LIÈGE université

THE QUATERNARY BIOGEOGRAPHIC HISTORY OF BRYOPHYTES: A WINDOW INTO THEIR ABILITY TO FACE GLOBAL CHANGE

Insights from European and Amazonian bryophyte populations

fnrs
LA LIBERTÉ DE CHERCHER

Alice Ledent

Belgium, June 2019

II

The Quaternary biogeographic history of bryophytes: a window into their ability to face global change

Alice Ledent

Department of Biology, Ecology and Evolution
University of Liege
June 2019

Alice Ledent

To Yuri and Ambre,
may my work inspire you to find
and pursue your own goals in life

Acknowledgements

VII

Abstract

Studying the influence of past climate changes on the distribution of species contributes to our understanding of the evolution of life on earth. Among past climate changes, the Quaternary period (from 2.4 Myrs to present time), characterized by high amplitude climatic oscillations, is considered one main determinant of current species distributions. Europe has long served as a model region to study the impact of past climate changes on extant biodiversity patterns. Its landscape is characterized by the presence of E-W-oriented mountain ranges, acting as effective barriers to migration for many organisms. Explicit historical scenarios for the post-glacial recolonization of Europe from distinct *refugia* have been discussed at length in the literature. In contrast, the impact of past climate changes on species distributions in tropical areas has been much less documented. In Amazonia, where the landscape is homogeneous without any apparent geographic barrier to migration, available fossil evidence describes range contractions and expansions of the evergreen rainforest during the Quaternary period.

Bryophytes are poïkilohydric and therefore appear as extremely sensitive to climate changes. Bryophytes disperse by means of spores or asexual diaspores, which are involved in frequent long-distance dispersal (LDD) events. These high dispersal capacities have cast doubts on the possibility to find signatures of historical events from analyses of the extant spatial patterns of genetic structure and diversity of their populations. In Amazonia in particular, recent ecological work suggests that dispersal does not show geographical structure across the area.

In the present thesis, we assembled and analyzed large molecular datasets at the level of the species range to determine how bryophytes responded to major Quaternary climate changes in environments characterized by different 'resistance' to migration and environmental heterogeneity, especially in Europe and Amazonia (ongoing study in the case of the latter). More specifically, the aims are to: (1) Test whether, due to the high dispersal capacities of bryophytes in general – and in particular in homogeneous environments without any apparent geographic barrier to migration –, the inverse isolation hypothesis – according to which any signal of isolation-by-distance (IBD) is erased beyond the limits of short-distance dispersal (SDD) by the intensity of LDD events – applies, erasing any historical signal in the extant spatial patterns of genetic structure and diversity of bryophyte populations; (2) Test the relevance of other differentiation mechanisms promoting speciation and, in particular isolation-by-environment (IBE), across a relatively homogeneous environment without any apparent geographic barrier to migration; (3) Infer the post-glacial history of bryophytes, in environments characterized by the presence (Europe) or the absence (Amazonia, ongoing study) of apparent geographic barriers to migration, from analyses of the extant spatial patterns of genetic structure and diversity of their populations.

The results strongly suggest that the LDD capacities of bryophytes did not homogenize the genetic structure of their populations, neither in an environment characterized by apparent geographic barriers to migration – such as the E-W-oriented mountain ranges in Europe –, nor in a much more homogeneous environment as in the Amazonian rainforest. In contradiction with the idea that the inverse isolation hypothesis applies in Amazonian bryophytes in particular, the IBD signal observed in 8 out of the 10 Amazonian bryophyte species consistently remained significant beyond the range of SDD, evidencing LDD limitations. This consistent persistence of the IBD signal contrasts with the result of a recent meta-analysis on IBD patterns in bryophytes and suggests that Amazonian bryophyte species experience more

dispersal limitations than species from other biomes. As a comparison, we showed that, within the same Amazonian environment, the spatial genetic structures observed in bryophytes are comparable to that of angiosperm species producing much larger seeds.

While a significant IBD signal characterizes the genetic structure of the vast majority of the Amazonian bryophyte species investigated here, our results are not consistent with the idea that isolation-by-resistance (IBR) and IBE contributed to the observed spatial patterns of genetic variation. Nevertheless, a low (0.059) but significant (P=0.004) F_{st} was found between sympatric specimens of the sibling *Syrrhopodon annotinus* and *S. simmondsii*, and their average kinship coefficients along a geographic gradient were consistently higher for conspecific comparisons than for interspecific comparisons, pointing to reproductive isolation between those two sympatric species characterized by different habitat requirements. Even if this single empirical result does not challenge the global idea that IBE does not prevail in extant patterns of genetic diversification in Amazonian bryophytes, it nonetheless contributes to growing evidence for genetic divergence observed along environmental gradients, suggesting that adaptation could play a more important role in shaping genetic patterns than previously thought.

Rejection of the hypothesis that high dispersal capacities of bryophytes erased any historical signal in the extant spatial patterns of genetic structure and diversity of their populations and, in particular rejection of the inverse isolation hypothesis, indicate that the data generated in the present thesis are suitable for demographic inference. We applied coalescence-based approaches to infer the post-glacial history of bryophyte populations from contrasting environments characterized by the presence (Europe) or the absence (Amazonia, ongoing study) of apparent geographic barriers to migration. In Europe, our analyses revealed that the post-glacial assembly of bryophytes likely involved a complex history. Extant European populations originated from multiple sources with a contribution from allochthonous migrants representing 90-100% in about half of the 15 investigated species, which demonstrates the importance of LDD for the post-glacial recolonization of Europe by bryophytes and is unparalleled in any previous phylogeograhic study on other organisms.

X

Résumé

L'étude de l'influence des changements climatiques passés sur la répartition des espèces contribue à notre compréhension de l'évolution de la vie sur terre. Parmi l'ensemble des changements climatiques passés, la période du Quaternaire (de 2,4 Ma à nos jours), caractérisée par des oscillations climatiques de forte amplitude, est considérée comme l'un des déterminants principaux de la répartition actuelle des espèces. L'Europe a longtemps servi de région modèle pour l'étude de l'impact des changements climatiques passés sur les patterns de biodiversité actuels. Son paysage est façonné par l'orientation O-E de ses chaines de montagnes agissant comme barrières à la migration pour de nombreux organismes. Des scénarios historiques explicites décrivant la recolonisation post-glaciaire en Europe depuis différentes régions refuges ont été longuement discutés dans la littérature. Par contre, l'impact des changements climatiques passés sur la répartition des espèces dans les régions tropicales a été beaucoup moins documenté. En Amazonie, où le paysage est homogène et vierge de barrières géographiques apparentes à la migration, les preuves fossiles disponibles décrivent une succession d'épisodes de contractions et d'expansions de la forêt à feuilles pérennes durant le Quaternaire.

Les bryophytes étant poïkilohydriques, elles apparaissent comme des organismes extrêmement sensibles aux changements climatiques. Les bryophytes dispersent par le biais de spores ou de diapsores asexuées, impliquées dans des événements de dispersion à longues distances (LDD) fréquents. Ces fortes capacités dispersives ont semé le doute quant à la possibilité de trouver des traces d'événements historiques dans l'analyse des patterns spatiaux actuels de structure et de diversité génétiques de leurs populations. En Amazonie, en particulier, une étude écologique récente suggère que la dispersion ne révèle pas de structure géographique au sein de la région.

Dans cette thèse, nous avons assemblé et analysé de larges sets de données moléculaires, à travers l'entièreté de l'aire de répartition des espèces, pour déterminer comment les bryophytes ont répondu aux changements climatiques majeurs du Quaternaire dans des environnements caractérisés par différentes « résistances » à la migration et différentes hétérogénéités environnementales, en particulier en Europe et en Amazonie (étude en cours dans le cas de cette dernière). Plus spécifiquement, les objectifs sont de : (1) Tester si, en raison des fortes capacités dispersives des bryophytes en général – et en particulier dans des environnements homogènes et vierges de barrières géographiques apparentes à la migration –, l'hypothèse d'isolement inverse – selon laquelle tout signal d'isolement par la distance (IBD) est effacé au-delà des limites de la dispersion à courtes distances (SDD) par l'intensité des événements de dispersion à longues distances (LDD) – s'applique, effaçant ainsi tout signal historique des patterns spatiaux actuels de structure et de diversité génétiques de leurs populations ; (2) Tester la pertinence d'autres mécanismes de différentiation promouvant la spéciation et, en particulier, celui de l'isolement par l'environnement (IBE), au sein d'un environnement relativement homogène et vierge de barrières géographiques apparentes à la migration ; (3) Inférer l'histoire post-glaciaire des bryophytes, dans des environnements caractérisés par la présence (Europe) ou l'absence (Amazonie, étude en cours) de barrières géographiques apparentes à la migration, par l'analyse des patterns spatiaux actuels de structure et de diversité génétiques de leurs populations.

Les résultats suggèrent fortement que les capacités de dispersion à longues distances (LDD) des bryophytes n'ont pas homogénéisé la structure génétique de leurs populations, ni dans un environnement caractérisé par la présence de barrières géographiques apparentes à la

migration – comme celles que composent les chaines de montagnes européennes orientées selon l'axe O-E –, ni dans un environnement beaucoup plus homogène comme celui de la forêt tropicale amazonienne. En contradiction avec l'idée selon laquelle l'hypothèse d'isolement inverse s'applique pour les bryophytes amazoniennes en particulier, le signal d'isolement par la distance (IBD) observé chez 8 des 10 espèces de bryophytes amazoniennes est constamment resté significatif au-delà de la gamme de dispersion à courtes distances (SDD), mettant ainsi en évidence des limitations de dispersion à longues distances (LDD). Cette constante persistance du signal d'isolement par la distance (IBD) contraste avec le résultat d'une méta-analyse récente traitant des patterns d'isolement par la distance (IBD) chez les bryophytes et suggère que les espèces de bryophytes amazoniennes rencontrent d'avantage de limitations dispersives que les espèces d'autres biomes. À fin de comparaison, nous avons démontré que, au sein du même environnement qu'est l'Amazonie, les structures génétiques spatiales observées chez les bryophytes sont comparables à celles d'espèces d'angiospermes produisant des graines beaucoup plus larges.

Tandis qu'un signal d'isolement par la distance (IBD) caractérise la vaste majorité des espèces de bryophytes amazoniennes étudiées ici, nos résultats ne sont pas compatibles avec l'idée que l'isolement par la « résistance » (IBR) et l'isolement par l'environnement (IBE) ont contribué aux patterns spatiaux de variation génétique observés. Cependant, une valeur de F_{st} faible (0,059) mais significative (P=0,004) a été trouvée entre les individus sympatriques de *Syrrhopodon annotinus* et *S. simmondsii*, et leurs coefficients « kinship » moyens calculés le long d'un gradient géographique étaient constamment plus élevés pour les comparaisons conspécifiques que pour les comparaisons interspécifiques, mettant en évidence de l'isolement reproducteur entre ces deux espèces sympatriques caractérisées par des exigences d'habitat différentes. Même si ce résultat empirique unique ne remet pas en question l'idée globale que l'isolement par l'environnement (IBE) ne prédomine pas dans les patterns de diversification génétique des bryophytes, il contribue néanmoins à accréditer les preuves croissantes de divergence génétique observées le long de gradients environnementaux, suggérant ainsi que l'adaptation pourrait jouer un rôle plus important qu'on ne l'imaginait dans le façonnement des patterns génétiques.

Le rejet de l'hypothèse selon laquelle les fortes capacités dispersives des bryophytes ont effacé tout signal historique des patterns spatiaux actuels de structure et de diversité génétiques de leurs populations et, en particulier, le rejet de l'hypothèse d'isolement inverse indiquent que les données générées dans cette présente thèse sont adaptées pour des études d'inférence démographique. Nous avons employé des approches basées sur la coalescence pour inférer l'histoire post-glaciaire des bryophytes dans des environnements contrastés caractérisés par la présence (Europe) ou l'absence (Amazonie, étude en cours) de barrières géographiques apparentes à la migration. En Europe, nos analyses ont révélé que l'assemblage post-glaciaire des bryophytes impliquait probablement une histoire complexe. Les populations européennes actuelles ont des origines multiples parmi lesquelles la contribution des migrants allochtones représentait 90-100% dans près de la moitié des 15 espèces étudiées, ce qui démontre l'importance de la dispersion à longues distances (LDD) dans la recolonisation post-glaciaire de l'Europe par les bryophytes et est sans parallèle avec tout autre étude phylogéographique portant sur d'autres organismes.

List of contributions

Papers/manuscripts

I. Ledent, A., Pereira, M., Overson, R.P., Laenen, B., Mardulyn, P., Gradstein, S.R., Zartman, C.E., & Vanderpoorten, A. Strong spatial genetic structure in Amazonian bryophytes raises concerns about their ability to overcome fragmentation. *Manuscript*.

II. Pereira, M.R.*, Ledent, A.*, Mardulyn, P., Zartman, C.E., & Vanderpoorten, A. (2019). Maintenance of genetic and morphological identity in two sibling *Syrrhopodon* species (Calymperaceae, Bryopsida) despite extensive introgression. *Journal of Systematics and Evolution*, in press.

III. Ledent, A.*, Désamoré, A.*, Laenen, B.*, Mardulyn, P., McDaniel, S.F., Zanatta, F., Patiño, J.*, & Vanderpoorten, A*. (2019). No borders during the post-glacial assembly of European bryophytes. *Ecol Lett*, *22*(6), 973–986. https://doi.org/10.1111/ele.13254

Ongoing study

IV. Ledent, A., Dellicour, S., Pereira, M.R., Zartman, C.E., Overson, R.P., Mardulyn, P.*, & Vanderpoorten, A.*. Comparing the impact of past climate changes and deforestation on Amazonian bryophyte populations.

*These authors contributed equally to the study.

Table of contents

XV

List of figures

List of abbreviations

yrs/Kyrs/Myrs/Byrs	years/thousand years/Million years/Billion years
Ma (french)	Million d'années
BP	Before Present
LGM/PLGM	Last Glacial Maximum/Penultimate Last Glacial Maximum
c.	*circa*/around
i.e.	*id est*/that is
e.g.	*exempli gratia*/for example
IC species	species currently distributed in areas that were covered in ice at LGM
IF species	species currently distributed in areas that were ice-free at LGM
E-W/E/W/NE	East-West/East/West/NorthEast
O-E (french)	ouest-est
°/°E/°N	degree/degree of longitude East/degree of latitude North
DNA	DeoxyriboNucleic Acid
cpDNA/nDNA	chloroplastic DNA/nuclear DNA
aDNA/sedaDNA	ancient DNA/sedimentary ancient DNA
ddNTP	Dideoxynucleotide
bp	base pair
PCR	Polymerase Chain Reaction
NGS	Next Generation Sequencing
RADseq/ddRAD	Restriction site-Associated DNA sequencing/double digest RADseq
GBS	Genotyping-By-Sequencing
SNP	Single Nucleotide Polymorphism
Pc	Probability of coalescence
N	Effective population size
SDM	Species Distribution Model
GB	Geographic Background
ABC	Approximate Bayesian Computation
BF	Bayes Factor
F_{st}	fixation index
F_{ij}	kinship coefficient
F_{is}	inbreeding coefficient
He	expected Heterozygosity
PrS	number of Private alleles
π	nucleotide diversity
LDD/SDD	Long-Distance Dispersal/Short-Distance Dispersal
IBD	Isolation-By-Distance
IBE	Isolation-By-Environment
IBR	Isolation-By-Resistance
P	P-value
ha/yr	hectare per year
m/km/km^2	meter/kilometer/square kilometer
°C	degree Celsius

Introduction

Climate changes through time

Studying past climate is extremely important to understand how it influenced the evolution of life on earth through time. Paleoclimatic studies are typically based on the analysis of cores from different soil types to find markers of climate changes and determine when these changes occurred. Current techniques (mostly based on the analysis of oxygen isotopes from cores) allow to trace back the earth climate history since the Archean Eon (3.9 to 2.5 billion years (Byrs) ago, Hessler, 2011). Tracing back the earth's climate history revealed how much it has fluctuated through time. For instance, the earth experienced four large-scale episodes of glaciations through time: the two *Snowball Earth* (the first one *c.* 2.4 Byrs ago and the second one from 800 to 650 Million years (Myrs) ago, Kirschvink et al., 2000), the Permian glaciations (335 to 260 Myrs ago, Montañez & Poulsen, 2013) and the Quaternary glaciations (2.4 Myrs ago to present time, Hewitt, 2000). Those changes all impacted species distributions. For example, during the long global cooling that started more than 50 Myrs ago and ended with the Quaternary glaciations (Fig. 1, Hansen, Sato, Russell, & Kharecha, 2013), palm-like tree floras north of the Arctic circle (Pross et al., 2012) went extinct with the advance of ice-sheets which ended up covering large parts of North America and Eurasia in the northern hemisphere (Fig. 2, Clark et al., 2009). Among all past climate changes, the Quaternary period, being the last large-scale episode of glaciations that the earth experienced, is considered to have had the most dramatic impact on current species distributions (Hewitt, 2000).
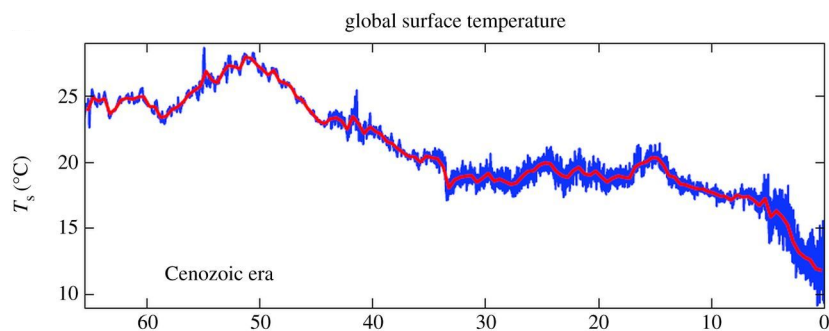
Figure 1: Surface temperature estimate (T_s) in Celsius degrees for the past 65.5 Myrs. The red curve has a 500 Kyrs resolution. Modified from: Hansen, Sato, Russell, & Kharecha, 2013.

The Quaternary period is characterized by high amplitude climatic oscillations leading to glacial/interglacial cycles on both hemispheres, rapid sea-level changes (up to 150m), small continental drift (<100 km), the redistribution of rocks and sediment through ice-rafting, and megafaunal extinctions on all continents (Pillans & Gibbard, 2012). The duration of each glacial/interglacial cycle varied through time. From the beginning of the Quaternary (2.4 Myrs ago) to 0.9 Myrs ago, the duration of a cycle was *c.* 41 thousand years (Kyrs). It then expanded to a *c.* 100 Kyrs, which is the duration still observed now. Each cycle included a long glacial period and a short interglacial period (Fig. 4). During the entire Quaternary period, there were *c.* 40 to 50 glacial periods. The areas located beyond 65° of latitude towards the poles were always covered in ice, while the areas right below those latitudes, were, more or less widely, covered in ice only during glacial periods (Berger, Mesinger, & Sijacki, 2012). In the southern hemisphere, the Antarctic and the Andes were always covered in ice during glacial periods. In the northern hemisphere, the ice-sheets were much more extended during glacial periods and

1

covered about half of North America, most of Fennoscandia, and the southern mountains (see LGM ice-sheets as an example in Fig. 2).
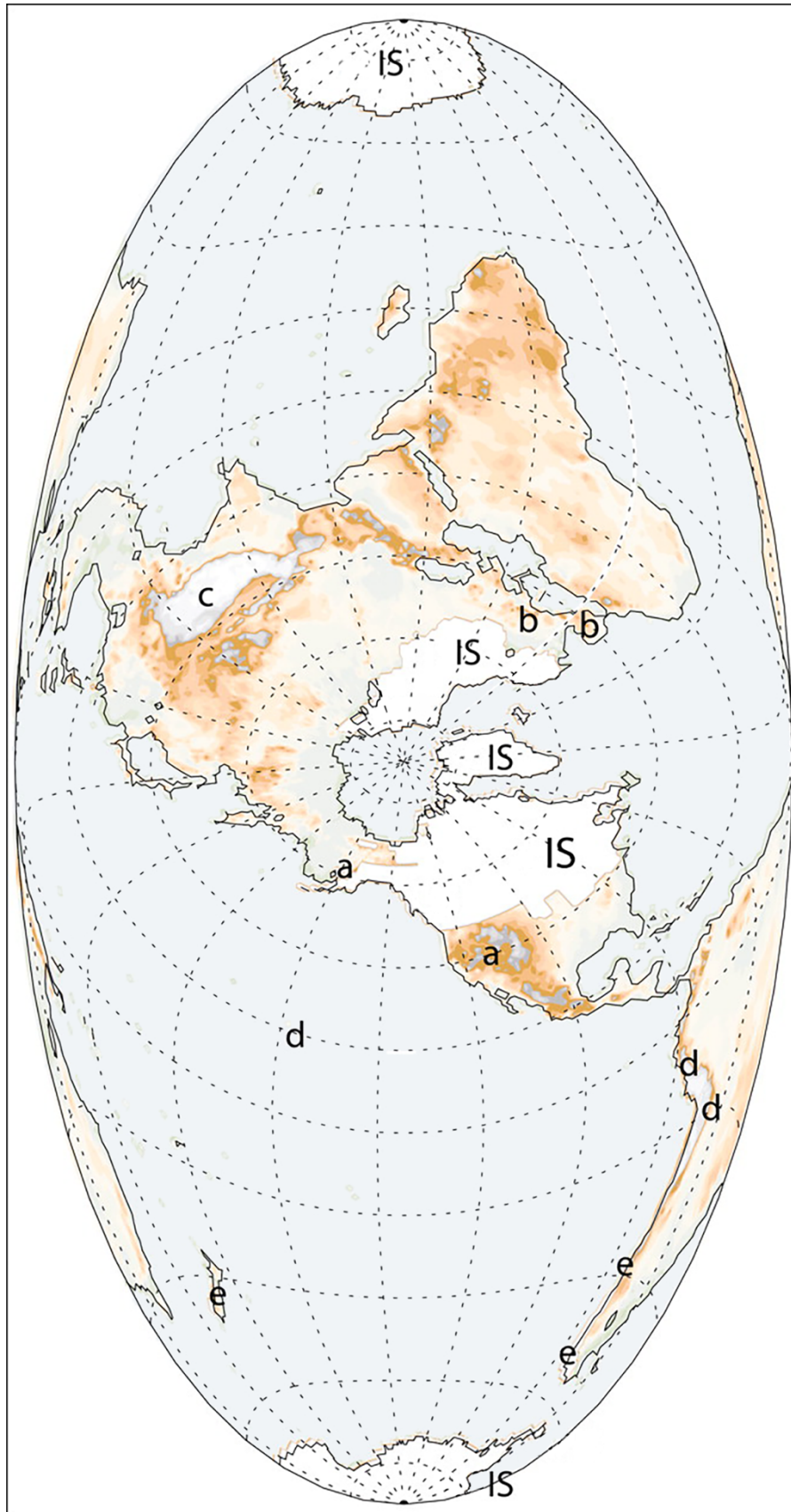
Figure 2: Distribution of ice-sheets at the Last Glacial Maximum. IS, Ice-Sheet. Also shown are areas of mountain glaciation (a, western North America; b, Europe; c, Tibet; d, tropics and subtropics; e, Southern Hemisphere). Modified from: Clark et al., 2009.

Each glacial/interglacial cycle was different, influencing the limit of the ice-sheets during different glacial periods. In Europe for example, among the four last glacial periods (the Elster glacial period: 478-415 Kyrs ago, the Saale glacial period: in two parts between 385 and 130 Kyrs ago, and the Weichsel glacial period: 115-11.7 Kyrs ago), the Saale glacial period was the one with the globally largest ice-sheets, but the ice-sheets of the Elster glacial period covered some parts of Europe that were not covered by the Saale glacial period (Fig. 3, Andersen & Borns, 1994). Not only was each glacial period different, but within a glacial period, there were also a lot of variations such as in temperatures (Fig. 4), ice-sheets extents and sea-levels (Hansen et al., 2013). During each glacial period, the moment at which the ice-sheets reached their maximum size is referred to as the "glacial maximum". Although the Penultimate Last Glacial Maximum (*c.* 140 Kyrs ago, PLGM, Schneider, Schmitt, Köhler, Joos, & Fischer, 2013) involved globally larger ice-sheets, current species distributions were mostly shaped by the Last Glacial Maximum (*c.* 22 Kyrs ago, LGM, Clark et al., 2009; Hewitt, 2000). Indeed, the LGM was not only characterized by large ice-sheets, but also by a very dry climate around the globe. In Eurasia, LGM paleovegetation reconstructions based on pollen records depict this part of the world as mostly treeless, with a dominance of steppe, tundra and other xeric types of vegetation (Tzedakis, Emerson, & Hewitt, 2013; Wu, Guiot, Brewer, & Guo, 2007). In tropical regions, palaeodunes were present at LGM, for instance, in regions like Amazonia (Filho, Schwartz, Tatumi, & Rosique, 2002).
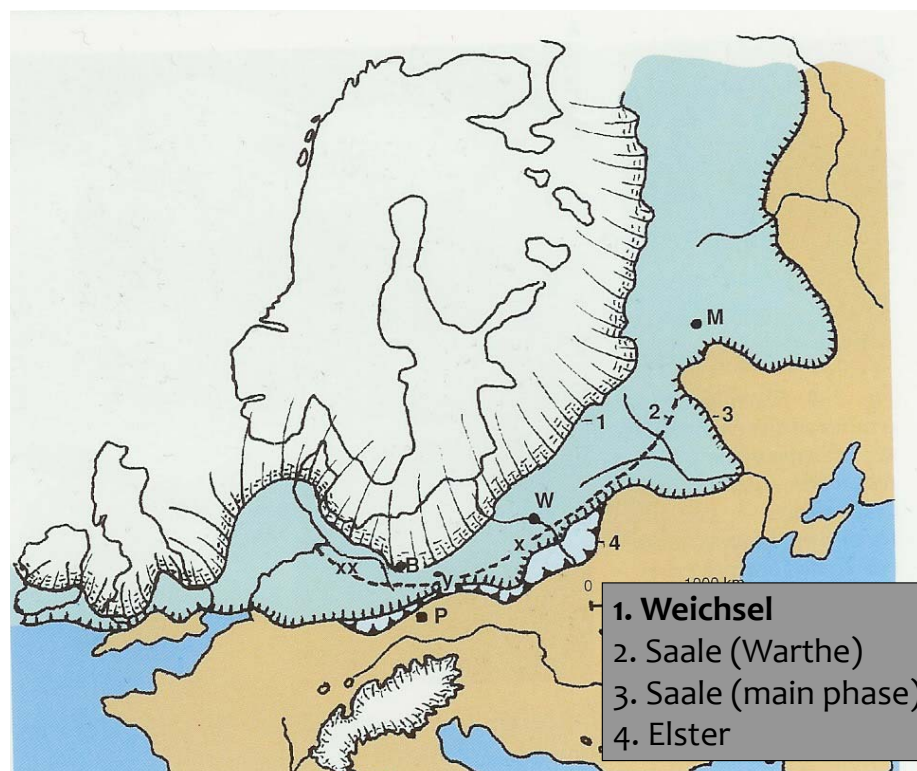
Figure 3: European ice-sheets extent during the four last glaciations of the Quaternary period. Modified from: Andersen & Borns, 1994.
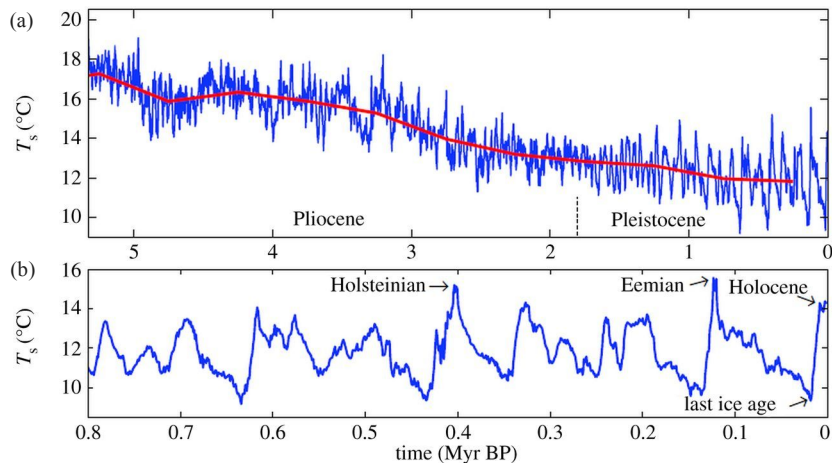
Figure 4: Surface temperature estimate (T$_s$) in Celsius degrees for (a) the Pliocene and Pleistocene including an expanded time scale for (b) the past 800 Kyrs. The red curve has a 500 Kyrs resolution. Modified from: Hansen et al., 2013.

The impact of Quaternary climatic oscillations on the distribution of species

1. Temperate biome ice-free (IF) during the Last Glacial Maximum

The European temperate biome has long served as a model region to study the impact of past climate changes on extant biodiversity patterns (Lumibao, Hoban, & McLachlan, 2017). The classical historical scenario describing the impact of Quaternary glacial/interglacial cycles on European species currently distributed in areas that were ice-free at LGM (hereafter IF species), the "southern *refugium* scenario", is based on paleontological and phylogeographic evidence and suggests that IF species persisted in Mediterranean *refugia* during glacial periods, from which they recolonized northern area during interglacial periods (Hewitt, 1996, 1999, 2000, 2004; Médail & Diadema, 2009). In this scenario, a decrease in genetic diversity along the northern migration routes was expected (Hewitt, 1999) until Petit (2003) showed that the genetic diversity peaked at mid-latitudes, at the level of the contact zone among recolonization waves from different southern *refugia*, and decreased at high latitudes due to Long-Distance Dispersal (LDD) events and associated founder effects. The "northern micro-*refugium* scenario" alternatively proposes that IF species persisted within micro-*refugia* located between the northern ice-sheet and the main southern mountain ranges during glacial periods (Bhagwat & Willis, 2008). According to this scenario, populations expanded from northern micro-*refugia* during warmer periods, like the current interglacial. The inferred post-glacial migration routes of IF species were often, but not always, blocked by the E-W-oriented mountain ranges of the Pyrenees and the Alps (Avise, 2000; Mátyás & Sperisen, 2001; Petit et al., 2002; Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998). But what happened to those species for which these mountain ranges were not barriers to migration, but served as their habitats?

2. Arctic-Alpine biomes covered in ice (IC) during the Last Glacial Maximum

Darwin (1859) and Hooker (1861) proposed that European species currently distributed in areas that were covered in ice at LGM (hereafter IC species) migrated towards lowland areas with the advancing ice-sheets during glacial periods. Although this hypothesis received some support from molecular phylogeographic analyses (Schönswetter, Popp, & Brochmann, 2006; Skrede, Eidesen, Portela, & Brochmann, 2006), paleontological evidence suggests that lowland areas south of the northern ice-sheet experienced a cold and dry climate that was unsuitable for IC floras (Abbott & Brochmann, 2003). Hultén (1937) alternatively suggested that Beringia, a region encompassing Northeast Russia and Northwest America, which remained ice-free

during Quaternary glaciations, served as a source for the back-colonization of Europe (Eidesen et al., 2013). Other molecular evidence raised the possibility that some IC species survived in local micro-*refugia* within the ice-sheets such as in the European southern mountain ranges (Schönswetter, Paun, Tribsch, & Niklfeld, 2003; Schönswetter, Stehlik, Holderegger, & Tribsch, 2005) – from where they potentially back-colonized northern areas – and in the Arctic area (Westergaard et al., 2011).

3. Tropical biomes

Tropical biomes have been less studied with regard to the impact of past climate changes on species distributions. The lack of pollen records in these regions largely explains the scarcity of references (Ledru, Bertaux, Sifeddine, & Suguio, 1998). As opposed to Europe, lowland tropical biomes are relatively homogeneous environments without any apparent geographic barrier to migration for wind-dispersed spores, as evidenced by the low beta-diversity among floristic assemblages distant of several thousands of kilometers (Condit et al., 2002; Pitman, Terborgh, Silman, & Nuñez V., 1999). Available paleontological evidence suggests that all the tropical regions of the world experienced severe vegetation shifts during the glacial/interglacial cycles of the Quaternary period, with forest contractions (bottlenecks) during glacial periods. African rainforests were substantially more impacted than Neotropical ones (Parmentier et al., 2007). In particular, pollen and geochemical evidence suggests that, during the LGM, the African rainforest area was reduced by *c.* 84%, whereas the Amazon humid forest area probably shrank to only 54% of its present-day extent (Anhuf et al., 2006). In Amazonia, paleoenvironment reconstructions based on fossil records lead to the emergence of two main competing scenarios to describe the impact of Quaternary climate changes on species distributions: (i) the invasion of the lowland evergreen continuous forest by cold-adapted Andean taxa during glacial periods (Bush et al., 2001; Colinvaux, De Oliveira, Moreno, Miller, & Bush, 1996; Haberle & Maslin, 1999); and (ii) the replacement of the lowland evergreen forest by xeric taxa during the entire Quaternary period due to a globally drier environment (Behling & Hooghiemstra, 2001; van der Hammen & Absy, 1994). These two scenarios are, however, not mutually exclusive. In fact, recent pollen evidence and dynamic vegetation models revealed both the invasion of cold-adapted species across the entire Amazonian lowland forest during the mid-Pleniglacial (*c.* 60,000 yrs BP) and the replacement of this vegetation by a xeric open one in southern lowland Amazonia during the increase of aridity that took place between the late-Pleniglacial (*c.* 28,000 yrs BP) and the LGM (*c.* 22,000 yrs BP, Mayle, Beerling, Gosling, & Bush, 2004; Reis et al., 2017). However the constitution of palaeodunes during this late-Pleniglacial to LGM transition in northern lowland Amazonian areas, and especially on the riverbanks of the Rio Negro suggests, even if not a widely-distributed, a locally or regionally drier climate across the entire lowland Amazonian rainforest during the Quaternary period (Filho et al., 2002). The impacts of Quaternary climate changes in Amazonia described by those different paleoenvironment reconstructions all imply range contractions (bottlenecks) and expansions of the evergreen continuous lowland Amazonian forest during the Quaternary period. However, the exact intensity of these demographic changes remains unclear.

The impact of climate changes on bryophytes

Bryophytes are the most ancient lineage of extant land plants (Embryophyta, Edwards, Morris, Richardson, & Kenrick, 2014). They are composed of three groups: the Hepaticophyta (liverworts), with *c.* 6000 species; the Anthocerotophyta (hornworts), a small group of *c.* 300 species; and the Bryophyta (mosses)*,* currently the largest group with *c.* 12,000 species.

Bryophytes are non-vascular plants, meaning that they cannot pump up water from the soil, but instead absorb water and nutrients from their entire surface directly from rainfall. Unlike vascular plants, which are drought-resistant and aim at maintaining sufficient levels of water during periods of drought thanks to the presence of a waterproof cuticle and the regulation of gas exchange through their stomata, bryophytes are poïkilohydric, *i.e.* drought-tolerant. This means that they dry-out with ambient air, becoming dormant under dry conditions and then resume physiological activity upon moistening. Bryophytes are therefore very sensitive to variations of the precipitation regime (He, He, & Hyvönen, 2016).

Bryophytes globally exhibit a very high cold tolerance. Recent evidence points to their ability of in-vitro regeneration after hundreds to thousands of years in the ice (La Farge, Williams, & England, 2013; Roads, Longton, & Convey, 2014). Conversely, bryophytes exhibit comparatively lower temperature optima than angiosperms (Furness & Grime, 1982). Because of their low tolerance to hot temperatures and their sensitivity to variations of the precipitation regime, bryophytes appear as extremely sensitive organisms to climate changes (He et al., 2016), leading Tuba, Slack, & Stark (2011) to qualify them as the "pinsons in the coal mine" in the context of climate changes. As a result, bryophyte distributions are expected to have quickly responded to climate changes in the past, as confirmed by studies of range dynamics (Frahm & Klaus, 2001; Zechmeister, Moser, & Milasowszky, 2007) and stratigraphic analyses of macroremains preserved in peat (Ellis & Tallis, 2000; Jonsgard & Birks, 1995).

Except for macroremains preserved in peat, the Quaternary fossil record of bryophytes is, however, extremely poor as compared to the comparatively rich fossil record in vascular plants (see Patiño & Vanderpoorten, 2018 for review), calling for a molecular approach to reconstruct their biogeographic history from information on the extant genetic structure and diversity of their populations. In contrast to higher plants and animals, however, there has been no comprehensive effort to reconstruct the Quaternary history of bryophytes to date. Especially, even if several species-level phylogeographies have been published for the European flora (see Kyrkjeeide, Stenøien, Flatberg, & Hassel, 2014 for review), virtually no study has yet attempted to investigate the impact of past climate changes in tropical environments.

Bryophytes are haplo-diplophasic organisms with a dominant haploid phase. They disperse by means of spores or asexual diaspores. Spore density quickly decreases with increasing distance from the mother sporophyte. The majority of the spores, however, disperse across large distances (Lönnell, Hylander, Jonsson, & Sundberg, 2012; Sundberg, 2005). In *Sphagnum*, for example, only 6.8-22.4% of the spores are deposited within a 3.2m range (Sundberg, 2005). Bryophyte spore deposition curves are therefore typically fat-tailed, resulting in high, distance-dependent colonization probabilities near the source, and much lower, but still substantial, distance-independent colonization probabilities once spores are airborne away from the source (Lönnell et al., 2012, 2015). Spore production strongly depends on sexual systems because sexual reproduction involves that a male gamete is able to swim to a female archegonium in a continuous film of water. Monoecious species (*c.* 1/3 of the moss and liverwort species) are therefore frequently fruiting because male and female gametangia can be found on the same individual. Although self-incompatibility was recently reported in mosses (Stark & Brinda, 2013), monoecious bryophyte species are assumed to be capable of self-fertilization. In fact, high F_{is} values typically characterize the sporophytic phase of monoecious species (Hutsemékers, Hardy, & Vanderpoorten, 2013; Johnson & Shaw, 2015; Klips, 2015). In dioecious species, conversely, fertilization is hampered by the distance

separating male and female plants. Therefore, dioecious species produce, on average, significantly fewer sporophytes than monoecious ones (Laaka-Lindberg, Hedderson, & Longton, 2000; Longton, 1997). Based on these observations, monoecious bryophyte species have therefore been perceived as better dispersers than dioecious ones (Longton & Schuster, 1983; but see Laenen et al., 2016). Lowland tropical bryophytes are precisely characterized by "higher than normal levels of monoecism" (Longton & Schuster, 1983). Increased rates of monoecy, combined with other traits such as precocious germination of spores, while still in the capsule, have led to the hypothesis that tropical bryophytes are particularly well-equipped for LDD and rapid and efficient establishment (Longton & Schuster, 1983). Using null model analyses based on metacommunity concepts for Amazonian epiphytic bryophyte communities, Mota de Oliveira & ter Steege (2015) concluded that "dispersal did not show geographical structure across the area". Metacommunity analyses thus raise the intriguing notion that Amazonian bryophytes are, due to their high dispersal capacities, homogeneous across very large spatial scales, leading Mota de Oliveira & ter Steege (2015) to assume that, at the scale of the Amazon basin, an area of 6 million square kilometers, "epiphytic bryophytes behave as one single metacommunity". In such conditions, an inverse isolation effect is predicted to develop (Barbé, Fenton, & Bergeron, 2016; Sundberg, 2005). An inverse isolation effect involves a higher genetic diversity of colonizing propagules with increasing isolation (*i.e.* distance from the mother sporophyte), thus counteracting genetic differentiation. Consequently, no Isolation-By-Distance (IBD) is expected beyond a distance corresponding to short-distance dispersal (SDD) events owing to the well-mixed and diverse propagule pool, except perhaps at very large scales, at which other factors (including geographic barriers and historical factors) might operate (Szövényi, Sundberg, & Shaw, 2012).

The high dispersal capacities of bryophytes have major consequences for the evolutionary biology and biogeographic history of the group. First, high dispersal capacities would not be compatible with genetic differentiation, and ultimately, allopatric speciation at the regional scale due to the erosion of IBD by LDD, especially in such highly homogeneous environments as lowland tropical areas. This raises the question of whether other mechanisms, such as isolation-by-environment (IBE), could promote speciation in such environments. Second, the application of the inverse isolation hypothesis, according to which any signal of IBD is erased beyond the limits of SDD, casts doubts on the possibility to find signatures of historical events from analyses of the extant spatial genetic variation of bryophyte populations (Van Der Velde & Bijlsma, 2003). Third, the high long-distance capacities of bryophytes raise the intriguing idea that post-glacial recolonization could have taken place from geographically remote areas. In Europe for instance, mounting evidence points to the relevance of NE Atlantic islands (Hutsemékers et al., 2011; Laenen et al., 2011) and North America (Stenøien, Shaw, Shaw, Hassel, & Gunnarsson, 2011) as sources of recolonizing propagules.

Research aims

In the present thesis, we assembled and analyzed large molecular datasets at the level of the species range to determine how bryophytes responded to major Quaternary climate changes in environments characterized by different 'resistance' to migration and environmental heterogeneity, especially in Europe and lowland Amazonia.

More specifically, the aims are to:

1. Test whether, due to the high dispersal capacities of bryophytes in general – and in particular in homogeneous environments without any apparent geographic barrier to migration –, the inverse isolation hypothesis – according to which any signal of isolation-by-distance (IBD) is erased beyond the limits of short-distance dispersal (SDD) by the intensity of long-distance dispersal (LDD) events – applies, erasing any historical signal in the extant spatial patterns of genetic structure and diversity of bryophyte populations (H1, Papers I and III).
2. Test the relevance of other differentiation mechanisms promoting speciation and, in particular isolation-by-environment (IBE), across a relatively homogeneous environment without any apparent geographic barrier to migration (H2, Papers I and II).
3. If H1 is rejected, infer the post-glacial history of bryophytes, in environments characterized by the presence (Europe, Paper III) or the absence (lowland Amazonia, ongoing study IV) of apparent geographic barriers to migration, from analyses of the extant spatial patterns of genetic structure and diversity of their populations (H3).

Methods

Taxonomic and molecular sampling

In the study focusing on Europe (Paper III), 3 and 12 species currently distributed in areas that were covered in ice (IC) or ice-free (IF) at the Last Glacial Maximum (LGM), respectively, have been investigated. IC species include *Amphidium lapponicum*, *Timmia austriaca*, and *T. bavarica*. IF species include *Amphidium mougeotii*, *Calypogeia fissa*, *Diplophyllum albicans*, *Homalothecium sericeum*, *Metzgeria conjugata*, *M. furcata*, *Orthotrichum affine*, *O. lyellii*, *Plagiothecium denticulatum*, *P. undulatum*, *Plagiomnium undulatum*, and *Scorpiurium circinatum*. Specimens were sampled across their entire distribution range, but with a focus on Europe and North America due to previous evidence for the existence of genetic connections between them (Désamoré et al., 2016; Kyrkjeeide et al., 2014; Stenøien et al., 2011; Szövényi, Terracciano, Ricca, Giordano, & Shaw, 2008). Specimens were then assigned to each of 3 or 5 regions – which correspond to the definition of source and sink areas in historical scenarios referenced in the literature and describing the post-glacial recolonization of Europe (see "Introduction" section of this thesis) – for IF and IC species, respectively (Table S1, Fig. 3, Paper III). The genetic markers (Sanger *loci*) included specific chloroplastic (cpDNA) and nuclear (nDNA) regions selected for their suitable range of variation as evidenced by previous phylogeographic studies (see Patiño & Vanderpoorten, 2018 for review). Specimens of liverworts (5 species) were sequenced at 2-3 cpDNA *loci* and specimens of mosses (10 species) were sequenced at up to 3 cpDNA *loci* and 1-2 nDNA *loci* (Tables 1 and 2, Paper III). Chloroplastic *loci* were concatenated and treated as one *locus* due to the linkage of chloroplastic genes (absence of recombination), whereas nuclear loci were each treated individually. DNA alignment matrices are available from Figshare (see "Data accessibility statement" section in Paper III for DOIs).

In the studies focusing on lowland Amazonia (Papers I and II and ongoing study IV), the following bryophyte species have been investigated: *Archilejeunea fuscescens*, *Bazzania hookeri*, *Leucobryum martianum*, *Micropterygium trachyphyllum*, *Octoblepharum albidum*, *O. pulvinatum*, *Syrrhopodon annotinus*, *S. helicophyllus*, *S. hornschuchii*, *S. simmondsii*, and *Thysananthus amazonicus*. Specimens were sampled across a 42,640 km^2 area of lowland rainforest in the Rio Negro Basin, north of Manaus (Fig. 1, Table S1, Paper I). Sanger *loci* typically used in bryophytes for phylogeographical analyses were entirely monomorphic among the sampled specimens. We performed a screening for individual variation among a set of commonly used *loci* (a combination of five cpDNA *loci* and one nDNA *locus*) and did not find any variation for any *locus* (unpublished work). In particular, the two sibling species *Syrrhopodon annotinus* and *S. simmondsii* were recently found to be nested within the same clade but their phylogenetic identity was unresolved, calling for the development of more variable markers (Pereira et al., 2019). We thus designed and performed a specific approach of genome-wide sequencing, called "Restriction site-Associated DNA sequencing" (RADseq), which is one of the Next Generation Sequencing (NGS) technologies, to recover genetic markers for this set of Amazonian bryophyte species (see "Molecular protocols" and "Bioinformatics processing" sections in Paper I). RADseq raw data (fastq files) were submitted to the NCBI Sequence Read Archive (SRA). The biosample accession number for each specimen can be found in Table S1, Paper I. An overview of both Sanger and NGS sequencing technologies, the latter focusing on RADseq, can be found in Box 1.

Box 1: Overview of Sanger and NGS sequencing technologies

Sequencing genetic markers involves a chain of wet and dry lab manipulations leading to the recovery of the sequence of each of these markers. The first massively used sequencing technique (often referred to as "first generation sequencing") is called "Sanger sequencing" (Sanger, Nicklen, & Coulson, 1977). This technique only allows to sequence a short DNA fragment of *c.* 1000 base pairs (bp) at a time. It involves two amplification steps: one PCR and one sequencing reaction. It uses specific PCR primers to first amplify the given fragment. Then, during the sequencing reaction, copies from the same fragment are terminated after a different number of nucleotides and a fluorescent dideoxynucleotide (ddNTP) is added at the end of each copy by a DNA polymerase, so that fluorescence detection through time (under capillary electrophoresis) allows to recover the full sequence of the specific DNA fragment/genetic marker, nucleotide by nucleotide. This technique first emerged in 1977 and dominated the sequencing field for about three decades. Even if this technique allowed the sequencing of the first complete genome: the Human Genome Project in 2003 (Collins, Morgan, & Patrinos, 2003), recovering a full genome fragment by fragment consumes both a lot of time and money. Sanger sequencing is thus mostly used to sequence selected fragments, especially genetic markers, across the genome of a given species.

Next Generation Sequencing (NGS) refers to new techniques which allow to massively sequence DNA fragments in parallel. Parallel sequencing has been mainly made possible through the miniaturization of the sequencing reactions, the immobilization of the fragments on a solid surface or stand and the enhancement of the detection system (Kulski, 2016). For example, one of the latest sequencers released, the Illumina NextSeq© 500, allow to read, in high output mode, up to 400,000,000 fragments (afterwards called "reads") on a single sequencing lane (in Illumina technologies, the surface on which DNA fragments are immobilized is called a "flow cell" and is divided into lanes that are read independently by the sequencer during sequencing runs). Furthermore, new techniques no longer necessarily require a PCR amplification (Metzker, 2010), so that the amount of information collected in a single lane is enough to recover the full genome of most living organisms, provided that during the lab preparation, all genomic fragments were kept. Moreover, not only multiple fragments can be sequenced at the same time, but fragments from different individuals and even different organisms can be sequenced at the same time when indexes are used because further computational analyses allow to group together reads containing the same index. NGS technologies finally opened the door for metagenomics (see Sleator, Shortall, & Hill, 2008 for review): the sequencing of all DNA fragments present in an environmental sample such as soil, water or intestine, for example.

Even if, per sequenced fragment, NGS is much less expensive than Sanger sequencing, because it sequences an enormous amount of DNA fragments at the same time, the price of an NGS sequencing run is very high (*e.g.* around 3000 USD for one lane on an Illumina NextSeq© 500 in high output mode). This is why many techniques are being developed to only recover the sequences of some DNA fragments across whole genomes, such as target-enrichment (see Mamanova et al., 2010 for review), transcriptome sequencing (see Tachibana, 2015 for review), Restriction site-Associated DNA sequencing (RADseq, see Andrews, Good, Miller, Luikart, & Hohenlohe, 2016 for review), or metabarcoding, for environmental DNA (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014). Among these techniques, RADseq, which is the NGS technique used in this thesis, has increasingly been used as a low-cost method to recover genetic markers. The way it works is very simple: the genome is cut into small fragments by one or two restriction enzymes during the process called "digestion", then, during

the process called "ligation", each fragment is linked to an adapter, which contains both an individual index and a specific site that allows the fragment to bind and be immobilized on the sequencer's reaction surface. After ligation, RADseq techniques almost always involve a PCR amplification and always involve a size selection so that only fragments of desired sizes get sequenced. The size of the fragments to be sequenced is important because it influences the sequencing efficiency: the highest efficiency is obtained with fragment sizes of *c.* 150 to 400bp on an Illumina NextSeq© 500, for example. Some specific RADseq techniques, such as double digest RAD sequencing (ddRADseq), use a direct size-selection step based on magnetic beads or agarose gel-cutting to ensure that the fragments have the size resulting in the highest sequencing efficiency for the specific sequencer used; other techniques, such as Genotyping-By-Sequencing (GBS), use an indirect size-selection through the limitation of the PCR amplification: PCRs do not amplify too large fragments. Once sequences have been recovered from the prepared RADseq libraries of DNA fragments, dry lab techniques are used to keep, among all those reads, only genetic markers. The main advantage of RADseq is that knowledge on the studied species' genomes is not required, because most restriction enzymes find restriction sites in any genome to bind and cut this genome, so that RADseq can even be used in non-model organisms. However, previous information on the genome composition help to predict where and how often different restriction enzymes would cut this genome. The disadvantage of RADseq techniques is that there is no information regarding where each read comes from, not even from which genome (nuclear, chloroplastic or mitochondrial). When reference genomes of the studied species are available, it is possible to map the reads on the genomes in order to localize them, but, without reference genomes, RADseq can only be used to recover a set of genetic markers, regardless of where those markers come from, it thus cannot be used, for example, to target specifically given genetic markers or to reconstruct whole genomes.

Coalescent simulations

In this thesis, a specific analytical approach to analyze genetic markers has been used: coalescent simulations. Coalescent simulations were used to test the hypothesis that the high dispersal capacities of bryophytes erased any historical signal in the extant spatial patterns of genetic structure and diversity of bryophyte populations in Europe (H1, Paper III) and to infer the post-glacial history of bryophytes, in environments characterized by the presence (Europe, Paper III) or the absence (lowland Amazonia, ongoing study IV) of apparent geographic barriers to migration, from analyses of the extant spatial patterns of genetic structure and diversity of their populations (H3). Beyond the qualitative description of summary statistics or gene trees, coalescent simulations have increasingly appeared as a promising tool to infer demographic histories from genetic markers (Thomé & Carstens, 2016).

The coalescent model

The coalescent model was first described by Kingman (1982). Reviews and books on the coalescent model include Hein, Schierup, & Wiuf, 2004; Hudson, 1990; Marjoram & Joyce, 2011; Nordborg, 2001; and Wakeley, 2008. I briefly describe here its important aspects. The coalescent model aims at reconstructing the genealogy of all sampled gene copies of a given genetic marker in a population going backward through time until reaching the last common ancestral copy of all of them. The coalescent model is purely demographic, which means that coalescent simulations are totally independent of the mutation process. The separation between the demographic and the mutation process is only possible under the assumption of neutral evolution (*i.e.* the absence of selection) in which all gene copies have the same probability to

be transferred to the next generation (*i.e.* independent of the allelic type). The coalescent model relies on the fact that the genealogy of any genetic marker is modified by the demographic events that the population went through.

The coalescent model defines the probability of coalescence, which corresponds to the probability that two sampled gene copies among all sampled gene copies of a given population share a common ancestor at a given generation, and depends on both the sampling size (n) and the effective population size (N) of this population, following this relation: $Pc \approx n(n-1)/2N$. Because the probability of coalescence depends on the effective population size of a given population, it is influenced by any demographic change in this population. For example, if a population underwent a bottleneck, then its effective population size (N) dropped whereas its sampled size was likely to not be affected, so that the probability of coalescence in the population increased, resulting in shorter branches of its genealogy during that period (Fig. 5). The coalescent model thus allows to simulate the gene genealogy of any sequenced genetic marker of a given population undergoing changes in its demography (*e.g.* a bottleneck), provided that the effective population size, at different periods, and the sampling size of the population are defined. Such simulations are called "coalescent simulations".
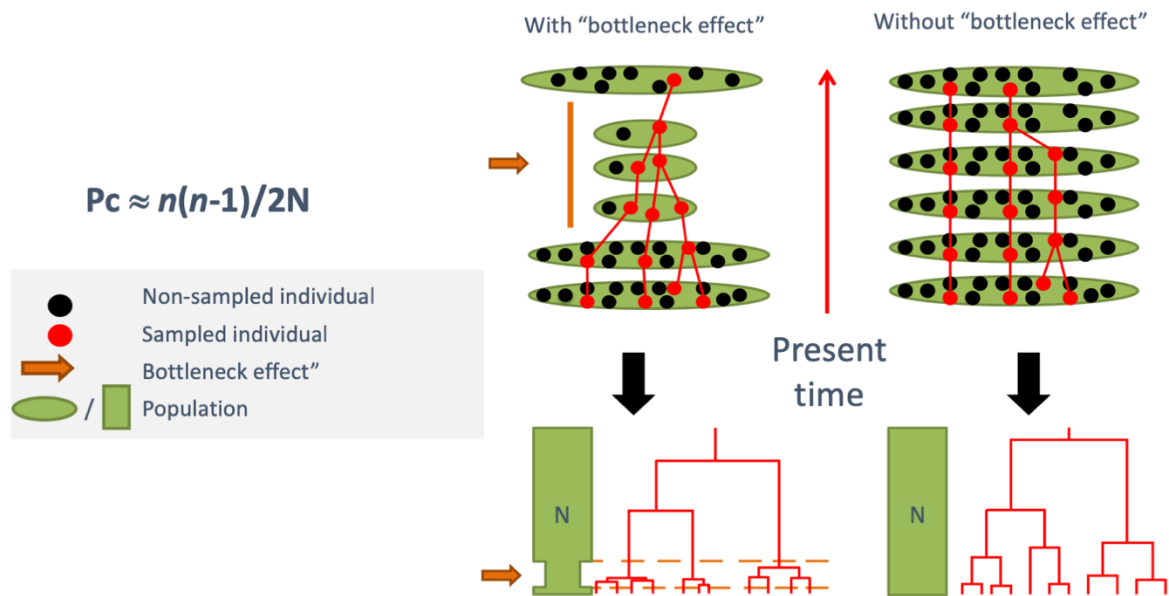
Figure 5: Putative impact of a bottleneck on a gene genealogy built under the coalescent model. Pc: probability of coalescence; n: sampling size; N: effective population size.

The standard coalescent model simulates gene genealogies of a unique given population. However, populations of organisms are usually subdivided into subpopulations connected, to some extent, by migrations (Wakeley, 2001). The structured coalescent model aims at reconstructing, for a given genetic marker, the gene genealogy of all sampled gene copies across a meta-population divided into subpopulations among which there is some pattern of migration (Hudson, 1991). If one considers a meta-population subdivided into two subpopulations where at each generation, a small fraction of each subpopulation is made of migrants from the other subpopulation, then the probability of coalescence between any two sampled gene copies coming from the same subpopulations among all sampled gene copies of the meta-population (Pc(s)) and the probability of coalescence between any two sampled gene copies coming from the two different subpopulations among all sampled gene copies of the meta-population (Pc(d)) could be established at a given generation. Pc(s) would be computed as the probability of coalescence in a unique panmictic population, while Pc(d) would depend on the migration rate between the two subpopulations and is likely to be much smaller than

Pc(s). The time of coalescence, which corresponds to the time required to reach the common ancestor of two sampled gene copies among all sampled gene copies in a given population, could be established as well for any two sampled gene copies coming from the same subpopulation and any two sampled gene copies coming from the two different subpopulations. The expectation of those times (Et) would follow these two relations: $Et(s) = n$ and $Et(d) = n + \frac{n-1}{M}$; with Et(s), the expectation of the time of coalescence between any two sampled gene copies coming from the same subpopulation; Et(d), the expectation of the time of coalescence between any two sampled gene copies coming from the two different subpopulations; n, the sampling size of the meta-population; and M, the migration rate between the two subpopulations. Following these equations, if M is small, Et(d) would be much larger than Et(s), so that the differentiation between the two subpopulations would be high (Fig. 6), while if M is large, Et(d) would be close to Et(s), so that the two subpopulations would be similar (Fig. 6, Hudson, 1991). In a meta-population context, the migration rate between subpopulations is thus another important parameter, along with the sampling size and the effective population size of populations, that will impact the gene genealogies of any genetic marker simulated under the coalescent model.
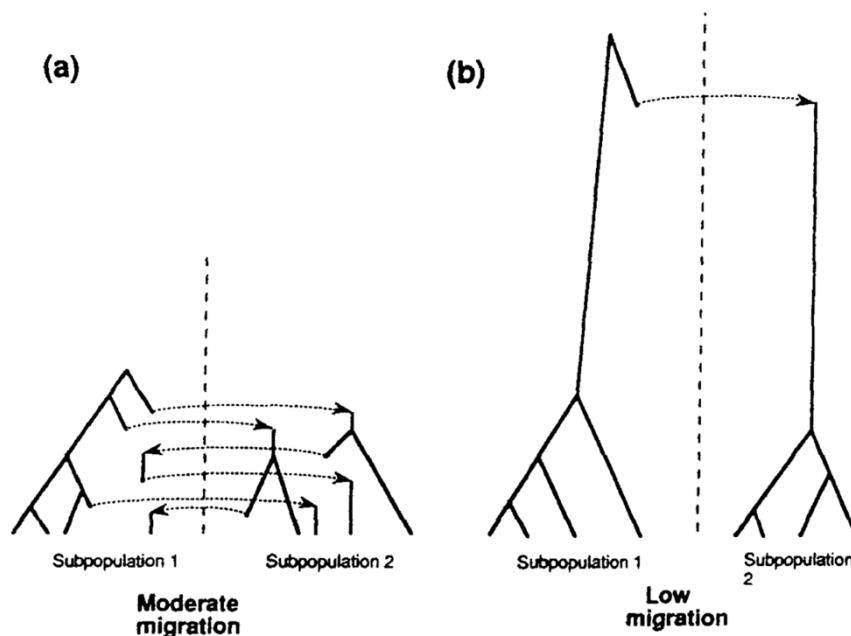
Figure 6: (a) An example of a coalescent gene genealogy for a sampling size of 8, 4 from each of 2 subpopulations, when migration rate is moderately high. Each migration event is indicated by a dotted line with an arrow that indicates the actual direction of movement of an individual migrant. In this case, there would be relatively little differentiation of the two subpopulations. (b) An example of a coalescent gene genealogy with low migration rate. In this genealogy there is a single migration event. Alleles from within a subpopulation will be much more similar than alleles from different subpopulations. From: Hudson, 1991.

In this thesis, we are implementing two different types of structured coalescent models: a classic coalescent population model in Paper III, using the software SIMCOAL (Excoffier, Novembre, & Schneider, 2000), and a spatially explicit coalescent model in ongoing study IV, using the software PHYLOGEOSIM (Dellicour, Kastally, Hardy, & Mardulyn, 2014). Classic coalescent population simulations describe the demographic history of pre-defined putative large panmictic subpopulations while spatially explicit models assess the demographic history of individuals within continuous portions of the species range. Especially, in spatially explicit models, the studied area is represented by a matrix composed of pixels of a desired size, each considered as a small panmictic subpopulation, so that the effective population size and the sampling size are defined in each pixel of the matrix and the migration rate is defined between

each pair of pixels of the matrix. This is opposed to classic coalescent population simulations where the effective population size and the sampling size are only defined within, and migrations are only allowed between, large putative panmictic populations. In Europe, explicit historical scenarios describing the post-glacial recolonization of Europe by IF and IC species are referenced in the literature and were presented within the introduction section of this thesis. These scenarios define clearly distinct source and sink populations for both IF and IC species (Table S1, Fig. 3, Paper III). The use of classic coalescent population models – simulating the demographic history of pre-defined populations – was thus relevant in Paper III which aimed at reconstructing the post-glacial history of these pre-defined source and sink European bryophyte populations. In Amazonia conversely, metacommunity analyses raised the intriguing notion that Amazonian bryophytes are, due to their high dispersal capacities, homogeneous across very large spatial scales. This led Mota de Oliveira & ter Steege (2015) to assume that, at the scale of the Amazon basin, "epiphytic bryophytes behave as one single metacommunity". The use of spatially explicit models – simulating the demographic history of individuals within continuous portions of the species range – is thus relevant in ongoing study IV which aims at reconstructing the post-glacial history of this putative single continuous Amazonian bryophyte population. Because patterns of isolation-by-distance (IBD) are often observed within continuous portions of a species' range rather than among clearly defined panmictic populations of this species, spatially explicit models can lead to much more realistic simulations than classic coalescent population models (Dellicour, Fearnley, et al., 2014). However, they require much more computation time to reach, for a given genetic marker and under a given demographic scenario, the last common ancestral gene copy of all sampled gene copies in a given population (Dellicour, Kastally, et al., 2014), especially if the studied area is large, as in Paper III where the studied area is composed of the entire Holarctic (Fig. 2, Paper III). This reinforced the choice, in Paper III, of classic coalescent population models instead of spatially explicit coalescent models.

Because one of the most important parameters to define when performing coalescent simulations is the effective population size (N), and because this information is unknown for the studied bryophyte species, we implemented Species Distribution Models (SDMs) to infer past and present effective population sizes in Paper III and ongoing study IV. SDMs result in the estimation of the suitability of a given species across any geographic area where environmental data are available. This geographic area is represented by a grid composed of pixels of a defined size. At the end of the SDM process, each pixel of this grid is associated with a value of suitability. Those suitability values can further be used as estimators for the effective population size within each pixel and/or across the overall geographic area (Patiño et al., 2015). In this thesis, in order to infer the effective population size from the suitability values resulting from the SDM process, we either binarized the values by defining a threshold, to get only suitability values of 0 or 1 in each pixel (Paper III) or we normalized the values, to get a range of suitability values going from 0 to 1 in each pixel (ongoing study IV). We then multiplied these suitability values in each pixel by the number of different genotypes reported in a population whose size corresponds to the size of the pixels in order to obtain the effective population size in each pixel. We finally re-scale these effective population size values by pixel accordingly to the size of the populations defined in our demographic scenarios. An overview of SDMs can be found in Box 2.

Box 2: Overview of Species Distribution Models (SDMs)

Major reviews on Species Distribution Models (SDMs) include Elith & Leathwick, 2009; and Guisan, Thuiller, & Zimmermann, 2017. I briefly describe here their important aspects. SDMs are purely ecological and rely on the niche theory. They model the distribution of a species in regions where selected environmental factors are known by extrapolating the information contained in the combinations of factors found where the species is present and, in some models, the combination of factors found where the species is absent. The goal of SDMs is to determine, on a map, potentially suitable habitats for a given species at any given time at which the selected environmental factors are known.

The first step of this technique is to determine the rules that define which combinations of environmental factors are suitable or not for a given species within the studied area. To do so, the Geographic Background (GB) has to be defined, it is the area on which the rules are computed, usually, it corresponds to the size of the whole studied area. The GB is represented by a grid composed of pixels equal in size. The resolution at which the suitability will be predicted is defined by the size of those pixels. In order to define the rules of the model, two datasets are required: the presence points of the species (and absence points, in presence-absence models), called the "dependent variables" (these points have to represent the most part of the variance that resides in the combinations of environmental factors found where the species is present, and absent, if absence points are implemented), and the environmental factors for each of those points, called the "independent variables". The selection of the environmental factors is extremely important since the simulations will only take them into account to define suitable habitats for the species. Their combination thus has to represent the most part of the variance which resides in any potential combination of environmental factors that explain the distribution of the given studied species. Most of the time, SDMs imply climatic factors only. When implemented, the absence points are either given by the user, in presence-absence models, or sampled randomly across the GB, in presence-pseudo-absence models such as Maxent (see Elith et al., 2011 for review). The method used to define the rules of the model is based on machine learning and relies on the extraction of the values of the independent variables (the environmental factors) in each of the species presence point (dependent variable), and, in non-presence-only models, in each of the species absence point, to predict the combinations of environmental factors that are suitable or not for a given species. Different algorithms predicting the potential presence or absence of a given species based on the combination of environmental factors present in a given pixel across the GB are tested based on their reliability to provide good predictions. Those tests require that some of the presence (and absence points, if implemented) of the species are not used for the definition of the rules, but are instead kept to further test those rules. The algorithms that passed the tests are used to build the rules that define which combinations of environmental factors are suitable or not for a given species across the GB. Those rules represent the model of species distribution.

The second step of SDMs is the projection of the model onto the desired area. This area could be either the same area as the GB or a completely different area (different in size or in its geographic localization). The projection could also be done in the past or in the future. The only requirement is that the environmental factors that have been beforehand chosen to build the model are known, at the chosen resolution, within that area, and for that specific time-scale. The result of projected SDMs is a grid, matching the map of the projected area, where each pixel is associated with a specific value of suitability, computed based on the model's rules. Such a grid is called a "raster". These suitability values in each pixel of the grid indicate the probability for the species to inhabit those pixels, SDMs are thus probabilistic models. The

index of suitability represents the complete range of suitability values associated with a given grid. The range of this index is different for each model and each projection.

Since SDMs only rely on ecological factors they are limited because they do not consider, for example, dispersal limitations, nor species interactions. Instead, they model the suitable habitats of a species but not necessarily its realized habitats. Recently, new software have been developed to consider both the suitable habitats of a species, as described by the projected SDMs, and the dispersal limitations of a species (*e.g.* MigClim, Engler, Hordijk, & Guisan, 2012, and CATS, Lehsten et al., 2014) to predict the distribution of species through time. Such software are called "cellular automaton" since they model on a grid (the projected SDMs) how individuals of a species migrate from pixel to pixel through time, taking into account their dispersal capacities. Low dispersal capacities could eventually lead, for a given species, to not occupy some parts of its habitat predicted as suitable by the projected SDMs. These models, combining SDMs and dispersal limitations, are very accurate to predict the impact of climate changes on species distributions (Engler & Guisan, 2009).

Evolution of DNA sequences along coalescent gene genealogies

In this thesis, we used structured coalescent models to simulate gene genealogies under different LGM demographic scenarios, aiming at identifying the scenario which is the more likely to explain the spatial genetic variation observed within the sequenced genetic markers. Especially, in Paper III, we compared explicit demographic scenarios describing the post-glacial recolonization routes of European IF and IC species as referenced in the literature and presented in the introduction section of this thesis (for scenarios' description, see Fig. 2, Paper III), and in ongoing study IV, since all scenarios referenced in the literature and aiming at reconstructing the Quaternary history of the lowland Amazonian forest imply more or less important range contractions (bottlenecks) during glacial periods (see "Introduction" section of this thesis), we will compare demographic scenarios implementing different intensities of bottleneck at LGM. In both cases, in order to be able to determine the demographic scenario which is the more likely to explain the spatial genetic variation observed within the sequenced genetic markers, simulating the evolution of DNA sequences along each of the simulated coalescent gene genealogy is needed. During this process, referred to as the "mutation process", models of nucleotide substitutions are used. Mutations are randomly added to a given gene genealogy, with the probability of occurrence along a branch being proportional to its length. Since demographic coalescent simulations result in simulated gene genealogies with different branch lengths for different demographic scenarios (Fig. 5), at the end of the mutation process, gene genealogies simulated under different demographic scenarios end up with different sets of DNA sequences (Fig. 7). It is thus possible to compare the set of DNA sequences resulting from each demographic scenario with the set of observed DNA sequences of any given genetic marker to determine the scenario which is the more likely to explain the spatial genetic variation residing in the observed dataset of that given genetic marker. When more than one genetic marker is included in the study, the observed set of DNA sequences is composed of a matrix combining the sets of sequences of each genetic marker and the mutation process is repeated for each of the genetic markers included in the study.

With "bottleneck effect"

12 substitutions
5 different alleles

Without "bottleneck effect"

12 substitutions
9 different alleles

1 2 2 2 3 3 3 4 5 5 5

1 1 2 3 4 5 5 6 7 8 9

● Substitution in one site
➙ "Bottleneck effect"
1,2,...9 Different alleles
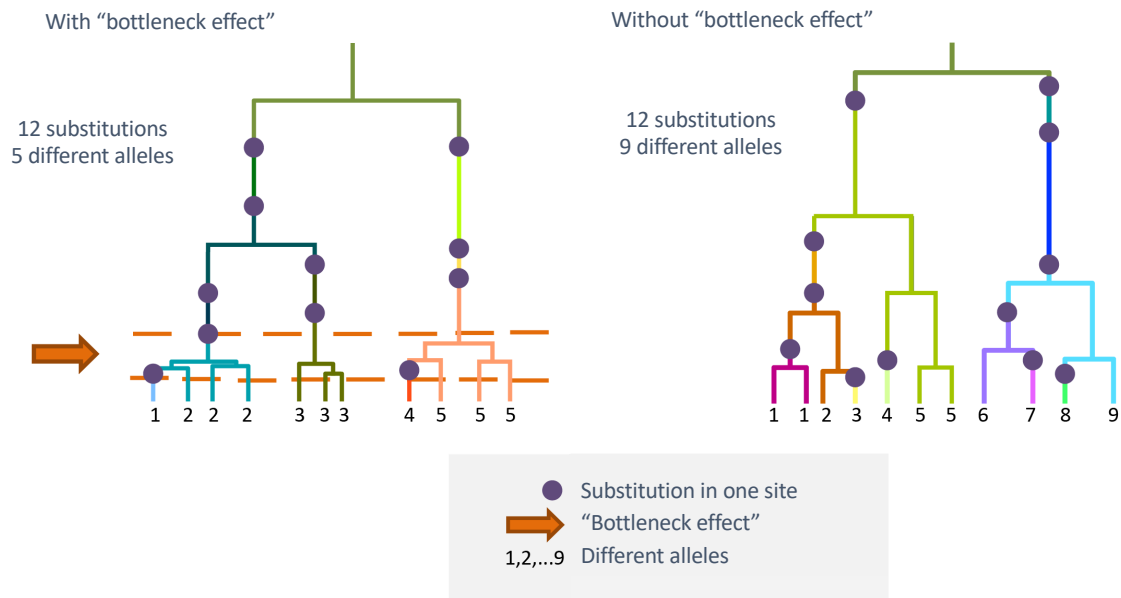
Figure 7: Putative impact of a bottleneck on the number of distinct alleles among sequences of gene copies generated along a gene genealogy built under the coalescent model. Distinct branch colors correspond to distinct alleles.

Approximate Bayesian Computation (ABC)

Major reviews on Approximate Bayesian Computation (ABC) in the context of demographic inference include Beaumont, 2010; Bertorelle, Benazzo, & Mona, 2010; and Csilléry, Blum, Gaggiotti, & François, 2010. I briefly describe here its important aspects, emphasizing the aspects used in this thesis. In order to get a statistically relevant comparison of different demographic scenarios, it is important to generate, for each demographic scenario, a large number of coalescent simulations and to use computational methods that allow a fairly rapid selection of the scenario which is the more likely to explain the spatial genetic variation observed within the sequenced genetic markers. Computing likelihood estimates for any given simulation is extremely time-consuming (Marjoram & Tavaré, 2006). Thus, to circumvent the need for likelihood computations, ABC is used (Beaumont, Zhang, & Balding, 2002). ABC is a Bayesian statistical methodology that approximates the likelihood (Csilléry et al., 2010). In this thesis, both the classic coalescent population model – implemented in Paper III – and the spatially explicit coalescent model – being implemented in ongoing study IV – are integrated within an ABC pipeline based on the basic rejection algorithm. This algorithm allows (i) to use *prior* information on data (represented as probability distribution of parameters) to generate simulations under distinct scenarios, (ii) to use descriptive statistics to summarize the observed and the simulated sets of DNA sequences, (iii) to compute Euclidean distances (based on regression equations) between the observed and each of the simulated sets of DNA sequences, (iv) to select the simulations that are closest to the observed set of DNA sequences, (v) to extract *posterior* information (represented as probability distribution of parameters) from those selected closest simulations, and (vi) to determine the best fit scenario (*i.e.* the scenario with the highest proportion of simulations among those selected closest simulations, this proportion is called the "*posterior* probability" of a given scenario, Pritchard, Seielstad, Perez-Lezaun, & Feldman, 1999). In ongoing study IV, different types of distances will be computed between the observed and the simulated sets of DNA sequences (step iii), and the type of distance that allows the best to distinguish the different scenarios will be retained for the next steps. In Paper III, step v and vi were performed using a recent formulation called "ABC-GLM", this algorithm

differs from the basic rejection algorithm especially by the selection of the best-fit model based on its Bayes Factor (BF) instead of its *posterior* probability (Leuenberger & Wegmann, 2010).

Not only does using ABC methods, instead of likelihood estimates, allow to select statistically a best-fit scenario among distinct demographic scenarios in a reasonable amount of time, but it also allows to define scenarios based on *prior* distributions of parameter values instead of fixed values and further evaluate the uncertainty associated with those parameters by refining those large *prior* distributions into smaller *posterior* distributions based on the values observed within the closest simulations. Different kinds of *prior* information on the population(s) studied and the demographic events it(they) underwent can be used to define the *prior* distribution of values for each parameter such as the information obtained from the computation of descriptive summary statistics on the observed set of DNA sequences or general knowledge of the studied organism. The *prior* distribution of values for each parameter is sampled once per simulation, resulting in one gene genealogy per combination of sampled parameter values. The number of simulations to perform for a given scenario thus have to be large enough to explore most of the parameter space (*i.e.* most of the variance which resides in all potential combinations of parameter values, Wegmann, Leuenberger, & Excoffier, 2009). Another important choice to make in ABC is the number of summary statistics to use in order to summarize the observed and the simulated sets of DNA sequences (step ii of the basic rejection algorithm). This choice should be done so that the statistics summarize the relevant information on spatial genetic variation contained within the observed and the simulated sets of DNA sequences without adding statistics that provide non-useful or redundant information (Bertorelle et al., 2010). In Paper III, we used the following set of summary statistics: expected heterozygosity (He) and number of private alleles (PrS) in each population, global and pairwise F_{st} among populations, and average nucleotide diversity within and among populations (π) (Tables 1 and 2, Paper III). In ongoing study IV, we will use spatially explicit vectors of allele frequencies as summary statistics.

Paper I
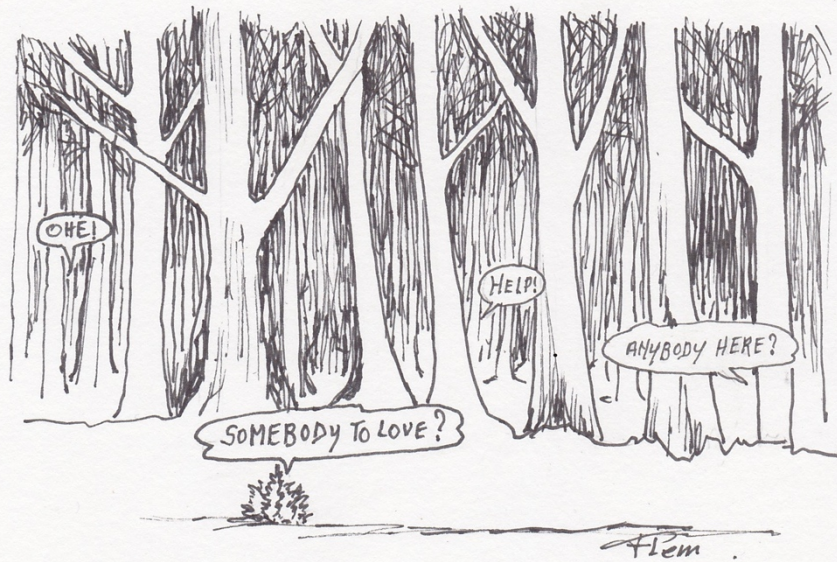
Strong spatial genetic structure in Amazonian bryophytes raises concerns about their ability to overcome fragmentation

Ledent, A.[1,2], Pereira, M.[3], Overson, R.P.[4], Laenen, B.[5], Mardulyn, P.[6], Gradstein, S.R.[7], Zartman, C.E.[3] and Vanderpoorten, A.[1]

[1]Institute of Botany, Université de Liège, Sart Tilman, 4000 Liege, Belgium
[2]Bioinformatics and genetics, Naturhistoriska riksmuseet, P. O. Box 50007, SE-104 05 Stockholm, Sweden
[3]National Institute for Amazonian Research, Department of Biodiversity, Petrópolis, CEP 69060-001, Manaus, Amazonas, Brazil
[4]Global Institute of Sustainability, Arizona State University, Tempe, AZ 85281, USA
[5]Department of Ecology, Environment, and Plant Sciences, Science for Life Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden
[6]Evolutionary Biology & Ecology, Université Libre de Bruxelles, 1050 Brussels, Belgium
[7]Systematics and Evolution, Museum National d'Histoire Naturelle, Case Postale 39, 57 rue Cuvier, 75231 Paris cedex 05, France

Abstract

- Bryophytes, and especially lowland tropical species, have been perceived as excellent dispersers. In such groups with efficient dispersal capacities, the inverse isolation hypothesis proposes that any spatial genetic structure is erased beyond the limits of short-distance dispersal (SDD), raising the hypothesis that they may overcome habitat fragmentation in a region experiencing the fastest rates of deforestation in the world. Here, we determine whether environmental variation, geographic barriers, and geographic distance contribute to explain the spatial genetic structure of Amazonian bryophyte species.

- Single Nucleotide Polymorphism data were produced from a Restriction site-Associated DNA sequencing (RADseq) protocol for 10 Amazonian bryophyte species. F-statistics and Mantel tests were employed to test the hypothesis that isolation-by-resistance (IBR), isolation-by-environment (IBE) or isolation-by-distance (IBD) shape the observed patterns of genetic variation.

- The slope of the Mantel tests between kinship coefficients and geographic distance were significant for 8 out of the 10 investigated species and remained significant beyond the scale of SDD, evidencing significant IBD patterns and invalidating the application of the inverse isolation hypothesis. The spatial genetic structures reported here in bryophytes with tiny spores of *c.* 20 μm are comparable to those documented for angiosperm species producing much larger seeds.

- In contrast to Wallace's and the gradient hypotheses of a strong imprint of the Amazonian hydrographic network and variation in soil conditions, respectively, on the genetic structure of Amazonian biota, our results are not consistent with the idea that IBR and IBE have shaped the genetic structure of bryophyte species. Most importantly, our results suggest that Amazonian bryophytes exhibit substantial spatial genetic structures that point to dispersal limitations that were not expected based on their apparently homogeneous distribution in the landscape, suggesting that even organisms perceived as extremely efficient dispersers like bryophytes are severely exposed to forest fragmentation.

Keywords: Amazonia, tropical rainforest, bryophytes, spatial genetic structure, Restriction site-Associated DNA sequencing (RADseq), isolation-by-distance (IBD).

Introduction

The extraordinary diversity in Amazonia results from long and complex interactions between geologic, climatic and ecological processes (Hoorn &Wesselingh 2010; Smith et al. 2014). By far the most impressive feature of the geography of the Amazon basin is the presence of large rivers such as the Amazon and the Rio Negro, the first two most important rivers in terms of annual discharge in the world. Such rivers may, as Wallace (1852) already hypothesized, act as barriers to gene flow between populations inhabiting opposite river banks, promoting speciation through isolation-by-resistance (IBR, McRae 2006). This hypothesis has subsequently been supported by molecular evidence in a wide range of organisms including primates, anurans, and squamates (see Ortiz et al. 2018 for review), but questioned in others characterized by higher dispersal capacities (Santorelli et al. 2018), whose distributions are expected to be more strongly associated with local environmental conditions than with geographic distance or the presence of barriers (Dambros et al. 2017). In higher plants for instance, the gradient hypothesis proposes that strong environmental gradients, like white-sand forests adjacent to other more nutrient-rich forests called "terra firme", promote isolation-by-environment (IBE, Wang and Bradburd 2014), and hence, high levels of both phylogenetic and taxonomic beta-diversity across habitats at small geographic scales (Guevara et al. 2016), but high alpha- and low beta-diversities at larger scales due to efficient dispersal across large distances (Condit et al. 2002; Latimer et al. 2005).

High alpha- and low beta-diversities are particularly expected in organisms with high dispersal capacities, counteracting isolation-by-distance (IBD) and IBR, and low levels of genetic specialization, counteracting IBE. Among land plants, such features are precisely found in bryophytes, which are characterized by their failure to adaptively radiate (Patino et al. 2014) and exhibit, due to the extremely small size of their diaspores, extremely high dispersal capacities (Patino & Vanderpoorten 2019). Using null model analyses based on metacommunity concepts for Amazonian epiphytic bryophyte communities, Mota de Oliveira & ter Steege (2015) in fact reported a correlation between floristic and geographical distances between 15 and 2835 km close to 0, which they did not consider biologically relevant, and concluded that, at the massive scale of this region of 6 million square kilometers, "epiphytic bryophytes behave as one single metacommunity". Mari et al. (2016) similarly concluded that microsite availability, as opposed to dispersal limitation, is the most important mechanism structuring white-sand forest vascular epiphyte communities. This finding has substantial consequences because it would suggest that epiphytes have the dispersal capacities to overcome the rapid fragmentation of the Amazonian rainforest. Mota de Oliveira & ter Steege's hypothesis (2015) is supported by spore-trapping experiments according to which, although spore densities quickly decrease with distance from the source, a higher proportion of spores originates from sources farther away than the nearest sources with increasing isolation (Sundberg 2005), so that the tail of the dispersal kernel, beyond 500m-1 km, is distance-independent (Lönnell et al. 2012). In such conditions, an inverse isolation effect is predicted to develop (Sundberg 2005; Barbé et al. 2016). An inverse isolation effect involves a higher genetic diversity of colonizing propagules with increasing isolation, thus counteracting differentiation. Consequently, no IBD is expected beyond a distance corresponding to short-distance dispersal (SDD) events owing to the well-mixed and diverse propagule pool (Szövényi et al. 2012). Such predictions have important consequences because rapid forest destruction in the Brazilian Amazon alone left an area 1.5 times larger than the area cleared as either forest fragments <100 km^2 in size, or as areas prone to edge effects (Skole and Tucker, 1993), so that the Amazon Basin is presently experiencing rates of absolute deforestation higher than any other region on the planet (Laurance, 1998).

In the present paper, we determine whether environmental variation, geographic barriers, and geographic distance contribute to explain the spatial genetic structure of Amazonian bryophyte species. Given the failure of bryophytes to radiate and the high dispersal capacities of the group, especially evidenced in the Amazonian context (Mota de Oliveira & ter Steege 2015), we expect that the spatial genetic structure of Amazonian bryophytes is not explained by environmental variation. Furthermore, if the inverse isolation hypothesis applies, we expect that geographic barriers do not contribute to the observed spatial genetic structure and that any signal of IBD is erased beyond the range of SDD, and hence, that Amazonian bryophytes have the dispersal capacities to overcome the rapid fragmentation of their habitat.

Material and Methods

Study area and taxonomic sampling

Ten bryophyte species, including four liverworts (*Archilejeunea fuscescens, Bazzania hookeri, Micropterygium trachyphyllum*, and *Thysananthus amazonicus*) and six mosses (*Leucobryum martianum, Octoblepharum albidum, O. pulvinatum, Syrrhopodon annotinus* (including *S. simmondsii*, as both species behave like one common gene pool, Pereira et al. 2019)*, S. helicophyllus*, and *S. hornschuchii*) were used as models. Specimens of these species were sampled within an area of *c*. 50,000 km^2 in the Rio Negro basin, north of Manaus (Fig. 1, Table S1). The Rio Negro, which is there wide of up to *c*. 20 km, forms a natural geographic barrier, which served to test for IBR. 15 to 40 specimens were sampled per species in the course of a 30 km-long transect along the Rio Salomois, a 150 km-long transect along the Rio Negro up to (and including) Jau National Park, and a 150 km-long transect between Manaus, Presidente Figueiredo and Balbina. The sampling took place in two main kinds of habitats, including low-stature, open forests on sandy soil (white-sand forests) and non-flooded tropical lowland rain forests (terra firme, Adeney et al. 2016), which provided the habitat differentiation used to test for IBE. Specimens were collected in tubes and readily dried-out in silica gel.

Molecular protocols

Samples were frozen in liquid nitrogen for 5 minutes and DNA extracted using the DNeasy Plant Mini Kit (QIAGEN). SNP libraries were prepared based on a Restriction site-Associated DNA sequencing (RADseq) protocol modified from the Genotyping-By-Sequencing (GBS) protocol described by Elshire et al. (2011) and starting with the digestion of 100ng DNA with ApeKI. Modifications included a double size selection of DNA fragments of 150-400bp using SPRI beads to only target fragments of sequenceable size; an amplification with a Q5 Hot Start High-Fidelity DNA Polymerase (NEB) to enhance specificity and reduce amplification errors; a scalable complexity reduction using longer 3′ primers that cover the entire common adapter, the 3′ restriction site and extend 1 or 2 bases into the insert, as implemented in Sonah et al. (2013); and a purification of PCR products using AMPure XP beads instead of QIAquick PCR Purification Kit (QIAGEN). We gave each individual a forward and a reverse 4-8bp-long barcode (identical, one at the 5′-end and one at the 3′-end, for paired-end sequencing), such that each individual had a unique barcode and was multiplexed with other individuals. These barcodes were selected from the 384 barcodes specifically designed to be used with ApeKI (https://www.maizegenetics.net). The concentration of PCR products was assessed by fluorometry with the Quant-iT PicoGreen dsDNA Assay Kit before multiplexing to ensure the equimolarity of PCR products in final libraries. The distribution of fragment sizes for each library was analyzed by capillary electrophoresis with a QIAxcel (QIAGEN) to look for any remaining adapter dimer (*c*. 128bp). If present, adapter dimers were removed by selecting fragments of >150bp on a polyacrylamide

gel. Paired-end sequencing (2X75bp) of the libraries was performed with an Illumina NextSeq 500 sequencer in low-output mode (*i.e.* 130,000,000 reads per lane).

Bioinformatics processing

Sequences of the adapters as well as low-quality bases (Phred score <20) at both 3'- and 5'-ends of each read were removed with cutadapt 1.16 (https://cutadapt.readthedocs.io). ipyrad 0.7.28 (https://ipyrad.readthedocs.io) was then used to demultiplex the libraries and to cluster alleles that diverged by less than 15%, within individuals into allele clusters, and then among individuals into *loci*. Following Paris et al. (2017), we set the minimum number of raw reads required to form an allele (read depth filter) to 3. Due to the haploid condition of the target species, allele clusters with more than one allele per individual were discarded. To avoid linkage among individual SNPs, one SNP per *locus* was randomly selected. We finally discarded any SNP that was sequenced in less than 10 and 25% of the individuals and then any specimen that was genotyped in less than 10% of the SNPs to produce two datamatrices (M10 and M25, respectively) to test the impact of the number of individuals, of SNPs and of missing data.

Statistical analyses

To test for IBD, we regressed pairwise kinship coefficients F_{ij} (Loiselle et al. 1995) between individuals and the logarithm of pairwise geographic distances. The regression slopes were first computed across the entire geographic range of the study. The significance of the slopes was tested by 1000 random permutations of individuals among sampling points across the entire geographic range (Mantel test). To test the application of the inverse isolation hypothesis, we then considered only pairs of individuals separated by a distance of at least 1 km. In this case, the randomization test cannot be implemented and the significance of the slopes was assessed by a Jackknife test. The slopes were thus recalculated after successively pruning one SNP from the data at a time to estimate the standard deviation of the slope across SNPs, and hence, determine whether its 95% confidence interval (*i.e.* 1.96 x the standard deviation for a significance level of 0.05) encompasses 0, in which case the slope would be considered as non-significant. To test for IBE and IBR, we assigned specimens to one of two groups (terra firme vs white-sand forest for IBE and W or E of the Rio Negro for IBR) and computed the F_{st} between the two groups. The significance of the observed F_{st} values was tested by 1000 random permutations of individuals among groups.

To compare the fine-scale genetic structure of bryophytes with that of angiosperms, we also computed the Sp statistics, which characterizes the rate of decrease of pairwise kinship coefficients between individuals with the logarithm of the distance (Vekemans & Hardy, 2004). The Sp statistics varies as a function of the mating system and dispersal traits, low values typically characterizing organisms with high dispersal capacities. The Sp statistics is measured as $-\hat{b}_F/(1 - \hat{F}_1)$, where $-\hat{b}_F$ is the regression slope on the logarithm of distance and $\hat{F}_1$ is the mean kinship coefficient between individuals belonging to the first distance interval that includes all pairs of neighbors. All the computations were performed with Spagedi 1.5d (Hardy & Vekemans, 2002).

Results

RADseq raw data were submitted to the NCBI Sequence Read Archive (SRA) with reference number PRJNA530510. Accession numbers for each individual can be found in Table S1. An average total of 17.6±13% million raw reads per individual was obtained across

species (Table 1). From those reads, the clustering of alleles diverging by a maximum of 15% within individuals led to an average of 9.7±13% million allele clusters per individual. The read depth filtering led to the loss of an average of 80±2% of the allele clusters. The filtering-out of heterozygous allele clusters led to the loss of another 15±1% of them.

The clustering of alleles clusters diverging by a maximum of 15% among individuals led to an average of 6196±2726 *loci* genotyped per individual across species. After the random selection of a single SNP per *locus*, we ended up with matrices including an average of 17,674 (min 4828, max 41866) and 782 (min 143, max 2107) SNPs for M10 and M25, respectively. In these matrices, an average of 25±6% (min 18, max 36) and of 38±3% (min 34, max 42) of the SNPs were sequenced per specimen per species for M10 and M25, respectively (Table 1).

The slope of the Mantel tests between kinship coefficients and geographic distance and their p-value, the F_{st} between populations from terra firme and white-sand forest on the one hand, and between populations separated by the Rio Negro on the other, and their p-values, along with the Sp statistics are presented in Table 2. The slope of the Mantel tests between kinship coefficients and geographic distance across the entire geographic range was significant for 8 out of the 10 investigated species. Kinship coefficients progressively decreased with geographic distance (Fig. 2). All of these eight observed significant IBD patterns across the entire geographic range remained significant (*i.e.* the 95% interval of confidence of the slope at a 0.05 significance level did not encompass 0) when the slope was computed among individuals located at more than 1 km from each other. Significant IBD tests were mostly obtained with datamatrices including a lower number of markers but also a higher number of individuals and lower amounts of missing data (<75%, M25). Significant IBD slopes were also obtained, in two cases only, with datamatrices including a higher number of SNPs but a lower number of specimens and higher levels of missing data (<90%, M10). Evidence for IBE and IBR was very weak, as significant F_{st} between populations from terra firme and white-sand forest on the one hand, and between populations separated by the Rio Negro on the other, were observed in only one and two datamatrices, respectively (Table 2). The Sp statistics was 0.012 on average across species, with a minimum of 0.006 in *S. helicophyllus* and a maximum of 0.038 in *S. annotinus* (Table 2).

Discussion

We report here the results of one of the first studies to use RADseq techniques in bryophytes (Lewis et al. 2017), wherein these techniques appear as promising tools to generate large numbers of variable genetic markers when the sequencing of targeted cpDNA or nDNA *loci* results in monomorphic datasets (Pereira et al. 2019). The number of SNPs obtained here is comparable to the 63-1397 SNPs obtained for a dataset with >75% of the individuals and *c.* 37% missing data in the only other moss species that had been investigated to date using RADseq techniques, *Tetraplodon fuegianus* (Lewis et al., 2017). These numbers are somewhat lower than those obtained in angiosperms, wherein 2000-3000 independent SNPs have been recently reported (*e.g.* Prunier et al. 2017; and Bell et al. 2018), which may reflect a truly lower genetic diversity caused, among others, by high rates of clonality typically found in bryophytes, but may also call for protocol improvements in these non-model organisms. In particular, an important drop of the number of allele clusters occurred in the present study when applying the recommended minimum number of raw reads required to form an allele (Paris et al. 2017), which may have been caused by the use of a restriction enzyme with a high frequency of target sites in the target genomes, resulting in a high number of DNA fragments but a low number of reads per fragment (low read depth). Most importantly, however, the strong genetic structures

reported here at the landscape scale (see below) unambiguously show that the data produced are suitable to test the investigated hypotheses.

In line with our predictions and in contrast to Wallace's (Ortiz et al. 2018) and the gradient (Guevara et al. 2016) hypotheses of a strong imprint of the Amazonian hydrographic network and variation in soil conditions, respectively, on the genetic structure of Amazonian biota, our results are not consistent with the idea that IBR and IBE have shaped the genetic structure of bryophyte species. While, in angiosperms, a signature of IBE was found in a meta-analysis in *c.* 20% of the cases (Sexton et al. 2014), and while, within Amazonia specifically, evidence for ecotypic differentiation between white-sand forest and terra firme populations is suggestive of an adaptive mechanism of edaphic specialization (Fine et al. 2013; Fine & Baraloto 2016), the absence of significant differences of allele frequencies between terra firme and white-sand forest populations (absence of IBE) in the investigated bryophytes is in line with the idea that bryophytes exhibit 'multi-purpose' genotypes and fail to diversify in heterogeneous environments, accounting for their failure to radiate (Patino et al. 2014). In turn, the absence of significant differences in allele frequencies between populations separated by the Rio Negro (absence of IBR) shows that Amazonian bryophytes' spores have the dispersal capacity to overcome the natural fragmentation of the landscape by the hydrographic network, in line with growing evidence that even large geographic barriers such as the Rio Negro are not an impediment for migration in groups such as birds, butterflies and vascular plants (Dambros et al. 2017; Santorelli et al. 2018).

In sharp contrast to our primary hypothesis that Amazonian bryophytes exhibit high dispersal capacities eroding any signal of genetic structure at the landscape scale, however, a significant IBD was observed in 8 out of the 10 investigated species. Our results are in line with experimental demographic studies of Amazonian leaf-inhabiting epiphytes, which point to dispersal limitation at both regional- (<50 km) (Zartman & Shaw 2006) and fine- (<20m) scales (Zartman et al. 2012) in understory habitats, but strongly challenge the perception, based on the homogeneous distribution of species in the landscape, that Amazonian bryophyte species "behave as one single metacommunity" (Mota de Oliveira & ter Steege 2015).

In further contradiction with the inverse isolation hypothesis, IBD consistently remained significant beyond the range of SDD, evidencing long-distance dispersal (LDD) limitations. In a recent meta-analysis of the decay of the IBD signal caused by LDD events in bryophytes, Vanderpoorten et al. (2019) showed that LDD capacities were sufficient to erase any IBD signal beyond 1000m from the source in 30-50% of the cases. The consistent persistence of the IBD signal reported here thus suggests that Amazonian bryophyte species experience more dispersal limitations than species from other biomes.

As a comparison, the average Sp statistics across species of 0.012 lays in the range reported for angiosperm species characterized by wind dispersal (0.012±0.012), while the maximal and minimal values of the observed Sp (0.038 in *S. annotinus* and 0.006 in *S. helicophyllus*) are in the range reported for species with gravity-dispersed seeds (0.028±0.016) and animal-dispersed seeds (0.008±0.005), respectively (Vekemans & Hardy, 2004). It is, in fact, striking to consider that, within the same Amazonian environment, the average Sp in bryophytes, whose spores measure *c.* 20 μm, is actually comparable to that of the Brazil nut tree *Bertholletia excelsa* (Sp=0.01-0.03), wherein no significant fine-scale genetic structures were revealed in 5 out of 9 investigated populations (Sujii et al. 2015), and whose seeds are enclosed within a 10-16 cm globose, functionally indehiscent woody capsule whose dispersal by scatterhoarding agoutis and acouchis, and occasionally squirrels, is restricted to a few hundreds of meters (Thomas et al. 2014).

The strong IBD pattern revealed here is consistent with the dispersal traits of the studied species, together with extrinsic features of their environment of dense rainforests that is not prone to long-distance wind-dispersal. These dispersal traits include the absence of male expression (*S. annotinus*, Pereira, Dambros, & Zartman, 2016), the prevalence of dioicy in Calymperaceae associated with low sporophyte production (Pereira et al., 2019), the immersion of the sporophytes within perichaetial leaves or very short setae (*S. annotinus* and *S. helicophyllus*), and the absence of reduction of the peristome (*S. annotinus* and *S. hornschuchii*).

Altogether, our results thus suggest that Amazonian bryophytes exhibit substantial spatial genetic structures that point to dispersal limitations that were not expected based on their apparently homogeneous distribution in the landscape. Although these results would need to be complemented by analyses that allow for actual estimates of dispersal rates, such as spatially explicit coalescent analyses, they clearly suggest that even organisms perceived as extremely efficient dispersers like bryophytes are severely exposed to forest fragmentation.

Data accessibility

RADseq raw data (fastq files) were submitted to the NCBI Sequence Read Archive (SRA) with bioproject accession number PRJNA530510. Accession numbers for each individual can be found in Table S1.

Author contributions

AV designed the study. AL, MP, CZ, SG, and AV sampled and prepared the specimens. AL and RO produced the libraries. AL composed the SNPs datamatrices. AL, BL, and AV analyzed the data. AV wrote the manuscript with the assistance of all co-authors.

References

Adeney JM, Christensen N, Vincenti A & Cohn-haft M. 2016. White-sand ecosystems in Amazonia. Biotropica 48: 7–23.

Barbé M, Fenton NJ & Bergeron Y. 2016. So close and yet so far away: long-distance dispersal events govern bryophyte metacommunity reassembly. Journal of Ecology, 104, 1707–1719.

Bell N, Griffin PC, Hoffmann AA & Miller AD. 2018. Spatial patterns of genetic diversity among Australian alpine flora communities revealed by comparative phylogenomics. Journal of Biogeography 45: 177-189.

Condit R, Pitman N, Egbert GL Jr, Chave J, Terborgh J, Foster RB, Núñez P, Aguilar S, Valencia R, Villa G, Muller-Landau HC, Losos E & Hubbell SP. 2002. Beta-diversity in tropical forest trees. Science 295:666-669.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES & Mitchell SE. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6: e19379.

Dambros CL. et al. 2017. Isolation by distance. not rivers. control the distribution of termite species in the Amazonian rain forest. Ecography 40: 1242–1250. 2017.

Fine PVA & Baraloto C. 2016. Habitat endemism in white-sand forests: Insights into the mechanisms of lineage diversification and community assembly of the Neotropical flora. Biotropica 48: 24–33.

Fine PVA, Zapata F, Daly DC, Mesones I, Misiewicz TM, Cooper HF & Barbosa CEA. 2013. The importance of environ- mental heterogeneity and spatial distance in generating phylogenetic structure in edaphic specialist and generalist tree species of Protium (Burseraceae) across the Amazon Basin. J. Biogeogr. 40: 646–661.

Guevara JE et al. 2016. Low Phylogenetic Beta Diversity and Geographic Neo-endemism in Amazonian White-sand Forests. BIOTROPICA 48(1): 34–46 2016.

Hardy OJ & Vekemans X. 2002. SPAGeDi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2: 618–620.

Hoorn C & Wesselingh FP. 2010. Amazonia: landscape and species evolution. A look into the past. Wiley, Oxford.

Latimer AM, Silander JA Jr & Cowling RM. 2005. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. Nature 309: 1722-1725.

Laurance WF. 1998. A crisis in the making: responses of Amazonian forests to land use and climate change. Trends in Ecology and Evolution 13, 411–412.

Lewis LR, Biersma EM, Carey SB, Holsinger K, McDaniel SF, Rozzi R & Goffinet B. 2017. Resolving the northern hemisphere source region for the long-distance dispersal event that gave rise to the South American endemic dung moss Tetraplodon fuegianus. American Journal of Botany 104: 1651–1659.

Loiselle BA, Sork VL, Nason J & Graham C. 1995. Spatial genetic structure of a tropical understory shrub, Psychotria officinalis (Rubiaceae). American Journal of Botany 82:1420–1425.

Mari MLG, Toledo JJ, Nascimento HEM & Zartman CE. 2016. Regional and fine scale variation of holoepiphyte community structure in Central Amazonian white-sand forests. Biotropica 48: 70-80.

McRae, BH. 2006. Isolation by resistance. Evolution 60:1551–1561.

Mota de Oliveira S & ter Steege H. 2015. Bryophyte communities in the Amazon forest are regulated by height on the host tree and site elevation. J. Ecol. 102: 441-450.

Ortiz DA et al. 2018. Environmental transition zone and rivers shape intraspecific population structure and genetic diversity of an Amazonian rain forest tree frog. Evol Ecol (2018) 32:359–378.

Paris JR, Stevens JR & Catchen JM. 2017. Lost in parameter space: a road map for STACKS. Methods in Ecology and Evolution 8: 1360-1373.

Patiño J & Vanderpoorten A. 2019. Bryophyte biogeography. Critical Reviews in Plant Sciences 37: 175-209.

Patiño J, Carine MA, Fernández-Palacios JM, Otto R, Schaefer H & Vanderpoorten A. 2014. The anagenetic world of spore-producing land plants. New Phytologist 201: 305-311

Pereira MR, Dambros CS & Zartman CE. 2016. Prezygotic resource-allocation dynamics and reproductive trade-offs in Calymperaceae (Bryophyta). American Journal of Botany 103:1-9.

Pereira MR, Ledent A, Mardulyn P, Zartman CE & Vanderpoorten A. Maintenance of genetic and morphological identity in two sibling Syrrhopodon species (Calymperaceae, Bryopsida) despite extensive introgression. Journal of Systematics and Evolution, in press.

Prunier R, Akman M, Kremer CT, Aitken N, Chuah A, Borevitz J & Holsinger KE. 2017. Isolation by distance and isolation by environment contribute to population differentiation in Protea repens (Proteacee L.), a widespread South African species. American Journal of Botany 104: 674–684.

Santorelli S, Magnusson WE & Deus CP. 2018. Most species are not limited by an Amazonian river postulated to be a border between endemism areas. Sci Rep 8:2294

Sexton JP, Hangartner SB & Hoffmann AA. 2014. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution, 68, 1–15.

Skole D & Tucker CJ. 1993. Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. Science 260, 1905–1910.

Smith BT, McCormack JE & Cuervo AM et al. 2014. The drivers of tropical speciation. Nature 515:406–409.

Sonah H, Bastien M, Iquira E, Tardivel A, Légaré G, Boyle B, Normandeau E, Laroche J, Larose S, Jean M & Belzile F. 2013. An improved Genotyping by Sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. PLoS ONE 8: e54603.

Sujii PS, Martins K, de Oliveira Wadt LH, Renno Azevedo VC & Solferini VN. 2015. Genetic structure of Bertholletia excelsa populations from the Amazon at different spatial scales. Conserv Genet 16:955–964.

Sundberg S. 2005. Larger capsules enhance short-range spore dispersal in Sphagnum, but what happens further away? Oikos, 108, 115–124.

Szövényi P, Sundberg S, & Shaw AJ. 2012. Long-distance dispersal and genetic structure of natural populations: an assessment of the inverse isolation hypothesis in peat mosses. Molecular Ecology, 21, 5461–5472.

Thomas E, Alcázar Caicedo C, Loo J & Kindt R. 2014. The distribution of the Brazil nut (Bertholletia excelsa) through time: from range contraction in glacial refugia, over human-mediated expansion, to anthropogenic climate change. Bol. Mus. Para. Emílio Goeldi. Cienc. Nat., Belém, 9(2): 267-291.

Vanderpoorten A, Patiño J, Désamoré A, Laenen B, Gorski P, Papp B, Hola E, Korpelainen H & Hardy OJ. 2019. To what extent are bryophytes efficient dispersers? Journal of Ecology, in press.

Vekemans X & Hardy OJ. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular Ecology 13: 921–935.

Wallace AR. 1852. On the monkeys of the Amazon. Proc Zool Soc Lond 20:107–110.

Wang IJ & Bradburd GS. 2014. Isolation by environment. Mol. Ecol. 23:5649–5662.

Zartman CE. & Shaw AJ. 2006. Metapopulation extinction thresholds in rain forest remnants. Am. Nat. 167: 177–189.

Zartman CE, Nascimento HEM, Cangani KG, Alvarenga LDP & Snäll T. 2012. Fine-scale changes in connectivity affect the metapopulation dynamics of a bryophyte confined to ephemeral patches. J. Ecol. 100: 980-986.
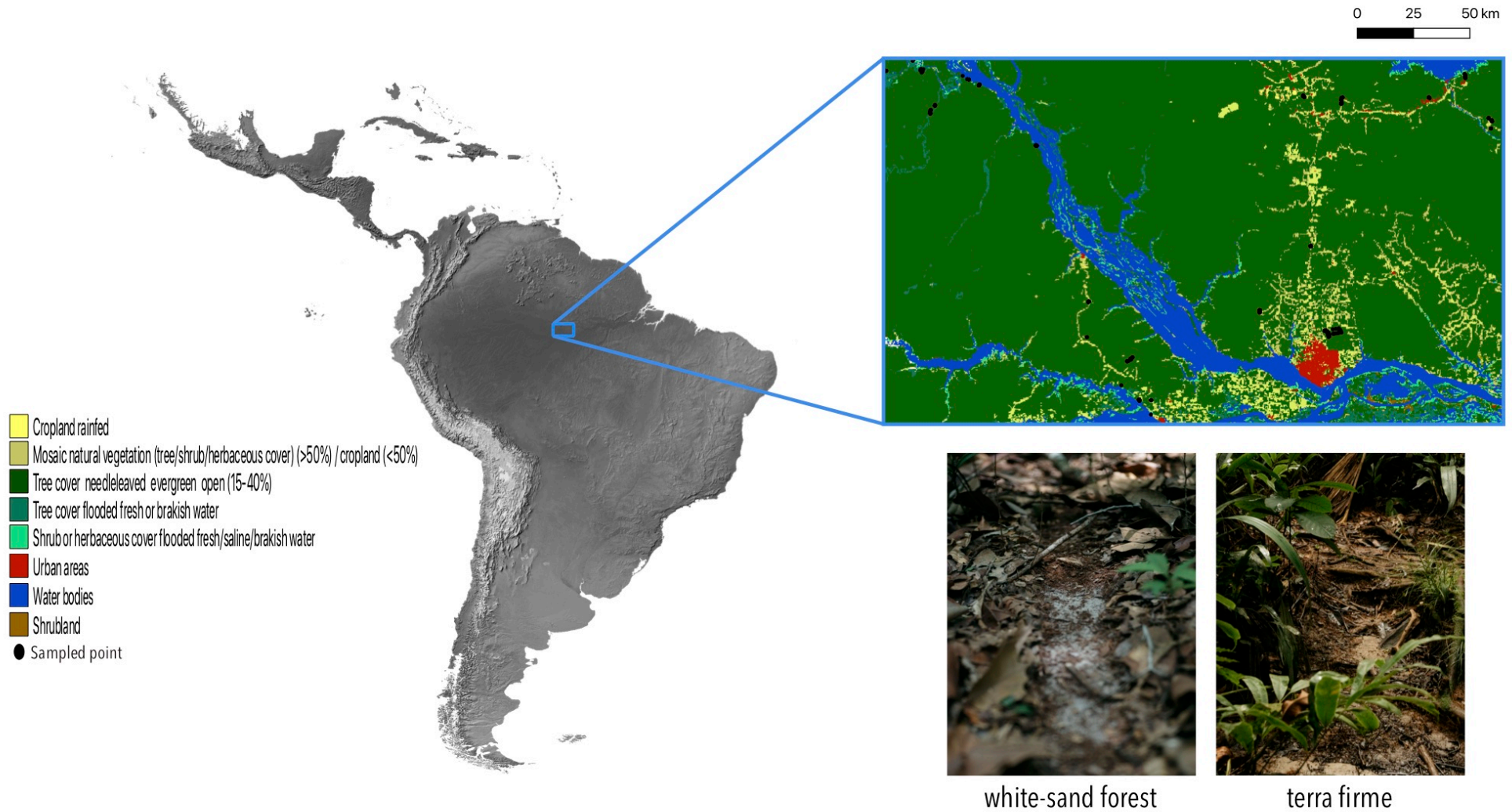
Fig. 1. Study area and sampling design.
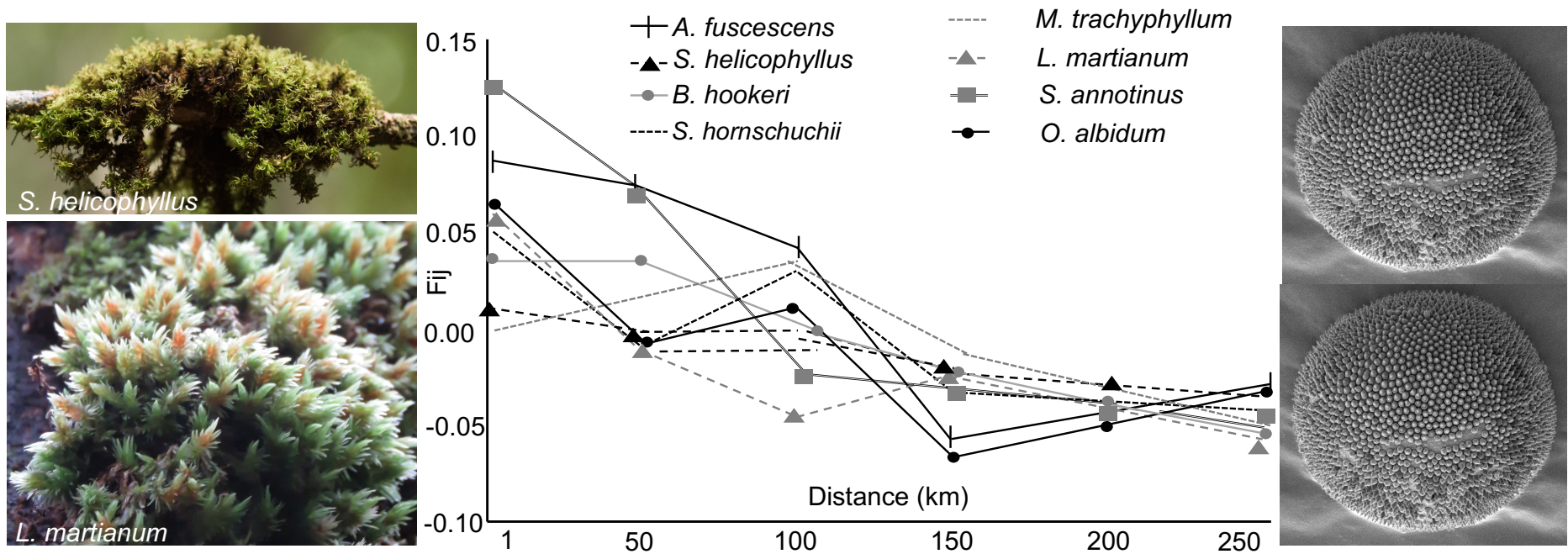
Fig. 2. Spatial autocorrelogram of the variation of kinship coefficients F_{ij} derived from SNP variation between pairs of individuals (computed from datamatrices with >25% sequenced specimens per SNP, M25) as a function of geographic distance in 8 bryophyte species showing significant isolation-by-distance slopes (see Table 2 for slope and p-values) in central Amazonia.

Table 1. Average number of raw reads, number of allele clusters, percentage of allele clusters filtered out to ensure a minimum of 3 reads per allele (% Read depth), percentage of heterozygous allele clusters filtered out (% Heterozygous) and number of *loci*, among individuals, along with number of SNPs (N), average percentage of sequenced SNPs per individual (%SNPs) and number of individuals (n) in datamatrices with >10 and >25% sequenced specimens per SNP (M10 and M25, respectively) for 10 Amazonian bryophyte species, as well as mean, standard deviation (SD) and percentage of the standard deviation (%SD) of those same variables among species.

	Raw reads	Allele clusters	% Read depth	% Heterozygous	Loci	N(M25)	N(M10)	%SNPs(M25)	%SNPs(M10)	n(M25)	n(M10)
A. fuscescens	1474403	794027	78	16	7814	325	33284	38	18	38	23
B. hookeri	2188823	1182346	79	14	5278	860	14648	34	29	23	14
L. martianum	1767423	992325	80	14	10909	143	12162	40	22	41	29
M. trachyphyllum	1677308	945706	79	16	3157	2107	6820	42	36	13	12
O. albidum	1647361	917278	81	15	4910	268	17666	40	26	26	15
O. pulvinatum	1783417	964060	81	15	10111	453	41866	40	22	31	18
S. annotinus	1538315	770311	76	14	5825	132	10339	37	21	40	25
S. helicophyllus	2076985	1136967	82	14	6811	984	22168	34	19	27	24
S. hornschuchii	1691457	960142	80	15	4605	892	12954	36	24	22	18
T. amazonicus	1756805	1040260	81	15	2537	1660	4828	39	30	15	15
MEAN	1760230	970343	80	15	6196	782	17674	38	25	28	19
SD	221087	130276	2	1	2755	666	11774	3	6	10	6
%SD	13	13	2	5	44	85	67	7	23	36	29

Table 2. Summary statistics of the genetic structure of 10 Amazonian bryophyte species. b-log and b(1)-log are the slopes of the regression between F_{ij} and geographic distance across the entire range (and their p-value) and at distances >1 km (and their standard deviation, SD), respectively. FST_{IBR} and FST_{IBR} are the F_{st} (and their p-value) between populations separated by the Rio Negro and occurring on different forest types (terra firme vs white-sand forest), respectively. sp is the Sp statistics (Vekemans & Hardy 2004), the mean Sp among species is also provided (MEAN). M10 and M25 refer to the datamatrices with >10 and >25% sequenced specimens per SNP, respectively. Significant statistics are in bold. NA: missing data.

	b-log (p)	b(1)-log±SD	$FST_{IBR}(p)$	$FST_{IBE}(p)$	Sp
A.fuscescens (M10)	0.0005 (p=0.508)	**-0.0062±0.0016**	0.00 (p=0.456)	**0.14 (p=0.004)**	NA
A.fuscescens (M25)	**-0.0189 (p=0.000)**	**-0.0241±0.0083**	0.03 (p=0.122)	0.01 (p=0.229)	0.021
B.hookeri (M10)	-0.0039 (p=0.303)	0.0069±0.0041	0.00 (p=0. 519)	NA	NA
B.hookeri (M25)	**-0.0181 (0.008)**	**-0.0340±0.0118**	0.00 (p=0. 463)	NA	0.019
L.martianum (M10)	**-0.0150 (p=0.005)**	**-0.0120±0.0027**	0.02 (p=0.118)	0.01 (p=0.344)	0.016
L.martianum (M25)	-0.0111 (p=0.150)	-0.0084±0.0198	**0.06 (p=0.004)**	0.00 (p=0.623)	NA
M.trachyphyllum (M10)	**-0.0255 (p=0.004)**	**-0.0367±0.0030**	0.01 (p=0.380)	0.00 (p=0.515)	-0.026
M. trachyphyllum (M25)	**-0.0153 (p=0.012)**	**-0.0273±0.0056**	0.01 (p=0.443)	0.00 (p=0.515)	0.015
O.albidum (M10)	-0.0076 (p=0.232)	-0.0094±0.0083	0.00 (p=0.411)	0.00 (p=0.898)	NA
O.albidum (M25)	**-0.0240 (p=0.01)**	--0.0320±0.0200	-0.0192 (p=0.6953)	0.00 (p=0.972)	0.0188
O.pulvinatum (M10)	-0.0086 (p=0.178)	-0.0100±0.0016	0.01 (p=1.000)	**0.29 (p=0.002)**	NA
O.pulvinatum (M25)	0.0023 (p=0.644)	0.0108±0.0067	0.00 (p=0.445)	0.02 (p=0.240)	NA
S.annotinus (M10)	**-0.0108 (p=0.033)**	**-0.0370±0.0100**	0.01 (p=0.336)	NA	0.011
S.annotinus (M25)	**-0.0337 (p=0.000)**	**-0.0432±0.0164**	0.03 (p=0.118)	NA	0.038
S.helicophyllus (M10)	-0.0003 (p=0.444)	0.0056±0.0026	0.00 (p=0.570)	NA	NA
S. helicophyllus (M25)	**-0.0064 (p=0.023)**	**-0.0135±0.0084**	0.00 (p=0.550)	NA	0.0065
S.hornschuchii (M10)	**-0.0089 (p=0.030)**	**-0.0065±0.0018**	0.00 (p=0.627)	0.02 (p=0.234)	0.0093
S.hornschuchii (M25)	0.0005 (p=0.578)	-0.0016±0.0036	0.00 (p=0.447)	0.00 (p=0.477)	NA
T.amazonicus (M10)	0.0022 (p=0.636)	0.0030±0.0118	0.00 (p=0.704)	NA	NA
T.amazonicus (M25)	0.0798 (p=0.991)	0.0798±0.0153	0.00 (p=0.704)	NA	NA
MEAN					0.012

Table S1. Voucher information and Sequence Read Archive (SRA) biosample accession numbers from bioproject PRJNA530510. All vouchers are hosted at the herbaria of the University of Liege (ULiege) and of the *Instituto Nacional de Pesquisas da Amazônia* (INPA). WSF: white-sand forest; TF: terra firme; NA: missing data.

Collection id	Latitude	Longitude	Environment	Biosample accession
Archilejeunea fuscescens				
104,4	-2.952	-59.957	WSF	SAMN11316664
110,1	-2.950	-59.956	WSF	SAMN11316665
122,3	-2.941	-59.938	WSF	SAMN11316666
125,1	-2.939	-59.929	TF	SAMN11316667
135,1	-2.927	-59.911	TF	SAMN11316668
144,6	-3.209	-60.672	WSF	SAMN11316669
148,2	-3.045	-60.753	WSF	SAMN11316670
148,3	-3.045	-60.753	WSF	SAMN11316671
154,3	-3.206	-60.721	TF	SAMN11316672
161,10	-2.185	-61.131	WSF	SAMN11316673
164,1	-2.183	-61.134	TF	SAMN11316674
165,4	-2.813	-60.924	WSF	SAMN11316675
172,5A	-1.919	-61.404	WSF	SAMN11316676
172,5B	-1.919	-61.404	WSF	SAMN11316677
182,6	-1.884	-61.737	TF	SAMN11316678
185,6	-1.842	-61.629	TF	SAMN11316679
186,1	-1.842	-61.629	TF	SAMN11316680
187,1	-1.842	-61.629	TF	SAMN11316700
21,1	-1.911	-59.409	TF	SAMN11316701
220,4A	-2.054	-61.558	WSF	SAMN11316702
222,1	-2.055	-61.558	WSF	SAMN11316703
225,1	-2.055	-61.559	WSF	SAMN11316704
227,6	-2.024	-61.540	TF	SAMN11316751
27,2	-1.903	-59.409	WSF	SAMN11316750
28,2	-1.901	-59.409	WSF	SAMN11316869
38,4	-2.074	-59.315	TF	SAMN11316870
42,1	-2.087	-59.301	WSF	SAMN11316871
44,4	-2.116	-59.305	WSF	SAMN11316872
46,1	-1.985	-60.060	WSF	SAMN11316873
49,2B	-1.989	-60.056	WSF	SAMN11316874
56,2	-2.014	-59.906	WSF	SAMN11316875
58,1	-2.014	-59.905	WSF	SAMN11316876
59,1	-2.013	-59.905	WSF	SAMN11316877
67,3	-1.991	-59.902	WSF	SAMN11316878
68,1	-1.993	-59.903	TF	SAMN11316879
77,3	-2.589	-60.030	WSF	SAMN11316880
78,6	-2.590	-60.029	WSF	SAMN11316881
81,3	-2.929	-59.974	TF	SAMN11316882
94,2	-2.934	-59.961	NA	SAMN11316883
97,1	-2.939	-59.962	NA	SAMN11316884
97,2b	-2.939	-59.962	NA	SAMN11316885
98,1	-2.948	-59.961	NA	SAMN11316886
99,2	-2.951	-59.958	WSF	SAMN11316887

Collection id	Latitude	Longitude	Environment	Biosample accession
Bazzania hookeri				
104,3	-2.952	-59.957	WSF	SAMN11316888
110,4	-2.950	-59.956	WSF	SAMN11316889
122,1	-2.941	-59.938	WSF	SAMN11316890
145,1	-3.052	-60.768	WSF	SAMN11316891
165,7	-2.813	-60.924	WSF	SAMN11316892
178,3	-1.943	-61.363	WSF	SAMN11316893
188,1	-1.842	-61.629	TF	SAMN11316894
205,3A	-1.886	-61.591	WSF	SAMN11316895
214,4	-2.052	-61.559	WSF	SAMN11316896
215,1	-2.052	-61.559	WSF	SAMN11316897
215,1BIS	-2.052	-61.559	WSF	SAMN11316898
221,6	-2.054	-61.558	WSF	SAMN11316899
234,2	-2.847	-60.234	WSF	SAMN11316900
26,2	-1.904	-59.409	WSF	SAMN11316901
34,2	-1.903	-59.408	WSF	SAMN11316902
37,4	-1.991	-59.555	WSF	SAMN11316903
41,3	-2.083	-59.304	WSF	SAMN11316904
46,11	-1.985	-60.060	WSF	SAMN11316905
47,2B	-1.986	-60.059	WSF	SAMN11316906
53,2	-2.015	-59.908	WSF	SAMN11316907
69,1	-1.992	-59.906	TF	SAMN11316908
76,3	-2.588	-60.031	WSF	SAMN11316909
78,1	-2.590	-60.029	WSF	SAMN11316910
79,1	-2.589	-60.029	WSF	SAMN11316911
97,2a	-2.939	-59.962	NA	SAMN11316912
99,1	-2.951	-59.958	WSF	SAMN11316913

Collection id	Latitude	Longitude	Environment	Biosample accession
Leucobryum martianum				
104,2	-2.952	-59.957	WSF	SAMN11316929
105,5	-2.953	-59.958	WSF	SAMN11316930
110,5	-2.950	-59.956	WSF	SAMN11316931
111,1	-2.948	-59.956	WSF	SAMN11316932
122,5	-2.941	-59.938	WSF	SAMN11316933
144,5	-3.209	-60.672	WSF	SAMN11316934
145,2	-3.052	-60.768	WSF	SAMN11316935
145,6	-3.052	-60.768	WSF	SAMN11316936
149,5	-3.042	-60.754	WSF	SAMN11316937
154,1	-3.206	-60.721	TF	SAMN11316938
156,3	-3.207	-60.719	WSF	SAMN11316939
163,4	-2.184	-61.135	TF	SAMN11316940
165,5	-2.813	-60.924	WSF	SAMN11316941
166,2	-2.812	-60.922	WSF	SAMN11316942
167,1B	-2.811	-60.921	WSF	SAMN11316943
172,2	-1.919	-61.404	WSF	SAMN11316944
177,2	-1.943	-61.363	WSF	SAMN11316945
178,6	-1.943	-61.363	WSF	SAMN11316946
18,2	-1.913	-59.408	TF	SAMN11316947
184,2	-1.884	-61.737	TF	SAMN11316948
186,2	-1.842	-61.629	TF	SAMN11316949
187,2	-1.842	-61.629	TF	SAMN11316950
195,1C	-1.881	-61.594	WSF	SAMN11316951
195,3	-1.881	-61.594	WSF	SAMN11316952
197,2	-1.881	-61.599	WSF	SAMN11316953
199,2	-1.879	-61.598	WSF	SAMN11316954
210,9	-2.044	-61.558	WSF	SAMN11316955
214,5	-2.052	-61.559	WSF	SAMN11316956
216,7	-2.052	-61.559	WSF	SAMN11316957
217,3B	-2.053	-61.559	WSF	SAMN11316958
217,5	-2.053	-61.559	WSF	SAMN11316959
221,9	-2.054	-61.558	WSF	SAMN11316960
223,2	-2.056	-61.559	WSF	SAMN11316961
225,2	-2.055	-61.559	WSF	SAMN11316962
227,4	-2.024	-61.540	TF	SAMN11316963
228,1	-2.023	-61.540	TF	SAMN11316964
230,2	-2.022	-61.539	TF	SAMN11316965
231,2	-2.846	-60.235	TF	SAMN11316966
234,1	-2.847	-60.234	WSF	SAMN11316967
234,1BIS	-2.847	-60.234	WSF	SAMN11316968
239,2	-2.856	-60.235	WSF	SAMN11316969
24,1	-1.904	-59.410	TF	SAMN11316970
26,1Bis	-1.904	-59.409	WSF	SAMN11316971
36,1	-1.992	-59.554	WSF	SAMN11316972
38,2	-2.074	-59.315	TF	SAMN11316973
41,2	-2.083	-59.304	WSF	SAMN11316974
44,3	-2.116	-59.305	WSF	SAMN11316975
46,5	-1.985	-60.060	WSF	SAMN11316976
55,1	-2.015	-59.906	WSF	SAMN11316977
61,1	-2.012	-59.906	WSF	SAMN11316978
65,1	-2.003	-59.907	WSF	SAMN11316979
67,1	-1.991	-59.902	WSF	SAMN11316980
68,4	-1.993	-59.903	TF	SAMN11316981
78,4	-2.590	-60.029	WSF	SAMN11316982
79,3	-2.589	-60.029	WSF	SAMN11316983
87,1	-2.927	-59.969	TF	SAMN11316984
P423A	-2.046	-61.559	WSF	SAMN11316985

Collection id	Latitude	Longitude	Environment	Biosample accession
Micropterygium trachyphyllum				
127,1	-2.938	-59.918	TF	SAMN11316986
140,2	-2.919	-59.917	TF	SAMN11316987
167,2	-2.811	-60.921	WSF	SAMN11316988
178,2	-1.943	-61.363	WSF	SAMN11316989
195,2	-1.881	-61.594	WSF	SAMN11316990
205,3B	-1.886	-61.591	WSF	SAMN11316991
214,1	-2.052	-61.559	WSF	SAMN11316992
216,6	-2.052	-61.559	WSF	SAMN11316993
217,1	-2.053	-61.559	WSF	SAMN11316994
218,1	-2.053	-61.559	WSF	SAMN11316995
220,1	-2.054	-61.558	WSF	SAMN11316996
227,1	-2.024	-61.540	TF	SAMN11316997
230,1	-2.022	-61.539	TF	SAMN11316998
240,2	-2.856	-60.235	WSF	SAMN11316999
27,1	-1.903	-59.409	WSF	SAMN11317000
41,6	-2.083	-59.304	WSF	SAMN11317001
54,1	-2.016	-59.907	WSF	SAMN11317002
56,1	-2.014	-59.906	WSF	SAMN11317003
68,3	-1.993	-59.903	TF	SAMN11317004
96,1	-2.938	-59.962	TF	SAMN11317005

Collection id	Latitude	Longitude	Environment	Biosample accession
Octoblepharum albidum				
100,2	-2.950	-59.958	WSF	SAMN11353161
105,4	-2.953	-59.958	WSF	SAMN11353162
109,1	-2.951	-59.956	TF	SAMN11353163
117,1	-2.942	-59.957	TF	SAMN11353164
122,7	-2.941	-59.938	WSF	SAMN11353165
144,11	-3.209	-60.672	WSF	SAMN11353166
148,1	-3.045	-60.753	WSF	SAMN11317006
151,2	-3.037	-60.747	WSF	SAMN11317007
158,1	-3.148	-60.791	NA	SAMN11317008
159,1	-3.211	-60.718	TF	SAMN11317009
160,1	-3.266	-60.672	TF	SAMN11317010
161,3	-2.185	-61.131	WSF	SAMN11317011
162,1	-2.186	-61.135	TF	SAMN11317012
165,1B	-2.813	-60.924	WSF	SAMN11317013
182,1	-1.884	-61.737	TF	SAMN11317014
183,1	-1.884	-61.737	TF	SAMN11317015
185,2	-1.842	-61.629	TF	SAMN11317016
189,1	-1.878	-61.588	WSF	SAMN11317017
19,1	-1.912	-59.408	TF	SAMN11317018
202,1	-1.891	-61.594	WSF	SAMN11317019
210,5	-2.044	-61.558	WSF	SAMN11317020
212,2	-2.046	-61.558	WSF	SAMN11317021
214,3	-2.052	-61.559	WSF	SAMN11317022
224,2	-2.056	-61.559	WSF	SAMN11317023
226,1	-2.054	-61.559	WSF	SAMN11317024
227,5	-2.024	-61.540	TF	SAMN11317025
23,1	-1.906	-59.410	TF	SAMN11317026
35,1	-1.994	-59.552	TF	SAMN11317027
38,1	-2.074	-59.315	TF	SAMN11317028
43,1	-2.085	-59.303	WSF	SAMN11317029
44,1	-2.116	-59.305	WSF	SAMN11317030
46,4	-1.985	-60.060	WSF	SAMN11317031
81,1	-2.929	-59.974	TF	SAMN11317032
84,1	-2.928	-59.972	TF	SAMN11317033
95,1	-2.935	-59.961	NA	SAMN11317034

Collection id	Latitude	Longitude	Environment	Biosample accession
Octoblepharum pulvinatum				
131,1	-2.936	-59.907	TF	SAMN11317035
144,2	-3.209	-60.672	WSF	SAMN11317036
147,1	-3.045	-60.754	WSF	SAMN11317037
150,1	-3.038	-60.748	WSF	SAMN11317038
153,1	-2.955	-60.929	TF	SAMN11317039
156,2	-3.207	-60.719	WSF	SAMN11317040
161,9	-2.185	-61.131	WSF	SAMN11317041
163,2	-2.184	-61.135	TF	SAMN11317042
165,2	-2.813	-60.924	WSF	SAMN11317043
172,3	-1.919	-61.404	WSF	SAMN11317044
172,4A	-1.919	-61.404	WSF	SAMN11317045
172,4B	-1.919	-61.404	WSF	SAMN11317046
173,1	-1.918	-61.410	WSF	SAMN11317047
176,1	-1.939	-61.361	TF	SAMN11317048
177,1	-1.943	-61.363	WSF	SAMN11317049
18,3	-1.913	-59.408	TF	SAMN11317050
182,5	-1.884	-61.737	TF	SAMN11317051
182,8	-1.884	-61.737	TF	SAMN11317052
185,4	-1.842	-61.629	TF	SAMN11317053
186,3	-1.842	-61.629	TF	SAMN11317054
190,5	-1.878	-61.589	WSF	SAMN11317055
195,1B	-1.881	-61.594	WSF	SAMN11317056
201,4	-1.876	-61.595	WSF	SAMN11317057
210,10	-2.044	-61.558	WSF	SAMN11317058
214,6	-2.052	-61.559	WSF	SAMN11317059
217,3A	-2.053	-61.559	WSF	SAMN11317060
221,3	-2.054	-61.558	WSF	SAMN11317061
223,1	-2.056	-61.559	WSF	SAMN11317062
225,3	-2.055	-61.559	WSF	SAMN11317063
231,3	-2.846	-60.235	WSF	SAMN11317064
240,3	-2.856	-60.235	WSF	SAMN11317065
25,1	-1.904	-59.410	WSF	SAMN11317066
37,2	-1.991	-59.555	WSF	SAMN11317067
38,3	-2.074	-59.315	TF	SAMN11317068
46,8	-1.985	-60.060	WSF	SAMN11317069
55,2	-2.015	-59.906	WSF	SAMN11317070
68,2	-1.993	-59.903	TF	SAMN11317071
70,1	-1.993	-59.906	TF	SAMN11317072
76,4	-2.588	-60.031	WSF	SAMN11317073
81,2	-2.929	-59.974	TF	SAMN11317074
P420A	-2.045	-61.559	WSF	SAMN11317075
P423B	-2.046	-61.559	WSF	SAMN11317076

Collection id	Latitude	Longitude	Environment	Biosample accession
Syrrhopodon annotinus				
144,10	-3.209	-60.672	WSF	SAMN11317077
144,14	-3.209	-60.672	WSF	SAMN11317078
144,9	-3.209	-60.672	WSF	SAMN11317079
145,3	-3.052	-60.768	WSF	SAMN11317080
161,5	-2.185	-61.131	WSF	SAMN11317081
165,1A	-2.813	-60.924	WSF	SAMN11317082
173,4	-1.918	-61.410	WSF	SAMN11317083
178,1	-1.943	-61.363	WSF	SAMN11317084
205,4	-1.886	-61.591	WSF	SAMN11317085
212,1	-2.046	-61.558	WSF	SAMN11317086
217,4	-2.053	-61.559	WSF	SAMN11317087
231,9	-2.846	-60.235	WSF	SAMN11317088
232,2	-2.846	-60.235	WSF	SAMN11317089
233,1	-2.846	-60.234	WSF	SAMN11317090
238,4	-2.855	-60.235	WSF	SAMN11317091
28,1	-1.901	-59.409	WSF	SAMN11317092
31,3	-1.897	-59.412	WSF	SAMN11317093
41,7	-2.083	-59.304	WSF	SAMN11317094
41,8	-2.083	-59.304	WSF	SAMN11317095
43,2	-2.085	-59.303	WSF	SAMN11317096
45,1	-1.983	-60.060	WSF	SAMN11317097
50,1	-1.990	-60.056	WSF	SAMN11317098
55,3	-2.015	-59.906	WSF	SAMN11317099
77,2	-2.589	-60.030	WSF	SAMN11317100
P424A	-2.047	-61.559	WSF	SAMN11317101
122,6	-2.941	-59.938	WSF	SAMN11317174
144,12	-3.209	-60.672	WSF	SAMN11317175
145,5	-3.052	-60.768	WSF	SAMN11317176
149,2	-3.042	-60.754	WSF	SAMN11317177
161,7	-2.185	-61.131	WSF	SAMN11317178
167,1A	-2.811	-60.921	WSF	SAMN11317179
178,7	-1.943	-61.363	WSF	SAMN11317180
203,1	-1.888	-61.592	WSF	SAMN11317181
216,1	-2.052	-61.559	WSF	SAMN11317182
217,2A	-2.053	-61.559	WSF	SAMN11317183
228,3	-2.023	-61.540	TF	SAMN11317184
238,5	-2.855	-60.235	WSF	SAMN11317185
31,2	-1.897	-59.412	WSF	SAMN11317186
41,4	-2.083	-59.304	WSF	SAMN11317187
45,3	-1.983	-60.060	WSF	SAMN11317188
51,1	-1.988	-60.058	WSF	SAMN11317189
62,1	-2.012	-59.906	WSF	SAMN11317190
65,2	-2.003	-59.907	WSF	SAMN11317191
67,5	-1.991	-59.902	WSF	SAMN11317192
77,4	-2.589	-60.030	WSF	SAMN11317193
79,5	-2.589	-60.029	WSF	SAMN11317194
80,1	-2.588	-60.030	WSF	SAMN11317195
81,4	-2.941	-59.974	TF	SAMN11317196

Collection id	Latitude	Longitude	Environment	Biosample accession
Syrrhopodon helicophyllus				
101,1	-2.951	-59.957	WSF	SAMN11317115
105,1	-2.953	-59.958	WSF	SAMN11317116
105,3	-2.953	-59.958	WSF	SAMN11317117
107,2	-2.953	-59.958	WSF	SAMN11317118
110,2	-2.950	-59.956	WSF	SAMN11317119
146,1	-3.051	-60.763	WSF	SAMN11317120
149,1A	-3.042	-60.754	WSF	SAMN11317121
149,1B	-3.042	-60.754	WSF	SAMN11317122
165,6	-2.813	-60.924	WSF	SAMN11317123
166,4	-2.812	-60.922	WSF	SAMN11317124
221,2	-2.054	-61.558	WSF	SAMN11317125
224,1	-2.056	-61.559	WSF	SAMN11317126
235,2	-2.846	-60.234	WSF	SAMN11317127
238,2	-2.855	-60.235	WSF	SAMN11317128
238,6	-2.855	-60.235	WSF	SAMN11317129
25,3	-1.904	-59.410	WSF	SAMN11317130
33,1Bis	-1.903	-59.408	WSF	SAMN11317131
41,1	-2.083	-59.304	WSF	SAMN11317132
46,10	-1.985	-60.060	WSF	SAMN11317133
46,7Bis	-1.985	-60.060	WSF	SAMN11317134
46,9	-1.985	-60.060	WSF	SAMN11317135
49,1	-1.989	-60.056	WSF	SAMN11317136
53,3A	-2.015	-59.908	WSF	SAMN11317137
65,4	-2.003	-59.907	WSF	SAMN11317138
71,2	-1.999	-59.908	WSF	SAMN11317139
75,2	-2.015	-59.908	WSF	SAMN11317140
75,3	-2.015	-59.908	WSF	SAMN11317141
76,1	-2.588	-60.031	WSF	SAMN11317142
76,5	-2.588	-60.031	WSF	SAMN11317143
77,1	-2.589	-60.030	WSF	SAMN11317144
78,3	-2.590	-60.029	WSF	SAMN11317145
83,1	-2.928	-59.972	TF	SAMN11317146

Collection id	Latitude	Longitude	Environment	Biosample accession
Syrrhopodon hornschuchii				
108,1	-2.952	-59.955	NA	SAMN11317147
115,1	-2.944	-59.957	TF	SAMN11317148
118,1	-2.942	-59.957	TF	SAMN11317149
119,1	-2.938	-59.958	TF	SAMN11317150
125,3	-2.939	-59.929	TF	SAMN11317151
125,4	-2.939	-59.929	TF	SAMN11317152
135,5	-2.927	-59.911	TF	SAMN11317153
136,1	-2.926	-59.914	TF	SAMN11317154
143,2	-2.923	-59.941	TF	SAMN11317155
151,1	-3.037	-60.747	WSF	SAMN11317156
153,2	-2.955	-60.929	TF	SAMN11317157
154,4	-3.206	-60.721	TF	SAMN11317158
155,1	-3.207	-60.720	TF	SAMN11317159
157,1	-3.207	-60.718	WSF	SAMN11317160
162,2	-2.186	-61.135	TF	SAMN11317161
163,1	-2.184	-61.135	TF	SAMN11317162
163,3	-2.184	-61.135	TF	SAMN11317163
164,2	-2.183	-61.134	TF	SAMN11317164
164,3	-2.183	-61.134	TF	SAMN11317165
190,2	-1.878	-61.589	WSF	SAMN11317166
211,1	-2.044	-61.558	WSF	SAMN11317167
37,1	-1.991	-59.555	WSF	SAMN11317168
42,2	-2.087	-59.301	WSF	SAMN11317169
67,8	-1.991	-59.902	WSF	SAMN11317170
82,2	-2.928	-59.972	TF	SAMN11317171
83,3	-2.928	-59.972	TF	SAMN11317172
84,2	-2.928	-59.972	TF	SAMN11317173

Collection id	Latitude	Longitude	Environment	Biosample accession
Thysananthus amazonicus				
105,8	-2.953	-59.958	WSF	SAMN11316587
107,1	-2.953	-59.958	WSF	SAMN11316588
112,1	-2.948	-59.956	WSF	SAMN11316604
144,1	-3.209	-60.672	WSF	SAMN11316589
146,2	-3.051	-60.763	WSF	SAMN11316625
149,3	-3.042	-60.754	WSF	SAMN11316638
166,3	-2.812	-60.922	WSF	SAMN11316637
184,1	-1.884	-61.737	TF	SAMN11316642
235,3	-2.846	-60.234	WSF	SAMN11316644
41,5	-2.083	-59.304	WSF	SAMN11316643
46,3	-1.985	-60.060	WSF	SAMN11316646
46,6	-1.985	-60.060	WSF	SAMN11316645
49,2A	-1.989	-60.056	WSF	SAMN11316648
53,1	-2.015	-59.908	WSF	SAMN11316647
63,1	-2.014	-59.905	WSF	SAMN11316649
75,1	-2.015	-59.908	WSF	SAMN11316651
76,6	-2.588	-60.031	WSF	SAMN11316650
78,5	-2.590	-60.029	WSF	SAMN11316658

Paper II

Alain Vanderpoorten   ORCID iD: 0000-0002-5918-7709

Research Article

Maintenance of genetic and morphological identity in two sibling *Syrrhopodon* species (Calymperaceae, Bryopsida) despite extensive introgression

running title: introgression in mosses

Marta R. Pereira[1+], Alice Ledent [2+], Patrick Mardulyn[3], Charles E. Zartman[1+],

Alain Vanderpoorten[2+*]

[1]National Institute for Amazonian Research, Department of Biodiversity, Petrópolis, CEP 69060-001, Manaus, Amazonas, Brazil

[2]University of Liège, Institute of Botany, B22 Sart Tilman, 4000 Liège, Belgium

[3]Evolution Biologique et Ecologie, Université Libre de Bruxelles, 1050 Brussels, Belgium

[+]These authors contributed equally to the work.

*Author for correspondence. E-mail: a.vanderpoorten@uliege.be

Abstract

Bryophytes are a group of land plants wherein the role of hybridization has long been challenged. Using Genotyping by Sequencing to circumvent the lack of molecular variation at selected loci previously used for phylogeny and morphology, we determine the level of genetic and morphological divergence and reproductive isolation between the sibling *Syrrhopodon annotinus* and *S. simmondsii* (Calymperaceae, Bryopsida) that occur in sympatry but in different habitats in lowland Amazonian rainforests. A clear morphological differentiation and a low (0.06), but significant Fst derived from the analysis of 183 SNPs were observed between the two species. Conspecific pairs of individuals consistently exhibited higher average kinship coefficients along a gradient of geographic isolation than interspecific pairs. The weak, but significant genetic divergence observed is consistent with growing evidence that ecological specialization can lead to genetic differentiation among bryophyte species. Nevertheless, the spatial genetic structures of the two species were significantly correlated, as evidenced by the significant slope of the Mantel test based on kinship coefficients between pairs of interspecific individuals and the geographic distance separating them. Interspecific pairs of individuals are thus more closely related when they are geographically closer, suggesting that isolation-by-distance is stronger than the interspecific reproductive barrier and pointing to interspecific gene flow. We

conclude that interspecific introgression, whose role has long been questioned in bryophytes, may take place even in species wherein sporophyte production is scarce due to dioicy, raising the question as to what mechanisms maintain differentiation despite weak reproductive isolation.

1 Introduction

Interspecific hybridization has long been recognized as a widespread and important phenomenon in plant evolution (Payseur & Rieseberg, 2016; Alix et al., 2017). Alloploidization, which involves whole genome duplication with the chromosome numbers of the two parents being summed in the hybrid species, has long been perceived as the dominant form of hybrid speciation (Abbott et al., 2010). Homoploid speciation, the evolution of a species without change in chromosome number, is thought to be a rarer form of speciation, although its detectability might increase with the advances of genomics (Abbott et al., 2010).

In fact, genomic methods have dramatically improved our ability to detect introgression and have expanded the number of taxa amenable to a detailed study of hybridization (Goulet et al., 2017). By far the most widely-used approach to detect hybridization with molecular data is the use of model-based methods to infer global (genome-average) and local (locus-specific) ancestry from population variation data, such as STRUCTURE (Porras-Hurtado et al., 2013). These methods rely, however, on a series of assumptions, including not or

weakly linked loci, which are not met in clonal organisms. Alternatively, Hardy & Vekemans (2001) introduced a method to evidence interspecific gene flow by contrasting the effect of reproductive barriers between species and isolation by distance within species on population genetic structure. Although this method does not make any assumptions on genetic structure, an important requirement is that samples of each species show similar geographic distributions to ensure that, for all types of pairwise comparisons (i.e., within and between species), a full range of geographic distances between samples is present. Since this method integrates geographic distance with genetic relatedness, it allows for determining the spatial scale at which interspecific gene flow occurs.

Here, we apply this approach to address the question of hybridization in bryophytes, wherein, despite substantial evidence for allopolyploidization (e.g., Barbulescu et al., 2017; Karlin & Smouse, 2017; Nieto-Lugilde et al., 2018) and associated shifts in sexual systems (Perley & Jesson, 2015), hybridization has long been perceived as nothing more than 'an ephemeral and evolutionarily insignificant phenomenon' (Natcheva & Cronberg, 2004). The reason for this, perhaps, is the fact that bryophytes are typically seen as mostly clonal organisms. Two-third of the moss species are dioicous and rarely produce sporophytes. Monoicous species, in turn, are characterized by high rates of selfing (Eppley et al., 2007; Hutsemékers et al., 2013). Such features do not, in principle, promote

52

hybridization. Actual evidence for recombination of nuclear markers associated with sympatric occurrence of phenotypically intermediate specimens is indeed extremely limited (Shaw, 1998; Natcheva & Cronberg, 2007), and hybridization in bryophytes has been inferred from equivocal evidence such as intermediate morphology (see Natcheva & Cronberg, 2004 for review) and incongruence between chloroplast and nuclear DNA sequences (e.g., Hedenäs, 2015, 2017). In the present paper, we focus on the species pair formed by *Syrrhopodon simmondsii* and *S. annotinus* (Calymperaceae), which appears as an ideal model to test for reproductive isolation in mosses. These two species are sympatric in their Amazonian range where they occupy different micro-habitats, *S. annotinus* mostly occurring on mineral substrates such as rock and soil, whereas *S. simmondsii* tends to occur on organic substrates such as dead wood, humus, and trees (Reese, 1993), avoiding the common problem that ecological differentiation is confounded with distance and/or barriers to dispersal (Weber et al., 2017). Specimens of *S. annotinus* and *S. simmondsii* clustered within a clade in phylogenetic analyses based on a combination of five chloroplast and one nuclear region (Pereira et al., 2019), but the lack of molecular variation within that clade called for the use of more variable markers. Single Nucleotide Polymorphisms (SNPs) have been successfully employed in fine-scale population genetics of mosses (Rosengren et al., 2015, 2016) and we therefore

built a SNP library from a modified Genotyping by Sequencing (GBS) protocol to test the morphological species concept of *S. annotinus* and *S. simmondsii,* address the following questions and test the following hypotheses: To what extent are the spatial genetic and morphological structures of the two species correlated? If *S. annotinus* and *S. simmondsii* are reproductively isolated, we expect that reproductive barriers are stronger than isolation-by-distance effects, and hence, that there is no relationship between the genetic and morphological similarity of pairs of interspecific individuals and the geographic distance separating them (H1). If the two species are conspecific, we expect that the regression curves between conspecific and interspecific pairs of individuals and geographic distance are overlapping (H2). If interspecific gene flow occurs between two genetically diverging species, we expect that the genetic and morphological similarity of pairs of individuals is higher in conspecific than in interspecific comparisons, but that the regression slope between interspecific genetic, and potentially also morphological, similarity and geographic distance is significant (H3).

2 Material and Methods

2.1 Sampling and molecular protocols

Forty-eight specimens, including 25 of *Syrrhopodon annotinus* and 23 of *S. simmondsii*, were sampled from a 42 640 km^2 area of lowland (< 100 m) rainforest in the Rio Negro Basin North of Manaus (Table S1; Data S1). Climatic conditions, characterized by yearly average temperatures of 27.5 °C and 2145 mm of rainfall, are homogeneous across the study area. Main forest types include dense rainforests, more open forest types developing on white sand, and seasonally inundated forests. *Syrrhopodon annotinus* and *S. simmondsii* were only found in the white-sand forest (Sierra et al., 2018). The study area lays in the core distribution area of these two species, *S. annotinus* being an Amazonian endemic while *S. simmondsii* is endemic to northern South America (Reese, 1993). The sampling was organized to include a range of geographic distance among specimens from 0 to 250 km, in local sympatry (specimens of the two species at a distance of less than 25 m) or in allopatry (only specimens from one species found within the nearest 25 m). Specimens were collected in tubes and readily dried-out in silica gel.

Samples were frozen in liquid nitrogen for 5 minutes and DNA extracted using the DNeasy Plant Mini Kit (QIAGEN). SNP libraries were prepared based on a

GBS protocol (Elshire et al., 2011) modified as follows: (i) 100 ng DNA were digested with ApeKI and (ii) a double size selection of DNA fragments of 150-400 bp was performed using SPRI beads to only target fragments of sequenceable size. Fragments were amplified with a Q5 Hot Start High-Fidelity DNA Polymerase NEB to enhance specificity and reduce amplification errors. A scalable complexity reduction was achieved by using longer 3′ primers that cover the entire common adapter, the 3′ restriction site and extend 1 or 2 bases into the insert following Sonah et al. (2013). PCR products were purified using AMPure XP beads. We gave each individual a forward and a reverse 4–8 bp long barcode (one at the 5′ end and one at the 3′ end), such that each individual had a unique barcode and could be multiplexed with all other individuals. These barcodes were selected from the 384 barcodes specifically designed to be used with ApeKI (http://www.maizegenetics.net/genotyping-by-sequencing-gbs). The concentration of PCR products was assessed by fluorometry with the Quant-iT PicoGreen dsDNA Assay Kit before multiplexing to ensure the equimolarity of PCR products in final libraries. Distribution of fragment sizes for each library was analyzed by capillary electrophoresis with a QIAxcel to look for any remaining adaptor (around 128 bp). If present, adaptors were removed by selecting fragments of > 150 bp on a polyacrylamide gel. Paired-end sequencing (2X75 bp) of the libraries was performed with an Illumina NextSeq. 500

sequencer in low-output mode (i.e., 130 000 000 reads per line).

Sequences of the adaptors at both 3' and 5' ends of each read as well as low-quality sequences (Phred score <20) at both ends were removed with cutadapt 1.16 (https://cutadapt.readthedocs.io). ipyrad 0.7.28 (https://ipyrad.readthedocs.io) was then used to demultiplex the libraries and to cluster alleles that diverged by less than 15% within and then among individuals. Following Paris et al. (2017), we set the minimum number of raw reads required to form an allele to 3. Due to the haploid condition of the target species, loci with more than one allele per individual were discarded. We then discarded any locus that was sequenced in less than 30% of the individuals.

2.2 Statistical analyses

To test the morphological species concept of *S. annotinus* and *S. simmondsii,* we performed analyses at the level of allelic combinations on the one hand, and allele frequencies on the other. We therefore performed a Principal Component Analysis of the SNP data matrix to identify combinations of alleles that best discriminate specimens and then employed a one-way analysis of variance to determine whether there were significant differences of the average score of specimens morphologically assigned to either species, and hence, determine whether the observed genetic segregation corresponds to the distinction between

57

the two species. We further computed the Fst between groups of specimens morphologically assigned to either *S. annotinus* or *S. simmondsii* to seek for genetic differences at the level of allele frequencies. Significance of Fst was assessed by 1000 random permutations of individuals among species. To test the hypothesis that the two species are reproductively isolated, we computed Mantel tests between kinship coefficients Fij (Loiselle et al., 1995) (i) for pairs of conspecific individuals, and (ii) for pairs of interspecific individuals, and the geographic distance separating them. The significance of the slope associated with the Mantel test among conspecific pairs was assessed by random permutations of 1000 individuals among localities. For comparisons between species, care must be taken in the interpretation of significance tests when spatial autocorrelation occurs within each species (Hardy & Vekemans, 2001). This problem arises because randomization of spatial positions not only uncouples the spatial structures of both species, as needed, but also suppresses the spatial autocorrelation within each species. Thus, for comparisons between species, the randomization tests may overestimate the association between geographic distances and genetic distances between interspecific individuals. Therefore, to assess the significance of the slope of the Mantel tests among non-conspecific species pairs, we implemented a Jacknife test, wherein the slope was recalculated after successively pruning one locus from the data at a time, to

estimate the standard deviation of the slope across loci, and hence, determine whether its 95% confidence interval encompasses 0. One-way analysis of variance was employed to test the hypothesis that average kinship coefficients between pairs of conspecific individuals were higher than between interspecific pairs of specimens. All computations were performed with Spagedi 1.5 (Hardy et al., 2002).

To compare the fine-scale genetic structure of *S. annotinus* and *S. simmondsii* with that of angiosperms, we also computed the *Sp* statistics, which characterizes the rate of decrease of pairwise kinship coefficients between individuals with the logarithm of the distance (Vekemans & Hardy, 2004). The *Sp* statistics varies as a function of the mating system and dispersal traits, low values typically characterizing organisms with high dispersal capacities and outbreeding mating systems. The *Sp* statistics is measured as $-\hat{b}_F/(1 - \hat{F}_1)$, where $-\hat{b}_F$ is the regression slope on the logarithm of distance and F_1 is the mean kinship coefficient between individuals belonging to the first distance interval that includes all pairs of neighbours, was computed for comparison with the values reported in angiosperms (Vekemans & Hardy, 2004).

2.3 Morphological analyses

Fourteen variable gametophytic characters (Table 1) were scored on the specimens used for molecular analysis. Ten randomly selected leaves were sampled from each of three randomly selected shoots per collection. To determine whether there is a continuous morphological range of variation between the two species, a Principal Component Analysis was performed, and significant differences in the average score of specimens assigned to each species, respectively, on the first two axes, were sought using one-way analysis of variance. To test the hypothesis that morphological differentiation is consistently higher in interspecific than in intraspecific comparisons along a gradient of geographic distance (H2), we computed Euclidian distances of both pairs of conspecific and interspecific individuals separately. We then computed the correlation coefficient between Euclidian distances for both intra- and interspecific comparisons and geographic distance, and employed one-way analysis of variance to test that average morphological differences were significantly higher in inter- than intraspecific comparisons.

60

3 Results

An average total of 1.5 million reads per individual was obtained. From those 1.5 million of reads per individual, the clustering of alleles diverging by a maximum of 15% within individuals led to an average of 750 000 allele clusters per individual. The minimum number of raw reads required to form an allele (set to 3) led to the loss of an average of 86% of the allele clusters. The filtering-out of heterozygous allele clusters led to the loss of another 2% of them. From the resulting matrix of 14 000 loci on average per specimen, we ended-up, after clustering loci diverging by a maximum of 15% among individuals, filtering-out loci and individuals with >75 and 90% of missing data, respectively, with a matrix including 40 specimens listed in Table 1 from an initial number of 48 and 183 loci. In that matrix (Data S1), 37.2±16.1% of the loci were sequenced on average per specimen.

The analyses seeking for significant differences between *S. annotinus* and *S. simmondsii* at the level of allele combinations on the one hand, and allele frequencies on the other, revealed an extremely weak genetic divergence between the two species. Thus a weak and marginally significant ($F = 4.1$, $P = 0.05$) average difference was found between the scores of individuals morphologically assigned to either *S. annotinus* or *S. simmondsii* along the first axis of the Principal Component Analysis of the SNP datamatrix (Fig. 1). The

*F*st between specimens morphologically assigned to either *S. annotinus* or *S. simmondsii* was 0.059 (*P* = 0.004), reflecting the weak, but significant difference of allele frequencies between the two species.

Average kinship coefficients along a geographic gradient were consistently higher for conspecific comparisons than for interspecific comparisons, and significantly so in 3 out of the 6 comparisons (Fig. 2). The *Sp* statistics derived from the Mantel test based upon comparisons of conspecific pairs of individuals was 0.04 (0.043 for *S. annotinus* and 0.024 for *S. simmondsii*). Nevertheless, the spatial genetic structures of the two species was significantly correlated. In fact, the interval of confidence of the regression slopes between pairs of conspecific individuals (*b*-log slope = -0.033±0.011, *P* < 0.001), but also between pairs of interspecific individuals (*b*-log slope = -0.034±0.010) did not encompass 0, evidencing a significant pattern of isolation-by-distance in both cases.

The distribution of specimens along the first axes of the PCA of morphological characters, which accounted for 55 and 16% of the total variance, respectively, is shown in Fig. 3A. There were significant morphological differences between *S. annotinus* and S. *simmondsii*, as revealed by significant ANOVAS of the average score of specimens assigned to each species along PCA1 (*P*<0.001), but not PCA2 (*P*=0.26). PCA1 was mostly correlated negatively with chlorophyllose cell width, leaf serration, and presence/absence of chlorocyst papillae, and

positively with plant size, leaf length, costa width, and stem thickness (Fig. 3B). None of the intra- and interspecific Euclidian distances were correlated with geographic distance (Fig. 4). Differences between pairs of specimens of different species were consistently significantly higher than between pairs of conspecific individuals across the geographic range (Fig. 4).

4 Discussion

The significant Fst between sympatric specimens assigned to *Syrrhopodon annotinus* and *S. simmondsii* indicates that reproductive isolation led to differences in allele frequencies between two sympatric species characterized by different habitat requirements. Shaw et al. (1987) similarly found significant differences in a sympatric pair of moss species (*Climacium americanum* and *C. dendroides*) from different habitats. Although it is not possible to determine whether habitat differentiation triggered or followed speciation, these observations are consistent with growing evidence that ecological specialization can lead to genetic differentiation within bryophyte species. These observations contrast with the hypothesis that physiological plasticity prevails over genetic specialization in bryophytes based on experimental work suggesting that, in contrast with the vast majority of angiosperm species, bryophyte species do not tend to develop ecotypes, but rather display an inherent broad ability to cope with environmental variation (Shaw, 1992). The results presented here contribute

to growing evidence for genetic divergence (Szövényi et al., 2009; Hutsemekers et al., 2010; Pisa et al., 2013; Mikulaskova et al., 2015; Magdy et al., 2016) and speciation (Johnson et al., 2015) along environmental gradients in bryophytes, suggesting that adaptation could play a more important role in shaping genetic patterns than previously thought. In particular, an increased role for ecological speciation contrasts with the hypothesis that the failure of bryophytes to radiate is caused by the limited importance of isolation-by-environment in the group (Patiño et al., 2014).

The slight, but significant genetic divergence between *S. annotinus* and *S. simmondsii* (6% of the total variance of allele frequencies) and the significantly higher kinship coefficients in intraspecific than in interspecific pairs of individuals suggest that the sharp morphological differences between the two species cannot be simply explained by plasticity. Furthermore, no correlation between morphological variation and geographic distance was observed in *S. annotinus* and *S. simmondsii*. Interspecific morphological differences were consistently significantly higher than intraspecific differences across the geographic range investigated. This indicates that, in contrast to *S. leprieurii* (Pereira et al., 2013), the observed morphological differentiation between *S. annotinus* and *S. simmondsii* cannot be explained by an alternative hypothesis of geographic variation within a single species (Zapata & Jimenez, 2012).

However, despite the fact that pairs of conspecific individuals consistent-ly exhibited higher average kinship coefficients than pairs of interspecific indi-viduals, the slope of the Mantel test based on kinship coefficients between inter-specific pairs was significantly spatially correlated. This indicates that isolation-by-distance is stronger than the interspecific reproductive barrier. Interspecific pairs of individuals are therefore more closely related when they are geograph-ically closer, evidencing interspecific gene flow and strongly supporting Natche-va & Cronberg's (2004) suggestion that hybridization may be a common, yet largely overlooked mechanism in bryophytes.

The strong isolation-by-distance pattern revealed here is consistent with the dispersal traits of the studied species, which are characterized by immersed sporophytes within perichaetial leaves, fairly large spores of > 30 µm, the ab-sence of a peristome, and infrequent gemmae and sporophyte production (Reese, 1993), together with extrinsic features of their environment of dense rainforests that is not prone to long-distance wind-dispersal. The *Sp* statistics of 0.04 lays in the upper limit of the range reported for species with gravity-dispersed seeds (0.028±0.016) and is higher than the range reported for species characterized by wind (0.012±0.012) and animal (0.008±0.005)-dispersed seeds (Vekemans & Hardy, 2004). Such a poor dispersal capacity does not point to the transportation of diaspores by animals, which has been increasingly documented in bryophyte

species from dense forest environments with low wind connectivity (Heinken et al., 2001; Parsons et al., 2007; Rudolphi, 2009; Wilkinson et al., 2017). The Sp statistics was, in turn, lower than that reported for plant species characterized by selfing (and even more, clonal) mating systems (0.14±0.08), which is consistent with the hypothesis that interspecific gene flow may occur in sympatry. Individuals of *S. annotinus* and *S. simmondsii* in fact grow in the close vicinity of each other, so that sperm cells may reach the archegonia of a non-conspecific female individual. In a one-year monitoring survey of sex expression, Pereira et al. (2016) failed, however, to find expressed males and sporophytes in *S. annotinus*, while expressed males and sporophytes were found in <1% and 5% of investigated individuals, respectively, in *S. simmondsii*. Since gametangia production was positively correlated with precipitation (Pereira et al., 2016), we hypothesize that interspecific gene flow reported here dates back to historical periods of increased sexual reproduction associated with historically wetter climates.

Our results point to three main conclusions. First, based on the fact that there is (i) a weak, but significant *F*st between sympatric specimens assigned to *S. annotinus* on the one hand, and *S. simmondsii* on the other, and that kinship coefficients are significantly higher in interspecific than in intraspecific comparisons and (ii) a sharp morphological differentiation, we tentatively suggest that *S. annotinus* and *S. simmondsii* represent two distinct species.

However, the weakness of the reproductive barrier between them, evidenced by the significance of the regression slope of Fij in interspecific comparisons, suggests that *S. annotinus* and *S. simmondsii* diverged recently. During the speciation process, various aspects of lineage divergence in fact arise (De Queiroz, 2007). Sister species progressively diverge from each other over time, but the acquisition of the different properties defining them (when they become phenotypically diagnosable, reciprocally monophyletic for one or multiple loci, reproductively incompatible, ecologically distinct, etc.) is not simultaneous, potentially leading to conflicting assessments of species identity before and after the acquisition of any one of those properties.

Second, significant patterns of genetic structure were observed despite the comparatively low number of SNPs that could be expected with GBS techniques and the fairly large amounts of missing data. By comparison with the results obtained in other plant studies using GBS or Rad-seq approaches, > 1000 SNPs were obtained for each of 663 individuals in *Protea repens* (Prunier et al., 2017), an average of 2778 independent SNP loci was obtained in 7 Australian Alpine species (Bell et al., 2018), and 17 982 SNPs were obtained in the fir *Keteleeria davidiana* with 50% missing data (Shih et al., 2018). The 183 SNPs obtained here are more comparable to the 63 SNPs obtained in the only other moss species that had been investigated to date using GBS or Rad-seq

techniques, *Tetraplodon fuegianus*, in its northern hemisphere range (Lewis et al., 2017). The lower number of SNPs obtained in mosses may reflect a truly lower genetic diversity caused, among others, by high rates of clonality, but may also call for protocol improvements in these non-model organisms. Nevertheless, the present results, together with those of Lewis et al. (2017), suggest that SNP data are a promising tool for shallow systematics (Fernández-Mazuecos et al., 2018; Ding et al., 2019) and fine-scale genetic structure (Attard et al., 2018) in bryophytes when traditional markers such as low-copy nuclear genes and chloroplast genes are not polymorphic.

Third, the maintenance of low, but significant levels of genetic divergence and a high morphological differentiation despite weak reproductive isolation (in fact, weaker than the strength of isolation-by-distance) is puzzling. The rarity of intermediate hybrid phenotypes was previously reported in mosses, wherein it was interpreted in terms of a higher viability of spores that had inherited most of the genome from one or the other parent, or the low survival of the recombinant gametophytes (Cronberg, 1996; Cronberg & Natcheva, 2002). In the case of *S. annotinus* and *S. simmondsii*, we suggest that two mechanisms may contribute to the observed differentiation. First, the intensity of recombination may be limited by the resistance of large parts of the genome against heterospecific genes, maintaining the genetic distinctness of the species

(Natcheva & Cronberg, 2007). Second, habitat specialization may play a key role through the counter-selection of hybrids, calling for the implementation of experimental transplantations and fitness measurements.

Acknowledgements

References

Abbott RJ, Hegarty MJ, Hiscock SJ, Brennan AC. 2010. Homoploid speciation in action. *Taxon* 59: 1375-1386.

Alix K, Gérard PR, Schwarzacher T, Heslop-Harrison JSP. 2017. Polyploidy and interspecific hybridization: Partners for adaptation, speciation and evolution in plants. *Annals of Botany* 120: 183-194.

Attard CRM, Beheregaray LB, Möller LM. 2018. Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. *Molecular Ecology* 18: 381-390.

Barbulescu EVI, Patzak SDF, Feldberg K, Schäfer-Verwimp A, Rycroft DS, Renner MAM, Heinrichs J. 2017. Allopolyploid origin of the leafy liverwort Plagiochila britannica (Plagiochilaceae). *Botanical Journal of the Linnean Society* 183: 250-259.

Bell N, Griffin PC, Hoffmann AA, Miller AD. 2018. Spatial patterns of genetic diversity among Australian alpine flora communities revealed by comparative phylogenomics. *Journal of Biogeography* 45: 177-189.

Chmielewski MW, Eppley SM. 2019. Forest passerines as a novel dispersal vector of viable bryophyte propagules. *Proceedings of the Royal Society B:*

Biological Sciences 286: 20182253.

Cronberg N. 1996. Clonal structure and fertility in a sympatric population of the peat mosses, *Sphagnum rubellum* and *S. capillifolium*. *Canadian Journal of Botany* 74: 1375-1385.

Cronberg N, Natcheva R. 2002. Hybridization between the peat mosses, *Sphagnum capillifolium* and *S. quinquefarium* (Sphagnaceae, Bryophyta) as inferred by morphological characters and isozyme markers. *Plant Systematics and Evolution* 234: 53-70.

De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* 56: 879–886.

Ding X, Xiao JH, Li L, Conran GJ, Li J. 2019. Congruent species delimitation of two controversial gold☐thread nanmu tree species based on morphological and restriction site☐associated DNA sequencing data. *Journal of Systematics and Evolution*. doi: 10.1111/jse.12433.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6: e19379.

Eppley SM, Taylor PJ, Jesson LK. 2007. Self-fertilization in mosses: a

comparison of heterozygote deficiency between species with combined versus separate sexes. *Heredity* 98: 38–44.

Fernández-Mazuecos M, Mellers G, Vigalondo B, Sáez L, Vargas P, Glover BJ. 2018. Resolving recent plant radiations: Power and robustness of Genotyping-by-Sequencing. *Systematic Biology* 67: 250–268.

Goulet EB, Roda F, Hopkins R. 2017. Hybridization in plants: Old ideas, new techniques. *Plant Physiology* 173: 65–78.

Hardy OJ, Vekemans X. 2001. Patterns of allozyme variation in diploid and tetraploid *Centaurea jacea* at different spatial scales. *Evolution* 55: 943–954.

Hardy OJ, Vekemans X. 2002. SPAGeDi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618–620.

Hedenäs L. 2015. Molecular and morphological incongruence among the genera around *Sarmentypnum* (Bryophyta: Calliergonaceae). *Nova Hedwigia* 100: 279–292.

Hedenäs L. 2017. Three molecular markers suggest different relationships among three *Drepanocladus* species (Bryophyta: Amblystegiaceae). *Plant Systematics and Evolution* 303: 521–529.

Heinken T, Lees R, Raudnitschka D, Rung S. 2001. Epizoochorous dispersal of bryophyte stem fragments by roe deer (*Capreolus capreolus*) and wild boar (*Sus scrofa*). *Journal of Bryology* 23: 293–300.

Hutsemékers V, Hardy O, Mardulyn P, Shaw A, Vanderpoorten A. 2010. Macroecological patterns of genetic structure and diversity in the aquatic moss *Platyhypnidium riparioides*. *New Phytologist* 185: 852–864.

Hutsemékers V, Hardy OJ, Vanderpoorten A. 2013. Does water facilitate gene flow in spore-producing plants? Insights from the fine-scale genetic structure of the aquatic moss *Rhynchostegium riparioides*. *Aquatic Botany* 108: 1-6.

Johnson MG, Granath G, Tahvanainen T, Pouliot R, Stenøien HK, Rochefort L, Rydin H, Shaw AJ. 2015. Evolution of niche preference in *Sphagnum* peat mosses. *Evolution* 69: 90–103.

Karlin EF, Smouse PE. 2017. Allo-allo-triploid *Sphagnum × falcatulum*: Single individuals contain most of the Holantarctic diversity for ancestrally indicative markers. *Annals of Botany* 120: 221-231.

Lewis LR, Biersma EM, Carey SB, Holsinger K, McDaniel SF, Rozzi R, Goffinet B. 2017. Resolving the northern hemisphere source region for the long-distance dispersal event that gave rise to the South American endemic dung moss

Tetraplodon fuegianus. American Journal of Botany 104: 1651–1659.

Loiselle BA, Sork VL, Nason J, Graham C. 1995. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* 82:1420–1425.

Magdy M, Werner O, McDaniel SF, Goffinet B, Ros RM. 2016. Genomic scanning using AFLP to detect loci under selection in the moss *Funaria hygrometrica* along a climate gradient in the Sierra Nevada Mountains, Spain. *Plant Biology* 18: 280–288.

Mikulaskova E, Hajek M, Veleba A, Johnson MG, Hajek T, Shaw AJ. 2015. Local adaptations in bryophytes revisited: the genetic structure of the calcium-tolerant peatmoss *Sphagnum warnstorfii* along geographic and pH gradients. *Ecology & Evolution* 5: 229–242.

Natcheva R, Cronberg N. 2004. What do we know about hybridization among bryophytes in nature? *Canadian Journal of Botany* 82: 1687–1704.

Natcheva R, Cronberg N. 2007. Recombination and introgression of nuclear and chloroplast genomes between the peat mosses, *Sphagnum capillifolium* and *Sphagnum quinquefarium. Molecular Ecology* 16: 811-818.

Nieto-Lugilde M, Werner O, McDaniel SF, Koutecký P, Rizk SM, Ros RM. 2018. Peripatric speciation associated with genome expansion and female-biased

sex ratios in the moss genus *Ceratodon*. *American Journal of Botany*

105: 1009-1020.

Paris JR, Stevens JR, Catchen JM. 2017. Lost in parameter space: a road map for

STACKS. *Methods in Ecology and Evolution* 8: 1360-1373.

Parsons JG, Cairns A, Johnson CN, Robson SKA, Shilton LA, Westcott DA.

2007. Bryophyte dispersal by flying foxes: a novel discovery. *Oecologia* 152:

112–114.

Patiño J, Carine MA, Fernández-Palacios JM, Otto R, Schaefer H,

Vanderpoorten A. 2014. The anagenetic world of spore-producing land plants.

New Phytologist 201: 305-311.

Payseur BA, Rieseberg LH. 2016. A genomic perspective on hybridization and

speciation. *Molecular Ecology* 25: 2337–2360.

Pereira MR, Dambros CS, Zartaman CE. 2013. Will the real *Syrrhopodon

leprieurii* please stand up? The influence of topography and distance on

phenotypic variation in a widespread Neotropical moss. *The Bryologist* 116: 58-

64.

Pereira MR, Dambros CS, Zartman CE. 2016. Prezygotic resource-allocation

dynamics and reproductive trade-offs in Calymperaceae (Bryophyta). *American

Journal of Botany* 103:1-9.

Pereira MR, Amorim BS, Sierra AM, McDaniel S, Payton A, Carey SB, Câmara

PEAS, Zartman CE. 2019. Advances in Calymperaceae (Dicranidae, Bryophyta): intra-familial phylogenetic relationships, geographical origin and divergence times. *The Bryologist*, in press.

Perley DS, Jesson LK. 2015. Hybridization is associated with changes in sexual system in the bryophyte genus *Atrichum*. *American Journal of Botany* 102: 555-565.

Pisa S, Werner O, Vanderpoorten A, Magdy M, Ros RM. 2013. Altitudinal genetic differentiation of the cosmopolitan moss *Bryum argenteum* in the mountains of Sierra Nevada (Spain, SW Europe). *American Journal of Botany* 100: 2000-2008.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu MV. 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in Genetics* 29: 98

Prunier R, Akman M, Kremer CT, Aitken N, Chuah A, Borevitz J, Holsinger KE. 2017. Isolation by distance and isolation by environment contribute to population differentiation in *Protea repens* (Proteaceae L.), a widespread South African species. *American Journal of Botany* 104: 674–684.

Reese WD. 1993. Calymperaceae. *Flora Neotropica* 58: 1-101.

Rosengren F, Hansson B, Cronberg N. 2015. Population structure and genetic diversity in the nannandrous moss *Homalothecium lutescens*: does the dwarf

male system facilitate gene flow? *BMC Evolutionary Biology* 15:1-12.

Rosengren F, Cronberg N, Hansson B. 2016. Balance between inbreeding and outcrossing in a nannandrous species, the moss *Homalothecium lutescens*. *Heredity* 116: 107-113.

Rudolphi J. 2009. Ant-mediated dispersal of asexual moss propagules. *The Bryologist* 112: 73–79.

Shaw JT, Meagher TR, Harley P. 1987. Electrophoretic evidence of reproductive isolation between two varieties of the moss, *Climacium americanum. Heredity* 59: 337–343.

Shaw AJ. 1992. The evolutionary capacity of bryophytes and lichens. In: Bates JW, Farmer AM eds. *Bryophytes and lichens in a changing environment*. London: Clarendon Press. 362–380.

Shaw AJ. 1998. Genetic analysis of a hybrid zone in *Mielichhoferia* (Musci). In: Bates JW, Ashton NW, Duckett JG eds. *Bryology for the twenty-first century*. Leeds, UK: Maney Publishing and the British Bryological Society. 161–174.

Shih KM, Chang CT, Chung JD, Chiang YC, Hwang SY. 2018. Adaptive genetic divergence despite significant isolation-by-distance in populations of Taiwan Cow-Tail Fir (*Keteleeria davidiana* var. *formosana*). *Frontiers in Plant Sciences* 9: 92.

Sierra AM, Vanderpoorten A, Gradstein SR, Pereira MR, Passos Bastos CJ,

Zartman CE. 2018. Bryophytes of Jau´ National Park (Amazonas, Brazil): Estimating species detectability and richness in a lowland Amazonian megareserve. *The Bryologist* 121: 571–588.

Sonah H, Bastien M, Iquira E, Tardivel A, Légaré G, Boyle B, Normandeau E, Laroche J, Larose S, Jean M, Belzile F. 2013. An improved Genotyping by Sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE* 8: e54603.

Szövényi P, Hock Z, Korpelainen H, Shaw AJ. 2009. Spatial pattern of nucleotide polymorphism indicates molecular adaptation in the bryophyte *Sphagnum fimbriatum*. *Molecular Phylogenetics and Evolution* 53: 277–286.

Weber JN, Bradburd GS, Stuart YE, Stutz WE, Bolnick DL. 2017. Partitioning the effects of isolation by distance, environment, and physical barriers on genomic divergence between parapatric threespine stickleback. *Evolution* 71: 342-356.

Vekemans X, Hardy OJ. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* 13: 921–935.

Wilkinson DM, Lovas-Kiss A, Callaghan DA, Green AJ. 2017. Endozoochory of large bryophyte fragments by waterbirds. *Cryptogamie Bryologie* 38: 223–228.

Zapata F, Jiménez I. 2012. Species delimitation: Inferring gaps in morphology across geography. *Systematic Biology* 61: 179-194.

Figures

Fig. 1. Principal Component Analysis of genetic variation at 183 SNPs in the sibling moss species *Syrrhopodon annotinus* and *S. simmondsii.* Values in parentheses indicate the proportion of explained variance of the first two axes.
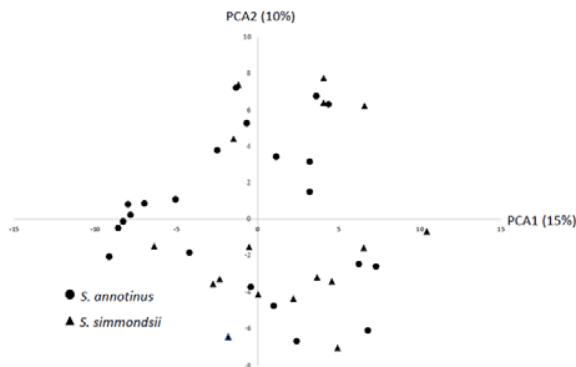
Fig. 2. Spatial autocorrelogram showing average kinship coefficients Fij (±SD) derived from the analysis of 183 SNPs per distance class between pairs of conspecific (black line) and interspecific (grey line) individuals of *Syrrhopodon annotinus* and *S. simmondsii* as a function of the distance between pairs of individuals. Stars indicate the confidence level (* <0.05) of the anova of intra- and interspecific comparisons at each distance class.
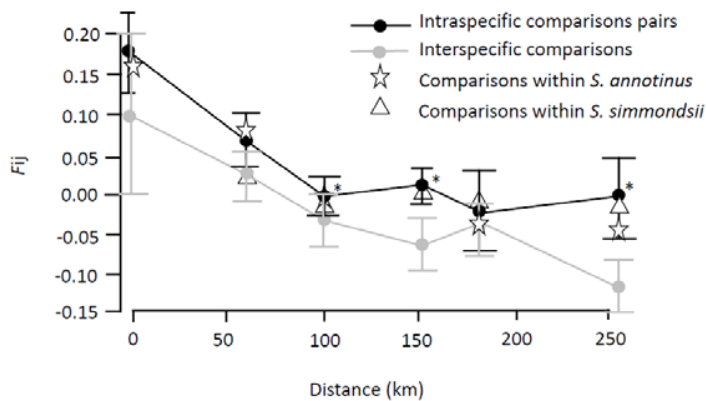
Fig. 3. Principal Components Analysis of 14 gametophytic characters in *Syrrhopodon annotinus* and *S. simmondsii*. **A,** Projection of the individuals onto the first two axes. **B,** Correlation between the variables (see Table 1 for abbreviations) and the axes.
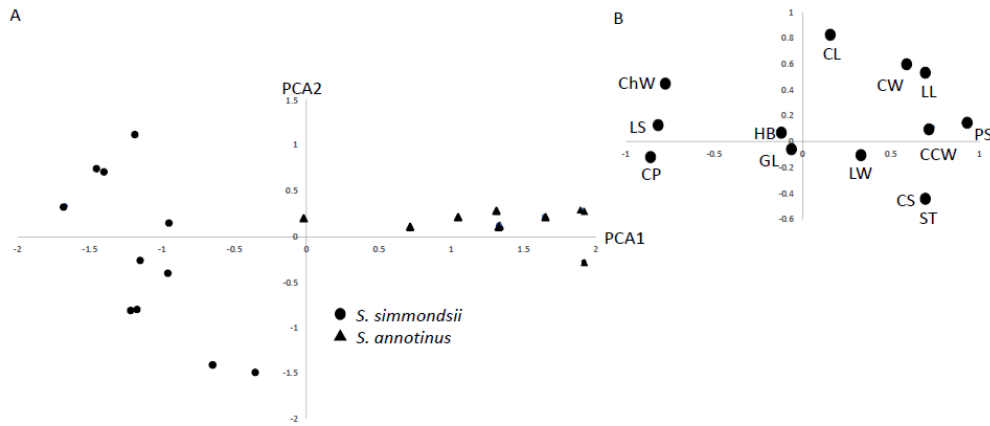
Fig. 4. Average±SD of the Euclidian distance of morphological characters (Table 1) between pairs of conspecific (black line) and interspecific (grey line) individuals of *Syrrhopodon annotinus* and *S. simmondsii* as a function of the distance between pairs of individuals. Stars indicate the confidence level (**, $P < 0.01$; ***, $P < 0.001$) of the anova of intra- and interspecific comparisons at each distance class.
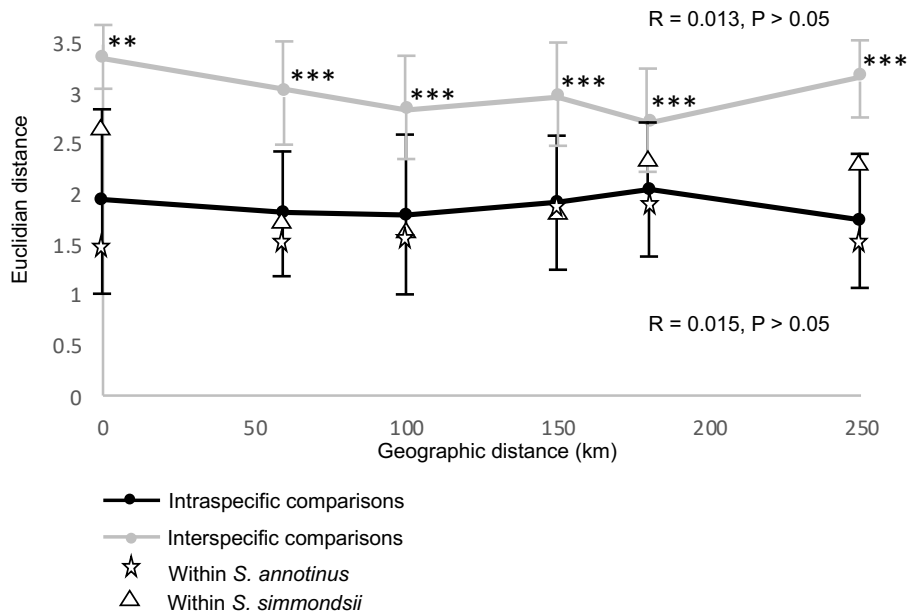
Table 1 Morphological characters scored to describe the differentiation between

Syrrhopodon annotinus and *S. simmondsii* and among conspecific specimens

Characters	Description
	LEAF CHARACTERS
1 Leaf length (LL)	0 = short (< 2.5 mm); 1 = long (2.5 - 5 mm); 2 = very long (> 5 mm)
2 Leaf width (LW)	0 = thin (< 0.5 mm); 1 = wide (> 0.5 mm)
3 Leaf serration (LS)	0 = absent; 1 = < 10 teeth present on either leaf side; 2 = with more than 10 noticeable teeth on each side of leaf
4 Costa width (CW)	0 = thin (10-60 µm); 1 = medium (60-200 µm); 2 = thick (> 200 µm)
5 Gemmiferous leaves (GL)	0 = absent; 1 = present

LEAF MARGIN

6 Hyaline border (HB)

0 = marginal; 1 = submarginal

CANCELLINAE

7 Cancellinae cell length (CL)

0 = short (15-50 µm); 1 = long (> 50

µm)

8 Cancellinae cell width (CCW)   0 = thin (7.5-12.5 µm); 1 = medium

(20-30 µm); 2 = wide (> 50 µm)

CHLOROPHYLLOSE CELLS

9 Chlorophyllose cell length (ChL)   0 = short (7- 9 µm); 1 = long (9- 12

µm)

10 Chlorophyllose cell width (ChW)   0 = thin; 1 = wide

11 Chlorocyst papillae (CP)   0 = absent; 1 = present

STEM CHARACTERS

12 Central strand (CS)   0 = absent; 1 = presente

13 Stem thickness (ST)   0 = small (1- 5 mm); 1 = large (5- 10

mm)

PLANT CHARACTERS

14 Plant size (PS)

0 = small (<1cm); medium sized (= 1-

2 cm); robust (≥ 2cm)

Paper III

No borders during the post-glacial assembly of European bryophytes

A. Ledent,[1,a] A. Désamoré,[2,a]
B. Laenen,[2,a] P. Mardulyn,[3]
S. F. McDaniel,[4] F. Zanatta,[1]
J. Patiño[5,6,a] and
A. Vanderpoorten[1,a]*

Abstract

Climatic fluctuations during the Last Glacial Maximum (LGM) exerted a profound influence on biodiversity patterns, but their impact on bryophytes, the second most diverse group of land plants, has been poorly documented. Approximate Bayesian computations based on coalescent simulations showed that the post-glacial assembly of European bryophytes involves a complex history from multiple sources. The contribution of allochthonous migrants was 95–100% of expanding populations in about half of the 15 investigated species, which is consistent with the globally balanced genetic diversities and extremely low divergence observed among biogeographical regions. Such a substantial contribution of allochthonous migrants in the post-glacial assembly of Europe is unparalleled in other plants and animals. The limited role of northern micro-refugia, which was unexpected based on bryophyte life-history traits, and of southern refugia, is consistent with recent palaeontological evidence that LGM climates in Eurasia were much colder and drier than what palaeoclimatic models predict.

Keywords

Bryophytes, climate change, dispersal, historical biogeography, Last Glacial Maximum, refugia.

Ecology Letters (2019) **22**: 973–986

INTRODUCTION

Pleistocene glacial cycles, ending *c.* 19 000 years ago at the Last Glacial Maximum (LGM), played a major role in structuring the distribution of biodiversity (Lumibao *et al.* 2017). During glacial maxima, ice made vast areas inaccessible for most plant and animal species to establish (Eidesen *et al.* 2013). In Europe, palaeontological and phylogeographic evidence suggest that species currently distributed in ice-free areas at LGM either persisted in southern refugia (Hewitt 1999, 2000, 2004; Médail & Diadema 2009) or in micro-refugia located in the steppe zone South of the ice sheet (Bhagwat & Willis 2008). During warmer periods, like the current interglacial, populations expanded northward from southern refugia or northern micro-refugia, generating admixed populations with high genetic diversity at mid-latitudes, and genetically depauperate populations at high latitudes resulting from long-distance dispersal events and associated founder effects (Hewitt 1999; Petit *et al.* 2003).

For species currently distributed in areas that were covered in ice at LGM, Darwin (1859) and Hooker (1862) similarly proposed that species migrated southwards with the advancing ice sheets. Although this hypothesis received some support (Schönswetter *et al.* 2006; Skrede *et al.* 2006), this explanation is incomplete. For example, palaeontological evidence suggests

that lowland areas south of the ice sheet experienced a cold and dry climate that was unsuitable for Arctic and Alpine floras (Abbott & Brochmann 2003). It is possible that some species survived in local micro-refugia within the ice-sheet area in the Arctic (Westergaard *et al.* 2011) or in southern mountain ranges, from where they potentially back-colonised northern areas (Schönswetter *et al.* 2003, 2005). However, Hultén (1937) alternatively suggested that ice-free regions of Beringia, a region encompassing north-east Russia and north-west America, served as a source for the recolonisation of Europe (Eidesen *et al.* 2013).

The ability of species to persist in northern refugia can be predicted from their life-history traits, which include, in woody plant species, short generation times, small seed sizes and vegetative reproduction under harsh environmental conditions (Bhagwat & Willis 2008). Such traits precisely characterise bryophytes, a diverse and conspicuous component of terrestrial ecosystems. In particular, the successful regeneration of subglacial bryophytes following hundreds to thousands of years of ice entombment (La Farge *et al.* 2013; Roads *et al.* 2014; Cannone *et al.* 2017), along with phylogeographic evidence for *in situ* persistence of moss species in Antarctica during the LGM (Pisa *et al.* 2013; Biersma *et al.* 2018), broadens the concept of glacial refugia, traditionally confined to survival of land plants beyond glacier margins. Bryophyte

[1]*Institute of Botany, University of Liege, Sart Tilman, 4000 Liege, Belgium*

[2]*Department of Ecology, Environment, and Plant Sciences, Science for Life Laboratory, Stockholm University, Stockholm, Sweden*

[3]*Evolutionary Biology & Ecology, Université Libre de Bruxelles, 1050 Brussels, Belgium*

[4]*Biology Department, University of Florida, Gainesville, FL 32611, USA*

[5]*Plant Conservation and Biogeography Group, Departamento de Botánica, Ecología y Fisiología Vegetal, Universidad de La Laguna, 38071 La Laguna, Spain*

[6]*Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), 38071 La Laguna, Spain*

Correspondence: E-mail: a.vanderpoorten@uliege.be

[a]*Equal contribution.*

diaspores are, however, capable of extreme long-distance dispersal, further raising the intriguing idea that the post-glacial recolonisation of Europe might have taken place from extra-European sources (Kyrkjeeide *et al.* 2014) such as North America (Stenøien *et al.* 2011) and Macaronesia (Hutsemékers *et al.* 2011; Laenen *et al.* 2011; Patiño *et al.* 2015).

The rich macrofossil record preserved in lake sediments and peat has provided a detailed picture of the Quaternary bog flora (Jakab & Sümegi 2011), but we lack a more global, community-scale understanding of the post-glacial recolonisation of Europe by bryophytes. Molecular phylogeography has increasingly appeared as a promising tool in historical biogeography, especially since model-based methods under a formal framework have overcome the limitations of the qualitative description of summary statistics or gene trees (Thomé & Carstens 2016). The high dispersal capacities of bryophytes have, however, casted doubt on the possibility to find signatures of historical demographic events from analyses of their extant spatial genetic patterns (Van der Velde & Bijlsma 2003). Spore-trapping experiments indeed revealed that a higher proportion of spores originates from sources farther away than the nearest sources, leading to an inverse isolation effect (Sundberg 2005) erasing any isolation-by-distance signal (Szövényi *et al.* 2012).

Here, we applied approximate Bayesian computations (ABCs) in the framework of the coalescent theory and combined them with predictions from species distribution models (Fig. 1) to address the following questions:

Can patterns of genetic structure and diversity be used to retrace the post-glacial history of European bryophytes, or did intense post-glacial migrations erase any spatial genetic structure, making it impossible to retrace the origin of migrants? If a significant genetic structure exists, which historical scenario best fits with the observed patterns of genetic structure and diversity? Given bryophyte life-history traits, we hypothesise that northern refugia or even populations buried under the ice sheet at LGM substantially contributed to the post-glacial recolonisation of Europe.

MATERIALS AND METHODS

### Population and molecular sampling

Twelve and three species with their core present distribution range in areas that were ice-free and covered in ice at LGM (hereafter, IF and IC species, respectively) were used as models (Tables 1 and 2). Specimens were sampled across the entire distribution range of the species, but with a focus on Europe

and North America due to previous evidence for the existence of genetic connections between them (Szövényi *et al.* 2008; Stenøien *et al.* 2011; Kyrkjeeide *et al.* 2016; Désamoré *et al.* 2016). Each specimen was assigned to each of three or five regions (Table S1, Fig. S1), which correspond to the definition of source and sink areas for IF and IC species respectively. We used here the circumscription of continental Europe of the European Environment Agency (https://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3) (Fig. 2). In the south-east, Caucasus, which marks the border between Europe and Asia, was excluded. The extra-European range included the Holarctic, but not the Southern Hemisphere, where some of the species exhibit scattered occurrences. Indeed, while the existence of bipolar ranges attests to discrete episodes of long-distance dispersal between the Southern and Northern Hemispheres over the past million years (see Biersma *et al.* 2017 for review), Southern Hemisphere populations are unlikely to have contributed to the post-glacial recolonisation of Europe, given that the Equator represents a strong geographic barrier to spore migrations (McDaniel & Shaw 2005; Vanderpoorten *et al.* 2010). For IF species, Europe was split into southern and northern ranges to delimitate areas that correspond to the location of northern and southern refugia, as defined, for example by Bhagwat & Willis (2008) and Médail & Diadema (2009) respectively. The definition of the southern refugia corresponds to the circumscription of the Mediterranean (*sensu* https://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3) but including the entire Iberian Peninsula. We also assigned a few specimens from northern Africa to the Mediterranean region. For IC species, the partitioning was based on the extent of the ice sheet at LGM (Abbott & Brochmann 2003), so that four regions were considered within Europe: 1. northern range under the ice sheet at LGM; 2a, b. ice-free and iced southern mountain ranges at LGM, respectively; 3. lowland range south of the ice sheet at LGM (Fig. 2).

We used published sequence data (Désamoré *et al.* 2016) and another 1941 newly produced sequences following the protocols of Désamoré *et al.* (2016) to expand the geographical and taxonomic sampling (Table S1). Specimens of liverworts were sequenced at 2–3 cpDNA loci including both coding and non-coding regions. For mosses, specimens were sequenced at up to 3 cpDNA loci and 1–3 intron-spanning nuclear loci selected for their suitable range of variation from McDaniel *et al.* (2013) (Tables 1 and 2, Table S1). Sequences were aligned with the 'muscle' algorithm (Edgar 2004) as implemented in Seaview 4.7 (Gouy *et al.* 2010). Gaps were treated as missing data. In all analyses, cpDNA data were

Figure 1 Graphical abstract of the integrative method employed to reconstruct the post-glacial history of European bryophytes. (1) Sampling of specimens across their distribution range. (2) Genotyping by Sanger sequencing at selected loci, producing a matrix of observed sequence data. (3) Simulation of sequence data under competing post-glacial recolonisation scenarios (3.1). Millions of allele genealogies are simulated under alternative coalescence models, for which demographic parameters are randomly sampled from prior probability distributions before each simulation. The prior distribution of one of these demographic parameters, population size, is derived from estimates inferred from species distribution models. DNA sequence data sets equal in size to the observed data are generated by implementing nucleotide substitutions along the simulated genealogies (3.2). Observed and simulated data are summarised using summary statistics characterising population structure and genetic diversity. The distances between observed and simulated summary statistics are computed and the simulations with the 1000 shortest distances are used to generate the posterior probability distributions for all demographic parameters and identify the scenario with the highest posterior probability (PP).
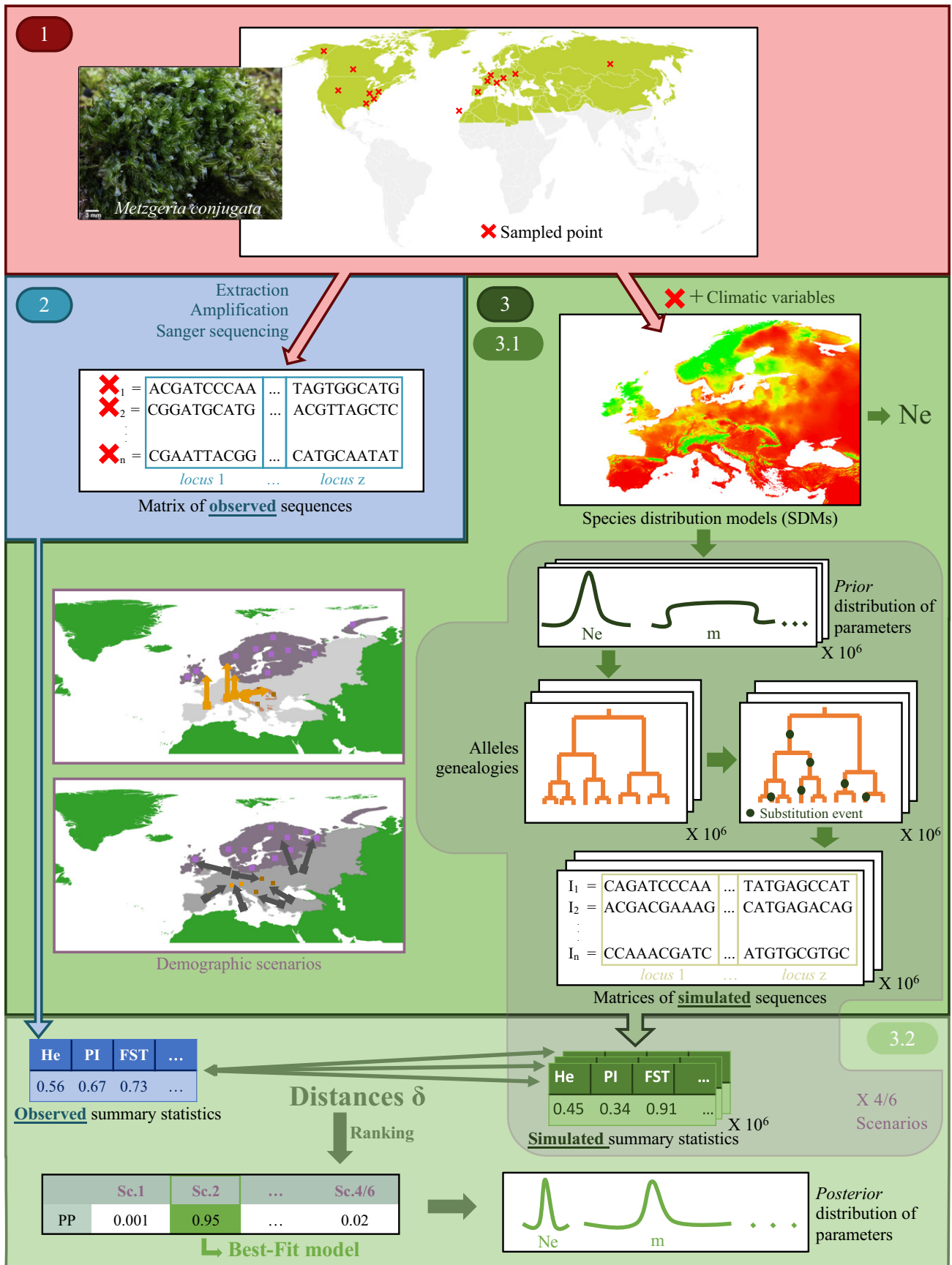
Table 1 Observed summary statistics (expected heterozygosity He, number of private alleles PrS, global and pairwise F_{ST} among regions, and average nucleotide diversity within and among regions (π)) of sequence variation at targeted nuclear (n) and chloroplast (cp) loci in three bryophyte species with a main present distribution range in areas covered in ice at LGM. *, ** and *** significance level (P < 0.05, 0.01 and 0.001) for F_{ST}. N: number of sequenced specimens. The number following the assignment of a locus to the nuclear (n) or chloroplastic (cp) genome is the number of sequences obtained for that locus. 0: extra-European range. 1: northern range under the ice sheet at LGM. 2a: southern mountain range ice-free at LGM. 2b: southern mountain range covered in ice at LGM. 3: Lowland range south of the ice sheet at LGM

	He0	He1	He2a	He2b	He3	PrS0	PrS1	PrS2a	PrS2b	PrS3	π0	π1	π2a	π2b	π3	F_{ST}
Amphidium lapponicum (Hedw.) Schimp. (AL, N = 86)																
1. AW098256–290 (n = 48)	0.850	0.515	/	1.000	/	16	0	/	0	/	5.34	0.00	/	0.00	/	0.35***
2. AW098158–317 (n = 74)	0.772	0.527	/	0.500	/	28	0	/	0	/	12.29	1.42	/	1.00	/	0.32***
3. atpIH (cp = 55)	0.703	0.392	/	0.667	/	18	0	/	0	/	8.33	0.42	/	1.33	/	0.45***
Timnia austriaca Hedw. (TA, N = 128)																
1. AW098393–199 (n = 119)	0.580	0.280	0.640	0.592	1.000	4	0	0	0	0	0.68	0.29	0.84	0.66	1.00	0.17***
2. AW086999–142 (n = 128)	0.182	0.414	0.118	0.441	0.000	1	0	0	0	0	0.19	0.41	0.12	0.44	0.00	0.47***
3. AW086525–514 (n = 67)	0.603	0.111	0.000	0.282	0.000	2	0	0	2	0	1.12	0.33	0.00	1.13	0.00	0.05
4. atpIH (cp = 88)	0.189	0.087	0.000	0.417	0.000	1	0	0	1	0	0.19	0.09	0.00	1.00	0.00	0.10*
Timnia bavarica Hessl. (TB, N = 73)																
1. AW087072–205 (n = 53)	0.956	0.714	0.893	0.495	0.524	12	1	3	0	0	7.69	4.86	7.86	2.65	6.95	0.29***
2. AW098078–239 (n = 71)	0.557	0.607	0.803	0.754	0.607	2	0	0	0	0	0.64	0.79	1.20	1.23	0.96	0.22***
3. trnG (cp = 47)	0.303	0.607	0.733	0.564	0.600	0	1	0	0	0	0.30	0.68	0.93	0.62	0.60	0.07

Pairwise F_{ST}

	Fst01	Fst02a	Fst02b	Fst03	Fst2a1	Fst2a2b	Fst2a3	Fst2b1	Fst2b3	Fst13
Amphidium lapponicum (Hedw.) Schimp. (N = 86)										
1	0.38***	/	0.38***	/	/	/	/	0.00	/	/
2	0.09	/	0.09	/	/	/	/	0.00	/	/
3	0.37***	/	0.37***	/	/	/	/	0.33	/	/
Timnia austriaca Hedw. (N = 128)										
1	0.06*	0.27***	0.24***	0.08	0.25**	0.00	0.00	0.23*	0.00	0.17
2	0.65***	0.00	0.21***	0.00	0.58***	0.12	0.00	0.28*	0.00	0.51
3	0.05*	0.10*	0.06*	0.00	0.00	0.099	0.000	0.038	0.000	0.00
4	0.04	0.03	0.23*	0.00	0.00	0.18	0.000	0.17	0.000	0.00
Timnia bavarica Hessl. (N = 73)										
1	0.27***	0.19***	0.34***	0.34***	0.021	0.02	0.24	0.08	0.50***	0.50***
2	0.37***	0.15***	0.22***	0.31***	0.31**	0.00	0.11	0.25***	0.26***	0.44**
3	0.12*	0.08	0.08	0.13	0.02	0.00	0.03	0.00	0.21*	0.27*

Pairwise π

	π01	π02a	π02b	π03	π2a1	π2a2b	π2a3	π2b1	π2b3	π13
Amphidium lapponicum (Hedw.) Schimp. (N = 86)										
1	4.44	/	4.440	/	/	/	/		/	/
2	4.44	/	4.440	/	/	/	/		/	/
3	11.58	/	11.577	/	/	/	/		/	/
Timnia austriaca Hedw. (N = 128)										
1	0.52		0.878	0.82	0.72	0.72	0.68	0.60	0.56	0.54
2	0.77		0.372	0.10	0.70	0.32	0.06	0.59	0.29	0.72
3	0.78		1.198	0.65	0.17	0.61	0.00	0.75	0.61	0.17
4	0.15		0.658	0.10	0.04	0.56	0.00	0.58	0.56	0.04
Timnia bavarica Hessl. (N = 73)										
1	8.76	8.95	7.685	1.06	5.80	4.43	8.37	3.58	8.57	10.49
2	1.10	1.05	1.181	1.11	1.46	1.20	1.23	1.39	1.52	1.56
3	0.54	0.67	0.500	0.50	0.82	0.74	0.80	0.60	0.77	0.87

Table 2 Observed summary statistics (expected heterozygosity He, number of private alleles PrS, global and pairwise F_{ST} among regions, and average nucleotide diversity within and among regions (π)) of sequence variation at targeted nuclear (n) and chloroplast (cp) loci in 12 bryophyte species having their main present distribution range in ice-free areas at LGM. *, ** and *** significance level ($P < 0.05$, 0.01 and 0.001) for F_{ST}. N: number of sequenced specimens. The number following the assignment of a locus to the nuclear (n) or chloroplastic (cp) genome is the number of sequences obtained for that locus. 0: extra-European range. 1: northern European range. 2: southern European range.

	He0	He1	He2	PrS0	PrS1	PrS2	F_{ST}	Fst10	Fst20	Fst21	π10	π20	π21	π0	π1	π2
Amphidium mougeotii (Schimp.) Schimp. (AM, N = 142)																
1. AW098448–78 (n = 121)	0.38	0.57	0.38	0	1	0	0.00	0.00	0.00	0.00	0.580	0.404	0.614	0.38	0.77	0.46
2. AW098158–317 (n = 122)	0.72	0.57	0.75	1	3	2	0.01***	0.06****	0.15***	0.11***	1.82	2.72	2.61	1.65	1.77	2.99
3. atpI-H (cp = 101)	0.17	0.46	0.36	0	3	0	0.03	0.07	0.00	0.00	0.76	0.53	0.86	0.35	1.02	0.72
Calypogeia fissa (L.) Raddi (CF, N = 89)																
1. psbD-atpI-H-trnG-trnL (cp)	0.92	0.92	0.85	7	5	23	0.10***	0.09***	0.11*	0.13*	9.69	11.24	8.89	12.90	7.07	11.1
Diplophyllum albicans (L.) Dumort. (DA, N = 88)																
1. psbD-atpI-H-trnG (cp)	0.70	0.67	0.90	3	5	0	0.21***	0.27***	0.14*	0.00	15.19	15.05	9.61	13.86	8.90	11.52
Homalothecium sericeum (Hedw.) Schimp. (HS, N = 129)																
1. trnG-rpl16-atpB-rbcL (cp)	0.71	0.59	0.81	2	2	5	0.18***	0.23***	0.59***	0.02***	16.15	5.70	12.35	1.75	21.40	2.74
Metzgeria conjugata Lindb. (MC, N = 69)																
1. atpI-H-trnG (cp)	0.49	0.44	0.57	21	0	0	0.25**	0.32***	0.05	0.00	3.14	3.14	0.00	5.47	0.00	0.00
Metzgeria furcata (L.) Dumort. (MF, N = 142)																
1. atpI-H-trnG (cp)	0.84	0.74	0.70	76	36	1	0.31***	0.41***	0.35***	0.08*	30.58	29.99	7.07	41.39	8.28	5.15
Orthotrichum affine Schrad. ex Brid. (OA, N = 91)																
1. AW086770–115 (n = 79)	/	0.61	0.75	/	1	6	0.05	/	/	0.05	/	/	4.68	/	3.64	5.61
2. AW086975–261 (n = 53)	/	0.80	0.78	/	5	15	0.00	/	/	0.00	/	/	5.24	/	5.43	5.12
3. AW098158–317 (n = 79)	/	0.59	0.86	/	2	6	0.07*	/	/	0.07	/	/	1.73	/	1.28	1.98
4. trnG-rpl32 (cp = 63)	/	0.42	0.43	/	1	1	0.00	/	/	0.00	/	/	1.37	/	1.27	1.53
Orthotrichum lyellii Hook. & Taylor (OL, N = 214)																
1. AW098786–108 (n = 154)	0.89	0.73	0.55	11	0	0	0.13***	0.16***	0.16***	0.02	4.00	4.01	2.15	5.75	2.26	3.64
2. AW098679–454 (n = 167)	0.92	0.53	0.68	22	0	1	0.17	0.21	0.16	0.00	9.36	9.84	2.79	12.95	3.40	5.53
3. AW098158–317 (n = 183)	0.91	0.46	0.57	67	2	2	0.12	0.11	0.13	0.13	23.70	25.63	9.17	34.78	7.87	9.01
4. trnG (cp = 181)	0.81	0.29	0.25	5	4	0	0.45***	0.42***	0.28***	0.62***	2.56	1.97	2.30	2.03	1.75	1.47
Plagiomnium undulatum (Hedw.) T.J.Kop (PU, N = 165)																
1. AW086783–208 (n = 141)	0.51	0.52	0.61	0	1	0	0.37***	0.44***	0.04*	0.45***	14.04	11.27	12.93	11.74	5.06	9.95
2. AW086566–259 (n = 150)	0.57	0.26	0.51	0	0	0	0.40****	0.59****	0.01	0.43***	16.29	13.96	12.65	13.73	2.27	13.77
3. trnG-trnT-rps4 (cp = 80)	0.55	0.50	0.62	1	2	0	0.43****	0.57****	0.00	0.47***	2.28	1.57	1.95	1.56	0.62	1.61
Plagiothecium denticulatum (Hedw.) Schimp. (PD, N = 123)																
1. AW86789–334 (n = 115)	0.85	0.89	1.00	4	11	0	0.02	0.03*	0.00	0.00	3.20	2.98	3.32	2.62	4.01	4.17
2. atpI-H (cp = 103)	0.46	0.80	1.00	0	3	0	0.10**	0.12**	0.08	0.00	2.29	2.28	1.66	2.13	1.94	2.00
Plagiothecium undulatum (Hedw.) Schimp. (PW, N = 110)																
1. AW086779–195 (n = 103)	0.84	0.71	0.33	9	1	0	0.09***	0.11****	0.00	0.00	1.40	1.02	0.71	1.53	0.98	0.33
2. AW086975–261 (n = 75)	0.90	0.63	0.50	11	1	0	0.01	0.03	0.00	0.00	1.98	1.71	0.87	2.68	1.22	0.50
Scorpiurium circinatum (Bruch) M.Fleisch. & Loeske (SCI, N = 52)																
1. AW098158–317 (n = 51)	0.81	0.71	0.84	29	0	5	0.19***	0.32***	0.23***	0.02	12.14	11.36	7.50	9.70	6.65	7.99
2. rpl32-atpI-H (cp = 52)	0.68	0.80	0.77	50	0	2	0.09***	0.04	0.14***	0.01	4.89	4.84	0.70	8.50	0.76	0.63
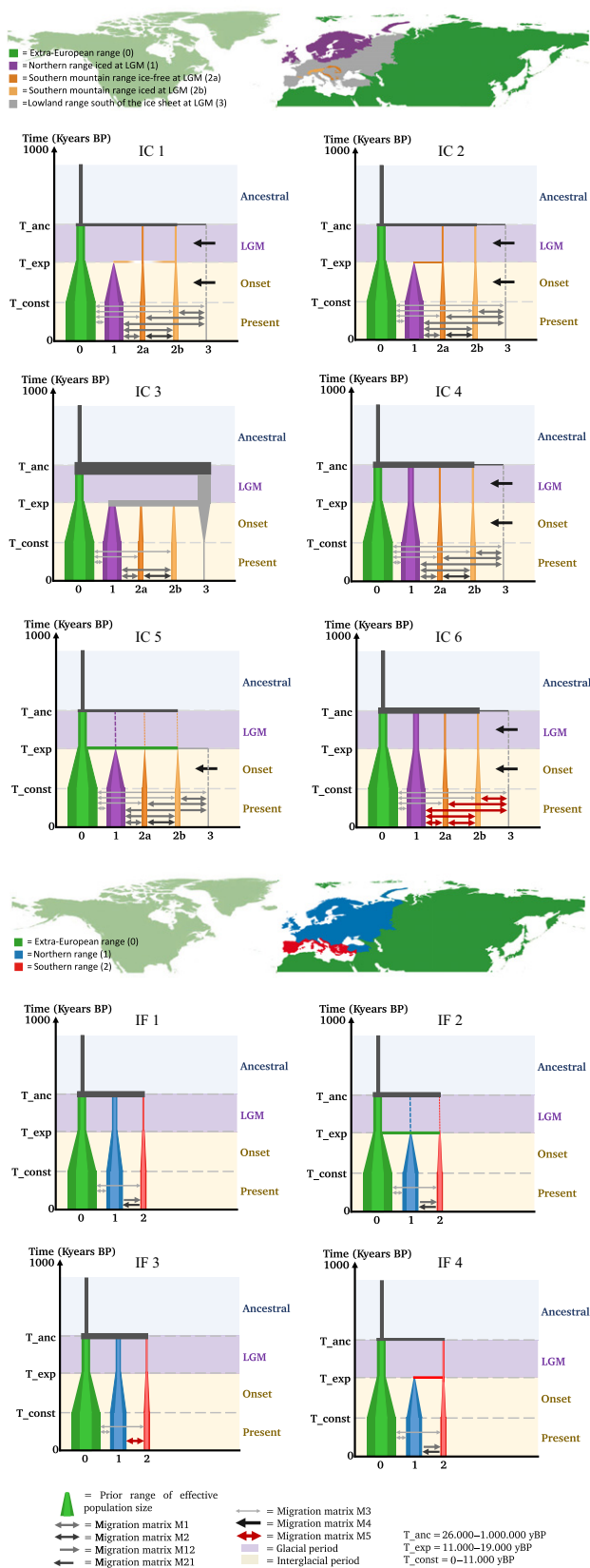
Figure 2 Hypothetical demographical scenarios for the post-glacial recolonisation of Europe since the Last Glacial Maximum in species with their present main distribution range in areas that were covered in ice and ice-free at LGM (scenarios IC and IF respectively). Arrows represent migration rates among regions. IC1: Southern mountain range refugia iced at LGM. IC2: Southern mountain range refugia ice-free at LGM. IC3: Lowland refugia south of the ice sheet at LGM. IC4: micro-refugia within the northern ice sheet and the southern mountain ranges. IC5: extra-European post-glacial recolonisation. IC6: no genetic structure due to intense post-glacial migrations. IF1: northern micro-refugia. IF2: extra-European post-glacial recolonisation. IF3: no genetic structure due to intense post-glacial migrations. IF4: southern refugia.

Statistical analyses

Characterising population structure and genetic diversity

Haplotype networks were computed for each species and locus to visualise the geographic partitioning of genetic variation and describe relationships among alleles from different regions with PEGAS (Paradis 2010). To characterise the genetic structure and diversity of each biogeographical region, the following summary statistics were computed with ARLSUMSTAT 3.5.2.2 (Excoffier & Lischer 2010) for each locus and species: expected heterozygosity (He), nucleotide diversity (π) and number of private polymorphic sites (PrS), global F_{ST}, F_{ST} and nucleotide diversity (π) between pairs of regions. In addition, we tested the hypothesis that the global F_{ST} was significantly different from 0 by randomly permuting specimens among regions 1000 times. Variation in summary statistics among species was described by a principal component analysis (PCA).

Design of the demographic scenarios

We determined whether sequence variation at each locus was compatible with competing historical demographic scenarios designed for IF and IC species, respectively, by implementing an ABC analysis (Lintusaari *et al.* 2017). Briefly, millions of allele genealogies are generated from coalescent simulations sampling a range of demographic parameters (effective population size Ne, migration rates). These genealogies are then used to map substitutions and generate simulated DNA alignments of the same size as the observed data. The simulations that exhibit the closest resemblance with the observed data are then selected to identify the demographic scenarios and their parameters that are the most likely to have generated the observed data (Fig. 1).

Our demographic scenarios involved four time slices (Fig. 2). The first time slice ranged between the final coalescence event and T_anc, at which all populations are merged, and at which Ne of the resulting population is divided by 1000 to allow the final coalescence of all remaining alleles. T_anc was sampled from a uniform distribution between 10^6 and 26 000 years BP. The second time slice ranged between T_anc and T_exp, which marks the beginning of the post-glacial expansion and was sampled from a uniform distribution between 19 000 and 11 000 years BP. During this second period, which encompasses the LGM, between 26 000 and 19 000 years BP, population sizes were constant. The third time slice corresponded to the post-glacial expansion, starting at T_exp and ending at T_const. The fourth time slice marked

concatenated and considered as one locus due to the linkage of chloroplastic genes, whereas each nuclear locus was considered individually.

the end of the expansion period, starting at T_const, which was sampled from a uniform distribution between 11 000–0 years BP, and ending at time present. During this period, Ne remained stable and migrations among all regions are allowed. This framework allowed to consider a large range of demographic scenarios, from extremely rapid expansions followed by stasis (T_exp and T_const both sampled at 11 000 years BP) to progressive and continuous expansion since the end of the LGM (T_exp sampled at 19 000 years BP, T_const sampled at time present).

For IF species, specific asymmetric migration rates described the movement of alleles between the southern and northern ranges (migration matrices M12 and M21, Fig. 2). For IC species, symmetric migrations were allowed between each European region (migration matrix M1, Fig. 2). A specific symmetric migration matrix (M2, Fig. 2) was implemented between the southern mountain ranges during the fourth time slice to allow these two populations to form a large, panmictic population. We controlled for migrations between the extra-European range and each of the European regions for all scenarios (symmetric migration matrix M3, Fig. 2).

For IC species, six demographic scenarios were designed (Fig. 2). In the southern mountain range refugium scenarios, populations from the northern range under the ice sheet at LGM originated at T_exp from refugia located in the southern glaciated (scenario IC1) or ice-free (scenario IC2) mountain ranges at LGM. The extant lowland population is considered to result from a recent founding event. To simulate this, all alleles in this lowland population were transferred into the northern areas and the southern mountain ranges by implementing a high migration rate (asymmetric migration matrix M4, Fig. 2) from T_const to T_anc. In scenario IC3, populations of the northern areas and the southern mountain ranges originated at T_exp from the lowland population. In this scenario, the latter represents the relicts of a large ancestral population at LGM. Populations of the northern areas and the southern mountain ranges underwent an expansion, whereas the lowland population underwent a bottleneck from T_exp to T_const. In scenario IC4, populations of the northern areas and the southern mountain ranges are the relicts of persisting populations in local micro-refugia since the LGM. Like in scenario IC1 and IC2, the lowland population is considered to result from a recent founding event from either the northern areas or the southern mountain ranges. Scenario IC5 involves that, although small populations might have persisted in Europe, the post-glacial recolonisation mainly took place from migrants of extra-European origin at T_exp, taking advantage of the empty niche space left in Europe after the LGM. We allowed a range ('Contri'), sampled from a uniform distribution between 80 and 100%, of individuals of extra-European origin, while the remaining 0–20% of individuals corresponded to relictual populations that persisted in Europe during the LGM. Like in scenarios IC1, IC2 and IC4, the population from the lowland areas was considered to result from a recent founding event from the northern areas or the southern mountain ranges. Finally, scenario IC6 represents our null hypothesis, according to which there is no significant genetic structure ($F_{ST} = 0$) due to high dispersal rates among European regions (migration matrix M5, Fig. 2) from T_const to time present.

For IF species, the northern refugium scenario (IF1) proposes that northern populations survived in situ in micro-refugia during LGM and were involved in the post-glacial recolonisation of northern Europe. Scenarios IF2 and IF3 describe an extra-European origin of the post-glacial recolonisation of Europe at T_exp and the absence of any spatial genetic structure, respectively, and their design is thus identical to the one of scenarios IC5 and IC6. In the southern refugium scenario (IF4), the post-glacial recolonisation of the southern range took place from Mediterranean refugia at T_exp.

Estimation and prior distributions of model parameters

We sampled uniform prior distributions conservatively bounded by the slowest and fastest rates observed across a range of species and loci in previous studies (Patiño et al. 2015; Désamoré et al. 2016). These prior distributions, which ranged between 10^{-7} and 10^{-4} substitutions/site/myr, were sampled independently, allowing for different posterior probability distributions to characterise the substitution rates of each locus.

Migration rates (matrices M1, M12, M21, M2 and M3) were independently sampled from a uniform distribution ranging between 0 and 10% of migrants per population per generation (Patiño et al. 2015; Désamoré et al. 2016) but, based on Désamoré et al. (2016), we enforced a rule according to which the intercontinental migration rate (M3) was lower than the migration rate within Europe (M1, M12, M21 and M2). The matrices describing high migration rates (matrices M4 and M5) were set at 0.2.

We used information on life strategies to determine the age at sexual maturity defining the generation time. All the species investigated in the present study are perennial and start producing sporophytes from about 10 years (During 1992). All bryophyte species are, however, capable of clonal reproduction, which may take place from the earliest developmental stages. Therefore, we set the generation time at 5 years, which represents a compromise between early asexual reproduction and delayed sexual reproduction. The generation time was fixed instead of sampled from a prior distribution because uncertainty in the time estimates was taken into account by allowing a range, instead of a fixed value, for the time periods defined (T_const, T_exp and T_anc). The fact that the mode of the posterior distributions of those parameters was either centred or shifted towards the highest or lowest values of the prior range (Fig. S2) suggests that our estimate of the generation time is realistic, as an over- or underestimation of it would lead to posterior distributions of those parameters that would consistently increase or decrease towards the boundaries of the prior distribution.

The present Ne values were sampled from uniform prior distributions that were defined for each species (Table S2a,b), except for the lowland population in scenarios IC1–6, which was associated with a uniform prior ranging between 1 and 50 individuals. These prior distributions were derived using species distribution models to predict the number of macroclimatically suitable pixels of 5 km^2. This number was subsequently multiplied by the expected value of Ne per pixel from Patiño et al. (2015) and Désamoré et al. (2016) (see

Appendix S1 for details). The fact that the mode of the posterior probability distributions of Ne for all regions and species (Fig. S2) was comprised within the boundaries of the prior distributions confirms that the priors that we employed were realistic. To simulate the large ancestral lowland population in scenario IC3, a uniform prior distribution was set for its LGM Ne, whose range was defined using species distribution models projected on LGM climatic variables.

To simulate bottlenecks, we constrained Ne at LGM to be a portion of Ne at time present by multiplying the latter by a factor R sampled from a uniform prior distribution ranging from 0.0001 to 0.01. To simulate the founding effect experienced by extra-European migrants at T_exp in scenarios IC5 and IF3, we sampled R in Europe from a uniform distribution ranging between 0.005 and 0.0005, and we further implemented a rule according to which the size of the European population was always at least 10 times lower than that of the extra-European one. To simulate the founding effect experienced by migrants during the colonisation of the northern range at T_exp in scenarios IF4, IC1 and IC2, respectively, Ne of these populations at T_exp was sampled from a uniform distribution ranging between 2 and 100. We further implemented a rule according to which Ne of these populations was at least ten times lower than the size of the populations of the southern (scenario IF4) or the southern mountain ranges (scenarios IC1-2).

ABC analyses and model selection
Approximate Bayesian computation analyses were conducted using ABCtoolbox2.0 (Wegmann *et al.* 2010) in combination with fastsimcoal2 (Excoffier *et al.* 2013). For each species and scenario, 10^6 coalescent simulations were conducted. In each simulation, parameter values were randomly sampled from prior distributions described above. The nucleotide substitution model used to simulate the sequence data from each tree and for each locus was a Kimura 2-parameter with a transition–transversion ratio set at 0.33. Each analysis resulted in 10^6 simulated sequence data sets identical in size (i.e. number and length of sequences) to the corresponding observed data set.

We computed the summary statistics listed above for all simulated data. Highly correlated statistics (correlation coefficient > 95%) were removed. Euclidean distances were calculated between the normalised observed and simulated summary statistics. For each scenario and species, we retained the best 1000 simulations (i.e. those with the smallest Euclidean distance between the simulated and observed summary statistics). A post-sampling regression adjustment (ABC-GLM) was finally performed to derive the posterior probability distributions of each scenario and model parameters (Leuenberger & Wegmann 2010) from this first selection. The posterior probability distributions were smoothed using the diracPeakWidth parameter, set at 0.01.

Two model validation procedures were implemented. First, we determined whether the summary statistics of the observed data fell within the range of the simulated summary statistics. We computed the marginal density *P*-value and implemented a PCA of the summary statistics of the observed and 1000 best simulated data of each of the competing demographic

scenarios. Second, we measured the proportion of 'false positives' that is the selection of the wrong model during model selection (Robert *et al.* 2011). We sampled a total of 1000 sets of summary statistics simulated under each demographic scenario. These sets of 'pseudo-observed' data were analysed using the same procedure of model selection as described above to compute the percentage of simulations erroneously assigned to each scenario.

RESULTS

Network analyses and distribution maps of alleles can be found in Fig. S1. The widespread distribution of many alleles, together with the sharing of alleles among specimens from different biogeographical regions, the close relationships among alleles sampled from different biogeographical regions and the presence of unrelated alleles in individuals from the same biogeographical region, all point to intense migrations among regions.

The comparison of the observed patterns of genetic structure and diversity in the 15 selected species with those expected under competing post-glacial recolonisation scenarios in an ABC framework in fact points to a complex post-glacial history with a substantial contribution from populations located in their extra-European range. The scenario of a post-glacial recolonisation of Europe by extra-European migrants (IF2) had the highest posterior probability in 7 of the 12 IF species investigated, followed by the southern (IF4, 3 species) and northern (IF1, 2 species) refugium scenarios (Fig. 3). The scenario, according to which recent migrations erased any genetic structure (IF3), was never selected. For the three IC species investigated, the scenarios of local micro-refugia (IC4) on the one hand, and of recolonisation from lowland areas located south of the ice sheet at LGM (IC3) on the other hand, were selected for two and one of the species respectively (Fig. 3).

The examination of the posterior probability distributions of the demographic parameters (Fig. S2) reveals that, for the species that conform to a recolonisation of Europe from allochthonous origin, the distribution of the proportion of migrants of allochthonous origin at the beginning of the expansion (parameter 'Contri' in Fig. S2) peaks at the highest values of the prior range (i.e. 95–100%), except in *Plagiothecium undulatum*. In the three investigated IC species, the mode of the posterior distribution of the ongoing migration rate between Europe and extra-European regions is shifted towards the right of the mode of the prior distribution, at values of 2–4% of migrants per generation.

The posterior distribution of the timing of the beginning of the post-glacial expansion (T_exp, Fig. S2) consistently peaked at the lowest values of the prior distribution (approximately around 11 000–12 000 years BP), pointing to a delay of the post-glacial expansion, which did not occur directly at the end of the LGM, around 19 000 years BP. In five out of the seven species that conformed to the scenario of post-glacial recolonisation of extra-European origin, the distribution of the timing of the end of the expansion (T_const, Fig. S2) peaked at the highest values of the prior range (around 10 000–11 000 years BP). This timing indicates that, for

(a)
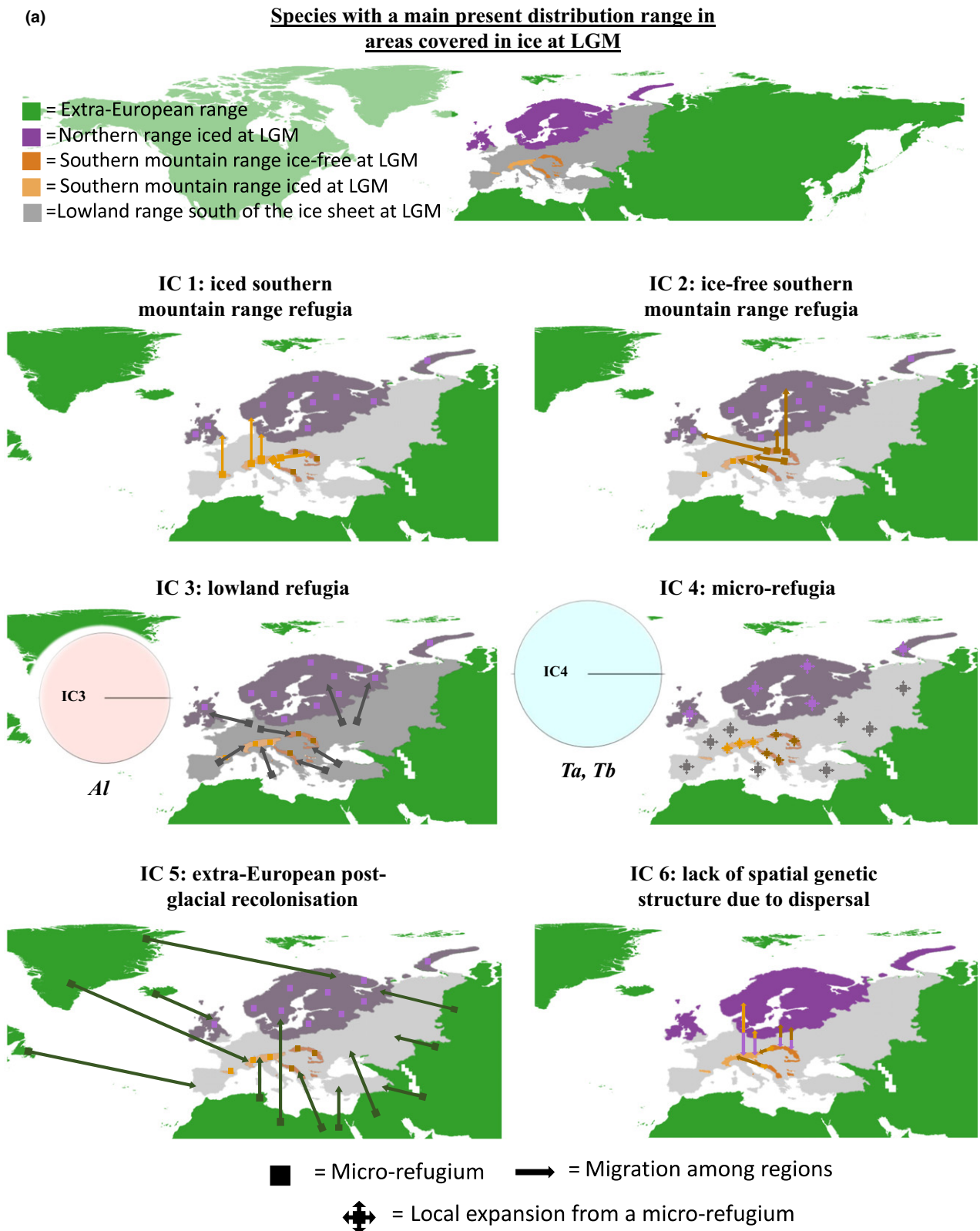
Figure 3 Support for competing post-glacial demographic scenarios in bryophyte species with their present main distribution range in areas that were covered in ice (a) and ice-free (b) at LGM (scenarios IC and IF, respectively) inferred by ABC model selection. Pie diagrams reflect the posterior probabilities of each scenario. Arrows represent migrations and squares represent hypothetical micro-refugia. *Al: Amphidium lapponicum; Am: Amphidium mougeotii; Cf: Calypogeia fissa; Da: Diplophyllum albicans; Hs: Homalothecium sericeum; Mc: Metzgeria conjugata; Mf: Metzgeria furcata; Oa: Orthotrichum affine; Ol: Orthotrichum lyellii; Pd: Plagiothecium denticulatum; Pw: Plagiothecium undulatum; Pu: Plagiomnium undulatum; Sci: Scorpiurium circinatum; Ta: Timmia autriaca; Tb: Timmia bavarica.*

(b)   **Species with a main present distribution range in ice-free areas at LGM**
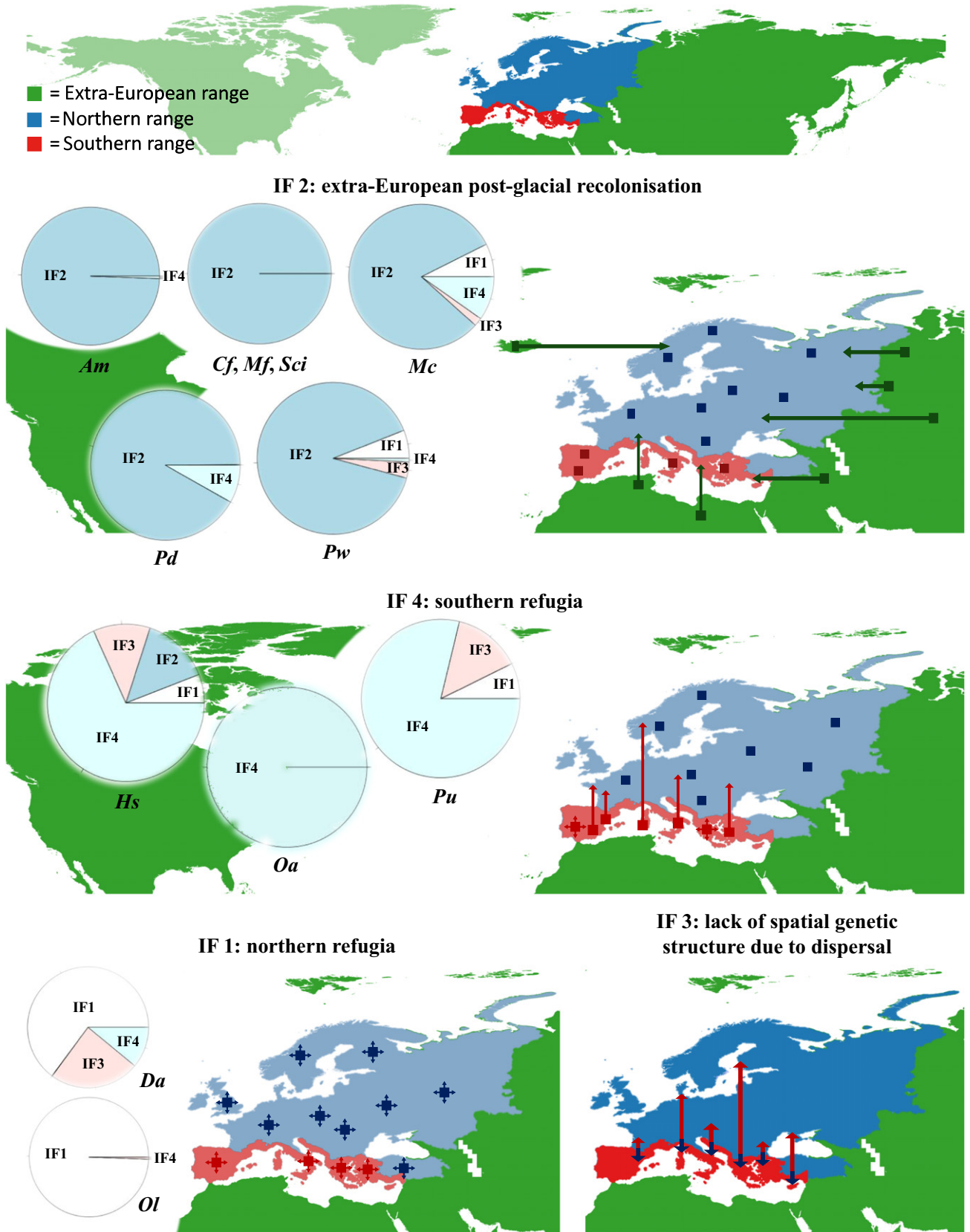

Figure 3 continued

these species, the post-glacial expansion was limited in time and followed by a long (about 10 000 years) period of stable population size until present time. In all the other species, the posterior distribution of the end of the post-glacial expansion peaked at the lowest values of the prior range (around time present), pointing to either a progressive expansion that ceased recently or is still underway, or to a very limited period of migrations among regions (since in all scenarios, migration among regions was permitted only after population expansion). The posterior probability distributions of migration rates between northern and southern areas in Europe after the expansion period consistently peaked at the lowest values of the prior range, pointing to the latter hypothesis.

A summary of the model validation procedures is presented in Table S3. The simulated data laid within the range of the observed data that is the values of the summary statistics computed from the observed data were included within the variation of those derived from the 1000 best simulations for each demographic scenario, as revealed on the first two axes of a PCA (Fig. S3a, b). In addition, the marginal density P-values of the best-fit models were higher than 0.05 in 10 out of the 15 species investigated, leading to the acceptance of the null hypothesis that the simulated data are

compatible with the observed data. The percentage of false positives during model selection was generally low, except for an obvious bias in the case of the northern micro-refugium scenario. The rate of false positive for the latter in *Diplophyllum albicans* and *Orthotrichum lyellii* suggests that patterns of genetic variation in these two species are compatible with different scenarios, strongly weakening the confidence for model selection in this case.

The observed population structure and genetic diversity summary statistics are displayed per species and per locus in Tables 1 and 2. The PCA of these statistics shows that species assigned to the same historical scenario of post-glacial recolonisation tend to exhibit similar genetic characteristics that differentiate them from species assigned to different scenarios (Fig. 4a). Thus, species assigned to the scenario of post-glacial recolonisation from extra-European origin, which occupy the upper right portion of the two first PCA axes, tend to exhibit lower global F_{ST} (Fig. 4b1), lower pairwise F_{ST} values (Fig. 4b2), and more similar genetic diversities (Table 2) between the southern and northern ranges than species assigned to other scenarios. Species assigned to the southern refugium scenario, which have negative coordinates along PCA2, exhibit high nucleotide diversities in the southern range, while species assigned to the northern refugium
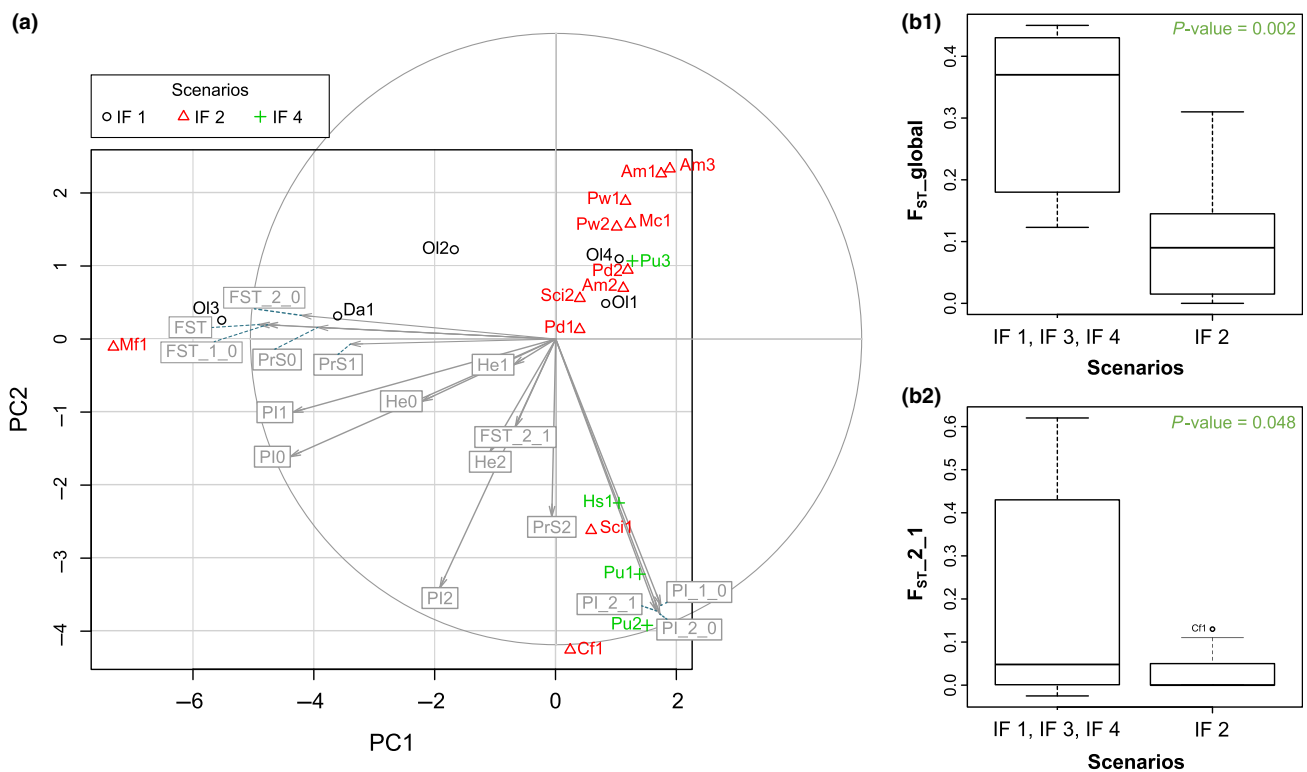
Figure 4 (a) Principal component analysis of the summary statistics characterising population structure and genetic diversity (expected heterozygosity He, nucleotide diversity π and number of private polymorphic sites PrS per region, global F_{ST}, F_{ST} and π between each pair of regions) per locus (see Table 2 for locus numbers) and region (0: extra-European range; 1: northern European range; 2: southern European range) in selected bryophyte species with their core present distribution in ice-free areas at LGM (see Table 2 for labels). Arrow lengths are proportional to the correlation coefficient of the corresponding summary statistic and the axes. Colours correspond to the assignment of each species to the scenario of post-glacial recolonisation of Europe (see Fig. 3), as selected by the ABC-GLM procedure. (b) Box plots and P-values of t-test comparisons of selected summary statistics whose relevance is highlighted by the PCA between species assigned to a scenario of allochthonous (IF2) or autochthonous (IF1, 3–4) post-glacial recolonisation of Europe respectively. (1) global F_{ST}. (2) F_{ST} between northern and southern Europe (2).

scenario exhibit high nucleotide diversities and high numbers of private segregating sites in northern areas.

DISCUSSION

Patterns of post-glacial recolonisation in European bryophytes are not consistent with the idea that the intensity of migrations erased any genetic structure. The diversity of scenarios involved highlights a complex post-glacial history of European bryophytes from multiple sources. This complexity contrasts with the prevailing model in which species that have their present distribution mainly in areas that were ice-free at LGM migrated northwards from southern refugia (Hewitt 2000; Petit *et al.* 2003). The validity of this historical demographic model has been recently challenged for North American trees (Lumibao *et al.* 2017). Our findings thus further challenge the taxonomic generality of the southern refugium scenario that has long been assumed to explain the distribution of genetic variation in bryophytes based on single-species analyses (Cronberg 2000; Grundmann *et al.* 2008).

The dominant scenario of post-glacial recolonisation of Europe by bryophyte species that are mainly distributed today in areas that were ice-free at LGM involved a major contribution of migrants of allochthonous origin. The globally balanced genetic diversities and extremely low divergence observed between southern and northern regions are fully consistent with such a scenario, in contrast to the gradual decrease in genetic diversity towards the north observed in seed plants (Petit *et al.* 2003).

Previous evidence in angiosperms demonstrated the post-glacial recolonisation of a remote Arctic archipelago from distant sources (Alsos *et al.* 2007). Individual instances of extra-European refugia (Hutsemékers *et al.* 2011; Laenen *et al.* 2011; Stenøien *et al.* 2011) and a *de novo* colonisation of oceanic areas from Macaronesian ancestors (Patiño *et al.* 2015) were further evidenced during the post-glacial history of European bryophytes. The present study is, however, the first to demonstrate and quantify, from a small but independent sample of 12 species that have their main present distribution ranges in ice-free areas at LGM, the substantial contribution of allochthonous migrants in the post-glacial assembly of European bryophyte floras, which is unparalleled in other plants and animals.

Although there was a slight shift of the mode of the posterior probability distribution towards high values of the migration rate between the European and extra-European range in species having their core present distribution range in iced areas at LGM, which is fully compatible with the substantial contribution of extra-European migrants in species distributed in ice-free areas at LGM, *Amphidium lapponicum, Timmia austriaca* and *Timmia bavarica* exhibited a strikingly different pattern of *in situ* persistence at their current locations, some of which were fully glaciated at LGM. Such a scenario is fully compatible with the reported ability of some species to remain viable after more than 1000 years in ice (La Farge *et al.* 2013; Roads *et al.* 2014).

The difference in the main origins of species that are currently distributed in areas that were ice-free and covered in ice at LGM is puzzling. This difference cannot, at first sight, be interpreted in terms of life-history traits. In *Amphidium*, for example, *A. lapponicum* produces large numbers of capsules with small spores, but recolonised Europe from autochthonous populations, whereas in *A. mougeotii*, which seldom produces sporophytes and does not produce specialised asexual diaspores, recolonisation took mostly place from allochthonous migrants.

Projections of the species distribution models calibrated under present climatic conditions onto LGM climatic layers (Fig. S4) predict that, with the exception of a few species (*Scorpiurium circinatum* and, to a lesser extent, *D. albicans* and *P. undulatum*), southern Europe and the southern range of northern Europe were extensively climatically favourable at LGM, pointing to the potential existence of sufficiently large refugia within Europe. Palaeontological evidence indicates, however, that the full-glacial landscape of Eurasia was largely treeless, with a dominance of steppe and other tundra types of vegetation, suggesting that palaeoclimatic reconstructions, on which our species distribution models are built, predicted a warmer and moister climate than it probably was (Tzedakis *et al.* 2013; Binney *et al.* 2017). Therefore, the limited contribution of the northern micro-refugium scenario, which contrasts with our primary expectations, and of the southern refugium scenario to explain the post-glacial history of the investigated species may be due to the fact that European refugia have been too small and too scattered as compared to the substantial waves of extra-European origin to actually contribute to the post-glacial recolonisation of the continent. As Semerikov *et al.* (2013) in fact suggested, it should not be firmly concluded that putative refugial populations necessarily contributed extensively to local modern populations, as the spread of new individuals from adjacent regions would have occurred over several millennia as climates changed (Binney *et al.* 2017).

A consistent signal for a delay in the post-glacial recolonisation of Europe since the end of the LGM (19 000 years BP) was evidenced across all investigated species by the shift of the posterior probability distribution marking the beginning of the expansion phase towards recent periods. Such a delay could be interpreted in terms of either unsuitable conditions at the beginning of the current interglacial period, and/or a delay in the recolonisation of suitable habitats. The critical transition from predominantly glacial to largely interglacial, moister climates inferred from major changes in fossil pollen records 14 000 years BP (Pearson 2006) supports the former hypothesis. A delay in the recolonisation of suitable habitats is, conversely, not supported by the match between the predicted and observed northern limit of distribution in the bryophyte species investigated here (Fig. S5), which contrasts with the absence of many angiosperm species in the North of their potential distribution area (Svenning *et al.* 2008; Normand *et al.* 2011). Striking range shifts have already been documented throughout Europe in bryophytes during the past decades (Bosanquet 2012), promoting the idea that major modifications are to be expected in the European bryophyte flora under ongoing climate change.

AUTHORSHIP

DATA ACCESSIBILITY STATEMENT

DNA sequence data are archived in GenBank. DNA alignment matrices and information on species distributions are available from Figshare Repositories: https://doi.org/10.6084/m9.figshare.7381130 and https://doi.org/10.6084/m9.figshare.7422794.

REFERENCES

Abbott, R.J. & Brochmann, C. (2003). History and evolution of the arctic flora: in the footsteps of Eric Hultén. *Mol. Ecol.*, 12, 299–313.

Alsos, I.G., Eidesen, P.B., Ehrich, D., Skrede, I., Westergaard, K., Jacobsen, G.H. *et al.* (2007). Frequent long-distance plant colonization in the changing Arctic. *Science*, 316, 1606–1609.

Bhagwat, S.A. & Willis, K.J. (2008). Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *J. Biogeogr.*, 35, 464–482.

Biersma, E.M., Jackson, J.A., Hyvönen, J., Koskinen, S., Linse, K., Griffiths, H. *et al.* (2017). Global biogeographic patterns in bipolar moss species. *Roy. Soc. Open Sci.*, 4, 170147.

Biersma, E.M., Jackson, J.A., Stech, M., Griffiths, H., Linse, K. & Convey, P. (2018). Molecular data suggest long-term *in situ* Antarctic persistence within Antarctica's most speciose plant genus, *Schistidium*. *Front. Ecol. Evol.*, 6, 77.

Binney, H., Edwards, M., Macias-Fauria, M., Lozhkin, A., Anderson, P., Kaplan, J.O. *et al.* (2017). Vegetation of Eurasia from the last glacial maximum to present: key biogeographic patterns. *Quat. Sci. Rev.*, 157, 80–97.

Bosanquet, S. (2012). Vagrant epiphytic mosses in England and Wales. *Field Bryol.*, 107, 3–17.

Cannone, N., Corinti, T., Malfasi, F., Gerola, P., Vianelli, A., Vanetti, I. *et al.* (2017). Moss survival through *in situ* cryptobiosis after six centuries of glacier burial. *Sci. Rep.*, 7, 4438.

Cronberg, N. (2000). Genotypic diversity of the epiphytic bryophyte *Leucodon sciuroides* in formerly glaciated versus nonglaciated parts of Europe. *Heredity*, 84, 710–720.

Darwin, C. (1859). *The Origin of Species*. John Murray, London.

Désamoré, A., Patiño, J., Mardulyn, P., Laenen, B., McDaniel, S., Zanatta, F. *et al.* (2016). High migration rates shape the postglacial history of amphi-Atlantic bryophytes. *Mol. Ecol.*, 25, 5568–5584.

During, H.J. (1992). Ecological classification of bryophytes and lichens. In *Bryophytes and Lichens in Changing Environment*. (eds Bates, J.W., Farmer, A.M.). Clarendon Press, Oxford, pp. 1–31.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5, 113.

Eidesen, P.B., Ehrich, D., Bakkestuen, V., Alsos, I.G., Gilg, O., Taberlet, P. *et al.* (2013). Genetic roadmap of the Arctic: plant dispersal highways, traffic barriers and capitals of diversity. *New Phytol.*, 200, 898–910.

Excoffier, L. & Lischer, H.E.L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.*, 10, 564–567.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C. & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genet.*, 9, e1003905.

Gouy, M., Guindon, S. & Gascuel, O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.*, 27, 221–224.

Grundmann, M., Ansell, S.W., Russell, S.J., Koch, M.A. & Vogel, J.C. (2008). Hotspots of diversity in a clonal world — the Mediterranean moss *Pleurochaete squarrosa* in Central Europe. *Mol. Ecol.*, 17, 825–838.

Hewitt, G.M. (1999). Post-glacial recolonisation of European biota. *Biol. J. Linn. Soc.*, 68, 87–112.

Hewitt, G.M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.

Hewitt, G.M. (2004). Genetic consequences of climatic oscillations in the quaternary. *Phil. Trans. Roy. Soc. B*, 359, 183–195.

Hooker, J.D. (1862). Outlines on the distribution of arctic plants. *Trans. Linn. Soc. London*, 23, 251–348.

Hultén, E. (1937). *Outline of the History of Arctic and Boreal Biota During the Quarternary Period*. Lehre J. Cramer, New York.

Hutsemékers, V., Shaw, A.J., Szvovenyi, P., Gonzalez-Mancebo, J.M., Muñoz, J. & Vanderpoorten, A. (2011). Islands are not sinks of biodiversity in spore-producing plants. *Proc. Natl Acad. Sci. USA*, 108, 18989–18994.

Jakab, G. & Sümegi, P. (2011). The role of bryophyte paleoecology in quaternary climate reconstructions. In *Bryophyte Ecology and Climate Change*. (eds Tuba, Z., Slack, N.G., Stark, L.R.). Cambridge University Press, Cambridge, pp. 335–358.

Kyrkjeeide, M.O., Stenøien, H.K., Flatberg, K.I. & Hassel, K. (2014). Glacial refugia and post-glacial colonization patterns in European bryophytes. *Lindbergia*, 37, 47–59.

Kyrkjeeide, M.O., Hassel, K., Flatberg, K.I., Shaw, A.J., Brochmann, C. & Stenøien, H.K. (2016). Long-distance dispersal and barriers shape genetic structure of peatmosses (*Sphagnum*) across the Northern Hemisphere. *J. Biogeogr.*, 43, 1215–1226.

La Farge, C., Krista, H., Williams, K.H. & England, J.H. (2013). Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. *Proc. Natl Acad. Sci. USA*, 110, 9839–9844.

Laenen, B., Désamoré, A., Devos, N., Shaw, A.J., Carine, M.A., Gonzalez-Mancebo, J.M. *et al.* (2011). Macaronesia: a source of hidden genetic diversity for post-glacial recolonization of western Europe in the leafy liverwort *Radula lindenbergiana*. *J. Biogeogr.*, 38, 631–639.

Leuenberger, C. & Wegmann, D. (2010). Bayesian computation and model selection without likelihoods. *Genetics*, 184, 243–252.

Lintusaari, J., Gutmann, M.U., Dutta, R., Kaski, S. & Corander, J. (2017). Fundamentals and recent developments in Approximate Bayesian Computation. *Syst. Biol.*, 66, e66–e82.

Lumibao, C.Y., Hoban, S.M. & McLachlan, J. (2017). Ice ages leave genetic diversity 'hotspots' in Europe but not in Eastern North America. *Ecol. Lett.*, 20, 1459–1468.

McDaniel, S.F. & Shaw, A.J. (2005). Selective sweeps and intercontinental migration in the cosmopolitan moss *Ceratodon purpureus* (Hedw.). *Brid. Mol. Ecol.*, 14, 1121–1132.

McDaniel, S.F., Baren, M.Jv, Jones, K.S., Payton, A.C. & Quatrano, R.S. (2013). Estimating the nucleotide diversity in *Ceratodon purpureus* (Ditrichaceae) from 218 conserved exon-primed, intron-spanning nuclear loci. *Appl. Plant Sci.*, 1, 1200387.

Médail, F. & Diadema, K. (2009). Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *J. Biogeogr.*, 36, 1333–1345.

Normand, S., Ricklefs, R.E., Skov, F., Bladt, J., Tackenberg, O. & Svenning, J.C. (2011). Postglacial migration supplements climate in determining plant species ranges in Europe. *Proc. R. Soc. Lond. B*, 278, 3644–3653.

Paradis, E. (2010). Pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26, 419–420.

Patiño, J., Carine, M., Mardulyn, P., Devos, N., Mateo, R.G., González-Mancebo, J.M. *et al.* (2015). Approximate Bayesian Computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. *Syst. Biol.*, 64, 579–589.

Pearson, R.G. (2006). Climate change and the migration capacity of species. *Trends Ecol. Evol.*, 21, 111–113.

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

Pisa, S., Biersma, E.M., Convey, P., Patiño, J., Vanderpoorten, A., Werner, O. *et al.* (2013). The cosmopolitan moss *Bryum argenteum* in Antarctica: back-colonization from extra-regional Pleistocene refugia or in-situ survival? *Polar Biol.*, 37, 1469–1477.

Roads, E., Longton, R.E. & Convey, P. (2014). Millenial timescale regeneration in a moss from Antarctica. *Curr. Biol.*, 24, R222–R223.

Robert, C.P., Cornuet, J.M., Marin, J.M. & Pillai, N.S. (2011). Lack of confidence in approximate Bayesian computation model choice. *Proc. Natl Acad. Sci. USA*, 108, 15112–15117.

Schönswetter, P., Paun, O., Tribsch, A. & Niklfeld, H. (2003). Out of the Alps: colonization of Northern Europe by East Alpine populations of the Glacier Buttercup *Ranunculus glacialis* L. (Ranunculaceae). *Mol. Ecol.*, 12, 3373–3381.

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.

Schönswetter, P., Popp, M. & Brochmann, C. (2006). Rare arctic-alpine plants of the European Alps have different immigration histories: the snow bed species *Minuartia biflora* and *Ranunculus pygmaeus*. *Mol. Ecol.*, 15, 709–720.

Semerikov, V.L., Semerikova, S.A., Polezhaeva, M.A., Kosintsev, P.A. & Lascoux, M. (2013). Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): a range-wide analysis of cytoplasmic markers. *Mol. Ecol.*, 22, 4958–4971.

Skrede, I., Eidesen, P.B., Portela, R.P. & Brochmann, C. (2006). Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Mol. Ecol.*, 15, 1827–1840.

Stenøien, H.K., Shaw, A.J., Shaw, B., Hassel, K. & Gunnarsson, U. (2011). North American origin and recent European establishment of the amphi-Atlantic peat moss *Sphagnum angermanicum*. *Evolution*, 65, 1181–1194.

Sundberg, S. (2005). Larger capsules enhance short-range spore dispersal in Sphagnum, but what happens further away? *Oikos*, 108, 115–124.

Svenning, J.C., Normand, S. & Kageyama, M. (2008). Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *J. Ecol.*, 96, 1117–1127.

Szövényi, P., Terracciano, S., Ricca, M., Giordano, S. & Shaw, A.J. (2008). Recent divergence, intercontinental dispersal and shared polymorphism are shaping the genetic structure of amphi-Atlantic peatmoss populations. *Mol. Ecol.*, 17, 5364–5377.

Szövényi, P., Sundberg, S. & Shaw, A.J. (2012). Long-distance dispersal and genetic structure of natural populations: an assessment of the inverse isolation hypothesis in peat mosses. *Mol. Ecol.*, 21, 5461–5472.

Thomé, M.T. & Carstens, B.C. (2016). Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. *Proc. Natl Acad. Sci. USA*, 113, 8010–8017.

Tzedakis, P.C., Emerson, B.C. & Hewitt, G.M. (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends Ecol. Evol.*, 28, 696–704.

Van der Velde, M. & Bijlsma, R. (2003). Phylogeography of five *Polytrichum* species within Europe. *Biol. J. Linn. Soc.*, 78, 203–213.

Vanderpoorten, A., Gradstein, S.R., Carine, M.A. & Devos, N. (2010). The ghosts of Gondwana and Laurasia in modern liverwort distributions. *Biol. Rev. Camb. Philos. Soc.*, 85, 471–487.

Wegmann, D., Leuenberger, C., Neuenschwander, S. & Excoffier, L. (2010). ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, 11, 116.

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.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.
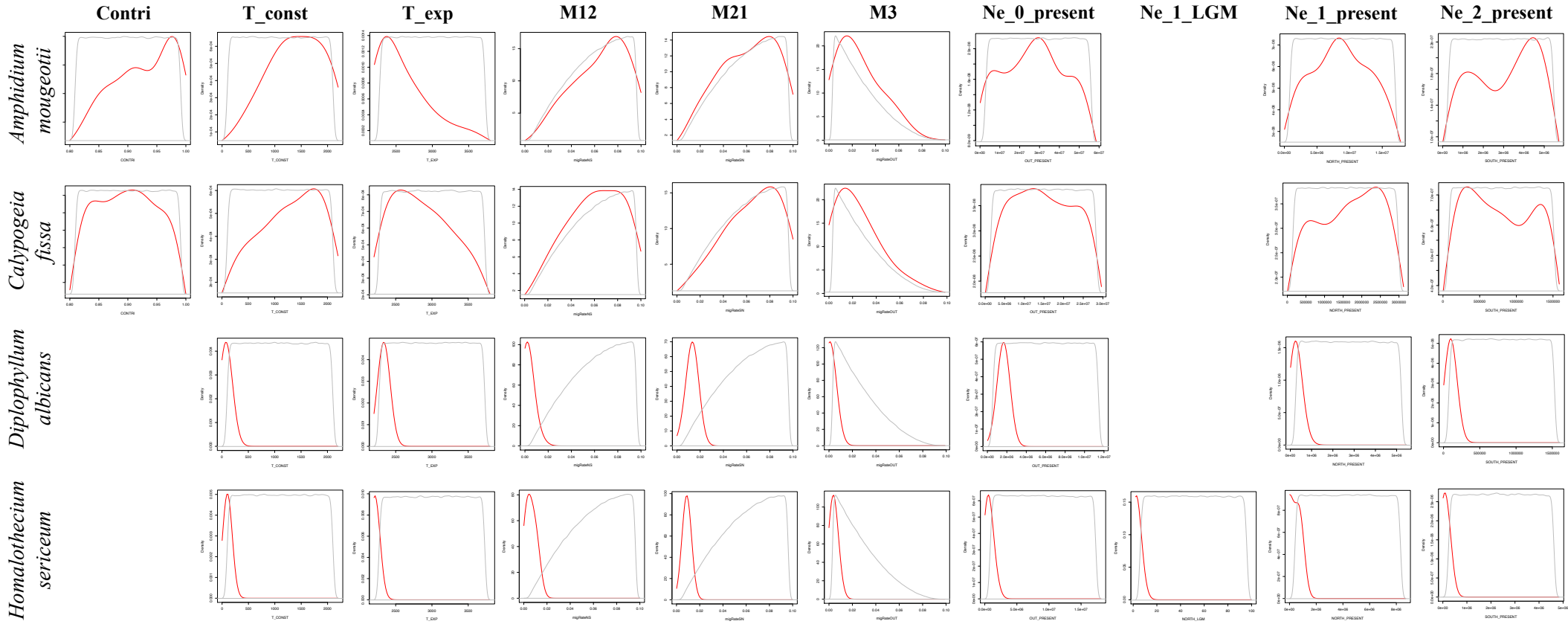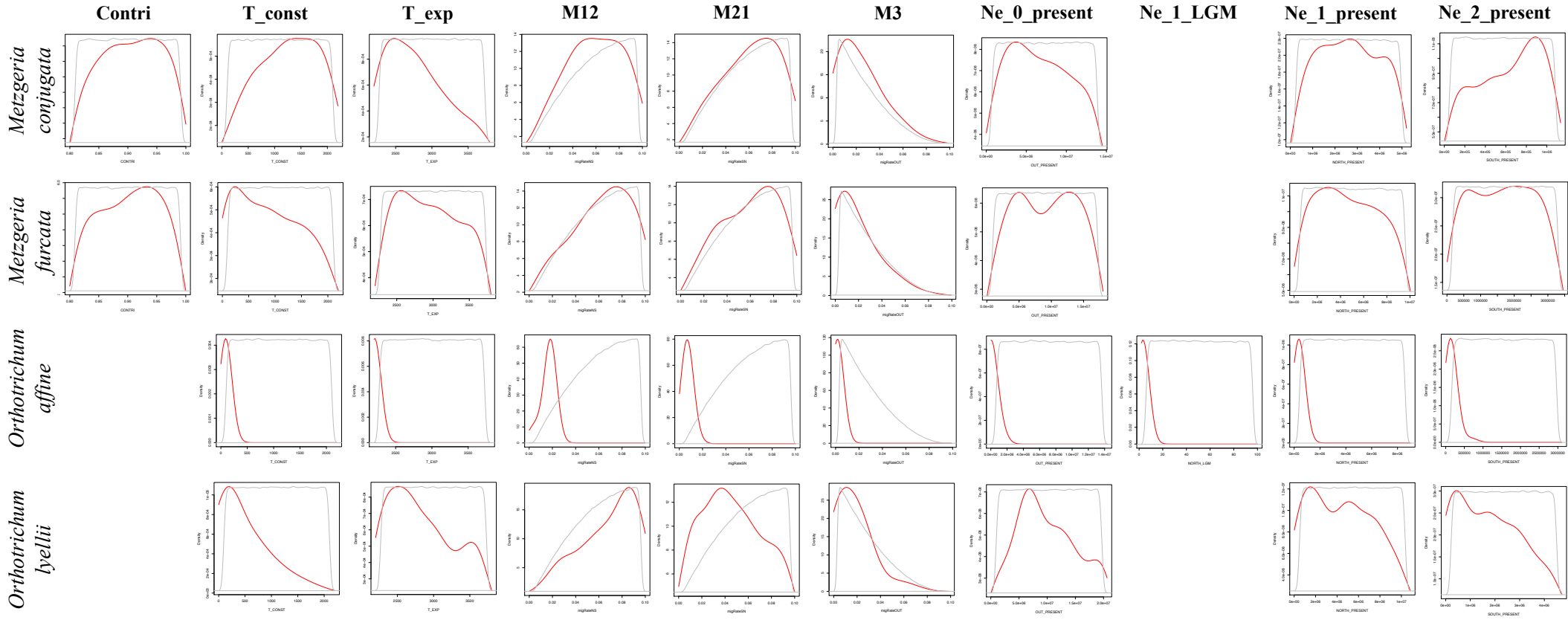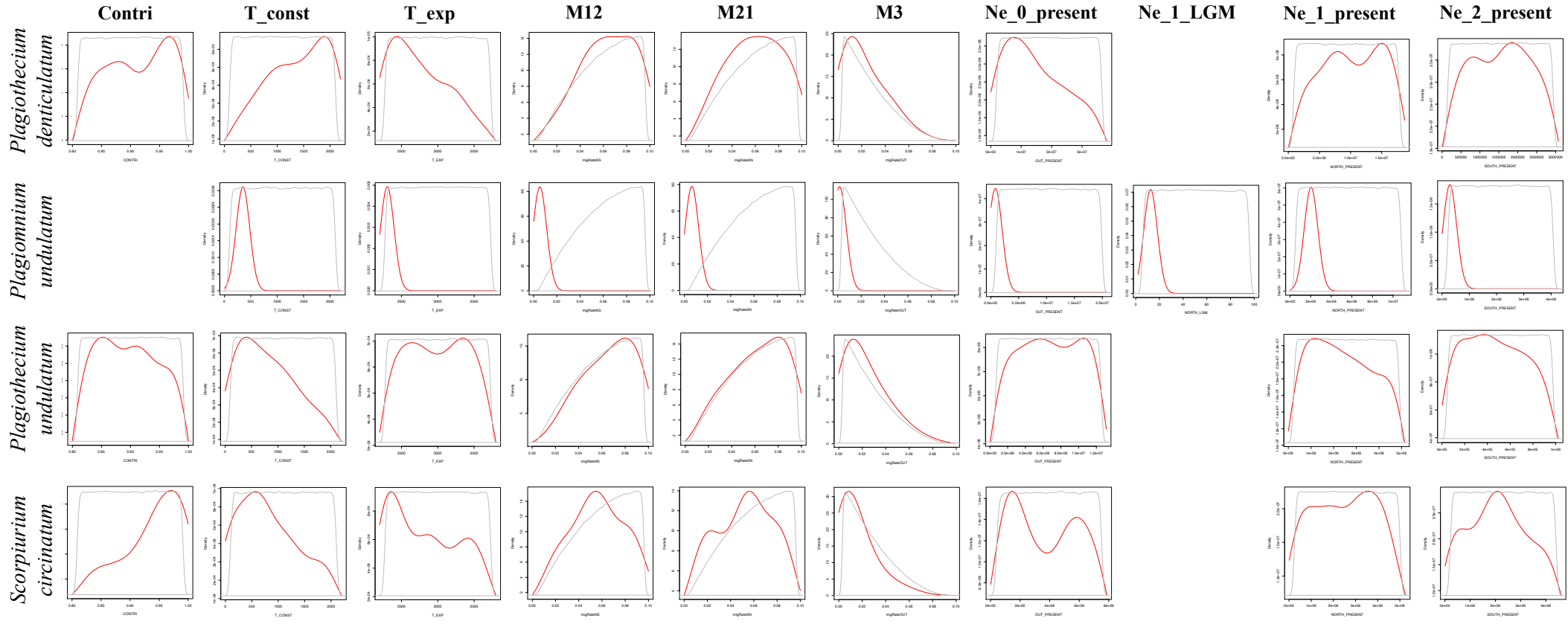
Figure S2 *Prior* (grey line) and *posterior* (red line) probability distributions of key demographic parameters under the best-fit scenario of post-glacial recolonization inferred from the "ABC-GLM" procedure. T_exp: beginning of the post-glacial expansion, sampled between 19,000 and 11,000 yrs BP (in generations). T_const: end of the post-glacial expansion and beginning of the period of stable population size, with migrations among European regions, sampled between 11,000 yrs BP and present time (in generations).

(pp. 102-104) Species with their main present distribution range in ice-free (IF) areas at LGM. Contri: proportion of migrants of extra-European origin recolonizing Europe at T_exp. M12, M21: migration rates between northern and southern Europe and vice versa between T_const and present time. M3: migration rates between European and extra-European ranges between T_const and present time. Ne_X_present: effective population size at present time in the following regions: 0, extra-European range; 1, northern European range; and 2, southern European range. Ne_1_LGM: effective population size at LGM in the northern European range.

(p. 105) Species with their main present distribution range in areas covered in ice (IC) at LGM. M1: migration rates between the northern European range under the ice-sheet at LGM and other European ranges between T_const and present time. M2: migration rates between the southern mountain European ranges ice-free and covered in ice at LGM between T_const and present time. M3: migration rates between European and extra-European ranges between T_const and present time. Ne_X_present: effective population size at present time in the following regions: 0, extra-European range; 1, northern European range under the ice-sheet at LGM; 2a, b, southern mountain European ranges ice-free and covered in ice at LGM, respectively; 3, Lowland European range south of the northern European ice-sheet at LGM. Ne_3_LGM: effective population size at LGM in the Lowland European range south of the northern European ice-sheet at LGM.

Contri | T_const | T_exp | M12 | M21 | M3 | Ne_0_present | Ne_1_LGM | Ne_1_present | Ne_2_present

Plagiothecium denticulatum

Plagiomnium undulatum

Plagiothecium undulatum

Scorpiurium circinatum

104

	T_const	T_exp	M1	M2	M3	Ne_0_present	Ne_1_present	Ne_2a_present	Ne_2b_present	Ne_3_present	Ne_3_LGM

Amphidium lapponicum

Timmia austriaca

Timmia bavarica
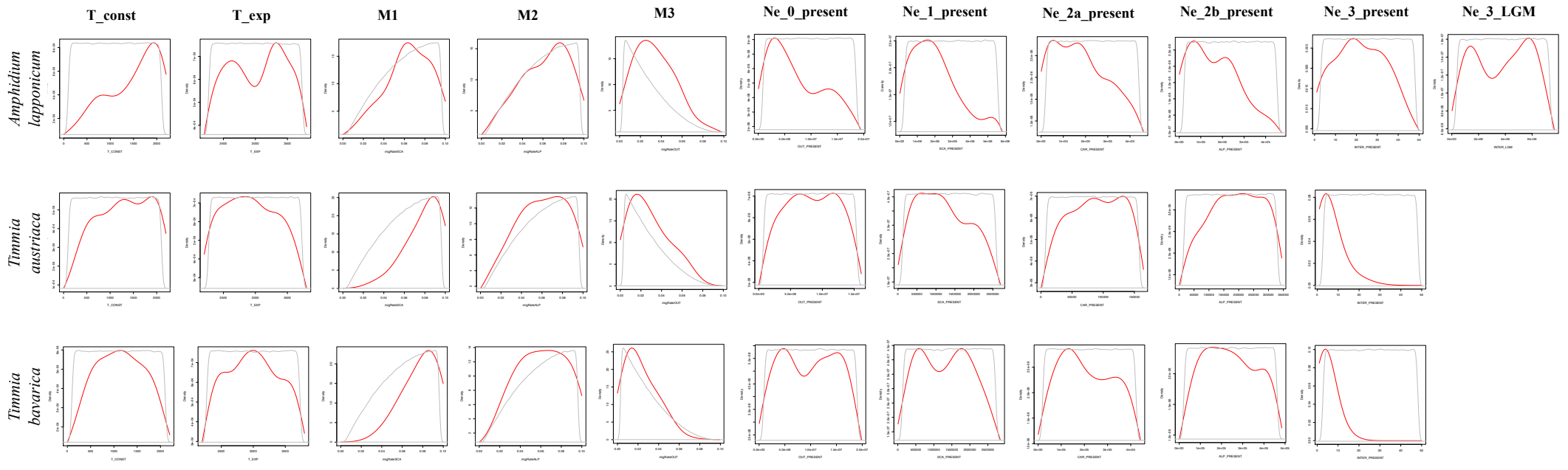
Figure S4 Projections of species distribution models (SDMs) calibrated on a set of sampling points (Table S1) for the 15 selected bryophyte species under present climate conditions onto LGM climatic layers. Climatic suitability increases from red to green colors.

Species with a main present distribution range in ice-free areas at LGM
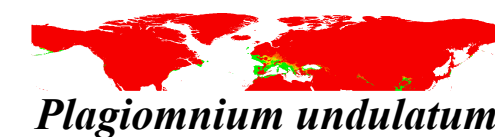
Amphidium mougeotii

Calypogeia fissa

Diplophyllum albicans

Homalothecium sericeum

Metzgeria conjugata

Metzgeria furcata

Orthotrichum affine

Orthotrichum lyellii

Plagiothecium denticulatum

Plagiothecium undulatum

Plagiomnium undulatum

Scorpiurium circinatum

Species with a main present distribution range in areas covered in ice at LGM

Amphidium lapponicum

Timmia austriaca

Timmia bavarica
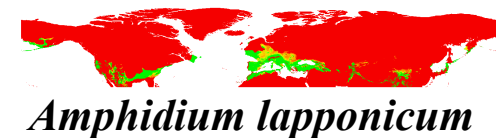
Figure S5 Observed range (dots) and range predicted (map) by species distribution models (SDMs) projected onto current macroclimatic layers for selected bryophyte species in Europe. Climatic suitability increases from red to green colors. Observed data and source information are available from Figshare with the following DOIs: 10.6084/m9.figshare.7376798 (*Plagiomnium undulatum*), 10.6084/m9.figshare.7376801 (*Homalothecium sericeum*), 10.6084/m9.figshare.7376804 (*Metzgeria furcata*), 10.6084/m9.figshare.7376834 (*M. conjugata*), 10.6084/m9.figshare.7376807 (*Orthotrichum affine*), 10.6084/m9.figshare.7376828 (*O. lyellii*), 10.6084/m9.figshare.7376810 (*Plagiothecium undulatum*), 10.6084/m9.figshare.7376825 (*P. denticulatum*), 10.6084/m9.figshare.7376813 (*Diplophyllum albicans*), 10.6084/m9.figshare.7376822 (*Calypogeia fissa*), 10.6084/m9.figshare.7376831 (*Amphidium mougeotii*), 10.6084/m9.figshare.7376843 (*A. lapponicum*), 10.6084/m9.figshare.7376837 (*Scorpiurium circinatum*), 10.6084/m9.figshare.7376840 (*Timmia austriaca*), 10.6084/m9.figshare.7376846 (*T. bavarica*).
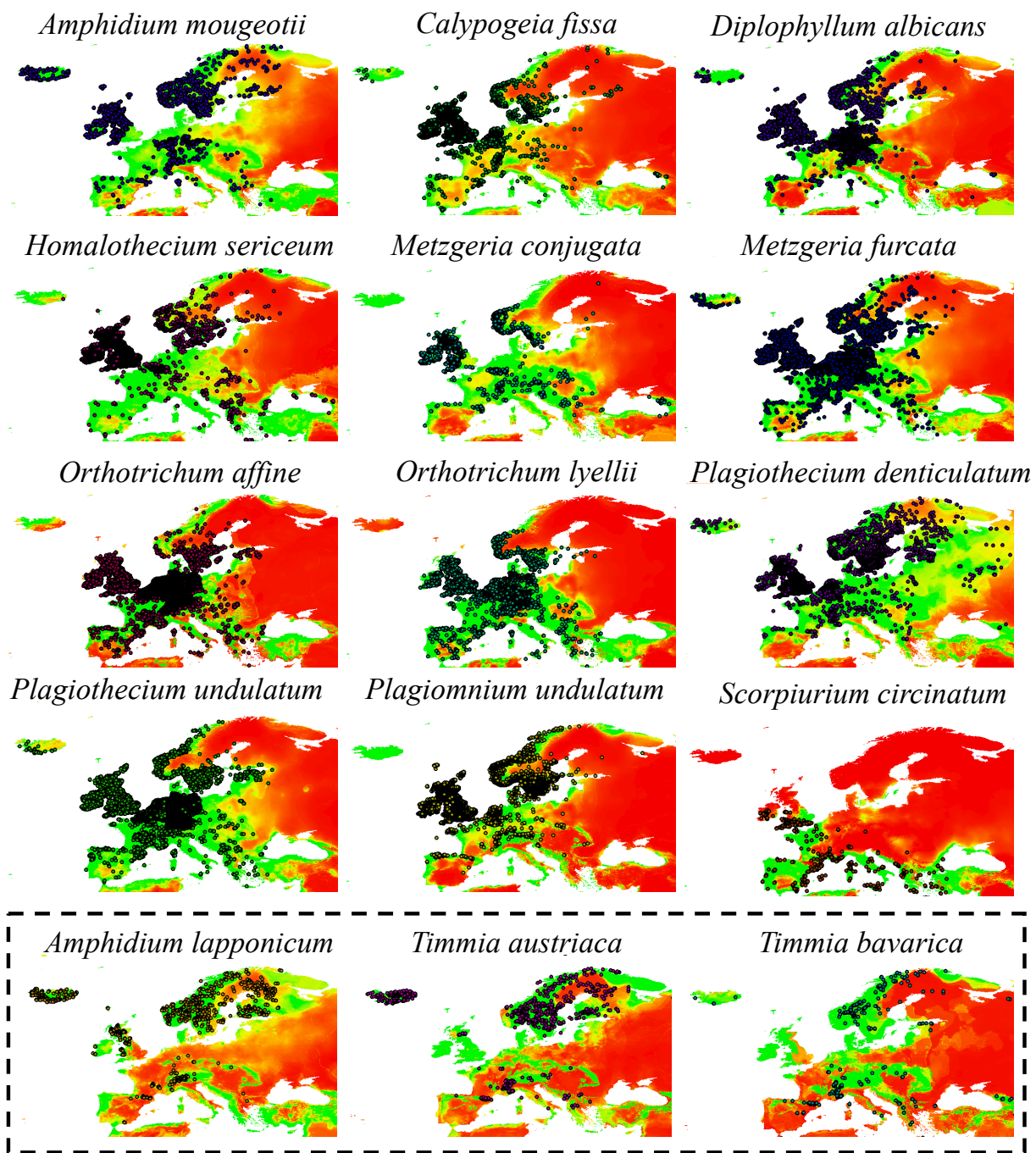
Table S1 Please find Table S1, along with additional supporting information online in the Supporting Information section at the end of the article. https://onlinelibrary.wiley.com/doi/full/10.1111/ele.13254

Discussion and perspectives

General discussion

1. <u>To what extent have the high dispersal capacities of bryophytes erased any historical signal in the extant spatial patterns of genetic structure and diversity of their populations (H1)?</u>

The results presented here strongly suggest that the long-distance dispersal (LDD) capacities of bryophytes did not homogenize the genetic structure of their populations, neither in an environment characterized by apparent geographic barriers to migration – such as the E-W-oriented mountain ranges in Europe –, nor in a much more homogeneous environment as in the lowland Amazonian rainforest. In Europe, the historical scenarios according to which recent migrations erased any genetic structure in the extant spatial patterns of genetic structure and diversity of bryophyte populations (scenarios "IF 3" and "IC 6" in Figs 2 and 3, Paper III) were never selected in a sample of 3 and 12 species that are currently distributed in areas that were covered in ice (IC) or ice-free (IF) at the Last Glacial Maximum (LGM), respectively (Fig. 3, Paper III). In Amazonia, most strikingly and in sharp contrast to our primary hypothesis that Amazonian bryophytes exhibit high dispersal capacities eroding any signal of isolation-by-distance (IBD) at the landscape scale, significant genetic structures were observed in 8 out of the 10 investigated species (Table 2, Paper I). Our results are in line with experimental demographic studies of Amazonian leaf-inhabiting epiphytes, which point to dispersal limitation at both regional- (<50 km) (Zartman & Shaw, 2006) and fine- (<20m) scales (Zartman, Nascimento, Cangani, Alvarenga, & Snäll, 2012) in understory habitats, but strongly challenge the perception, based on the homogeneous distribution of species in the landscape, that Amazonian bryophyte species "behave as one single metacommunity" (Mota de Oliveira & ter Steege, 2015). In further contradiction with the inverse isolation hypothesis, IBD consistently remained significant beyond the range of short-distance dispersal (SDD, Table 2, Paper I), evidencing LDD limitations.

In a recent meta-analysis of the decay of the IBD signal caused by LDD events in bryophytes, Vanderpoorten et al. (2019) showed that LDD capacities were sufficient to erase any IBD signal beyond 1000m from the source in 30-50% of the cases. The consistent persistence of the IBD signal reported here thus suggests that Amazonian bryophyte species experience more dispersal limitations than species from other biomes. As a comparison, the Sp statistics, which characterizes the rate of decrease of pairwise kinship coefficients between individuals with the logarithm of the distance (*i.e.* how strong the IBD signal is, Vekemans & Hardy, 2004), has an average value across Amazonian bryophyte species of 0.012 (Table 2, Paper I). Such a value lays in the range reported for angiosperm species characterized by wind dispersal (0.012±0.012), while the maximal and minimal values of Sp observed here (0.038 in *Syrrhopodon annotinus* and 0.006 in *S. helicophyllus*) are in the range reported for species with gravity-dispersed seeds (0.028±0.016) and animal-dispersed seeds (0.008±0.005), respectively (Vekemans & Hardy, 2004). It is, in fact, striking to consider that, within the same Amazonian environment, the average Sp in bryophytes, whose spores measure *c.* 20 μm, is actually comparable to that of the Brazil nut tree *Bertholletia excelsa* (Sp=0.01-0.03), wherein no significant fine-scale genetic structures were revealed in 5 out of 9 investigated populations (Sujii, Martins, Wadt, Azevedo, & Solferini, 2015), and whose seeds are enclosed within a 10-16 cm globose, functionally indehiscent woody capsule whose dispersal by scatterhoarding agoutis and acouchis, and occasionally squirrels, is restricted to a few hundreds of meters (Thomas, Caicedo, Loo Ii, & Kindt, 2014).

The strong IBD pattern revealed here is consistent with the dispersal traits of the studied species, together with extrinsic features of their environment of dense rainforests that is not prone to long-distance wind-dispersal. These dispersal traits include the absence of male expression (*S. annotinus*, Pereira, Dambros, & Zartman, 2016), the prevalence of dioicy in Calymperaceae associated with low sporophyte production (Pereira et al., 2019), the immersion of the sporophytes within perichaetial leaves or very short setae (*S. annotinus* and *S. helicophyllus*), and the absence of reduction of the peristome (*S. annotinus* and *S. hornschuchii*).

2. <u>What is the relevance of other differentiation mechanisms promoting speciation and, in particular IBE, across a relatively homogeneous environment without any apparent geographic barrier to migration (H2)?</u>

While a significant IBD signal characterizes the genetic structure of the vast majority of the Amazonian bryophyte species investigated here, our results are not consistent with the idea that isolation-by-resistance (IBR) and isolation-by-environment (IBE) contributed to the observed spatial patterns of genetic variation (Table 2, Paper I). The similar allele frequencies found on either side of the Rio Negro in the study area contrasts with Wallace's (1852) hypothesis, which proposes that the Amazonian hydrographic network acts as barriers to gene flow between populations inhabiting opposite river banks, promoting speciation through IBR (McRae, 2006). Wallace's (1852) hypothesis had been however subsequently supported by molecular evidence in a wide range of organisms including primates, anurans, and squamates (see Ortiz, Lima, & Werneck, 2018 for review). Regarding IBE, the lack of difference of allele frequencies in populations from the two distinct main forest types in Amazonia, *i.e.* the terra firme and the white-sand forest, contrasts with evidence for the gradient hypothesis (Guevara et al., 2016), which proposes that the strong environmental differences that exist between these two contrasting habitats promote IBE (Wang & Bradburd, 2014). While, in angiosperms, a signature of IBE was found in a meta-analysis in *c.* 20% of the cases (Sexton, Hangartner, & Hoffmann, 2014), and while, within Amazonia specifically, evidence for ecotypic differentiation between white-sand forest and terra firme populations is suggestive of an adaptive mechanism of edaphic specialization (Fine & Baraloto, 2016; Fine et al., 2013), the absence of significant differences of allele frequencies between terra firme and white-sand forest populations in the investigated bryophyte species is in line with the idea that bryophytes exhibit 'multi-purpose' genotypes and fail to diversify in heterogeneous environments, accounting for their failure to radiate (Patiño et al., 2014).

Nevertheless, a low (0.059) but significant (P=0.004) F_{st} was found between sympatric specimens of the sibling *Syrrhopodon annotinus* and *S. simmondsii* (see "Results" section in Paper II), and their average kinship coefficients along a geographic gradient were consistently higher for conspecific comparisons than for interspecific comparisons (Fig. 2, Paper II). This points to reproductive isolation between the two sympatric species, which are characterized by different habitat requirements: *S. annotinus* mostly occurring on mineral substrates such as rock and soil, whereas *S. simmondsii* tends to occur on organic substrates such as dead wood, humus, and trees (Reese, 1993). Shaw, Meagher, & Harley (1987) similarly found significant differences in a sympatric pair of moss species (*Climacium americanum* and *C. dendroides*) from different habitats. Although it is not possible to determine whether habitat differentiation triggered or followed speciation, and even if these empirical results do not challenge the global idea that IBE does not prevail in extant patterns of genetic diversification in Amazonian bryophytes, they nonetheless contribute to growing evidence for genetic divergence (Hutsemékers, Hardy, Mardulyn, Shaw, & Vanderpoorten, 2010; Magdy, Werner, McDaniel, Goffinet, & Ros, 2016; Mikulášková et al., 2015; Pisa, Werner, Vanderpoorten, Magdy, & Ros,

2013; Szövényi, Hock, Korpelainen, & Shaw, 2009) and speciation (Johnson & Shaw, 2015) observed along environmental gradients, suggesting that adaptation could play a more important role in shaping genetic patterns than previously thought.

3. <u>What do analyses of the extant spatial patterns of genetic structure and diversity of bryophyte populations reveal from their post-glacial history (H3)?</u>

Rejection of the hypothesis that high dispersal capacities of bryophytes erased any historical signal in the extant spatial patterns of genetic structure and diversity of their populations and, in particular rejection of the inverse isolation hypothesis (see §1 above), indicate that the data generated in the present thesis are suitable for demographic inference. We applied coalescence-based approaches to infer the post-glacial history of bryophyte populations from contrasting environments characterized by the presence (Europe) or the absence (lowland Amazonia) of apparent geographic barriers to migration. In the latter case, we are currently implementing a spatially explicit coalescent model (ongoing study IV). I thus focus the discussion here on the identification of the post-glacial recolonization patterns of bryophytes in Europe (Paper III).

The post-glacial assembly of European bryophytes likely involved a complex history from multiple sources, since different scenarios of post-glacial recolonization (for scenarios' description, see Fig. 2 in Paper III) clearly emerged as the best-fit scenario for different species, within IF species, as well as within IC species (Fig. 3, Paper III). This complexity contrasts with the prevailing scenario in which IF species migrated northwards from southern *refugia* (Hewitt, 2000; Petit, 2003). The validity of this historical scenario has been recently challenged for North American trees (Lumibao et al., 2017). Our findings thus further challenge the taxonomic generality of the "southern *refugium* scenario" that has long been assumed to explain the distribution of genetic variation in bryophytes based on single-species analyses (Cronberg, 2000; Grundmann, Ansell, Russell, Koch, & Vogel, 2008).

In IF species, however, the scenario of extra-European post-glacial recolonization (scenario "IF 2" in Figs 2 and 3, Paper III), clearly emerged as dominant (*i.e.* best-fit for 7 out of the 12 IF species, Fig. 3, Paper III). The globally balanced genetic diversities and extremely low divergence observed between southern and northern regions are fully consistent with such a scenario, in contrast to the gradual decrease of genetic diversity towards the north observed in seed plants (Petit, 2003). Previous evidence in angiosperms demonstrated the post-glacial recolonization of a remote Arctic archipelago from distant sources (Alsos et al., 2007). Individual instances of extra-European *refugia* (Hutsemékers et al., 2011; Laenen et al., 2011; Stenøien et al., 2011) and a *de novo* colonization of oceanic areas from Macaronesian ancestors (Patiño et al., 2015) were further evidenced during the post-glacial history of European bryophytes. The present study is, however, the first to demonstrate and quantify, from a small but independent sample of 12 IF species, the substantial contribution from allochthonous migrants in the post-glacial assembly of European bryophyte floras (*posterior* distribution of the proportion of allochthonous migrants in the post-LGM European assembly of 90-100% in all IF species that conform to a recolonization of Europe from allochthonous origin except *Plagiothecium undulatum*, see "Contri" in Fig. S2, Paper III). This demonstrates the importance of LDD for the post-glacial recolonization of Europe by bryophytes and is unparalleled in any previous phylogeograhic study on other plants and animals.

IC species exhibited a strikingly different pattern of in-situ persistence in Europe (Fig. 3, Paper III), sometimes in regions that were fully glaciated at LGM, which is fully compatible

with the reported ability of some species to remain viable after more than 1000 years in ice (La Farge et al., 2013; Roads et al., 2014). However, the mode of the *posterior* distribution of the ongoing migration rate between European and extra-European regions is shifted towards the high values of the range (*i.e.* at values of 2-4% of migrants per generation, see "M3" in Fig. S2, Paper III), which is fully compatible with the substantial contribution of extra-European migrants in IF species.

The difference in the main origins of European bryophyte species is puzzling. This difference cannot, at first sight, be interpreted in terms of life-history traits. In *Amphidium* for example, *A. lapponicum* produces large numbers of capsules with small spores, but recolonized Europe from autochthonous populations, whereas in *A. mougeotii*, which seldom produces sporophytes and does not produce specialized asexual diaspores, recolonization took mostly place from allochthonous migrants (Fig. 3, Paper III).

Projections of our species distribution models (SDMs) onto LGM climatic layers (Fig. S4, Paper III) predict that, with the exception of a few species (*Scorpiurium circinatum* and, to a lesser extent, *Diplophyllum albicans* and *Plagiothecium undulatum*), southern Europe and the southern range of northern Europe were extensively climatically favorable at LGM. This suggests the potential existence of sufficiently large *refugia* within Europe and challenges the idea of a substantial contribution from allochthonous migrants in the post-glacial assembly of European bryophytes. Paleontological evidence indicates, however, that the full-glacial landscape of Eurasia was largely treeless, with a dominance of steppe and other tundra types of vegetation, suggesting that paleoclimatic reconstructions, on which our SDMs are built, predicted a warmer and moister climate than it probably was (Binney et al., 2017; Tzedakis et al., 2013). Therefore, the limited contribution of the "northern micro-*refugium* scenario", on the one hand – which contrasts with our primary expectations based on the high cold tolerance of bryophytes – and of the classical "southern *refugium* scenario", on the other, to explain the post-glacial history of the investigated species may be due to the fact that European *refugia* have been too small and too scattered as compared to the substantial waves of extra-European origin to actually contribute to the post-glacial recolonization of the continent. As Semerikov, Semerikova, Polezhaeva, Kosintsev, & Lascoux (2013) in fact suggested, it should not be firmly concluded that putative refugial populations necessarily contributed extensively to local modern populations, as the spread of new individuals from adjacent regions would have occurred over several *millennia* as climates changed (Binney et al., 2017).

The consistent signal for a delay in the beginning of the post-glacial expansion of IF species in Europe since the end of the LGM (19,000 yrs BP), that was evidenced by the shift of the *posterior* distribution marking the beginning of the expansion phase towards recent periods (*c.* 11,000-12,000 yrs BP, see "T_exp" in Fig. S2, Paper III), further supports this idea that paleoclimatic reconstructions, on which our species distribution models are built, are too optimistic in the way they describe LGM and post-LGM European suitable habitats. Indeed, such a delay could be interpreted in terms of either unsuitable conditions at the beginning of the current interglacial period – supporting the idea of too optimistic paleoclimatic reconstructions–, and/or a delay in the recolonization of newly suitable habitats. The critical transition from predominantly glacial to largely interglacial, moister climates inferred from major changes in fossil pollen records 14,000 yrs BP (Pearson, 2006) supports the first hypothesis. A delay in the recolonization of newly suitable habitats is, conversely, not supported by the match highlighted between the predicted and the observed northern limit of distribution in the bryophyte species investigated here (Fig. S5, Paper III), which contrasts with the absence of many angiosperm species in the north of their potential distribution areas (Normand et al., 2011; Svenning, Normand, & Kageyama, 2008).

In this thesis, we show that the ability of European and Amazonian bryophytes to face global change is at odds with what could be expected in regard to the 'resistance' to migration of both environments. Indeed, in Europe – where the landscape is shaped by E-W-oriented mountain ranges acting as barriers to migration for many organisms–, bryophytes – thanks to their high dispersal capacities –, mostly recolonized the mainland from allochthonous migrants, whereas in Amazonia – where the landscape is homogeneous without any apparent barrier to migration –, dispersal is currently limited within the mainland – because of less efficient dispersal capacities furthermore impeded by the closed landscape of the dense rainforest –, strongly questioning the ability of Amazonian bryophytes to keep track of their suitable habitats in the context of ongoing climate changes. The restricted dispersal capacities of Amazonian bryophytes therefore suggest that they would be more severely exposed to current human-mediated global change than European bryophytes, wherein striking range shifts have already been documented during the past decades (Bosanquet, 2012).

Perspectives

Comparing the relative impact of past climate changes and current anthropogenic disturbances on extant patterns of spatial genetic structure and diversity of bryophyte populations

Current human-mediated climate changes imply temperature and precipitation modifications at an unparalleled rate. The diagram of temperature variation across the Holocene (Fig. 8) in fact shows that we reached, already in 2013, temperatures as high as they were at the Holocene maximum (*c.* 7 Kyrs ago, Marcott, Shakun, Clark, & Mix, 2013). While unparalleled changes on biota are therefore to be expected during the forthcoming decades (see Bellard, Bertelsmeier, Leadley, Thuiller, & Courchamp, 2012 for review), many other factors, and in particular land conversion, currently represent a perhaps even more important threat on biodiversity than climate changes themselves (Parmesan & Yohe, 2003). In ongoing study IV, we intend to compare the impact of LGM climate changes, arguably the most important historical factors that shaped current biodiversity patterns, to the one of recent anthropogenic disturbances. In particular, we will focus on the deforestation of the Amazonian forest, which is presently experiencing absolute rates of deforestation, higher than any other region on the planet (*i.e.* almost 2.4 million ha/yr in 2002 and 2003, Laurance, Albernaz, Fearnside, Vasconcelos, & Ferreira, 2004; Nelson et al., 2006).
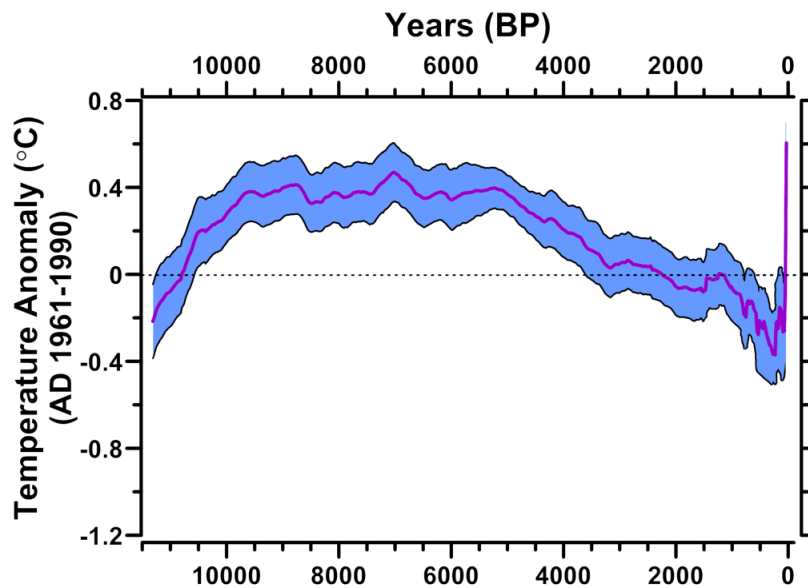
Figure 8: Reconstruction of global temperature anomalies relative to 1961-1990. Globally stacked temperature anomalies for a 5° × 5° area-weighted mean calculation (purple line) with its uncertainty (1σ, blue band). Modified from: Marcott, Shakun, Clark, & Mix, 2013.

The strong and unexpected signal of dispersal limitation presented in this thesis for organisms primarily considered as extremely efficient dispersers (Table 2, Paper I) clearly suggests that Amazonian bryophytes are severely exposed to forest fragmentation. In ongoing study IV, we will also assess, based on the *posterior* distribution of the migration rate between neighbor pixels, the actual dispersal rates of Amazonian bryophytes to determine how strongly and how quickly populations might suffer from ongoing deforestation. Finally, we will determine whether events of LDD occur across Amazonian bryophyte species, by comparing explicit demographic scenarios where LDD is implemented or not. Indeed, LDD might help to maintain connectivity among distant patches in the context of current fragmentation.

<u>Available tools to study the impact of past climate changes on the distribution of species</u>

Studying the influence of past climate changes on the distribution of species contributes to our understanding of the evolution of life on earth. In this thesis, we took advantage of demographic inference, especially coalescent simulations, to reconstruct past biogeographic histories of bryophyte populations. Although this technique increasingly appeared as a promising tool to infer the demographic history of populations from current genetic data (Thomé & Carstens, 2016), it has been tailored to the characterization of demographic changes in existing populations (Chikhi et al., 2018), therefore information from extinct populations is missing. Furthermore, because coalescent simulations infer demographic histories of populations from current spatial patterns of genetic structure and diversity, only historical events which have led to major modifications in populations are recovered. Coalescent simulations thus prevent the recovery of continuous long-term historical evolution of populations through time.

The fossil record, the only source of information providing absolute (geological) timing of historical events (Donoghue & Benton, 2007), has traditionally been used by biologists to study the long-term historical evolution of populations through time (Morlon, Parsons, & Plotkin, 2011). When the fossil record is rich, it is possible to trace back continuous

114

demographic changes in past populations and to point out past extinctions of populations. However, many groups of organisms lack an adequate fossil record (Morlon et al., 2011).

A promising tool to trace back the long-term historical evolution of populations through time is ancient DNA (aDNA, see Anderson-Carpenter et al., 2011 for review). Although not all fossils recovered contain preserved DNA suitable for current aDNA extraction techniques – given taphonomic and environmental factors can vary preservation rates dramatically (Knapp et al., 2012; Lord, Collins, DeFrance, LeFebvre, & Matisoo-Smith, 2018; Orlando et al., 2013) –, the aDNA recovered from only a couple of individuals may provide evolutionary information from the entire population, such as its size or its exposure to extinction at a given time (Palkopoulou et al., 2015). Moreover, a recent technique, called "sedimentary ancient DNA" (sedaDNA), recovers preserved aDNA from all organisms present within sediment samples, such as permafrost, cave sediments or lake sediments (Anderson-Carpenter et al., 2011; Willerslev, 2003; Willerslev & Cooper, 2005). This technique increases the number of fossil taxa found because it enables the identification of organisms that do not fossilize into macrofossils, or, easily recognizable microfossils (Birks & Birks, 2016).

As Hewitt (2000) declared, combining genetic and fossil data can greatly help our understanding of how organisms were affected by past climate changes. As an example, we could use the information extracted from fossil data and related aDNA, to enhance demographic inferences – such as coalescent simulations – by providing accurate *prior* estimations of past demographic parameters – including past effective population sizes and the timing of past demographic events. We could also combine demographic inferences with information on past extinct populations, recovered from aDNA, to improve the global picture of the long-term historical evolution of populations in organisms with a poor fossil record.

Global change and species conservation

On Monday the 6th of May, the United Nations' Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) released a 40-page summary of its findings for policymakers and the media, suggesting that *c.* 1 million species are currently facing extinction due to human activity (Díaz et al., 2019). Studying which conditions likely led to population extinction or survival and subsequent recovery in the past may help to predict which populations are currently the most endangered in the context of global change. Such predictions would help to take efficient and timely conservation measures.

Another tool increasingly proposed to support conservation decision making is species distribution models (SDMs) projected into putative future climatic conditions (Guisan et al., 2013). The validity of such models has been recently questioned, pointing to the many factors other than climate that play an important role in determining species distributions, such as biotic interactions and dispersal abilities (Pearson & Dawson, 2003). Ecologists are thus currently developing enhanced models that take into account such factors (Engler & Guisan, 2009; Lehsten et al., 2014; Thuiller et al., 2008).

As presented, we currently have the information and the tools to take adequate conservation measures in the ongoing context of global change. Let's now take adequate political decisions and choices in our everyday lives so that we might not be responsible for the extinction of 1 million species #WeChangeForLife. https://wechangeforlife.org

References

Abbott, R. J., & Brochmann, C. (2003). History and evolution of the arctic flora: in the footsteps of Eric Hulten. *Molecular Ecology*, *12*(2), 299–313. https://doi.org/10.1046/j.1365-294X.2003.01731.x

Alsos, I. G., Eidesen, P. B., Ehrich, D., Skrede, I., Westergaard, K., Jacobsen, G. H., … Brochmann, C. (2007). Frequent Long-Distance Plant Colonization in the Changing Arctic. *Science*, *316*(5831), 1606–1609. https://doi.org/10.1126/science.1139178

Andersen, B. G., & Borns, H. W. (1994). *The Ice Age World: An Introduction to Quaternary History and Research with Emphasis on North America and Northern Europe During the Last 2.5 Million Years*. Oxford University Press.

Anderson-Carpenter, L. L., McLachlan, J. S., Jackson, S. T., Kuch, M., Lumibao, C. Y., & Poinar, H. N. (2011). Ancient DNA from lake sediments: Bridging the gap between paleoecology and genetics. *BMC Evolutionary Biology*, *11*(1), 30. https://doi.org/10.1186/1471-2148-11-30

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Anhuf, D., Ledru, M.-P., Behling, H., Da Cruz, F. W., Cordeiro, R. C., Van der Hammen, T., … Da Silva Dias, P. L. (2006). Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *239*(3–4), 510–527. https://doi.org/10.1016/j.palaeo.2006.01.017

Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press.

Barbé, M., Fenton, N. J., & Bergeron, Y. (2016). So close and yet so far away: long-distance dispersal events govern bryophyte metacommunity reassembly. *Journal of Ecology*, *104*(6), 1707–1719. https://doi.org/10.1111/1365-2745.12637

Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics*, *41*(1), 379–406. https://doi.org/10.1146/annurev-ecolsys-102209-144621

Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian Computation in Population Genetics. *Genetics*, *162*(4), 2025–2035. Retrieved from http://www.genetics.org/content/162/4/2025.abstract

Behling, H., & Hooghiemstra, H. (2001). Neotropical Savanna Environments in Space and Time. In *Interhemispheric Climate Linkages* (pp. 307–323). Elsevier. https://doi.org/10.1016/B978-012472670-3/50021-5

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, *15*(4), 365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x

Berger, A., Mesinger, F., & Sijacki, D. (2012). A Brief History of the Astronomical Theories of Paleoclimates. In A. Berger, F. Mesinger, & D. Sijacki (Eds.), *Climate Change* (pp. 107–129). Vienna: Springer Vienna. https://doi.org/10.1007/978-3-7091-0973-1

Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, *19*(13), 2609–2625. https://doi.org/10.1111/j.1365-294X.2010.04690.x

Bhagwat, S. A., & Willis, K. J. (2008). Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *Journal of Biogeography*, *35*(3), 464–482. https://doi.org/10.1111/j.1365-2699.2007.01861.x

Binney, H., Edwards, M., Macias-Fauria, M., Lozhkin, A., Anderson, P., Kaplan, J. O., … Zernitskaya, V. (2017). Vegetation of Eurasia from the last glacial maximum to present: Key biogeographic patterns. *Quaternary Science Reviews*, *157*, 80–97. https://doi.org/10.1016/j.quascirev.2016.11.022

Birks, H. J. B., & Birks, H. H. (2016). How have studies of ancient DNA from sediments contributed to the reconstruction of Quaternary floras? *New Phytologist*, *209*(2), 499–506. https://doi.org/10.1111/nph.13657

Bosanquet, S. (2012). Vagrant epiphytic mosses in England and Wales. *Field Bryology*, *107*, 3–17.

Bush, M. B., Stute, M., Ledru, M.-P., Behling, H., Colinvaux, P. A., De Oliveira, P. E., … Webb, R. (2001). Paleotemperature Estimates for the Lowland Americas Between 30°S and 30°N at the Last Glacial Maximum. In *Interhemispheric Climate Linkages* (pp. 293–306). Elsevier. https://doi.org/10.1016/B978-012472670-3/50020-3

Chikhi, L., Rodríguez, W., Grusea, S., Santos, P., Boitard, S., & Mazet, O. (2018). The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. *Heredity*, *120*(1), 13–24. https://doi.org/10.1038/s41437-017-0005-6

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., … McCabe, A. M. (2009). The Last Glacial Maximum. *Science*, *325*(5941), 710–714. https://doi.org/10.1126/science.1172873

Colinvaux, P. A., De Oliveira, P. E., Moreno, J. E., Miller, M. C., & Bush, M. B. (1996). A Long Pollen Record from Lowland Amazonia: Forest and Cooling in Glacial Times. *Science*, *274*(5284), 85–88. https://doi.org/10.1126/science.274.5284.85

Collins, F. S., Morgan, M., & Patrinos, A. (2003). The Human Genome Project: Lessons from Large-Scale Biology. *Science*, *300*(5617), 286–290. https://doi.org/10.1126/science.1084564

Condit, R., Pitman, N., Leigh, E. G. J., Chave, J., Terborgh, J., Foster, R. B., … Hubbell, S. P. (2002). Beta-Diversity in Tropical Forest Trees. *Science*, *295*(5555), 666–669. https://doi.org/10.1126/science.1066854

Cronberg, N. (2000). Genetic diversity of the epiphytic bryophyte Leucodon sciuroides in formerly glaciated versus nonglaciated parts of Europe. *Heredity*, *84*(6), 710. https://doi.org/10.1046/j.1365-2540.2000.00719.x

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, *25*(7), 410–418. https://doi.org/10.1016/j.tree.2010.04.001

Darwin, C. R. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* (1st editio). London: John Murray.

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, *10*(9), 20140562–20140562. https://doi.org/10.1098/rsbl.2014.0562

Dellicour, S., Fearnley, S., Lombal, A., Heidl, S., Dahlhoff, E. P., Rank, N. E., & Mardulyn, P. (2014). Inferring the past and present connectivity across the range of a north american leaf beetle: Combining ecological niche modeling and a geographically explicit model of coalescence. *Evolution*, *68*(8), 2371–2385. https://doi.org/10.1111/evo.12426

Dellicour, S., Kastally, C., Hardy, O. J., & Mardulyn, P. (2014). Comparing Phylogeographic Hypotheses by Simulating DNA Sequences under a Spatially Explicit Model of Coalescence. *Molecular Biology and Evolution*, *31*(12), 3359–3372. https://doi.org/10.1093/molbev/msu277

Désamoré, A., Patiño, J., Mardulyn, P., Mcdaniel, S. F., Zanatta, F., Laenen, B., & Vanderpoorten, A. (2016). High migration rates shape the postglacial history of amphi-Atlantic bryophytes. *Molecular Ecology*, *25*(21), 5568–5584. https://doi.org/10.1111/mec.13839

Díaz, S., Settele, J., Brondízio, E., Ngo, H. T., Guèze, M., Agard, J., … Zayas, C. (2019). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*.

Donoghue, P. C. J., & Benton, M. J. (2007). Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution*, *22*(8), 424–431. https://doi.org/10.1016/j.tree.2007.05.005

Edwards, D., Morris, J. L., Richardson, J. B., & Kenrick, P. (2014). Cryptospores and cryptophytes reveal hidden diversity in early land floras. *New Phytologist*, *202*(1), 50–78. https://doi.org/10.1111/nph.12645

Eidesen, P. B., Ehrich, D., Bakkestuen, V., Alsos, I. G., Gilg, O., Taberlet, P., & Brochmann, C. (2013). Genetic roadmap of the Arctic: plant dispersal highways, traffic barriers and capitals of diversity. *New Phytologist*, *200*(3), 898–910. https://doi.org/10.1111/nph.12412

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*(1), 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Ellis, C. J., & Tallis, J. H. (2000). Climatic control of blanket mire development at Kentra Moss, north-west Scotland. *Journal of Ecology*, *88*(5), 869–889. https://doi.org/10.1046/j.1365-2745.2000.00495.x

Engler, R., & Guisan, A. (2009). MigClim: Predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions*, *15*(4), 590–601. https://doi.org/10.1111/j.1472-4642.2009.00566.x

Engler, R., Hordijk, W., & Guisan, A. (2012). The MIGCLIM R package - seamless integration of dispersal constraints into projections of species distribution models. *Ecography*, *35*(10), 872–878. https://doi.org/10.1111/j.1600-0587.2012.07608.x

Excoffier, L., Novembre, J., & Schneider, S. (2000). SIMCOAL : A General Coalescent Program for the Simulation of Molecular Data. *Journal of Heredity*, *91*(6), 506–510.

Filho, A. C., Schwartz, D., Tatumi, S. H., & Rosique, T. (2002). Amazonian Paleodunes Provide Evidence for Drier Climate Phases during the Late Pleistocene–Holocene. *Quaternary Research*, *58*(02), 205–209. https://doi.org/10.1006/qres.2002.2345

Fine, P. V. A., & Baraloto, C. (2016). Habitat endemism in white-sand forests: Insights into the mechanisms of lineage diversification and community assembly of the Neotropical flora. *Biotropica*, *48*(1), 24–33. https://doi.org/10.1111/btp.12301

Fine, P. V. A., Zapata, F., Daly, D. C., Mesones, I., Misiewicz, T. M., Cooper, H. F., & Barbosa, C. E. a. (2013). The importance of environmental heterogeneity and spatial distance in generating phylogeographic structure in edaphic specialist and generalist tree species of Protium (Burseraceae) across the Amazon Basin. *Journal of Biogeography*, *40*(4), 646–661. https://doi.org/10.1111/j.1365-2699.2011.02645.x

Frahm, J. P., & Klaus, D. (2001). Bryophytes as indicators of recent climate fluctuations in Central Europe. *Lindbergia*, *26*(2), 97–104.

Furness, S. B., & Grime, J. P. (1982). Growth Rate and Temperature Responses in Bryophytes: II. A Comparative Study of Species of Contrasted Ecology. *Journal of Ecology*, *70*(2), 525. https://doi.org/10.2307/2259920

Grundmann, M., Ansell, S. W., Russell, S. J., Koch, M. A., & Vogel, J. C. (2008). Hotspots of diversity in a clonal world - the Mediterranean moss Pleurochaete squarrosa in Central Europe. *Molecular Ecology*, *17*(3), 825–838. https://doi.org/10.1111/j.1365-294X.2007.03634.x

Guevara, J. E., Damasco, G., Baraloto, C., Fine, P. V. A., Peñuela, M. C., Castilho, C., … ter Steege, H. (2016). Low Phylogenetic Beta Diversity and Geographic Neo-endemism in Amazonian White-sand Forests. *Biotropica*, *48*(1), 34–46. https://doi.org/10.1111/btp.12298

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat Suitability and Distribution Models*. Cambridge University Press.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., … Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424–1435. https://doi.org/10.1111/ele.12189

Haberle, S. G., & Maslin, M. A. (1999). Late Quaternary Vegetation and Climate Change in the Amazon Basin Based on a 50,000 Year Pollen Record from the Amazon Fan, ODP Site 932. *Quaternary Research*, *51*(01), 27–38. https://doi.org/10.1006/qres.1998.2020

Hansen, J., Sato, M., Russell, G., & Kharecha, P. (2013). Climate sensitivity, sea level and atmospheric carbon dioxide. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(2001), 20120294–20120294. https://doi.org/10.1098/rsta.2012.0294

He, X., He, K. S., & Hyvönen, J. (2016). Will bryophytes survive in a warming world? *Perspectives in Plant Ecology, Evolution and Systematics*, *19*, 49–60. https://doi.org/10.1016/j.ppees.2016.02.005

Hein, J., Schierup, M., & Wiuf, C. (2004). *Gene Genealogies, Variation and Evolution: A primer in coalescent theory*. Oxford University Press.

Hessler, A. M. (2011). Earth's Earliest Climate. *Nature Education Knowledge*, *3*(10), 24.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, *58*(3), 247–276. https://doi.org/10.1111/j.1095-8312.1996.tb01434.x

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*(1–2), 87–112. https://doi.org/10.1006/bijl.1999.0332

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature, 405*(6789), 907–913. https://doi.org/10.1038/35016000

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1442), 183–195. https://doi.org/10.1098/rstb.2003.1388

Hooker, J. D. (1861). Outlines of the Distribution of Arctic Plants. *Transactions of the Linnean Society of London*, *23*(2), 251–348. https://doi.org/10.1111/j.1096-3642.1860.tb00131.x

Hudson, R. R. (1991). Gene genealogies and the coalescent process. In D. Futuyama & J. Antonovics (Eds.), *Oxford Surveys in Evolutionary Biology* (Vol. 7, pp. 1–44). Oxford University Press.

Hultén, E. (1937). *Outline of the history of arctic and boreal biota during the Quaternary period*. Stockholm: Bokförlags aktiebolaget Thule. Retrieved from file://catalog.hathitrust.org/Record/006168154

Hutsemékers, V., Hardy, O. J., Mardulyn, P., Shaw, A. J., & Vanderpoorten, A. (2010). Macroecological patterns of genetic structure and diversity in the aquatic moss Platyhypnidium riparioides. *New Phytologist*, *185*(3), 852–864. https://doi.org/10.1111/j.1469-8137.2009.03094.x

Hutsemékers, V., Hardy, O. J., & Vanderpoorten, A. (2013). Does water facilitate gene flow in spore-producing plants? Insights from the fine-scale genetic structure of the aquatic moss Rhynchostegium riparioides (Brachytheciaceae). *Aquatic Botany*, *108*, 1–6. https://doi.org/10.1016/j.aquabot.2013.02.001

Hutsemékers, V., Szovenyi, P., Shaw, A. J., Gonzalez-Mancebo, J.-M., Munoz, J., & Vanderpoorten, A. (2011). Oceanic islands are not sinks of biodiversity in spore-producing plants. *Proceedings of the National Academy of Sciences*, *108*(47), 18989–18994. https://doi.org/10.1073/pnas.1109119108

Johnson, M. G., & Shaw, A. J. (2015). Genetic diversity, sexual condition, and microhabitat preference determine mating patterns in S phagnum (Sphagnaceae) peat-mosses. *Biological Journal of the Linnean Society*, *115*(1), 96–113. https://doi.org/10.1111/bij.12497

Jonsgard, B., & Birks, H. H. (1995). Late-glacial mosses and environmental reconstructions at Krakenes, western Norway. *Lindbergia*, *20*(2–3), 64–82.

Kingman, J. F. C. (1982). The coalescent. *Stochastic Processes and Their Applications*, *13*(3), 235–248. https://doi.org/10.1016/0304-4149(82)90011-4

Kirschvink, J. L., Gaidos, E. J., Bertani, L. E., Beukes, N. J., Gutzmer, J., Maepa, L. N., & Steinberger, R. E. (2000). Paleoproterozoic snowball Earth: Extreme climatic and geochemical global change and its biological consequences. *Proceedings of the National Academy of Sciences*, *97*(4), 1400–1405. https://doi.org/10.1073/pnas.97.4.1400

Klips, R. A. (2015). DNA microsatellite analysis of sporophytes of the short-lived moss Physcomitrium pyriforme reveals a predominantly self-fertilizing mating pattern . *The Bryologist*, *118*(2), 200–211. https://doi.org/10.1639/0007-2745-118.2.200

Knapp, M., Horsburgh, K. A., Prost, S., Stanton, J.-A., Buckley, H. R., Walter, R. K., & Matisoo-Smith, E. A. (2012). Complete mitochondrial DNA genome sequences from the first New Zealanders. *Proceedings of the National Academy of Sciences*, *109*(45), 18350–18354. https://doi.org/10.1073/pnas.1209896109

Kulski, J. K. (2016). Next-Generation Sequencing — An Overview of the History, Tools, and "Omic" Applications. In *Next Generation Sequencing - Advances, Applications and Challenges* (pp. 3–60). InTech. https://doi.org/10.5772/61964

Kyrkjeeide, M. O., Stenøien, H. K., Flatberg, K. I., & Hassel, K. (2014). Glacial refugia and post-glacial colonization patterns in European bryophytes. *Lindbergia*, *2*, 47–59. https://doi.org/10.25227/linbg.01046

La Farge, C., Williams, K. H., & England, J. H. (2013). Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. *Proceedings of the National Academy of Sciences*, *110*(24), 9839–9844. https://doi.org/10.1073/pnas.1304199110

Laaka-Lindberg, S., Hedderson, T. A. J., & Longton, R. E. (2000). Rarity and reproductive characters in the British hepatic flora. *Lindbergia*, *25*, 78–84.

Laenen, B., Désamoré, A., Devos, N., Shaw, A. J., González-Mancebo, J. M., Carine, M. A., & Vanderpoorten, A. (2011). Macaronesia: a source of hidden genetic diversity for post-glacial recolonization of western Europe in the leafy liverwort Radula lindenbergiana. *Journal of Biogeography*, *38*(4), 631–639. https://doi.org/10.1111/j.1365-2699.2010.02440.x

Laenen, B., Machac, A., Gradstein, S. R., Shaw, B., Patiño, J., Désamoré, A., … Vanderpoorten, A. (2016). Geographical range in liverworts: does sex really matter? *Journal of Biogeography*, *43*(3), 627–635. https://doi.org/10.1111/jbi.12661

Laurance, W. F., Albernaz, A. K. M., Fearnside, P. M., Vasconcelos, H. L., & Ferreira, L. V. (2004). Deforestation in Amazonia. *Science*, *304*, 1109. https://doi.org/10.1515/9781400865277

Ledru, M.-P., Bertaux, J., Sifeddine, A., & Suguio, K. (1998). Absence of Last Glacial Maximum Records in Lowland Tropical Forests. *Quaternary Research*, *49*(02), 233–237. https://doi.org/10.1006/qres.1997.1953

Lehsten, D., Dullinger, S., Hülber, K., Schurgers, G., Cheddadi, R., Laborde, H., … Sykes, M. T. (2014). Modelling the Holocene migrational dynamics of Fagus sylvatica L. and Picea abies (L.) H. Karst. *Global Ecology and Biogeography*, *23*(6), 658–668. https://doi.org/10.1111/geb.12145

Leuenberger, C., & Wegmann, D. (2010). Bayesian Computation and Model Selection Without Likelihoods. *Genetics*, *184*(1), 243–252. https://doi.org/10.1534/genetics.109.109058

Longton, R. E. (1997). Reproductive biology and life-history strategies. *Advances in Bryology*, *6*, 65–101.

Longton, R. E., & Schuster, R. M. (1983). Reproductive biology. *New Manual of Bryology*, *1*, 386–462.

Lönnell, N., Hylander, K., Jonsson, B. G., & Sundberg, S. (2012). The Fate of the Missing Spores — Patterns of Realized Dispersal beyond the Closest Vicinity of a Sporulating Moss. *PLoS ONE*, *7*(7), e41987. https://doi.org/10.1371/journal.pone.0041987

Lönnell, N., Norros, V., Sundberg, S., Rannik, Ü., Johansson, V., Ovaskainen, O., & Hylander, K. (2015). Testing a mechanistic dispersal model against a dispersal experiment with a wind-dispersed moss. *Oikos*, *124*(9), 1232–1240. https://doi.org/10.1111/oik.01886

Lord, E., Collins, C., DeFrance, S., LeFebvre, M. J., & Matisoo-Smith, E. (2018). Complete mitogenomes of ancient Caribbean Guinea pigs (Cavia porcellus). *Journal of Archaeological Science: Reports*, *17*, 678–688. https://doi.org/10.1016/j.jasrep.2017.12.004

Lumibao, C. Y., Hoban, S. M., & McLachlan, J. (2017). Ice ages leave genetic diversity 'hotspots' in Europe but not in Eastern North America. *Ecology Letters*, *20*(11), 1459–1468. https://doi.org/10.1111/ele.12853

Magdy, M., Werner, O., McDaniel, S. F., Goffinet, B., & Ros, R. M. (2016). Genomic scanning using AFLP to detect loci under selection in the moss Funaria hygrometrica along a climate gradient in the Sierra Nevada Mountains, Spain. *Plant Biology*, *18*(2), 280–288. https://doi.org/10.1111/plb.12381

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., … Turner, D. J. (2010). Target-enrichment strategies for next-generation sequencing. *Nature Methods*, *7*(2), 111–118. https://doi.org/10.1038/nmeth.1419

Marcott, S. A., Shakun, J. D., Clark, P. U., & Mix, A. C. (2013). A reconstruction of regional and global temperature for the past 11,300 years. *Science*, *339*(6124), 1198–1201. https://doi.org/10.1126/science.1228026

Marjoram, P., & Joyce, P. (2011). Practical Implications of Coalescent Theory. In L. S. Heath & N. Ramakrishnan (Eds.), *Problem Solving Handbook in Computational Biology and Bioinformatics* (pp. 63–80). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-09760-2

Marjoram, P., & Tavaré, S. (2006). Modern computational approaches for analysing molecular genetic variation data. *Nature Reviews Genetics*, *7*(10), 759–770. https://doi.org/10.1038/nrg1961

Mátyás, G., & Sperisen, C. (2001). Chloroplast DNA polymorphisms provide evidence for postglacial re-colonisation of oaks (Quercus spp.) across the Swiss Alps. *Theoretical and Applied Genetics*, *102*(1), 12–20. https://doi.org/10.1007/s001220051613

Mayle, F. E., Beerling, D. J., Gosling, W. D., & Bush, M. B. (2004). Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1443), 499–514. https://doi.org/10.1098/rstb.2003.1434

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*(8), 1551. https://doi.org/10.1554/05-321.1

Médail, F., & Diadema, K. (2009). Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, *36*(7), 1333–1345. https://doi.org/10.1111/j.1365-2699.2008.02051.x

Metzker, M. L. (2010). Sequencing technologies — the next generation. *Nature Reviews Genetics*, *11*(1), 31–46. https://doi.org/10.1038/nrg2626

Mikulášková, E., Hájek, M., Veleba, A., Johnson, M. G., Hájek, T., & Shaw, J. A. (2015). Local adaptations in bryophytes revisited: the genetic structure of the calcium-tolerant peatmoss Sphagnum warnstorfii along geographic and pH gradients. *Ecology and Evolution*, *5*(1), 229–242. https://doi.org/10.1002/ece3.1351

Montañez, I. P., & Poulsen, C. J. (2013). The Late Paleozoic Ice Age: An Evolving Paradigm. *Annual Review of Earth and Planetary Sciences*, *41*(1), 629–656. https://doi.org/10.1146/annurev.earth.031208.100118

Morlon, H., Parsons, T. L., & Plotkin, J. B. (2011). Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences*, *108*(39), 16327–16332. https://doi.org/10.1073/pnas.1102543108

Mota de Oliveira, S., & ter Steege, H. (2015). Bryophyte communities in the Amazon forest are regulated by height on the host tree and site elevation. *Journal of Ecology*, *103*(2), 441–450. https://doi.org/10.1111/1365-2745.12359

Nelson, G. C., Bennett, E., Berhe, A. A., Cassman, K., DeFries, R., Dietz, T., … Zurek, M. (2006). Anthropogenic Drivers of Ecosystem Change: an Overview. *Ecology and Society*, *11*(2), art29. https://doi.org/10.5751/ES-01826-110229

Nordborg, M. (2001). Coalescent Theory. In D. J. Balding, M. Bishop, & C. Cannings (Eds.), *Handbook of Statistical Genetics* (pp. 843–877). John Wiley & Sons, Ltd. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470061619.ch25

Normand, S., Ricklefs, R. E., Skov, F., Bladt, J., Tackenberg, O., & Svenning, J.-C. (2011). Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1725), 3644–3653. https://doi.org/10.1098/rspb.2010.2769

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., … Willerslev, E. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, *499*(7456), 74–78. https://doi.org/10.1038/nature12323

Ortiz, D. A., Lima, A. P., & Werneck, F. P. (2018). Environmental transition zone and rivers shape intraspecific population structure and genetic diversity of an Amazonian rain forest tree frog. *Evolutionary Ecology*, *32*(4), 359–378. https://doi.org/10.1007/s10682-018-9939-2

Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., … Dalén, L. (2015). Complete Genomes Reveal Signatures of Demographic and Genetic Declines in the Woolly Mammoth. *Current Biology*, *25*(10), 1395–1400. https://doi.org/10.1016/j.cub.2015.04.007

Parmentier, I., Malhi, Y., Senterre, B., Whittaker, R. J., Alonso, A., Balinga, M. P. B., … WÖLL, H. (2007). The odd man out? Might climate explain the lower tree ?-diversity of African rain forests relative to Amazonian rain forests? *Journal of Ecology*, *95*(5), 1058–1071. https://doi.org/10.1111/j.1365-2745.2007.01273.x

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), 37–42. https://doi.org/10.1038/nature01286

Patiño, J., Carine, M., Fernández-Palacios, J. M., Otto, R., Schaefer, H., & Vanderpoorten, A. (2014). The anagenetic world of spore-producing land plants. *New Phytologist*, *201*(1), 305–311. https://doi.org/10.1111/nph.12480

Patiño, J., Carine, M., Mardulyn, P., Devos, N., Mateo, R. G., González-Mancebo, J. M., … Vanderpoorten, A. (2015). Approximate Bayesian Computation Reveals the Crucial Role of Oceanic Islands for the Assembly of Continental Biodiversity. *Systematic Biology*, *64*(4), 579–589. https://doi.org/10.1093/sysbio/syv013

Patiño, J., & Vanderpoorten, A. (2018). Bryophyte Biogeography. *Critical Reviews in Plant Sciences*, *37*(2–3), 175–209. https://doi.org/10.1080/07352689.2018.1482444

Pearson, R. (2006). Climate change and the migration capacity of species. *Trends in Ecology & Evolution*, *21*(3), 111–113. https://doi.org/10.1016/j.tree.2005.11.022

Pearson, R., & Dawson, T. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x

Pereira, M. R., Amorim, B. S., Sierra, A. M., McDaniel, S., Payton, A., Carey, S. B., … Zartman, C. E. (2019). Advances in Calymperaceae (Dicranidae, Bryophyta): intra-familial phylogenetic relationships, geographical origin and divergence times. *The Bryologist*, in press.

Pereira, M. R., Dambros, C. S., & Zartman, C. E. (2016). Prezygotic resource-allocation dynamics and reproductive trade-offs in Calymperaceae (Bryophyta). *American Journal of Botany*, *103*(10), 1838–1846. https://doi.org/10.3732/ajb.1600240

Petit, R. J. (2003). Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science*, *300*(5625), 1563–1565. https://doi.org/10.1126/science.1083264

Petit, R. J., Brewer, S., Bordács, S., Burg, K., Cheddadi, R., Coart, E., … Kremer, A. (2002). Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, *156*(1–3), 49–74. https://doi.org/10.1016/S0378-1127(01)00634-X

Pillans, B., & Gibbard, P. (2012). The Quaternary Period. In *The Geologic Time Scale* (pp. 979–1010). Elsevier. https://doi.org/10.1016/B978-0-444-59425-9.00030-5

Pisa, S., Werner, O., Vanderpoorten, A., Magdy, M., & Ros, R. M. (2013). Elevational patterns of genetic variation in the cosmopolitan moss Bryum argenteum (Bryaceae). *American Journal of Botany*, *100*(10), 2000–2008. https://doi.org/10.3732/ajb.1300100

Pitman, N. C. A., Terborgh, J., Silman, M. R., & Nuñez V., P. (1999). Tree species distributions in an upper Amazonian forest. *Ecology*, *80*(8), 2651–2661.

Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., & Feldman, M. W. (1999). Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, *16*(12), 1791–1798. https://doi.org/10.1093/oxfordjournals.molbev.a026091

Pross, J., Contreras, L., Bijl, P. K., Greenwood, D. R., Bohaty, S. M., Schouten, S., … Brinkhuis, H. (2012). Persistent near-tropical warmth on the Antarctic continent during the early Eocene epoch. *Nature*, *488*(7409), 73–77. https://doi.org/10.1038/nature11300

Reese, W. D. (1993). Calymperaceae. *Flora Neotropica*, *58*, 1–101. Retrieved from http://www.jstor.org/stable/4393834

Reis, L. S., Guimarães, J. T. F., Souza-Filho, P. W. M., Sahoo, P. K., de Figueiredo, M. M. J. C., de Souza, E. B., & Giannini, T. C. (2017). Environmental and vegetation changes in southeastern Amazonia during the late Pleistocene and Holocene. *Quaternary International*, *449*, 83–105. https://doi.org/10.1016/j.quaint.2017.04.031

Roads, E., Longton, R. E., & Convey, P. (2014). Millennial timescale regeneration in a moss from Antarctica. *Current Biology*, *24*(6), R222–R223. https://doi.org/10.1016/j.cub.2014.01.053

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5463–5467.

Schneider, R., Schmitt, J., Köhler, P., Joos, F., & Fischer, H. (2013). A reconstruction of atmospheric carbon dioxide and its stable carbon isotopic composition from the penultimate glacial maximum to the last glacial inception. *Climate of the Past*, *9*(6), 2507–2523. https://doi.org/10.5194/cp-9-2507-2013

Schönswetter, P., Paun, O., Tribsch, A., & Niklfeld, H. (2003). Out of the Alps: colonization of Northern Europe by East Alpine populations of the Glacier Buttercup Ranunculus glacialis L. (Ranunculaceae). *Molecular Ecology*, *12*(12), 3373–3381. https://doi.org/10.1046/j.1365-294X.2003.01984.x

Schönswetter, P., Popp, M., & Brochmann, C. (2006). Rare arctic-alpine plants of the European Alps have different immigration histories: the snow bed species Minuartia biflora and Ranunculus pygmaeus. *Molecular Ecology*, *15*(3), 709–720. https://doi.org/10.1111/j.1365-294X.2006.02821.x

Schönswetter, P., Stehlik, I., Holderegger, R., & Tribsch, A. (2005). Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, *14*(11), 3547–3555. https://doi.org/10.1111/j.1365-294X.2005.02683.x

Semerikov, V. L., Semerikova, S. A., Polezhaeva, M. A., Kosintsev, P. A., & Lascoux, M. (2013). Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (Larix sibirica): a range-wide analysis of cytoplasmic markers. *Molecular Ecology*, *22*, 4958–4971. https://doi.org/10.1111/mec.12433

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, *68*(1), 1–15. https://doi.org/10.1111/evo.12258

Shaw, J., Meagher, T. R., & Harley, P. (1987). Electrophoretic evidence of reproductive isolation between two varieties of the moss, Climacium americanum. *Heredity*, *59*(3), 337–343. https://doi.org/10.1038/hdy.1987.140

Skrede, I., Eidesen, P. B., Portela, R. P., & Brochmann, C. (2006). Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (Dryas octopetala L.). *Molecular Ecology*, *15*(7), 1827–1840. https://doi.org/10.1111/j.1365-294X.2006.02908.x

Sleator, R. D., Shortall, C., & Hill, C. (2008). Metagenomics. *Letters in Applied Microbiology*, *47*(5), 361–366. https://doi.org/10.1111/j.1472-765X.2008.02444.x

Stark, L. R., & Brinda, J. C. (2013). An experimental demonstration of rhizautoicy, self-incompatibility, and reproductive investment in Aloina bifrons (Pottiaceae). *The Bryologist*, *116*(1), 43–52. https://doi.org/10.1639/0007-2745-116.1.043

Stenøien, H. K., Shaw, A. J., Shaw, B., Hassel, K., & Gunnarsson, U. (2011). North American origin and recent European establishments of the amphi-Atlantic peat moss Sphagnum angermanicum. *Evolution*, *65*(4), 1181–1194. https://doi.org/10.1111/j.1558-5646.2010.01191.x

Sujii, P. S., Martins, K., Wadt, L. H. de O., Azevedo, V. C. R., & Solferini, V. N. (2015). Genetic structure of Bertholletia excelsa populations from the Amazon at different spatial scales. *Conservation Genetics*, *16*(4), 955–964. https://doi.org/10.1007/s10592-015-0714-4

Sundberg, S. (2005). Larger capsules enhance short-range spore dispersal in Sphagnum, but what happens further away? *Oikos*, *108*(1), 115–124. Retrieved from http://www.jstor.org/stable/3548496

Svenning, J.-C., Normand, S., & Kageyama, M. (2008). Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *Journal of Ecology*, *96*(6), 1117–1127. https://doi.org/10.1111/j.1365-2745.2008.01422.x

Szövényi, P., Hock, Z., Korpelainen, H., & Shaw, A. J. (2009). Spatial pattern of nucleotide polymorphism indicates molecular adaptation in the bryophyte Sphagnum fimbriatum. *Molecular Phylogenetics and Evolution*, *53*(1), 277–286. https://doi.org/10.1016/j.ympev.2009.06.007

Szövényi, P., Sundberg, S., & Shaw, A. J. (2012). Long-distance dispersal and genetic structure of natural populations: an assessment of the inverse isolation hypothesis in peat mosses. *Molecular Ecology*, *21*(22), 5461–5472. https://doi.org/10.1111/mec.12055

Szövényi, P., Terracciano, S., Ricca, M., Giordano, S., & Shaw, A. J. (2008). Recent divergence, intercontinental dispersal and shared polymorphism are shaping the genetic structure of amphi-Atlantic peatmoss populations. *Molecular Ecology*, *17*(24), 5364–5377. https://doi.org/10.1111/j.1365-294X.2008.04003.x

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., & Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*(4), 453–464. https://doi.org/10.1046/j.1365-294x.1998.00289.x

Tachibana, C. (2015). Transcriptomics today: Microarrays, RNA-seq, and more. *Science*, *349*(6247), 544–546. https://doi.org/10.1126/science.opms.p1500095

Thomas, E., Caicedo, C. A., Loo Ii, J., & Kindt, R. (2014). The distribution of the Brazil nut (Bertholletia excelsa) through time: from range contraction in glacial refugia, over human-mediated expansion, to anthropogenic climate change. *Bol. Mus. Para. Emílio Goeldi. Cienc. Nat*, *9*(2), 267–291.

Thomé, M. T. C., & Carstens, B. C. (2016). Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. *Proceedings of the National Academy of Sciences*, *113*(29), 8010–8017. https://doi.org/10.1073/pnas.1601064113

Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A., … Zimmermann, N. E. (2008). Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *9*(3–4), 137–152. https://doi.org/10.1016/j.ppees.2007.09.004

Tuba, Z., Slack, N. G., & Stark, L. R. (2011). *Bryophyte Ecology and Climate Change*. Cambridge University Press.

Tzedakis, P. C., Emerson, B. C., & Hewitt, G. M. (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, *28*(12), 696–704. https://doi.org/10.1016/j.tree.2013.09.001

van der Hammen, T., & Absy, M. L. (1994). Amazonia during the last glacial. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *109*(2–4), 247–261. https://doi.org/10.1016/0031-0182(94)90178-3

Van Der Velde, M., & Bijlsma, R. (2003). Phylogeography of five Polytrichum species within Europe. *Biological Journal of the Linnean Society*, *78*(2), 203–213. https://doi.org/10.1046/j.1095-8312.2003.00151.x

Vanderpoorten, A., Patiño, J., Désamoré, A., Laenen, B., Gorski, P., Papp, B., … Hardy, O. J. (2019). To what extent are bryophytes efficient dispersers? *Journal of Ecology*, in press.

Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, *13*(4), 921–935. https://doi.org/10.1046/j.1365-294X.2004.02076.x

Wakeley, J. (2001). The Coalescent in an Island Model of Population Subdivision with Variation among Demes. *Theoretical Population Biology*, *59*(2), 133–144. https://doi.org/10.1006/tpbi.2000.1495

Wakeley, J. (2008). *Coalescent Theory: An Introduction*. Freeman, W. H. & Company.

Wallace, A. R. (1852). On the Monkeys of the Amazon. *Proceedings of the Zoological Society of London*, 107–110. Retrieved from https://www.tandfonline.com/doi/full/10.1080/037454809494374

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649–5662. https://doi.org/10.1111/mec.12938

Wegmann, D., Leuenberger, C., & Excoffier, L. (2009). Efficient Approximate Bayesian Computation Coupled With Markov Chain Monte Carlo Without Likelihood. *Genetics*, *182*(4), 1207–1218. https://doi.org/10.1534/genetics.109.102509

Westergaard, K. B., Alsos, I. G., Popp, M., Engelskjon, T., Flat, K. I., & BROCHMANN, C. (2011). Glacial survival may matter after all: nunatak signatures in the rare European populations of two west-arctic species. *Molecular Ecology*, *20*(2), 376–393. https://doi.org/10.1111/j.1365-294X.2010.04928.x

Willerslev, E. (2003). Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments. *Science*, *300*(5620), 791–795. https://doi.org/10.1126/science.1084114

Willerslev, E., & Cooper, A. (2005). Review Paper. Ancient DNA. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1558), 3–16. https://doi.org/10.1098/rspb.2004.2813

Wu, H., Guiot, J., Brewer, S., & Guo, Z. (2007). Climatic changes in Eurasia and Africa at the last glacial maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling. *Climate Dynamics*, *29*(2–3), 211–229. https://doi.org/10.1007/s00382-007-0231-3

Zartman, C. E., Nascimento, H. E. M., Cangani, K. G., Alvarenga, L. D. P., & Snäll, T. (2012). Fine-scale changes in connectivity affect the metapopulation dynamics of a bryophyte confined to ephemeral patches. *Journal of Ecology*, *100*(4), 980–986. https://doi.org/10.1111/j.1365-2745.2012.01969.x

Zartman, C. E., & Shaw, A. J. (2006). Metapopulation Extinction Thresholds in Rain Forest Remnants. *The American Naturalist*, *167*(2), 177–189. https://doi.org/10.1086/499376

Zechmeister, H. G., Moser, D., & Milasowszky, N. (2007). Spatial distribution patterns of Rhynchostegium megapolitanum at the landscape scale - an expanding species? *Applied Vegetation Science*, *10*(1), 111–120. https://doi.org/10.1111/j.1654-109X.2007.tb00509.x