thirds of the total sporophore length; basal sporangia-bearing branches conspicuously distant from the other branches with reduced inter-branches distance after the third branches; in living plants, mature sporophores are twisted, projecting the sporangia-bearing branches outward from the trophophores (Figure 3).

Description:-- *Underground stems* (rhizomes) erect, unbranched, bearing 5+ fleshy roots. *Aboveground leaves* erect, glabrous, mostly with bright-green pinnae and green-yellowish rachises. Total leaf height from the ground to the top of the rachises (32)39-123(197) mm, with common stalks (7)12–73 mm long, (0.2)0.3-1 times the length of the mature sporophores. Mostly, the common stalks are half to about the same length as the mature sporophores. Junction between the sporophore and trophophore stalk is between or at the base of the pinna petioles of the proximal pinna pair; rarely higher or lower. *Trophophores* ovate to almost triangular with rounded apices, once pinnate, 12-40(45) mm long. Trophophore stalks mostly absent; when present, 1–2 mm long, mostly less than 0.1 times the entire length of the mature trophophores, which represent 2-6 % of the trophophore blade length. Trophophore blades 12–46 mm long, basally 4–12 mm wide. *Pinna outlines* mostly flabellate, appear rounded, stalked, with the stalk moderately large to thin and longer, or appear longer while thin, basal pinnae somewhat opposite, rather symmetrical, occasionally slightly extended lower side. Lateral margins shallowly to visibly concave, basiscopic margins rarely extended and not spreading, outer margins entire to slightly crenulate, more rarely cleft. Junction of outer and lateral margins with a rounded angle. Angles between basiscopic margin of the middle pinna and rachis about 36-80(124) °. Basal pinnae spanning an arc of (75)90-160°. Occasionally, pinnae bear a few sporangia on their margins, basal pinnae are replaced by additional sporophores, or rachises bear few sporangia or even small sprorophores. *Pinna pairs* (2)3-4(5), more or less parallel to the rachis, remote, rarely overlapping. Lower pinnae remote with upper pinnae remote to approximate, rarely every pinna approximate to overlapping; distance between proximal and 2nd pair 2–10 mm, distance between 2nd and 3rd pair 1–6 mm, the former distance 1.1–2.9 times larger than the latter. Basal pinna pairs slightly larger in size than the adjacent pairs, 9–26 mm long, 5–16 mm wide, stalked. The subsequent pairs gradually reduced in size, stalked; acroscopic margins of upper pinna pairs do not overlap the rachises, even in noticeably overlapping pinnae or immature plants. Pinnae tend to fold toward sporophores in living plants. *Pinna veins* dichotomous, 2-3 major veins entering the pinna base, 20+ ending near the margins. *Pinnae* with very thick texture, marginal veins poorly (even on nicely pressed herbarium specimens) to not visible at all. Rachises of thicker texture than the pinnae. *Sporophores* erect, 22-77 mm long, mature sporophores (0.9)1.3-2.4 times the entire length of the mature trophophores. Sporophore stalks 8-41 mm long, (0.2)0.4-1.2(1.4) times the entire length of the trophophores, fertile portion of sporophores 9-35 mm long and (0.6)0.9-2.4(3.4) times the length of sporophore stalks, branched, broadly to narrowly lanceolate in outline. At maturity, the fertile portion (40)50-70(80) % of the total mature sporophore length. Sporangia-bearing branches (4)5-6(7) pairs, difficult to estimate on small plants with subsessile sporangia, lowermost branches usually the longest, (1)3-11(19) mm long, 1-2 pinnate, moderately- to short-stalked, rarely subsessile, opposite, basal branches sometimes alternate. Distance between 1st branch pairs and 2nd pairs (0.9)1.3-2.6 times the distance between the 2nd and 3rd branch pairs, distance between subsequent branch appear much shorter compared to the two first inter-branches, and gradually decrease towards the apex. Basal sporangia-bearing branches perpendicular to the rachises, subsequent slightly ascending. In mature living plants, sprorophores twisted, projecting the sporangia-bearing branches outwards of the trophophore (Figure 3). *Sporangia* 11-110+ in plants with bushy fertile portion of sporophores, arranged by pairs, opposite, mostly paripinnate. Sporangia 0.8-1.5 mm in diameter. Sporophores longer at maturity than the entire length of the trophophores, with a stalk mostly shorter and rarely exceeding the trophophore. *Spores* 37–39 µm in longest diameter. 2C-value 20.18 pg; diploid.

Etymology:-- The name refers to characteristic rounded pinnae of this species.

Distribution:-- *Botrychium rotundum* is distributed across central and eastern Europe from France to Georgia, and from Scandinavia and the western part of Russia to Italy. There are a few records from the north of Siberia to central China, but we suspect that our specimens are hybrids.

Ecology:-- *Botrychium rotundum* grows from 110-2,818 m on neutral to basic soils, in dry meadows (Meso- to Xero-Bromion, Seslerion) and steppes, often on exposed slopes in rocky areas, sometimes with sparse vegetation cover, scree slopes, or stabilized calcareous rock slides. It is found associated with patches of *Alnus viridis*, *Pinus mugo*, *Juniperus communis*, and herbaceous plants such as *Thymus* spp., *Festuca* spp., *Globularia* spp., and *Hieracium* spp., with often a thin moss layer below the vegetation.

Sporangia open late May or early June at low elevations (below 500 meters) to mid-August at high elevations (above 2500 meters).

Additional specimens examined:-- Austria, VORALBERG, Klösterle, Albona, 1976 m, 12.8.2016, 47,128364, 10,182397, MG-lun63-1-AT; 1983 m, 12.8.2016, 47,127426, 10,180462, *MG-lun64-1-AT, MG-sim4-1-AT*; Zürs, 1720 m, 15.7.2014, 47,1534199105423, 10,1655507087707, *000148310* (NEU)

China, GANSU, Mt. Chagangliang, 2445 m, 12.8.2017, 33,53744399, 104,307415, *VM-lun55-1-CN*, *VM-lun55-2-CN*; INNER-MONGOLIA, 0, NA m, 08.7.1929, 40.364162, 106.472031, *01592701* (PE)

Croatia, PRIMORJE-GORSKI-KOTAR, In front of the mountain house, 1418 m, 02.7.2018, 45,429426, 14,621975, *VM-lun69-2-HR*

Estonia, LÄÄNE, Matsalu Nature reserve, peninsula pulse, 0 m, 5.7.1984, 58,784271362686, 23,5029963909154, *0209479* (MW)

France, HAUTES-PYRÉNÉES, Tourmalet pass, 2050 m, 7.8.2014, 42,909769, 0,1552221, *000148324* (NEU)

Georgia, RACHA-LECHKHUMI-AND-KVEMO-SVANETI, hiking trail to Mount Shkmeri, 1921,7 m, 03.7.2016, 42,5101440120488, 43,4080070257186, *VM-lun4-2-GE*, *VM-lun4-7-GE*

Germany, MECKLENBURG-VORPOMMERN, Buhlitz, 0 m, 16.6.2003, 54,452041, 13,384957, *DF4970* (NEU)

Italy, ABRUZZO, Rocca di Cambio Presso ex-miniera bauxite, 1639 m, 17.6.2017, 42,1736, 13,9224, *VF-lun60-2-IT*, *VF-lun60-3-IT*, *VF-lun60-4-IT*, *VF-lun60-5-IT*, *VF-lun60-9-IT*

Lithuania, VILNIUS, 360 m south of the bridge over the Merkys River. On the east side of the road. On the southern slope of a small hill, 196 m, 14.6.2017, 54,444481, 25,507406, *SJ-lun57-2-LT*, *SJ-lun57-3-LT*

Macedonia, SKOPJE, Side of the path to Solunska glava, above Alpine pine forest, 2011,8 m, 20.7.2018, 41,6775519680231, 21,4034769963473, *VM_lun77_1_MK*, *VM_lun77_4_MK*

Romania, PRAHOVA, between Cabana Valea Dorului and Varful cu Dor, 1830 m, 13.8.2016, 45,35514, 25,48783, *VM-lun20-4-RO*

Russia, KALÚGA, Polotnyanyy Zavod , 144 m, 17.8.2006, 54,7167, 35,9833, *0209500* (MW); KURSK, hillside with chalk outcrops, near the quarry, 156 m, 17.6.2001, 51,428964, 34,292429, *0209589* (MW); MORDOVIA, 1 km north of the village Russkie Dubrovki (Русские Дубровки), 235 m, 26.6.1990, 54,534189, 45,921497, *0209609* (MW); MORDOVIA, north of the village Gart, slope of the valley of the left bank of the river, 198 m, 15.6.1993, 54,151414, 45,637858, *0209614* (MW); REPUBLIC-OF-SAKHA, Floodplain R. Horongki, 0 m, 8.8.1958, 66,799399, 123,279983, *0003996* (MW) ; VORONEZH, near the farm Divnogorie (Дивногорье), 110 m, 6.6.1991, 50,967308, 39,292227, *0209594* (MW)

Switzerland, GRAUBÜNDEN, Aroser Weisshorn, 2620 m, 16.7.2014, 46,7894384165468, 9,63854491620804, *000100760* (NEU); SCHWYZ, Rigi, 1750 m, 28.6.2014, 47,897, 8,485, *000100747* (NEU); TICINO, Acquarossa, Ponto Valentino, 650 m, 20.5.2017, 46,4840296, 8,9408234, *AM-lun41-1-CH*, *AM-lun41-1-CH*, *AM-lun41-2-CH*, *AM-lun41-4-CH*; Mt Generoso, 1620 m, 22.6.2014, 45,9296548356533, 9,02030646800995, *000100768* (NEU); VALAIS, Sur le chemin pédestre montant au rocher en partant du parking à l'entrée de Fully, 544 m, 13.5.2017, 46,12347, 7,07589, *VM-lun1-2-CH; VM-lun1-3-CH; VM-lun1-4-CH* (paratypes); 540 meters, 2016, *VM-lun1-5-CH, VM-lun1-6-CH* (topotypes); Up to the Matterhorn glacier trail on the trail toward the Hornlichhutte, 2818 m, 22.8.2020, 45,9883634, 7,6949371, *VM-lun99-1-CH*, *VM-lun99-2-CH*

Notes:-- *Botrychium rotundum* was referred to as Species 3 in Mossion *et al.* (in prep.) and as Lunaria 3 in Dauphin *et al.* (2017) and Maccagni *et al.* (2017).It is sister species to *B. tunux*, which differs by having asymmetrical pinnae with an obviously extended lower side, lunate to lunulate, and often cleft; the junction between the sporophore and the trophophore below the proximal pinna pair; and a bulky sporophore projecting sporangia-bearing branches towards the trophophore so that the sporangia are facing the trophophore.

There are numerous specimens that we interpret as hybrids or introgressed specimens between *B. rotundum* and other species. Introgressed hybrids with *B. neolunaria* identified by allozyme markers in North America (*DF4970, DF8699, DF12039*) tend to have intermediate morphology with symmetrical to asymmetrical fan-shaped and cleft pinnae, with pinnae slightly ascending and then overlapping the rachises. Sporophores are longer (up to 82 mm), with sporophore stalks longer than fertile portion of sporophores. Trophophores tend to bear more pinna pairs (5-6). An additional number of specimens are here considered to represent possibly hybrids between *B. rotundum* and *B. neolunaria* (*VM-lun55-1-CN, VM-lun55-2-CN, 01592701 (PE), VF-lun60-3-IT, VF-lun60-4-IT*), *B. alaskense* (*0003996*) and an unidentified species (*0209500*) (Figure 12). These plants mostly show intermediate pinna morphology and sporophore traits. For example, *VM-lun55-1-CN* that exhibits a ratio between sporophore stalk and fertile portion of the sporophore reminiscent of *B. neolunaria*, and trophophore morphology closer to *B. rotundum*. One extremely robust plant has an overall monstrous morphology (*0209500*), with deeply incised pinnae, junction between the sporophore and trophophore very high on the trophophore, and a double sporophore. Some specimens are slightly more robust plants with very dense fertile sporophore portions, and often bear additional sporophores, but overall fit the morphology *B. rotundum*. However, they have average spore sizes of 43-45 µm, approaching the range of allotetraploid species in the genus, suggesting introgression with other species. Such putative hybrids were found clustered in two locations, one in Gansu province, China (*VM-lun55-1-CN; VM-lun55-2-CN*), and one in Italy (*VF-lun60-3-IT, VF-lun60-4-IT*), or as scattered individuals at various sites in North America, Inner Mongolia (China), and Russia.

A few individuals show unconventional appearance and pinna arrangement, such as *MG-lun63-1-AT* which has the appearance of a broccoli, and *VF-lun60-5-IT* which has broadly overlapping pinnae and excessively short sporophore stalk (Figure 12).

Botrychium tunux Stensvold & Farrar in Stensvold *et al.* (2002)

This species was described from Alaska based on evidence from allozyme studies and morphology (Stensvold et al. 2002). It is characterized by its short stature, short common stalk, frequently stalked, ovate trophophores, asymmetrical pinnae with their basiscopic side expanded, entire margins commonly cleft by shallow incisions with rounded sinuses, and sporophore stalks seldom exceeding the height of the trophophore (Stensvold et al. 2002). The fertile portions of the sporophores have thickened rachises and branches (Stensvold *et al.* 2002; V. Mossion, pers. obs.), and the sporangia face the trophophore in mature living plants (Stensvold *et al.* 2002 Figure 3.b.; V. Mossion, pers. obs.). Originally described from the Yakutat area in Alaska, *B. tunux* was later found to be more widespread in North America, from Alberta and Montana south to the Rocky Mountains (Nevada) and even in Scandinavia (Norway) (Stensvold 2008). More recently, it has been recorded in Europe (Italy, Montenegro, Switzerland), based on a broad molecular survey of European specimens of the *B. lunaria* group (Dauphin et al. 2017; Maccagni et al. 2017). With even further sampling, Mossion *et al.* (in prep.) have now also documented this species from Central Asia (Kyrgysztan).

Examined specimens:-- **Kyrgysztan**, NARYN, On the way to Chatry-Kul, in the pasture on the side of the road, 3223 m, 05.7.2017, 40,7971470337361, 75,1173890288919, *VM-lun46-1-KG*, *VM-lun46-2-KG*, *VM-lun46-3-KG*, *VM-lun46-5-KG*; BATKEN, Ak-Suu peak, 3800 m, 08.8.2019, 39,612543, 70,104812, *PS-lun96-1-KG*

Norway, SKIPHALSEN, 19 m, 18.8.2001, 64,19722222, 10,0925, *DF6219*

Switzerland, GRAUBÜNDEN, Aroser Weisshorn, 2620 m, 16.7.2014, 46,7894384165468, 9,63854491620804, *000100760* (NEU); TICINO, Pizzo di Claro, 2360 m, 21.7.2016, 46,30156, 9,06815, *AM-lun33-1-CH*; VALAIS, Zermatt, 2600 m, 20.6.2014, 45,9926459978971, 7,75480806827545, *000100748* (NEU)

Botrychium yaaxudakeit Stensvold & Farrar in Stensvold *et al.* (2002)

This species was described from Alaska based on allozyme markers and on morphological evidence. Presence of fixed heterozygosity and an average spore size (45 µm in the longest diameter) obviously larger than *B. lunaria* (36 µm in the longest diameter) suggested that *B. yaaxudakeit* is a tetraploid. Stensvold *et al.* (2002) identified *B. neolunaria* as one of the diploid parents which was later confirmed by plastid data to be the maternal donor (Dauphin et al. 2014). The identity of the second ancestral diploid remains unconfirmed, even though it has been proposed to be a Eurasian genotype of '*B. lunaria*'. Morphologically, *B. yaaxudakeit* is characterized by its tall stature, short common stalk, and more ovate trophophore than *B. lunaria*. It has strongly overlapped pinnae that also overlap the rachises and its basal pinnae are often short stalked with a span of over 180° with strongly recurved basiscopic inner margins (Stensvold et al. 2002). It remains the only polyploid species known from the *B. lunaria* group, which contrasts with the high frequency of polyploidization found in other groups of *Botrychium* (Wagner and Lord 1956; Wagner 1993; Williams and Waller 2012; Dauphin et al. 2018). *Botrychium yaaxudakeit* is only known from northwestern North America.

In the present study, we documented a single plant from Georgia (VM-lun6-14-GE, Mossion *et al.* in prep.) that had exceptionally large spores (average 43 µm) compared to other Eurasian plants of the *B. lunaria* group (average 36-38 µm). Morphologically, it is a small plant with long common stalk, a stalked trophophore, and overlapping pinnae with straight basiscopic inner margins. The sporophore stalk is very long and the fertile portion very short, as sometimes observed in *B. onondagense*. Pending further research, we are unable to name this plant. Nevertheless, we note that there is the possibility that a polyploid taxon of the *B. lunaria* group may occur in Eurasia.

Also, we have on several occasions found individuals of *Botrychium* that morphologically resemble *B. yaaxudakeit* in being very large with highly branched sporophores and visibly stalked basal pinnae. A selection of these were studied by flow cytometry and were found to be diploid (Mossion *et al.* in prep.). Some have malformed spores and could be considered as F1 hybrids between two species of the *B. lunaria* group. However, well-formed spores were also observed and those had a size (43-44 µm in average, ranging from 40-48 µm) approaching the spore size range of *B. yaaxudakeit*. We discuss this discrepancy between spore size and genome size in more detail under *B. orientale*.

Further undescribed species

The following two species were recovered as well-supported clades in the phylogenetic analysis of Mossion *et al.* (in prep.). However, we only have a limited number of samples available from each of them, and there is considerable morphological variability between the samples. We are thus at present unable to properly characterize them and refrain from a formal description. Nevertheless, we here provide some notes on the to call attention to these forms and to encourage further research.

Botrychium species 4

We know this species primarily from a series of specimens collected by B. Dauphin and L. Stenberg in Sweden in 2012. They were recovered as a well-supported clade (species 1) sister to the remainder of the *B. lunaria* group in previous studies (Dauphin et al. 2014, 2017), and in our study based on the same individuals and partly on re-sequenced DNA material (species 4 in Mossion *et al.* in prep). The genome size we estimated is a little smaller than *B. lunaria* with a range from 17.75 to 19.62 pg (*000147810* (NEU), 2C-value= 17.75 pg; *000147822* (NEU), 2C-value= 18.46; *000147805* (NEU), 2C-value= 18.63; *000147837* (NEU), 2C-value= 18.69; *000147845* (NEU), 2C-value= 19.62). The herbarium material available to us is composed a few immature and fragmentary (Figure 13) specimens and we thus refrained from a formal description. We notice, however, that these specimens all have broad lunate, entire to deeply incised pinnae reminiscent of *B. boreale*, with in most case regularly crenulate outer margins. Also, sporangia seem to be set on tiny petioles and not embedded in the branch rachises. This last criterion is promising as a morphological trait to recognize the species but must be verified on mature plants. Thus, more extensive collections are needed before the species can be properly characterized.

Botrychium species 8

This species is known from about 10 specimens from Europe (Great Britain, Austria, Slovakia, Romania, Lithuania, and Georgia). It was recovered by Mossion *et al.* (in prep.) as a well-resolved monophyletic clade nested between *B. onondagense* and a group of clades composed of *B. lunaria* (including *B. nordicum*), *B. orientale* and the *B. neolunaria* group. Morphologically, however, each of the specimens available to us is distinct and we are unable to define the species on morphological grounds (Figure 14). We suspect that this species may have evolved from possibly repeated hybridization between *B. lunaria* and *B. onondagense*.

Acknowledgments

For their help organising and carrying out field work, we are in debt to Alexandru Coltoiu and the Bucegi Natural Park, Calimani National Park administration, Claudia Danau and the Retezat National Park, Clément Duckert, Cun-Feng Zhao, Daniel Ston, Kenan Ćatović, Kuban Uulu Zholdoshbek and the administration of Sarychat-Ertash State Nature Reserve, Luca Lässer, Marine Ramirez, Quentin Dubois, Sabin Neatu and the Ceahlau National Park, Sandra Grünig, Sergei Aleksandrov and the Central Balkan National Park, Shamil Shetekauri, Siqi Liang, Tolkha Shetekauri and the team of Tbilisi Botanical Garden, Xianchun Zhang, and Zikriyo. We are also thankful to Ahmed Ouhammou (MARK), Alessio Maccagni, Alexey P. Seregin (MW), Andreas Beiser, Dauphin Benjamin, Donald R. Farrar (ISC), Filippo Prosser, Hedwig Meindl, Isabelle Chanaron (AIX), Jason Grant, Markus Grabher, Natalia Gamova, Robin Walls, Sigitas Juzėnas, Siri Birkeland, Teddy Dolstra, Vincenzo Ferri, Wittmann Helmunt, Xian-Yun Mu, and Xianchun Zhang (PE) for providing specimens. We thank Marcus Lehnert for correcting the Latin names, Amandine Pillonel for spore slide preparation, spore measurements, and preparation of the spore pictures, and Elke Kessler for conducting the flow cytometry analyses.

Literature chapter 3

Bubani P (1902) Flora Pyrenaea per Ordines Naturales gradatim digesta, Ulricus Hoeplius

Butters FK, Abbe EC (1953) A floristic study of Cook county, Northeastern Minnesota. Rhodora 55:94

Clarke JM, House HD (1923) Report of the state botanist for 1921. N Y State Mus Bull

Clausen RT (1938) A monograph of the Ophioglossaceae. Mem Torrey Bot Club 19:1–177

Clute WN (1905) A check list of the North American fernworts. Fern Bull 13:109–123

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 42:620–639. https://doi.org/10.1600/036364417X696483

Dauphin B, Grant J, Mráz P (2016) Ploidy level and genome size variation in the homosporous ferns *Botrychium* s.l. (Ophioglossaceae). Plant Syst Evol 302:575–584. https://doi.org/10.1007/s00606-016-1285-7

Dauphin B, Grant JR, Farrar DR (2020) Outcrossing mating system of the early-divergent moonwort fern (*Botrychium lunaria*, Ophioglossaceae) revealed in the european Alps. Int J Plant Sci 181:926–936. https://doi.org/10.1086/710491

Dauphin B, Grant JR, Farrar DR, Rothfels CJ (2018) Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. Mol Phylogenet Evol 120:342–353. https://doi.org/10.1016/j.ympev.2017.11.025

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. https://doi.org/10.3732/ajb.1300154

Dolezel J, Bartos J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot 95:99–110. https://doi.org/10.1093/aob/mci005

Farrar DR (2011) Systematics and taxonomy of genus *Botrychium*. https://www.herbarium.iastate.edu/moonwort-botrychium-systematics

Farrar DR, Gilman AV, Moran RC (2017) Ophioglossales. Online edition. R.F.C. Naczi, J. R. Abbott, and Collaborators, New manual of vascular plants of northeastern United States and adjacent Canada, NYBG Press. New York

Farrar DR, Johnson-Groh C (1991) A new prairie moonwort (*Botrychium* subgenus *Botrychium*) from northwestern Minnesota. Am Fern J 81:1–6

Fuchs L (1542) De historia stirpium commentarii insignes

Gilman AV, Farrar DR, Zika PF (2015) *Botrychium michiganense* sp. nov. (Ophioglossaceae), a new North American moonwort. J Bot Res Inst Tex 9:295–309

Gray SF (1821) A natural arrangement of British plants, according to their relations to each other. Baldwin, Cradock and Joy, Paternoster-row, London

Hauk WD, Haufler CH (1999) Isozyme variability among cryptic species of *Botrychium* subgenus *Botrychium* (Ophioglossaceae). Am J Bot 86:614–633. https://doi.org/10.2307/2656570

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. https://doi.org/10.1600/036364412X635377

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

Jarvis CE (2016) Dataset: Clifford Herbarium. Natural History Museum Data Portal (data.nhm.ac.uk)

Jonsell B, Jarvis CE (1994) Lectotypifications of Lìnnaean names for Flora Nordica Vol. 1 (Lycopodiaceae — Papaveraceae). Nord J Bot 14:145–164. https://doi.org/10.1111/j.1756-1051.1994.tb00581.x

Kato M, Sahashi N (1977) The Systematic position of *Botrychium lanuginosum*. Acta Phytotaxon Geobot 28:143–151. https://doi.org/10.18942/bunruichiri.KJ00001078240

Lellinger DB (2002) A modern multilingual glossary for taxonomic pteridology. Amerian Fern Society, Inc., Washington, D.C.

Linnaeus C (1753) Species Plantarum

Lyon HL (1905) A new genus of Ophioglossaceae. Bot Gaz 40:455–458

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. Alp Bot 127:185–196. https://doi.org/10.1007/s00035-017-0188-8

Masamune (1931). J Soc Trop Agric 3:

Milde J (1869) *Botrychium* Monographia. Zool-Bot Ges Wien Verhandlungen 19:55–190

Mossion V, Dauphin B, Grant J, et al (2021) Transcriptome-wide SNPs for *Botrychium lunaria* ferns enable fine-grained analysis of ploidy and population structure. Mol Ecol Resour 22(1), 254-271. https://doi.org/10.1111/1755-0998.13478

Mossion V, Koenen, E., Grant, J., Croll, D., Farrar, D. R., Kessler, M. (in prep.) Global diversification of the common moonwort ferns (*Botrychium lunaria* group, Ophioglossaceae) was mainly driven by Pleistocene climatic shifts.

Popovich SJ, Farrar DR, Gilman AV (2020) *Botrychium furculatum* (Ophioglossaceae), a new moonwort species from the Rocky Mountains of North America. Am Fern J 110:165–182. https://doi.org/10.1640/0002-8444-110.4.165

PPG I (2016) A community-derived classification for extant lycophytes and ferns. J Syst Evol 54:563–603. https://doi.org/10.1111/jse.12229

Rasband WS (1997) ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA

Reichstein T (1981) Hybrids in European Aspleniaceae (Pteridophyta): significance, recognition, genome analysis, and fertility; checklist of species and hybrids; description of some new hybrids and cytology of several already known hybrids. Bot Helvetica 89–139

Rieseberg LH, Brouillet L (1994) Are many plant species paraphyletic? TAXON 43:21–32. https://doi.org/10.2307/1223457

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089

Shinohara W, Nakato N, Yatabe-Kakugawa Y, et al (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. https://doi.org/10.1600/036364413X670232

Škoda B (1997) Taxonomic comments on the "Flora of NorthAmerica north of Mexico", vol. 2, with some new nomenclatural combinations for Pteridophyta. Preslia 68:341–359

Stensvold MC (2008) A taxonomic and phylogeographic study of the *Botrychium lunaria* complex. PhD thesis, 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. https://doi.org/10.1007/s12228-016-9455-z

Stensvold MC, Farrar DR, Johnson-Groh C (2002) Two new species of Moonworts (*Botrychium* subg. *Botrychium*) from Alaska. Am Fern J 92:150–160. https://doi.org/10.1640/0002-8444(2002)092[0150:TNSOMB]2.0.CO;2

Swartz O (1801) Filicum, ordine systematico redactarum. J Für Bot 1800:4–120

Syring J, Farrell K, Businský R, et al (2007) Widespread Genealogical Nonmonophyly in Species of Pinus Subgenus Strobus. Syst Biol 56:163–181. https://doi.org/10.1080/10635150701258787

Underwood LM (1903) An index to the described species of *Botrychium*. Bull Torrey Bot Club 30:47–48

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. https://doi.org/10.1093/aob/mcr267

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

Wagner WH, Grant JR (2002) *Botrychium alaskense*, a new moonwort from the interior of Alaska. Am Fern J 92:164–170

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

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. https://doi.org/10.2307/2482592

Wagner WH, Wagner FS (1983) Genus communities as a systematic Tool in the study of New World *Botrychium* (Ophioglossaceae). TAXON 32:51–63. https://doi.org/10.2307/1219850

Williams EW, Farrar DR, Henson D (2016) Cryptic speciation in allotetraploids: Lessons from the *Botrychium matricariifolium* complex. Am J Bot 103:740–753. https://doi.org/10.3732/ajb.1500281

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. https://doi.org/10.1086/664711

Xianchun Z, Lui Q, Sahashi N (2013) Ophioglossacae. In: Flora of China, Beijing: Science Press&St. Louis: Missouri Botanical Garden Press. Zhengyi, W., P. H. Raven & H. Deyuan, St. Louis

Yatskievych G (2003) Pteridophytes (Ferns). In: eLS. American Cancer Society

Zhang L, Fan X-P, Petchsri S, et al (2020) Evolutionary relationships of the ancient fern lineage the adder's tongues (Ophioglossaceae) with description of *Sahashia* gen. nov. Cladistics 36:380–393. https://doi.org/10.1111/cla.12408

Competing interests: none

Author contributions: VM, JG, and MK conceived this study. Field work was carried out by VM and XZ. VM, DF and MK took the taxonomic decisions. VM, MK, AM, DF, and XZ studied live specimens and provided the habitat documentation. VM realized the measurements. MK and VM wrote the first manuscript draft and all co-authors provided comments on drafts of the manuscript.

Figures

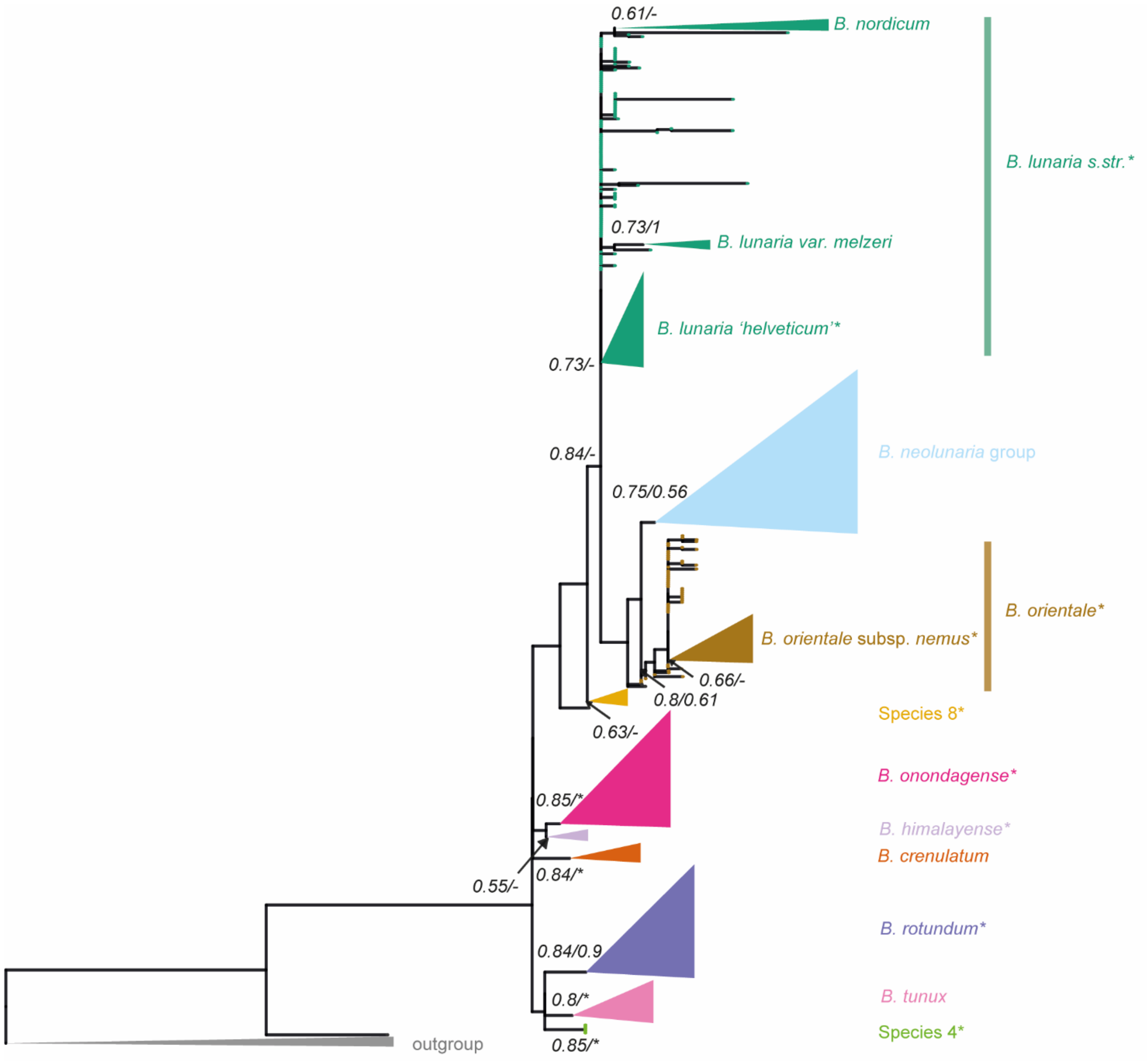


FIGURE 1: Simplified phylogram modified from Mossion *et al.* (in prep.). Maximum likelihood phylogenetic tree with some collapse clade and subclades represented by triangles. Triangle height relates the clade sizes, and the triangle width the maximum branch length. The colors refer to the clades as used in Mossion *et al.* (in prep). Clades and subclades statistical supports are given by bootstrap supports and posterior probabilities on the left and right, respectively (*i.e.*, * fully supported, - not applicable).

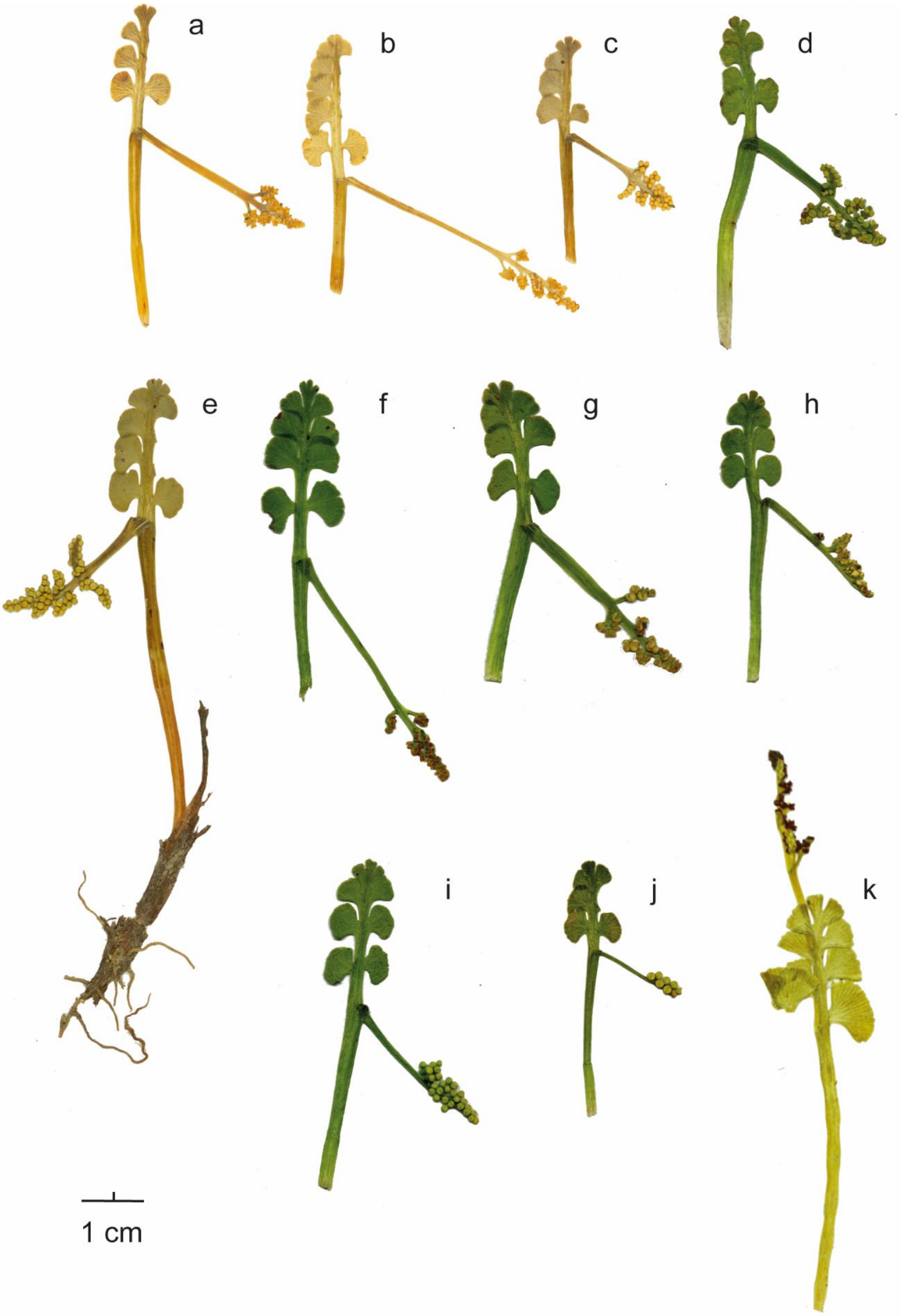


FIGURE 2: ***Botrychium himalayense***. (**a-i**) pressed type specimens from China, YUNNAN, Shangri-La, slope on the roadside, 4334 meters, 6 August 2017, *VM-lun54-3-CN*, *VM-lun54-4-CN*, *VM-lun54-5-CN*, *VM-lun54-2-CN*, *VM-lun54-1-CN*, *VM-lun54-7-CN*, *VM-lun54-8-CN*, *VM-lun54-9-CN*, *VM-lun54-10-CN* (holotypes, paratypes PE!). (**a**) holotype, (**b-i**) paratypes, (**j-k**) pressed specimens from across the distributional range in China. (**j**) *VM-lun52-1-CN*. (**k**) 01314569 (PE).

FIGURES 3: Live plants and habitats of the new species (**A-D**) *B. himalayense*, *locus typicus* China. (**B, E**) *B. rotundum*, Switzerland, Saas-Fee, *M. Kessler 15402* (Z/ZT). (**F**) *B. rotundum*, Macedonia, Skopje, Solunska glava, *VM-lun77-1-MK.* (**C-G**) *B. orientale*. (**C**) Bosnia Herzegovina, Herzegovina-Neretva, Veliki Vran. (**G**) Bulgaria, *locus typicus*. (**B**, **E** by M. Kessler; **A**, **C-D**, **F-G** by V. Mossion).

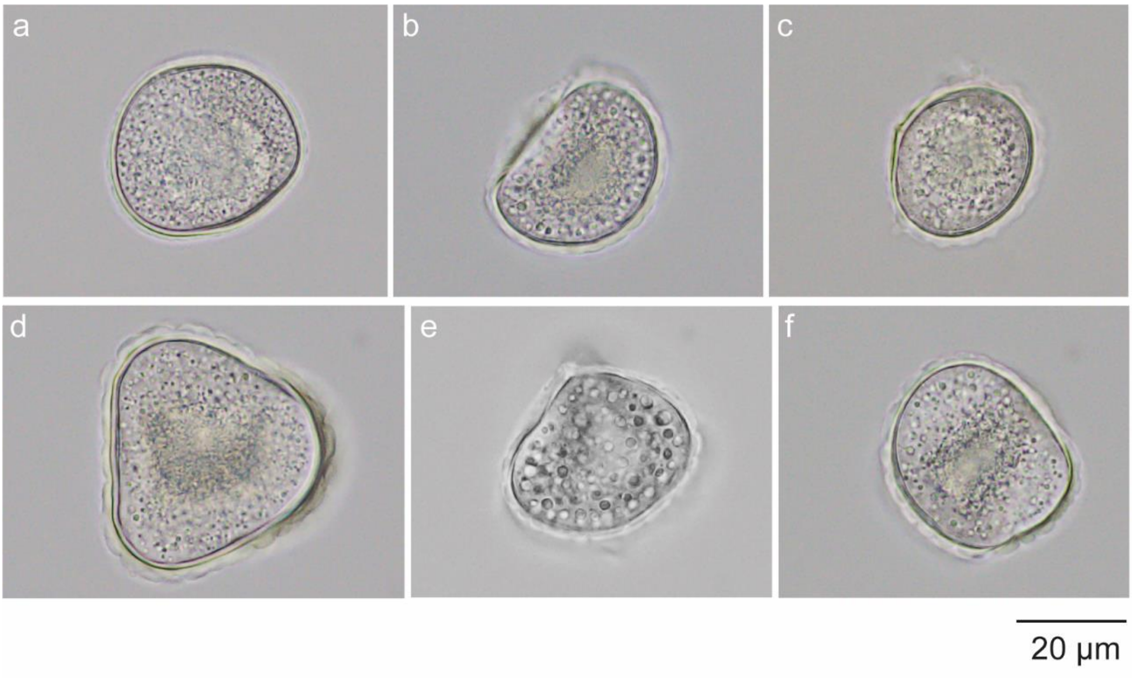


FIGURE 4: Photography of spores with a magnification of 400x. (**a-f**) *B. himalayense* (*VM-lun54-12-CN*), *B. lunaria* 'helveticum' (*VM-lun94-2-CH*), *B. onondagense* (*DF16506*), hybrids *B. onondagense* x indet (*VM-lun22-2-RO*), *B. orientale* (*VM-lun81-1-BG*), *B. rotundum* (*VM-lun1-1-CH*) (**a-d**, **f** by A. Pillonel; **e** by V. Mossion).

FIGURE 5: Habitats and live plants of *B. lunaria* from Switzerland. (**A-B**) BERN, Petit Chasseral; VALAIS, Fiddelberg. (**C-H**) Specimen from GRAUBÜNDEN, Nufenen and BERN, Chasseral (**C**, **E-G** by A. Maccagni; **D**, **H** by V. Mossion).

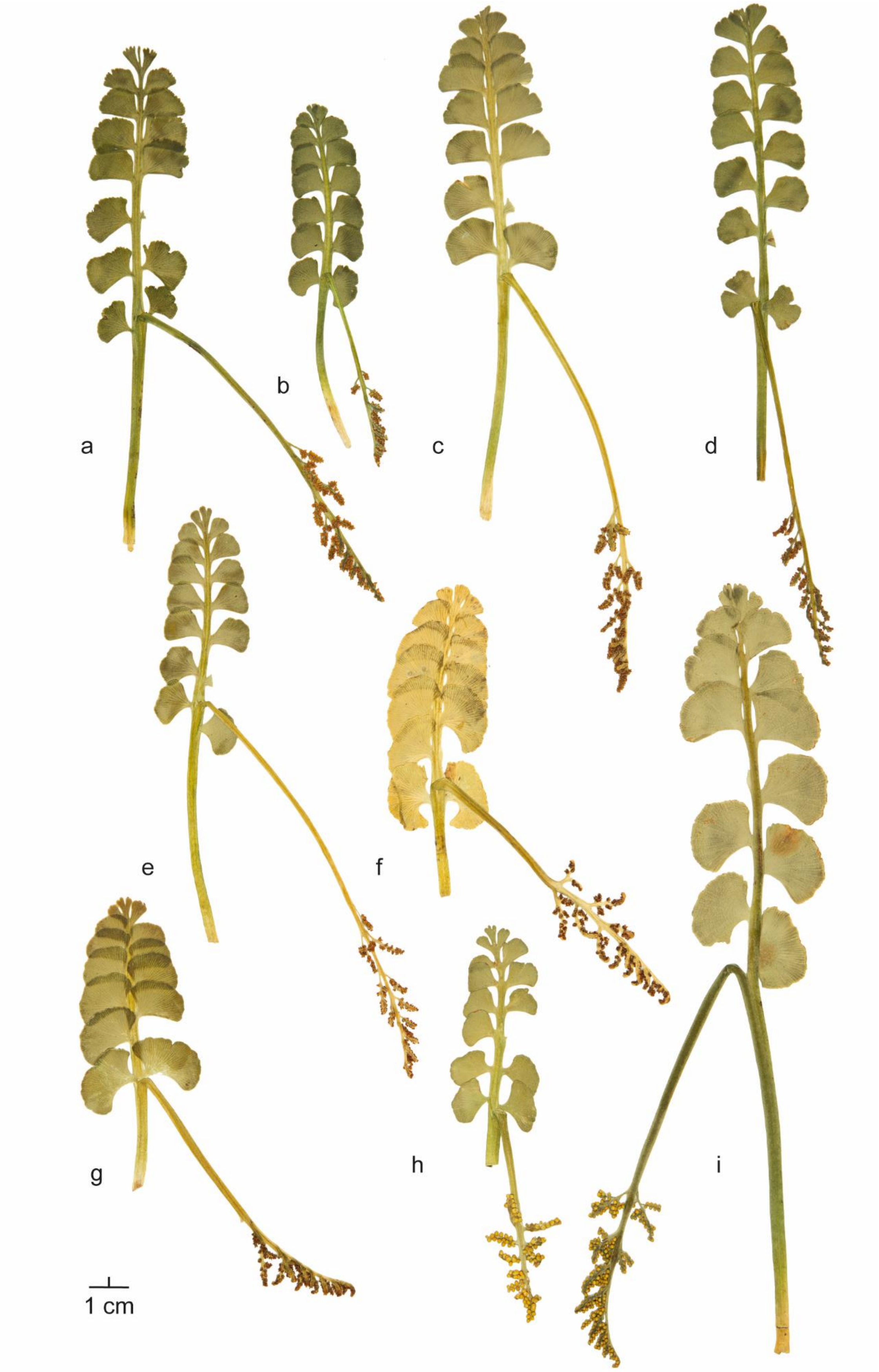


FIGURE 6: ***B. lunaria*** **'helveticum'**. (**a-i**) pressed specimens from Switzerland, *VM-lun93-4-CH*, *VM-lun 93-3-CH*, *VM-lun93-2-CH*, *VM-lun93-1-CH*, *VM-lun93-5-CH*, *VM-lun86-38-CH*, *VM-lun86-17-CH*, *VM-lun89-3-CH*, *VM-lun91-5-CH*.

FIGURE 7: ***B. onondagense***. (**a-n**) pressed specimens from across the distributional range including North America, Tajikistan, Austria and Romania. (**a-d**, **f-m**) DF5057, *DF16506*, *DF20315*, *VM-lun26-1-RO*, *DF20300*, *VM-lun26-2-RO*, *VM-lun21-1-RO*, *MH-lun39-1-AT*, *VM-lun23-1-RO*, *VM-lun23-4-RO*, *VM-lun21-2-RO*, *VM-lun22-1-RO*; (**e**) *0800417* (MW). (**n**) Potential hybrid with spore size enters the range of allotetraploid species, *VM-lun22-2-RO*.

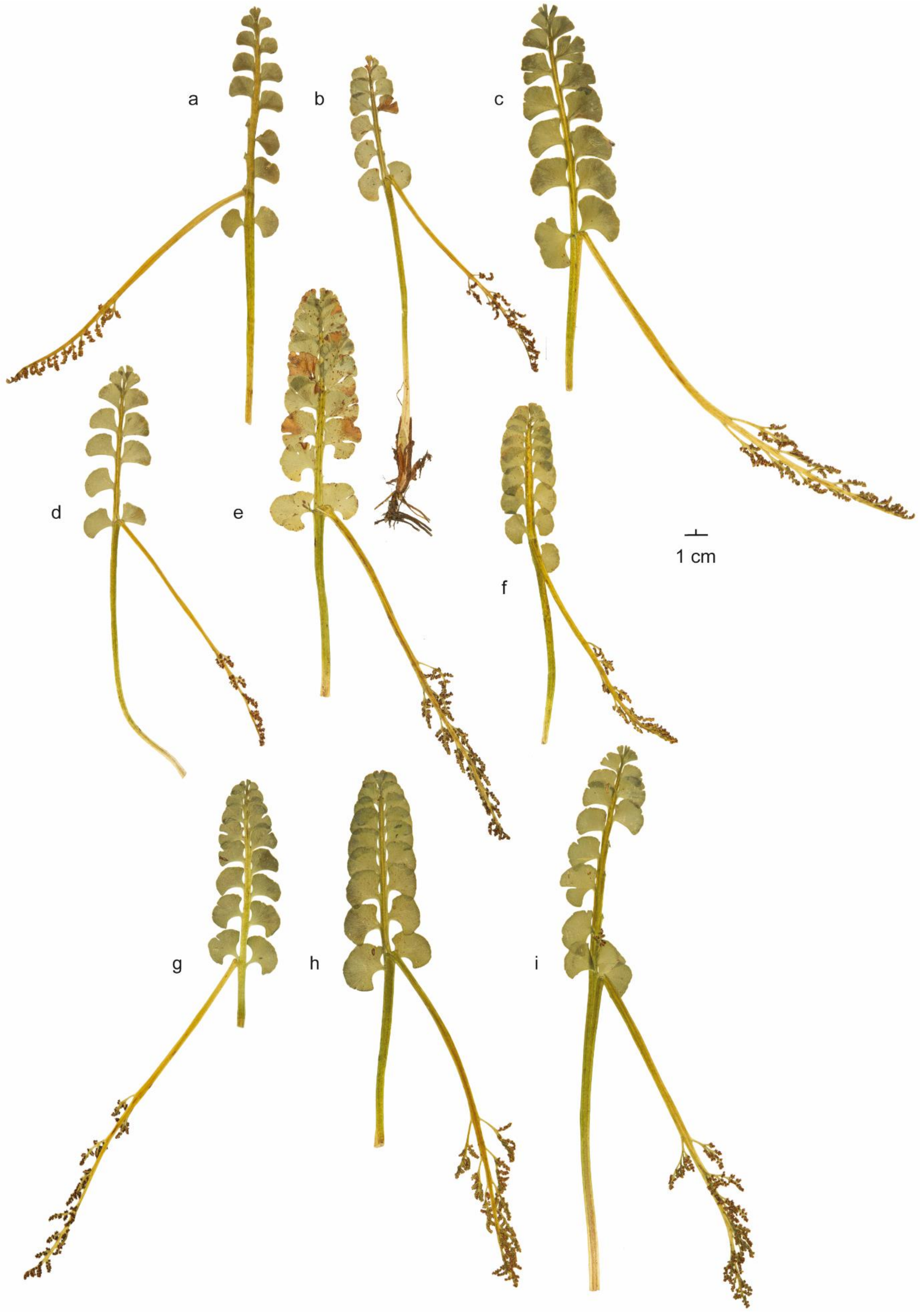


FIGURE 8: ***B. orientale*** (**a-i**) pressed type specimens from **Bulgaria**, BLAGOEVGRAD, on the side of hiking trail to Xaha Pametaka, 2134 meters, 29 July 2018, 41.77291397, 23.41994103, *VM-lun81-2-BG*, *VM-lun81-1-BG*, *VM-lun81-13-BG*, *VM-lun81-7-BG*, *VM-lun81-8-BG*, *VM-lun81-4-BG*, *VM-lun81-10-BG*, *VM-lun81-5-BG*, *VM-lun81-3-BG* (holotype, paratypes). (a-h) paratypes. (i) holotype.

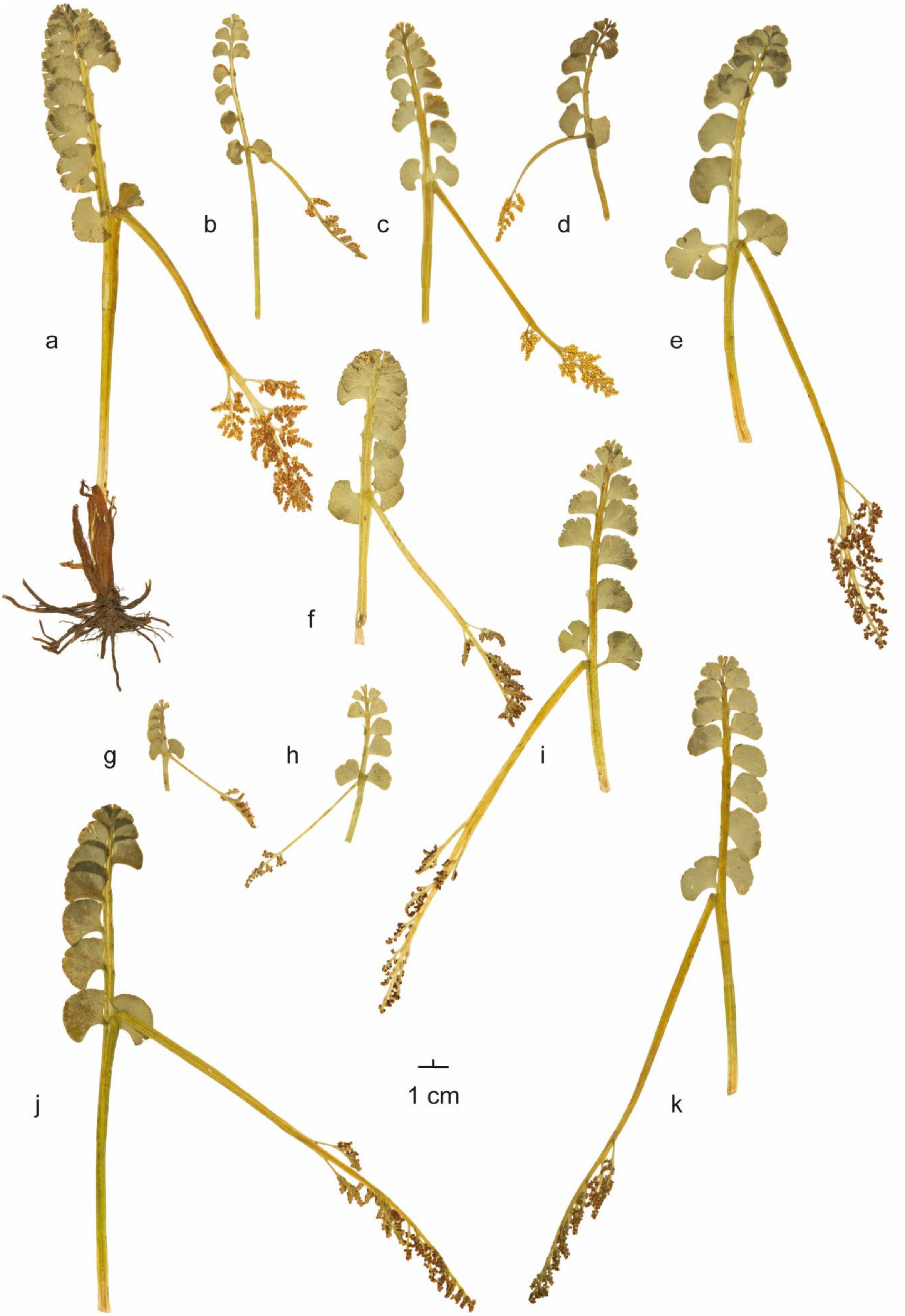


FIGURE 9: ***B. orientale*** (**a-i**) pressed specimens from across the distributional range including Bosnia Herzegovina, Bulgaria, Slovakia, Greece, Kyrgyzstan, *VM-lun74-1-BA*, *VM-lun75-4-BA*, *VM-lun72-8-BA*, *VM-lun14-2-UA*, *VM-lun73-2-BA*, *VM-lun9-1-SK*, *VM-lun50-2-KG*, *VM-lun83-7-BG*, *VM-lun79-7-GR*, *VM-lun82-3-BG*, *VM-lun74-2-BA*.

FIGURE 10: ***B. orientale*** **subsp.** ***nemus*** (**a-e, g-k**) pressed specimens from across the distributional range including xx, *0003998 (MW)*, *WS20140903001* (NEU), *VM-lun45-3-KG*, *VM-lun45-2-KG*, *VM-lun48-3-KG*, *RG3*, *MD12060*, *WS20140803001* (NEU), *0209622* (MW). (**f**) pressed type specimen from **Mongolia**, KHENTII, Norovlin, Eren-Daba ridge about 20 km NW from the village Norovlin (Норовлин), on the road to the river crossing the river Onon, 15 August 1989, Diriymaa S., Grubov V. I., Gubanov I. A., *0168258* (MW).

FIGURE 11: ***B. rotundum***. (**a-d**) pressed type specimens from **Switzerland**, VALAIS, Fully, on the trail side going from the parking at Fully entrance to the 'Rocher des Follatères', 544 meters, 13 May 2017, *VM-lun1-1-CH*, *VM-lun1-2-CH*, *VM-lun1-3-CH*, *VM-lun1-4-CH*. (**e-r**) pressed specimens from across the distributional range including Switzerland, Georgia, Romania, Macedonia, Lithuania, and Austria, *VM-lun99-1-CH*, *VM-lun4-7-GE*, *VM-lun4-2-GE*, *VM-lun20-4-RO*, *000100741* (NEU) [*AM-14-CHMOL-001-001-15]*, *VM-lun77-3-MK*, *VM-lun77-2-MK*, *VM-lun77-1-MK*, *SJ-lun57-3-LT*, *VM-lun77-6-MK*, *VM-lun77-5-MK*, *VM-lun77-4-MK*, *SJ-lun57-2-LT*, *MG-lun64-1*-AT.

FIGURE 12: ***B. rotundum***. (**a-i**) pressed specimens from across the distributional range including Italy, China, Austria, North America and Russia, genetically identified as introgressed hybrids (**f**) or suspected of being hybrids, *VF-lun60-4-IT*, *VM-lun55-2-CN*, *VM-lun55-1-CN*, VF-lun60-5-IT, *MG-lun64-1*-AT, *DF4970* (NEU), *VF-lun60-3-IT*, *0209500*, *0003996* (MW).

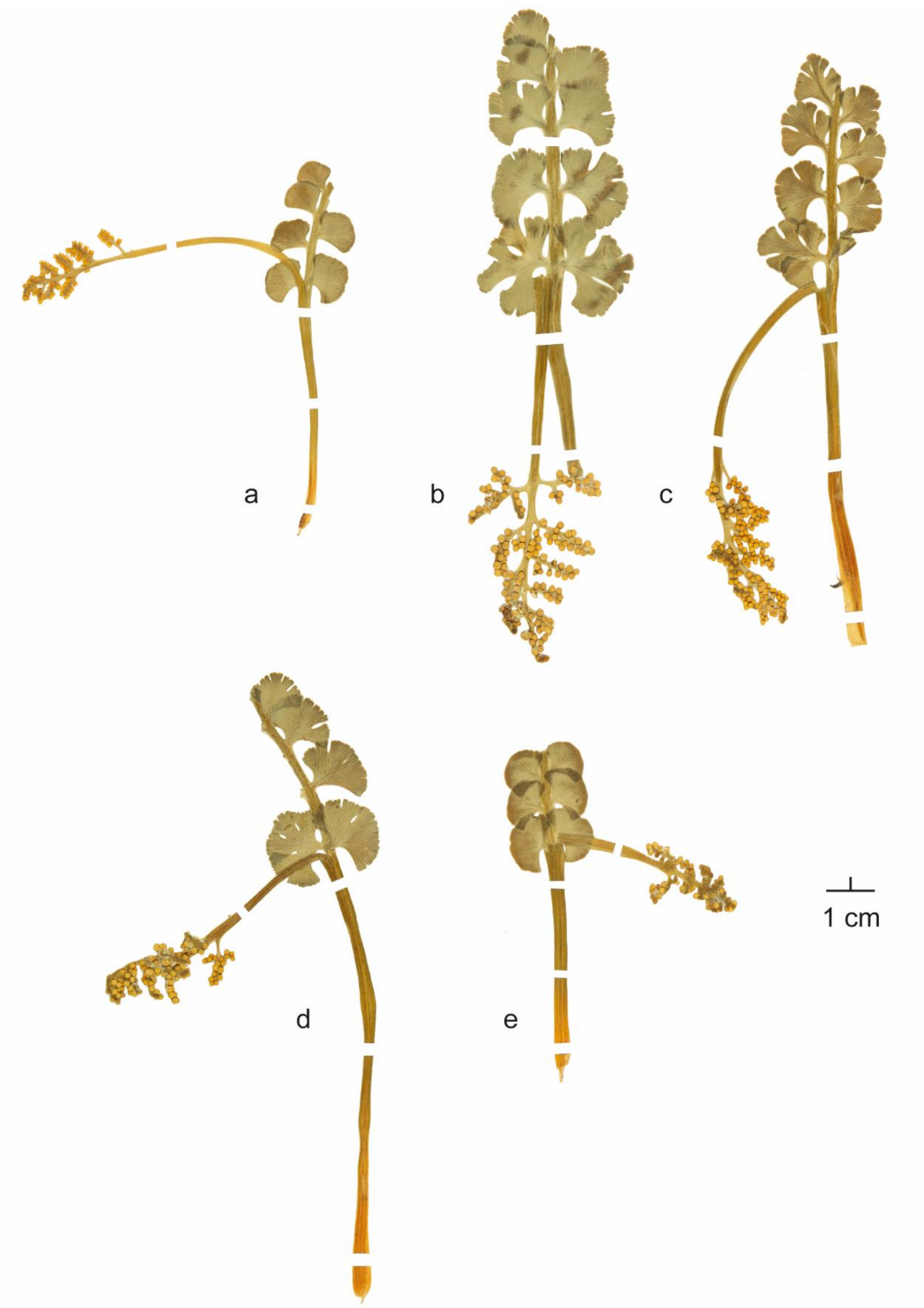


FIGURE 13: Species 4. (**a-f**) pressed specimens from Sweden, *000147856,000147858, 000147855, 000147857, 000147859* (NEU).

FIGURE 14: Species 8. (**a-h**) pressed specimens from Lithuania, Slovakia, Georgia, Romania, England, France. (**a-e, g-h)** *Pl1*, *Zel1*, *VM-lun-9-2-SK*, *VM-lun2-2-GE*, *VM-lun17-1-RO*, *VM-lun5-1-GE*, *3320*; **f** 014786 (AIX). (**i**) Living *VM-lun9-2-SK* (**c**).

Tables

TABLE 2: Summary of measured quantitative and qualitative traits and of selected ratios

Characters		Definition	*B. lunaria* '*helveticum*'	*B. orientale*	*B. onondagense*	*B. rotundum*	*B. himalayense*
Leaf length		From the bottom to the top of the longest part of the rachises	(30)80-178(210)	(13) 46-258	(17)41-125	(32) 39-123 (197)	62-100
Common stalk length		From the base to the junction[1]	19-81(94)	(4)8-126	4-50	(7) 12-73	18-51
Trophophore	Pinna pairs	number	(4)5-7(9)	(2)4-7(9)	(2)3-5(8)	(2)3-4(5)	3-5
	Total length	From the junction to the apex of the trophophore	34-78(100)	(8)20-98	12-55	12-40 (45)	15-28
	Blade length	From the bottom side of the first pinna petiole to the apex	36-77(95)	(8)21-103	11-58	12-46	12-22
	Stalk length	From the junction to the bottom side of the first pinna petiole	Absent, rarely present 1-4	Mostly absent, present 3-7(34)	(0.2)0.6-4(8)	Mostly absent, present 1-2	2-5(7)
	Base width	Distance between the basal pinnae lower sides[2]	5-23	3.3-30	3.20	4-12	5-8
	1st internode length	From the top of the petioles of the first pinna pair to the bottom of the petioles of the second pinna pairs	4-14(23)	3-18	2-10	2-10	2-5

	2^{sd} internode length	Same than above between second and third pinna pairs	3-12	2-13 (17)	1-5	1-6	1.5-2
Basal pinna pairs	Length	Distance between the middle of the pinna outer margins	6-31	(7)10-43	11-38	9-26	9-13
	Width	Distance between the upper and lower sides of the outer margins	10-20	(3)7-22	6-13	5-16	5-8
Sporophore	Total length	From the junction to the apices	37-113	(16)30-173	(12)22-100	22-77	15-47
	Stalk length	From the junction to the bottom of the lowest branch petioles	21-65	17-88	6-46	8-41	20-29
	Fertile portion length	From bottom of the lowest branch petioles to the apices	16-48	(9)13-82(92)	(6)10-60	9-35	11-21
	Branche pairs	Number	(5)6-8(10)	(4)6-9	(2)4-6	(4)5-6(7)	3-4
	Lower branch length	From the base of the petiole to the apex	3-13 (16)	3-26	2-17	(1) 3-11 (19)	4-7
	1^{st} branch internode length	Distance the top of the petioles of the first branch pair to the bottom of the petioles of the second branch pairs	3-11	2-19	2-13	4-18	2-5

	2^{sd} branch internode length	Same than above between second and third branch pairs	2-8	2-12	1-9	2-13	1-3
Basal pinnae span[2]			100-180	90-180	70-180	(75) 90-160	90-140
Middle pinnae angle[2]		Angle between the middle pinnae lower sides and the rachises	40-95	37-120	40-90	36-80 (124)	50-90
Sporangia diameter			0.9-1.2	0.6-1 (1.15)	0.7-1.2	0.8-1.5	0.9-1.5
Spore size		Spore longest diameter average and range	35-36	37 (36-39)	36 (33-39)	38 (37-39)	35 (34-36)
Color			Green ro bright yellow-green plants	Light yellow green pinnae and rachises	Bud green pinnae to light green rachises	Bright green pinnae, yellowish rachises	Bright green to yellow-green plants
SafeJunction		Below, in between or above the first pinna pairs	Between or at the base of the petioles of the first pinna pairs, rarely below	Above or between the petioles of the first pinna pairs, rarely below (subsp. *nemus*)	Below the petioles of the first pinna pairs	Mostly between or at the base of the petioles of the first pinna pairs	Conspicuously below the first pinna pairs
Trophophore	outline		Triangulate to oblong	Narrowly ovate to oblong	Ovate to narrowly oblong	Ovate to triangulate	Narrowly ovate

Pinnae	texture		Fleshy	Fleshy but not as thick as *B. rotundum*	Thick to thin	Very thick	Moderately thin, rachises thinner texture than pinnae
	outline	Flabellate (Fan-shaped) – orbiculate - cuneate – rhombic – lunate – symmetrical - asymmetrical	Cuneate to lunate, asymmetrical with extended lower side	Lunate, asymmetrical	Cuneate to flabellate	Flabellate, appear rounded, symmetrical	Flabellate to lunate, appear rounded, asymmetrical with oversized upper part
	orientation		Upward	Upward but less than *B. lunaria*	Downward to slightly ascending	downward	Downward to slightly ascending
	disposition	Opposite - alternate	Opposite to slightly alternate	alternate, basal pinnae sometimes opposite	Mostly alternate, basal somewhat opposite	opposite	opposite
Sporophore fertile	Outline		Narrowly lancaete	Broadly triangulate to ovate	lanceate	lanceolate	Narrowly ovate to triangulate
	Branches disposition		opposite	Opposite or alternate	Mostly opposite	Opposite, with sometimes alternate basal branches	opposite
Sporangia disposition			By pairs, opposite, paripinnate	By pair, opposite, mostly paripinnate	By pairs, opposite, imparipinnate	By pairs, opposite, mostly paripinnate	By pairs, opposite, lateral branches imparipinnate, apices paripinnate
Ratio	pinna internodes 1^{st} – 2^{sd}		0.8-2(2.4)	0.9-2.3	(0.9)1.1-3.3	1.1-2.9	1-3

sporophore total length – trophophore total length	1-2.2	1.2-2.2(3.6)	0.9-2.3	(0.9)1.3-2.4	1.4-2
sporophore stalk length – trophophore total length	0.6-1(1.4)	0.5-1.5(2)	0.5-1.5	(0.2)0.4-1.2(1.4)	0.9-1.2
sporophore fertile length – sporophore stalk length	0.5-1.1	0.4-1.2(1.4)	0.3-1.6	(0.6)0.9-2.4(3.4)	0.5-0.8
branch internodes 1st – 2sd	(0.8)1-2(2.4)	0.9-1.9	1-1.9(2.5)	(0.9)1.3-2.6	1.7-3
Sporophore fertile percentage of sporophore total length	30-50	(13)30-60	25-60	(40)50-70(80)	35-45

Note: All distances and lengths are given in millimeters, except spore size in micrometers, and angles in degrees. [1] Point where the sporophore and the trophophore separate.

General discussion

In the **first chapter**, we established a high-quality transcriptome for *Botrychium lunaria* ferns from a single reference individual and sequenced 11 additional transcriptomes of individual from two populations in Switzerland. The transcriptome-wide SNP dataset revealed a clear population structure, and substantial variation in heterozygosity among individuals. These findings suggest that sporophytic selfing or outcrossing are predominant processes which contrast with the general assumption of gametophytic selfing as dominant reproductive mode in *Botrychium* species. Those results highlight the power of transcriptome-wide datasets as a tool for population genetics of ferns and could allow to resolve fine grained population structure where other markers have lacked in resolution.

In the **second chapter**, we assessed the worldwide diversity within the *B. lunaria* group using plastid markers and recovered ten well-defined monophyletic clades, corresponding to eleven species-clade level of which three were not previously identified. We estimated species divergence times to 1.8-2.6 million years which coincides with the Pleistocene climatic oscillations. This finding suggests that the *B. lunaria* diversification was influenced by complex population dynamics through migration and extinction cycles. Quaternary glaciation cycles were also shown to have driven diversification through adaptation to specific soils and habitats. Our climatic niche analysis revealed large overlaps between species with some notable differences. We also found indications for soil preferences for some of the clades. Ecological factors most likely played a role in the group's diversification, however further studies are needed to ascertain the importance of such factors in the speciation process.

In the **third chapter**, we conducted morphological studies of herbarium vouchers and live plants, and we combined the results with the molecular and ecological evidence provided in the second chapter to describe three new species from Eurasia. Two species exhibited a large distribution range with *B. rotundum* being widespread throughout Europe and a few occurrences in Asia. *B. orientale* is occurring only in the easternmost part of Europe and across Asia. Finally, *B. himalayense* is restricted to Himalayas. The diagnostic traits turned out to be the pinnae outline and orientation, presence, or absence of a trophophore stalk, sporophore ratios, as well as live plant textures and habits. We also recognized the need to epitypify *B. lunaria* s.str. Indeed, the poor condition of the Linnaeus type challenges the assignment to a species-clade level on morphological grounds. The absence of locality information prevents any further collection for morphological or molecular analysis.

At least two large clades of the *B. lunaria* group (clade 1, *B. lunaria* s.str.; clade 10, *B. neolunaria* group) have remained poorly resolved despite of a dataset including many populations from all over the geographical distribution range. The issue likely stems from the limitations of the plastid markers used. Chloroplasts are transmitted only by the maternal parent in ferns; hence plastid phylogenies are blind to hybridization or introgression events. As discussed in the second chapter, hybrids between *B. neolunaria* and a 'B. lunaria' species were found to be widespread and common in mixed populations (Stensvold 2008; Stensvold and Farrar 2017) and a wide range of genome sizes was found among diploids of the *B. neolunaria* group. These findings suggest that reticulate evolution is involved in that clade diversification. The use of multiple, biparentally inherited markers were able to disentangle complex reticulate evolution of North America *Dryopteris* (Sessa et al. 2012). Species tree based on multiple independent genes often come with gene tree conflicting topologies. Methods were developed to integrate topological incongruences to assess the importance of hybridization and introgression events (Meng and Kubatko 2009; Woodhams et al. 2016). The minimum number of genes required for taking into account gene tree incongruences was estimated to be higher as branch lengths decrease (Shekhar et al. 2018). In our phylogenies, branch lengths of poorly resolved clades are extremely short, then extending the number of nuclear markers available for *Botrychium* would be a requirement to unravel the full picture of the clade's evolutionary history. To date, the nuclear markers available for the genus *Botrychium* are limited to a small number (Dauphin et al. 2018). However, single-copy nuclear markers were successfully develop using transcriptome mining (Rothfels et al. 2013), and this method could be applied to develop additional markers using our newly establish transcriptomics resources.

As defined in chapter 2, the differences recovered between species climatic niches were weak, but soil pH preference patterns emerged from the limited measurements available. We suggest that the detection of ecological differences was a matter of scale. Alternatively, would it be conceivable that ecological factors are involved in the diversification of symbionts rather than *B. lunaria* itself? In natural populations, *Botrychium* gametophytes do not survive after two or three cell divisions without a symbiosis with arbuscular mycorrhizal fungi (AMF) (Campbell 1911). Interestingly, it was shown that the geographical distribution of AMF diversity was mostly explained by temperature and soil pH. AMF studies of *B. crenulatum* and *B. lunaria* have shown that *Glomus* taxa were the dominant AMF (Winther and Friedman 2007; Sandoz et al. 2020). Only half of the *Glomus* phylotypes were found inside *B. crenulatum* gametophytes compared to the phylotype recovered from the sporophyte roots. A possibility would be that ecological niches of *B. lunaria* species are congruent with ecological niches of the AMF taxa present in the gametophytes. Unfortunately, the AMF diversity associated with *B. lunaria s. str*. is only known for sporophytes. Further studies looking at gametophyte AMF diversity of *B. tunux* or *B. rotundum* (only known from neutral and basic soils) and of *B. lunaria s.str*. (exclusively known from acidic soils) could enable to test this hypothesis. Similarly, one could hypothesize that species with a large distribution range and found on very diverse soil types, like *B. onondagense*, would have gametophytes associated with a large spectrum of potential symbionts. Testing those hypotheses might also shed light onto the implications of AMF in the group diversification and to better define the geographic distribution range of *B. lunaria* species even though establishing distribution patterns may remain complex as showed for the Montane unicorns (Hurlbert, 1990).

Traditionally, genetic lineages are translated into taxonomical units with a morphological background, which are most meaningful to identify species based on unambiguous morphological characters. However, in the *B. lunaria* group diagnostic traits are nearly universally very subtle (Farrar 2011; Farrar et al. 2017) and intermediate specimens (*i.e.* potential hybrids) are common in mixed populations (pers. observation, Wagner *et al.* 1985). Our morphological analyses were based on specimen measurements without proper morphometric analyses. More rigorous quantitative methods (Baur et al. 2014) would refine our understanding of the morphological differences between species. Regardless of such considerations, the proper identification of diversity has an important impact on conservation efforts. For example, among the new species we described, one species showed a narrow geographical distribution being in only a few localities in the Himalayas. As the pressure of human activity is constantly increasing, geographically restricted species are typically the most threatened aspect of biodiversity. Taxonomic work hence has meaning beyond the contributions to systematics and counteracts hopefully the silent disappearance of not yet well-identified species.

Perspectives

As discussed above many open questions remain regarding how *B. lunaria* has evolved and diversified. A rewarding question to tackle in future studies would be to assess the validity of the diagnostic traits we used to identify species on morphological grounds. In other words, what part of the phenotypic trait variation observed between species is genetically fixed among species pools? A well-tested approach to address such a question are common garden experiments. These consist of cultivating plants of distinct populations and species under the same environmental conditions. The morphology of the *Botrychium* sporophyte varies with abiotic factors and the plant's age (Stensvold 2008), so this would require growing plants from spores to control for age and maternal effects. To my knowledge, sporophytes of *Botrychium* species have not yet been successfully cultivated from spores. During my thesis, I developed based on Whittier's culture experiments (Whittier 1972, 1981, 2006) a protocol to grow *Botrychium* gametophytes in sterile media supplemented with sugars. My preliminary results indicate high rates of germination and growth of gametophytes for both *B. lunaria* species and *B. lanceolatum*. A step further would be to infect early stage (2-3 cells) gametophytes with AMF of the *Glomus* genus, which was found to be frequently associated with several species of *Botrychium* (Winther and Friedman 2007; Sandoz et al. 2020). AMF are known carbon providers for gametophytes (Winther and Friedman 2007) but could also influence gametophyte growth rates. In tomato plants, growth hormone signaling within plant roots is required to induce the AMF symbiosis (Hanlon and Coenen 2011). The importance of AMF growth hormones for plants are also suspected (Gogala 1991). We can reasonably hypothesize that gametophytes need AMF to mature, which could explain the absence of reproductive organs in my gametophytes even after four years of culturing efforts. Based on population surveys, the generation of gametophytes is thought to be a years-long

(Johnson-Groh 1998). Culture experiments which conserved living gametophytes of *Sceptridium dissectum* took up to 34 months (Whittier 2015). However, the actual gametophyte life span of *Botrychium* species remains poorly understood. The experimental setup required to grow *Botrychium* sporophytes from spores seems highly challenging. However, in case of success, it would allow, not only to perform common garden experiments, but also to study the reproductive mode of *Botrychium* under different conditions. Furthermore, this would allow to test for the strength of post-zygotic reproductive barriers using crossing experiments and possibly provide material for whole genome sequencing efforts. Gametophytes are haploid reducing drastically the genome size and by this complexity during the assembly process.

This thesis project extended our understanding of the *B. lunaria* group diversity and provided the first evidence of the implication of ecological factors in the group diversification, but also showed the limitations of the commonly used genetic markers. The genomic tools developed during the project could tackle these limitations and support future studies on reticulate evolution and population dynamics of this group and related ones.

Literature discussion

Baur H, Kranz-Baltensperger Y, Cruaud A, et al (2014) Morphometric analysis and taxonomic revision of Anisopteromalus Ruschka (Hymenoptera: Chalcidoidea: Pteromalidae) – an integrative approach. Syst Entomol 39:691–709. https://doi.org/10.1111/syen.12081

Campbell DH (1911) The eusporangiate: the comparative morphology of the Ophioglossaceae and the Marattiaceae., Carnegie Institution of Washington. Washington, D.C., USA.

Dauphin B, Grant JR, Farrar DR, Rothfels CJ (2018) Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. Mol Phylogenet Evol 120:342–353. https://doi.org/10.1016/j.ympev.2017.11.025

Farrar DR (2011) Systematics and taxonomy of genus *Botrychium*. https://www.herbarium.iastate.edu/moonwort-botrychium-systematics

Farrar DR, Gilman AV, Moran RC (2017) Ophioglossales. Online edition. R.F.C. Naczi, J. R. Abbott, and Collaborators, New manual of vascular plants of northeastern United States and adjacent Canada, NYBG Press. New York

Gogala N (1991) Regulation of mycorrhizal infection by hormonal factors produced by hosts and fungi. Experientia 47:331–340. https://doi.org/10.1007/BF01972074

Hanlon MT, Coenen C (2011) Genetic evidence for auxin involvement in arbuscular mycorrhiza initiation. New Phytol 189:701–709. https://doi.org/10.1111/j.1469-8137.2010.03567.x

Hurlbert SH (1990) Spatial distribution of the montane unicorn. Oikos 58:257–271. https://doi.org/10.2307/3545216

Johnson-Groh CL (1998) Population demographics, underground ecology and phenology of *Botrychium mormo*. in N. Berlin, P. Miller, J. Borovansky, U. S. Seal, and O. Byers, eds. Population and habitat viability assessment (PHVA) for the goblin fern (*Botrychium mormo*) Final Report. Conservation Biology Specialist Group, Apple Valley, MN., 103-108.

Meng C, Kubatko LS (2009) Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: a model. Theor Popul Biol 75:35–45. https://doi.org/10.1016/j.tpb.2008.10.004

Rothfels CJ, Larsson A, Li F-W, et al (2013) Transcriptome-Mining for Single-Copy Nuclear Markers in Ferns. PLOS ONE 8:e76957. https://doi.org/10.1371/journal.pone.0076957

Rothfels CJ, Li F-W, Sigel EM, et al (2015) The evolutionary history of ferns inferred from 25 low-copy nuclear genes. Am J Bot 102:1089–1107. https://doi.org/10.3732/ajb.1500089

Sandoz FA, Bindschedler S, Dauphin B, et al (2020) Biotic and abiotic factors shape arbuscular mycorrhizal fungal communities associated with the roots of the widespread fern *Botrychium lunaria* (Ophioglossaceae). Environ Microbiol Rep 12:342–354. https://doi.org/10.1111/1758-2229.12840

Sessa EB, Zimmer EA, Givnish TJ (2012) Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). BMC Evol Biol 12:104. https://doi.org/10.1186/1471-2148-12-104

Shekhar S, Roch S, Mirarab S (2018) Species Tree Estimation Using ASTRAL: How Many Genes Are Enough? IEEE/ACM Trans Comput Biol Bioinform 15:1738–1747. https://doi.org/10.1109/TCBB.2017.2757930

Stensvold MC (2008) A taxonomic and phylogeographic study of the *Botrychium lunaria* complex. PhD thesis, 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. https://doi.org/10.1007/s12228-016-9455-z

Whittier DP (1972) Gametophytes of *Botrychium dissectum* as grown in sterile culture. Bot Gaz 133:336–339

Whittier DP (2006) Red light inhibition of spore germination in *Ophioglossum crotalophoroides*. Can J Bot 84:1156–1158. https://doi.org/10.1139/b06-063

Whittier DP (2015) Delayed Growth in Mycoheterotrophic Gametophytes of Seedless Vascular Plants. Am Fern J 105:1–10. https://doi.org/10.1640/0002-8444-105.1.1

Whittier P (1981) Spore germination and young gametophyte development of *Botrychium* and *Ophioglossum* in axenic culture. Am Fern J 71:13–19. https://doi.org/10.2307/1546671

Winther JL, Friedman WE (2007) Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). Am J Bot 94:1248–1255. https://doi.org/10.3732/ajb.94.7.1248

Woodhams MD, Lockhart PeterJ, Holland BR (2016) Simulating and Summarizing Sources of Gene Tree Incongruence. Genome Biol Evol 8:1299–1315. https://doi.org/10.1093/gbe/evw065

Vinciane **Mossion**

Faubourg de la Gare, 5, 2000 Neuchâtel, Switzerland
☎ (+41) 79-705-21-70 | ✉ vinciane.mossion@gmail.com | 🐦 @VincianeMossion | 🎓 Mossion

Experience

Ph.D. candidate, *Evolutionary genetics laboratory, Institute of Biology, University of Neuchâtel (UNINE)*, Switzerland *2016 - 2022*

Advisor: Prof. Dr. Daniel Croll

- Developed genomic resources for *Botrychium* ferns (Ophioglossaceae) using a transcriptomic approach.
- Organised and carried extensive field work in Central (Switzerland, Slovakia, Romania, Poland) & Eastern Europe (Croatia, Bosnia and Herzegovina, North Macedonia, Greece, Bulgaria, Georgia) and Central (Kyrgyzstan, Tadjikistan) & Eastern Asia (China). Work on herbarium collections (MW, PE, AIX, NEU, SARA, Z/ZT).
- Investigated the diversity within the *B. lunaria* group at the global-scale distribution using phylogenetic, climatic, and morphological analysis combined with geographic distribution patterns.
- Designed experimental culturing of subterranean achlorophyllous fern gametophytes.

Research assistant, *Earth and Life Institute Biodiversity (ELIB), catholic University of Louvain (UCL)*, Belgium *2013 - 2015*

Advisor: Prof. Dr. Renate Wesselingh

- Established a reference collections of *Rhinanthus* species (Orobanchaceae) for molecular phylogenetic and taxonomical studies.
- Designed and carried extensive sampling covering the taxonomic diversity of the genus *Rhinanthus* including field work in the Western Alps (France, Switzerland, Italy) and South-Eastern Europe (Croatia, Bosnia and Herzegovina, Montenegro, Albania, Greece, Bulgaria) and visits of herbaria (BEO, KW, KWS, L, U, WAG).
- Prepared a DNA library for ddRAD sequencing.

Technician of computerisation and digitisation for the Global Plant Initiative project (GPI), *National Museum of Natural History (MNHN) (P)*, France *2013*

- Characterised, computerised (Sonnerat database) and digitised (Herbscan) *typus* specimens of the general phanerogamy collection.

Scientific collaborator, *Natural History Museum of Aix-en-Provence (AIX)*, France *2012*

- Researched *typus* specimens in the lichen and phanerogam collections, digitised botanical collections and wrote a popular science paper (La garance voyage, 100:32-35, 2012). Co-organised a symposium: 'les herbiers publics et privés de la région PACA'.

Education

Ph.D. in Plant Biology & Evolution *Neuchâtel, Switzerland*

Ph.D. thesis: 'Unravelling complex evolutionary histories by a multidisciplinary approach: The case of the common moonwort ferns'. Advisor: Prof. Dr. Daniel Croll. University of Neuchâtel *2016 - 2022*

M. S. in Ecology, Evolution, Biodiversity and Biometry *Lyon, France*

M.S. thesis: 'Effects of environmental factors and phenotypic attributes on the roe deer's diet quality'. Advisors: Dr. Jean-Michel Gaillard & Dr. Sonia Saïd. University Claude Bernard Lyon1 *2009 - 2011*

B.S. in Organism and Population Biology *Limoges, France*

University of Limoges, Faculty of Science and Technology *2005 - 2009*

Publication

Published

Mossion V., Dauphin B., Grant J., Kessler M., Zemp N., Croll D. Transcriptome-wide SNPs for Botrychium lunaria ferns enable fine-grained analysis of ploidy and population structure. *Molecular Ecology Resources*, 22 (1), 254-271, https://doi.org/10.1111/1755-0998.13478 *2021*

Presentations

Young Systematists Forum *online*

Mossion V., Koenen E., Croll D., Farrar D., Maccagni A. & Kessler M. Flash Talk: Unravelling the global diversity of the common moonwort ferns (*Botrychium lunaria*, Ophioglossaceae) *2021*

Botanica Sudalpina *online*

Mossion V., Grant J., Farrar D., & Kessler M. Talk: The multiple identities of the common moonwort, *Botrychium lunaria* (L.) Sw. *2021*

Annual PhD student meeting *Neuchâtel, Switzerland*

Mossion V. & Croll D. Talk: reference transcriptome for *Botrychium lunaria* reveals insights into the biology and evolution of the Lunaria clade of ferns *2019*

The 19th International Botanical Congress *Shenzhen, China*

Mossion V., Dauphin B., Maccagni A. & Grant J. Poster: Worldwide systematics and phylogeny of the *Botrychium lunaria* complex (Ophioglossaceae) *2017*

Award & Grants

2021	**Flash talk award in Botany**, The Systematics Association
2020	**Mobility - DPOB UNINE**, Visit in Iowa State University, Department of Ecology, Evolution and Organismal Biology, USA - not used due to sanitary situation (CHF 1,500)
2018	**Subvention égalité des chances - UNINE**, DNA sequencing & field work in Balkans (CHF 4,885)
2018	**Fonds des donations - UNINE**, Field work in Russia (CHF 3,565)
2017	**Student Travel Grants of Swiss Systematics Society & the Swiss Botanical Society**, 19th International Botanical Congress in Shenzhen, China (CHF 700, CHF 640)
2017	**Subvention égalité des chances - UNINE**, Field work in Tajikistan and Kyrgyzstan (CHF 4,886)
2017	**Systematics Research Funds 2016/17**, Field work session in Nepal (£ 950)
2016	**Fonds des donations - UNINE**, Field work in Eastern China (CHF 3,360)

Teaching activities

Practicals *2013 - 2021*

- Teaching assistant for the introduction to botany & plant histology, introduction to laboratory 1st BSc class, excursions of botany 2sd BSc class at the Institute of Biology - University of Neuchâtel, Switzerland (2016 - 2021)
- Teaching assistant for plants biology, physiology and systematics, 1st & 2sd BSc class, and molecular ecology 1st MSc class at the Earth and Life Institute Biodiversity (ELIB), catholic University of Louvain, Belgium (2013 - 2015)

Skills

Laboratory methods

Molecular biology, **DNA/RNA** extractions & quantifications, **PCR**, ddRADseq library preparation
Plant culture, sterile **culture media** for fern's gametophytes and unicellular green algae
Microscopy, permanent slides preparation using Euparal and Hoyer's mounting media

Informatic tools

Bio-informatics, TRINITY pipeline, basic of STACK pipeline (GENECO's course 2015), Bowtie, BUSCO, GATK, Geneious (Sanger data preparation)
Phylogenetics, Orthofinder, Astral, BEAST, RAxML-ng, ModelTest-ng, BayesPhylogenies2, PartitionFinder2
Systems, script, Linux, Windows NT, SH scripts, PHP, **R**
Graphic design, Office, Illustrator, Lightroom, Microsoft Office Pack, Zotero, LaTeX

LANGUAGE

French, mother tongue. **English**, B2 speaking, C1 comprehension & writing

Certification

INFO-FLORA: **Certificat 200 ”Bellis”**

Neuchâtel, Switzerland

2017

Referees

- **Prof. Dr. Daniel Croll**. Laboratory of Evolutionary Genetics, University of Neuchâtel, Switzerland.
  daniel.croll@unine.ch, +41 32 718 23 30
- **Dr. Michael Kessler**. Department of Systematic and Evolutionary Botany, University of Zürich, Switzerland.
  michael.kessler@systbot.uzh.ch, +41 44 634 84 32
- **Prof. Donald R. Farrar**. Department of Ecology, Evolution and Organismal Biology, Iowa State University of Science and Technology, United-States of America.
  dfarrar@iastate.edu

Associative activities

- **Entomological Society of Neuchâtel** (SNE) since 2019, Attend conferences and excursions.
- **Farnfreunde der Schweiz** since 2022. Attend excursions.
- **Ornithological groups of Baulmes and surroundings** (GOBE) and of **Neuchâtel** since 2017, Training in song birds identification. Participate in excursions and bird monitoring.
- **Pro-Natura** since 2017, Volunteer helping in maintenance work of natural reserve in Neuchâtel canton
- **Systematics Association** since 2016, Attend conferences.
- **Swiss Systematics Society** (SSS) and **Swiss Botanical Society** (SBS) since 2017, Attend conferences and workshops.