Genetic structure of the protist Physarum albescens (Amoebozoa) revealed by multiple markers and genotyping by sequencing

ORIGINAL ARTICLE

Genetic structure of the protist *Physarum albescens* (Amoebozoa) revealed by multiple markers and genotyping by sequencing

Oleg Shchepin[1,2] [ID] | Yuri Novozhilov[1] | Jan Woyzichovski[2] | Manuela Bog[2] | Ilya Prikhodko[1] | Nadezhda Fedorova[1,3] | Vladimir Gmoshinskiy[4,5] | Mathilde Borg Dahl[2,6] | Nikki H. A. Dagamac[2,7] | Yuka Yajima[8] | Martin Schnittler[2]

[1]Laboratory of Systematics and Geography of Fungi, Komarov Botanical Institute of the Russian Academy of Sciences, St. Petersburg, Russia

[2]General Botany and Plant Systematics, Institute of Botany and Landscape Ecology, University of Greifswald, Greifswald, Germany

[3]Faculty of Biology, Saint Petersburg State University, Saint Petersburg, Russia

[4]Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

[5]Polistovsky National Nature Reserve, Pskov Region, Russia

[6]Institute of Microbiology, Center for Functional Genomics of Microbes, University of Greifswald, Greifswald, Germany

[7]Department of Biological Sciences and Research Center for the Natural and Applied Sciences, University of Santo Tomas, Manila, Philippines

[8]Muroran Institute of Technology, Muroran, Japan

Correspondence

Oleg Shchepin, Laboratory of Systematics and Geography of Fungi, Komarov Botanical Institute of the Russian Academy of Sciences, St. Petersburg, Russia.

Email: oshchepin@gmail.com

Abstract

Myxomycetes are terrestrial protists with many presumably cosmopolitan species dispersing via airborne spores. A truly cosmopolitan species would suffer from outbreeding depression hampering local adaptation, while locally adapted species with limited distribution would be at a higher risk of extinction in changing environments. Here, we investigate intraspecific genetic diversity and phylogeography of *Physarum albescens* over the entire Northern Hemisphere. We sequenced 324 field collections of fruit bodies for 1–3 genetic markers (SSU, EF1A, COI) and analysed 98 specimens with genotyping by sequencing. The structure of the three-gene phylogeny, SNP-based phylogeny, phylogenetic networks, and the observed recombination pattern of three independently inherited gene markers can be best explained by the presence of at least 18 reproductively isolated groups, which can be seen as cryptic species. In all intensively sampled regions and in many localities, members of several phylogroups coexisted. Some phylogroups were found to be abundant in only one region and completely absent in other well-studied regions, and thus may represent regional endemics. Our results demonstrate that the widely distributed myxomycete species *Ph. albescens* represents a complex of at least 18 cryptic species, and some of these seem to have a limited geographical distribution. In addition, the presence of groups

of presumably clonal specimens suggests that sexual and asexual reproduction coexist in natural populations of myxomycetes.

KEYWORDS
cryptic species, DNA barcoding, phylogeography, protists, simulation, slime molds

1 | INTRODUCTION

Terrestrial soil protists have received much less scientific attention than prokaryotes, fungi, and viruses, despite their functional significance (Geisen et al., 2020). Their diversity and biogeography are generally understudied, but a few studies have found presumed cosmopolitan soil protist taxa to consist of multiple cryptic species (Pinseel et al., 2020; Ryšánek et al., 2015; Singer et al., 2019). Some of these cryptic species showed endemic distributions, and others proved to be cosmopolitan (Foissner, 2008). The fact that these endemic protist species exist implies that at least in some taxa, gene flow between remote populations is not frequent enough to prevent populations from diverging (Mann & Vanormelingen, 2013).

Among terrestrial protists, myxomycetes are especially convenient models to study biogeography and phylogeography. Myxomycetes are amoeboid protists with a complex life cycle that includes microscopic amoeboflagellates, resting microcysts, multinuclear plasmodia, and fruiting bodies (sporocarps) containing airborne spores. Myxomycete spores have been detected in the air (Kamono et al., 2009), but no direct evidences for a long-distance dispersal have been found. Nevertheless, their ability for a long-distance dispersal is indirectly confirmed by presumably cosmopolitan distribution of many species and a relatively diverse myxomycete assemblages found on remote islands (see a detailed discussion in Schnittler et al., 2021). Myxomycete sporocarps are often conspicuous, easy to find in nature and preserve as dried herbarium specimens. Herbarium specimens can be used for DNA isolation and 18S rDNA (SSU) barcoding (Borg Dahl, Brejnrod, et al., 2018; Schnittler et al., 2017). However, virtually all studies on myxomycete diversity and distribution have been conducted at the morphological species (morphospecies) level.

The few studies that investigated the genetic structure of particular myxomycete morphospecies produced similar results. Aguilar et al. (2014) studied the geographical distribution of SSU variants in *Badhamia melanospora* (here and throughout the manuscript, myxomycete nomenclature follows Lado, 2005–2021). They found two geographically structured groups of ribotypes, which were congruent with slight morphological differences in spores. The authors suggested that this taxon constituted a complex formed by at least two cryptic species. In *Physarum pseudonotabile*, a SSU and EF1A gene phylogeny revealed several well-separated genetic lineages (Novozhilov et al., 2013). Analysis of three independent genes (SSU, EF1A, and COI) and group I introns in SSU demonstrated the existence of three cryptic species in *Trichia varia* (Feng & Schnittler,

2015). Similarly, *Lepidoderma chailletii* turned out to be a polyphyletic complex, which consisted of three cryptic species, as was shown by the analysis of SSU, EF1A, and COI (Shchepin et al., 2016). Dagamac et al. (2017) analysed two-gene data (SSU and EF1A) and revealed four putative cryptic species in *Hemitrichia serpula* with subtle morphological differences. Finally, several large phylogroups were found within *Didymium nivicola* in SSU and EF1A phylogenies, with striking differences in genetic diversity between the Northern and the Southern Hemisphere (Janik et al., 2020).

However, many intriguing questions remain unanswered. How common is the presence of cryptic species in myxomycetes? How widely are such species distributed? Trying to answer these questions, we applied a three-gene approach and, for the first time in myxomycetes, genotyping by sequencing to investigate the genetic structure of populations of the myxomycete *Physarum albescens*.

Ph. albescens (Figure 1a,b) abundantly occurs in mountains of the Northern Hemisphere but is rarely reported from the Southern Hemisphere. It belongs to the ecological guild of snowbank (nivicolous) myxomycetes (Schnittler et al., 2015), which makes up for c. 10% of the morphospecies diversity in myxomycetes. Myxamoebae inhabit the uppermost soil layer; fructifications emerge in spring at the edge of melting snow patches in mountainous areas (Figure 1c; Ronikier & Ronikier, 2009). *Ph. albescens* is morphologically clearly distinctive from other members of the guild. Moreover, it is one of the most common nivicolous species in some regions. Its brightly coloured sporocarps are easily visible in nature, and no infraspecific taxa are described within it, which makes it an ideal candidate for a large-scale phylogeographic study.

2 | MATERIALS AND METHODS

2.1 | Material studied

We studied 324 herbarium specimens of *Ph. albescens*, which were collected in several mountain ranges of the Northern Hemisphere between 2010 and 2018 (Europe, Central Asia, East Asia, and North America). A specimen was defined as a colony of closely located sporocarps that share the same hypothallus and probably have originated from one plasmodium. Additionally, we studied 44 herbarium specimens that were received from other collectors. Based on the GPS coordinates (WGS84) of collection sites, all specimens were assigned to standardized localities of approximately 50 m diameter using the python package GEOPY and a custom python script
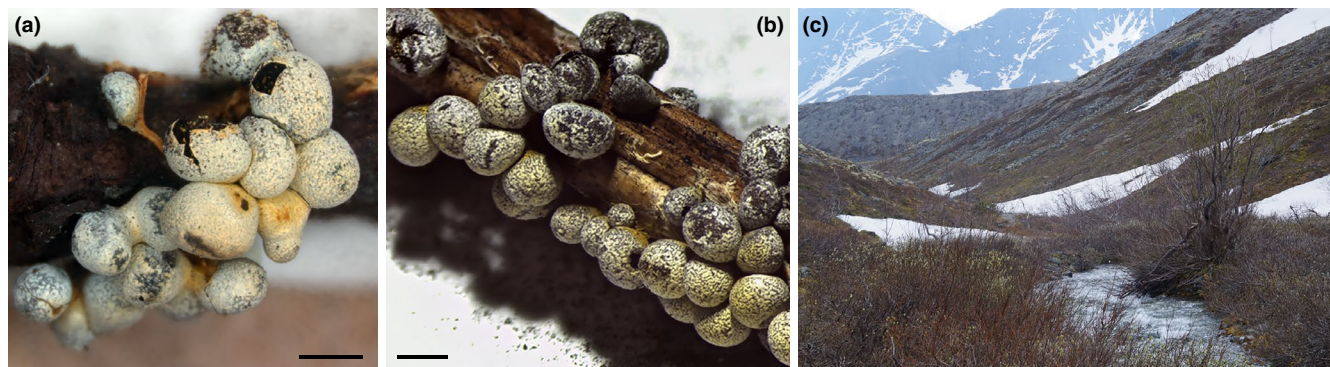
FIGURE 1 (a–b) Sporocarps of *Ph. albescens* from Khibiny Mts. (a, specimen LE305933) and Luvenga Tundra Mts. (b, MYX9295) at the Kola peninsula, Russia. (c) Typical habitat of *Ph. albescens* in Khibiny Mts. Scale bars = 1 mm

that performed greedy centroid-based clustering of collection sites. The resulting material included 368 specimens from 16 regions and 164 localities (Table S1).

2.2 | Genetic markers

Partial sequences of three independently inherited genetic markers were analysed. First, the beginning of the small ribosomal subunit gene (18S rDNA, SSU), including V2 region, located on nuclear minichromosomes. Second, the protein elongation factor subunit 1-alpha gene (EF1A) with nuclear chromosomal localization. Third, mitochondrial cytochrome c oxidase subunit one gene (COI). These markers proved to be highly variable in myxomycetes and valuable for species delimitation and phylogenetic inference (Schnittler et al., 2017).

For DNA extraction, approximately 2–5 sporocarps were sampled from a herbarium box and placed in 2 ml safe-lock tubes. Samples were homogenized either in a TissueLyser LT homogenizer (Qiagen) with 3 mm diameter steel balls or in a FastPrep-24 (MP Biomedicals) with sterilized sea sand and glass beads. DNA was extracted with PhytoSorb (Sintol) or with E.Z.N.A. Plant DNA Kit (Omega Bio-tek) according to the manufacturer's protocols.

The primers used were S1/SU19R or S3bF/S31R for SSU (Fiore-Donno et al., 2008, 2012; Hoppe & Schnittler, 2015; Schnittler et al., 2017), PB1F/PB1R for EF1A (Novozhilov et al., 2014), COIF1/COIR1 or COMF/COMRs for COI (Feng & Schnittler, 2015; Liu et al., 2015; Novozhilov et al., 2019). Amplification protocols were as follows: denaturation for 5 min at 95°C, 39 cycles (30 s at 95°C, 30 s at 56°C, 1 min at 72°C) and 5 min at 72°C with S1/SU19R; denaturation for 5 min at 95°C, 34 cycles (30 s at 95°C, 30 s at 57.6°C, 48 s at 72°C) and 5 min at 72°C with S3bF/S31R; denaturation for 5 min at 95°C, 35 cycles (30 s at 95°C, 30 s at 65.4°C, 1 min at 72°C) and 10 min at 72°C with PB1F/PB1R; denaturation for 5 min at 95°C, 34 cycles (30 s at 95°C, 50 s at 50.7°C or 52.0°C, 1 min at 72°C) and 10 min at 72°C with COIF1/COIR1 or COMF/COMRs.

Amplicons were sequenced on ABI 3500 automated DNA sequencer (Applied Biosystems) or by Macrogen Europe. Sequence

chromatograms were examined in BioEdit 7.2.5 (Hall, 1999) and base-calling errors were corrected manually. New sequences were submitted to GenBank (accession numbers MW691477–MW691847, MW692988–MW693025, MW701443–MW701877, Table S1).

2.3 | Genotyping by sequencing (GBS)

For genotyping, 108 samples were prepared (Table S1). They represented 98 herbarium specimens of *Ph. albescens* from seven phylogenetic groups. Of them, two specimens were prepared in triplicates and six specimens in duplicates for evaluation of technical errors. From each specimen, 25–100 sporocarps (depending on their size and condition) were transferred into a 2 ml microtube with sterilized sea sand and glass beads and homogenized with a FastPrep-24 (MP Biomedicals). DNA was extracted with the E.Z.N.A. Plant DNA Kit (Omega Bio-tek) according to the manufacturer's protocol.

At least 100 ng of high molecular weight DNA per sample was required for GBS. DNA integrity was checked using gel electrophoresis with Lambda DNA/HindIII Marker. DNA concentration was measured using a Qubit 4 fluorometer (Invitrogen). GBS was performed by LGC Genomics with MslI restriction endonuclease and paired-end sequencing (2 × 150 bp) on an Illumina NextSeq 500 platform (Illumina Inc.) with a sequencing depth of 0.5 million paired reads per sample. Preprocessing (demultiplexing, trimming sequencing adapters and restriction enzyme sites from raw sequence reads) was done by the sequencing facility. The raw data were submitted to the sequence read archive (PRJNA706537).

Preprocessed paired-end reads were de novo assembled and analysed using steps 3–7 of the IPYRAD 0.9.26 pipeline (Eaton & Overcast, 2020). In brief, reads were quality-filtered, dereplicated and clustered within samples with a specified sequence similarity threshold, maximum number of alleles per site, and minimum and maximum allowed sequencing depth. For each cluster, consensus sequence was produced. After that, consensus sequences from all samples were aligned together to obtain contigs. Only contigs that covered at least 10 samples were retained (value specified with the core parameter of step 7 "minimum number of samples per locus").

We optimized the core assembly parameter "clustering threshold for de novo assembly" (*ct*) by running the analysis with three different values of *ct* (0.85, 0.90, 0.95) on the data of 10 technical replicates. The value 0.95 was chosen since the lower values produced apparent assembly errors leading to increased genetic distances between replicated samples. The resulting alignment of single-nucleotide polymorphisms (SNPs) was subjected to phylogenetic analysis and phylogenetic network construction.

ABBA-BABA statistics were calculated in DSUITE 0.4 (Malinsky et al., 2021) using DTRIOS command for 61 specimens from five phylogroups (Ha-He) and 25 specimens from phylogroup B as an outgroup. Relationships between phylogroups were specified by a simple tree based on the well-resolved maximum likelihood tree obtained for the SNP alignment: (Outgroup,((Hc, Hb),(Hd,(Ha, He)))). DTRIOS calculates Patterson's *D* and f_4-ratio statistics and finds trios of phylogroups with significantly elevated *D* using *Z*-score, which is estimated with a block-Jackknife procedure that takes linkage among sites into account. The option Jknum (the number of jackknife blocks to divide the data set into) was left with its default value (20 blocks of 1278 SNPs), since different tested values (20, 30, 40, 50) did not alter the results. A Benjamini-Hochberg correction for multiple tests was applied to the resulting *p*-values to control the false-discovery rate. To aid the interpretation of correlated f_4-ratio results, *f*-branch statistic ($f_b(C)$) was calculated using FBRANCH command, and its results were plotted using dtools.py. For the phylogroup trios with significantly elevated *D*, a sliding-window analysis was performed using DINVESTIGATE command to assess whether the admixture signal was confined to specific contigs (window size = 2, step = 1).

To test for the presence of clonal specimens, genetic distances between all pairs of specimens and between technical replicates were calculated using the Hamming Distances method (Hamming, 1950) with default options in SPLITSTREE5 5.0.0_alpha (Huson, 1998; Huson & Bryant, 2006). Groups of specimens that showed genetic distances not exceeding the maximum distances between technical replicates were classified as putative clones.

2.4 | Phylogeny construction

Four sequence data sets were built: (1) a two-gene data set (SSU and exons of EF1A) for *Ph. albescens* and diverse members of Physarales to infer phylogenetic relationships of *Ph. albescens* with other members of the order, (2) an SSU data set for all specimens of *Ph. albescens* for initial screening of species-wide genetic diversity, (3) a three-gene data set (SSU, exons of EF1A, and COI) for a subset of specimens from each clade of the SSU phylogeny to test for the presence of cryptic species within the morphologically defined species, and (4) a SNP data set resulting from the analysis of GBS data to get a phylogeny with a higher resolution for a subset of specimens.

In the first three data sets, sequences of different genes were aligned separately with MAFFT online service with default gap penalties (Katoh et al., 2019). Two alignment strategies were applied: E-INS-I for SSU, G-INS-I for EF1A and COI (Katoh et al., 2005). All alignments were visually inspected, corrected if necessary, and trimmed. Poorly aligned regions with multiple indels in SSU alignment of the first data set were removed manually prior to the phylogenetic analyses (mask in Alignment S1 shows the remaining positions). Introns in all EF1A alignments were as well removed before the alanyses.

Maximum likelihood (ML) phylogenies for all four data sets were built using W-IQ-TREE 1.6.11 (Trifinopoulos et al., 2016) with 1000 ultrafast bootstrap replicates (Minh et al., 2013). For the SNP data set, an ascertainment bias correction model was applied. Optimal substitution models for ML analysis were chosen independently for each partition with MODELFINDER on W-IQ-TREE web server (Kalyaanamoorthy et al., 2017) according to the Bayesian information criterion.

Bayesian inference was computed with MRBAYES 3.2.1 on CIPRES Science Gateway (Huelsenbeck & Ronquist, 2001; Miller et al., 2010). Five simultaneous runs with the evolutionary model set to GTR+G4+I included one cold and three heated Monte Carlo Markov chains. The number of generations, sample frequencies, and burnin ratio were set to 50 M, 1000, and 0.25, respectively. The convergence of MCMC chains in Bayesian inference analyses was assessed using TRACER 1.7.1 and R package RWTY 1.0.2 (Warren et al., 2017). Clade confidence scores based on the results of Bayesian inference analyses were placed on ML trees using IQ-TREE (it implements an algorithm that, similar to sumt command in MRBAYES, counts the proportion of the Bayesian trees where every taxon bipartition appears).

The two-gene data set included SSU and EF1A sequences for 298 myxomycete specimens belonging to 47 species of different Physarales, with three *Lamproderma* species as an outgroup (Alignment S1, Table S1). Of them, 194 specimens belonged to *Ph. albescens*. 66 SSU and 55 EF1A sequences were obtained from GenBank. The concatenated alignment contained 1434 sites, 505 of them parsimony informative. Four partitions were defined (Chernomor et al., 2016): one for SSU (model SYM+R4) and three for different codon positions in EF1A (models HKY+F+R4, TVM+F+I+G4, and TVM+F+G4 for the first, second, and third position, respectively).

The SSU data set included 372 nucleotide sequences (Alignment S2, Table S1). Of these, 368 sequences of *Ph. albescens* were obtained in this study and four sequences fetched from GenBank, with *Physarum polycephalum* as an outgroup. Among 565 nucleotide positions, 56 were parsimony informative. For phylogenetic inference, a single partition was defined with TNe+R2 model.

The three-gene data set included SSU, EF1A, and COI sequences for 182 specimens of *Ph. albescens*. SSU and EF1A sequences of three specimens of *Badhamia foliicola* were added as an outgroup, based on the two-gene analysis results (Alignment S3, Table S1). The concatenated alignment contained 1946 sites, 302 of them parsimony informative. Five partitions were defined (Chernomor et al., 2016): one for SSU (model K2P+R2), three for different codon positions in EF1A (models F81+F+I, F81+F+I, and TPM2u+F+G4 for the first, second, and third position, respectively), and one for COI (model K3Pu+F+I+G4).

The SNP data set included SNP data for 108 samples (98 specimens and technical replicates) of *Ph. albescens* (Alignment S4, Table S1). The alignment contained 25,589 positions, 10,364 of them parsimony informative. For phylogenetic inference, a single partition was defined with TVM+F+R3 model.

2.5 | Phylogenetic network construction

A phylogenetic network was built for the three-gene data set and the SNP data set. The analysis was performed using SPLITSTREE5 5.0.0_alpha (Huson, 1998; Huson & Bryant, 2006). A distance matrix was calculated using the Hamming Distances method (Hamming, 1950) with default options. The neighbor-net method (Bryant & Moulton, 2004) was used to obtain splits. The Splits Network Algorithm (Dress & Huson, 2004) produced a splits network with 5716 nodes and 10,883 edges for the three-gene data set, and with 1650 nodes and 3002 edges for the SNP data set.

2.6 | Phylogroup delimitation

Several approaches were applied to delimit phylogroups representing putative reproductively isolated cryptic species. First, we used species delimitation approach based on multilocus "fields for recombination" (ml-FFRs) concept proposed by Doyle (1995). Following the logic described in detail by Flot et al. (2010), we tried to determine groups of specimens that were not necessarily reciprocally monophyletic in each of the gene phylogenies, but their pools of alleles of three independently inherited genes were mutually exclusive (mutual allelic exclusivity criterion). For this, unique sequence variants of SSU, EF1A, and COI in *Ph. albescens* were determined by clustering sequences with a 100% similarity threshold using the --cluster_size command in VSEARCH 2.14.2 (Rognes et al., 2016). Each sequence variant received a numeric code. A custom GUI python program was written to visualize the combinations of the variants of three independently inherited genes (SSU, EF1A, and COI) within individuals (herbarium specimens) of *Ph. albescens* (Shchepin, 2021a). The resulting groups of individuals that do not have any allele in common (ml-FFRs) were compared with the topologies of the three-gene phylogeny and the SNP-based phylogeny, and ml-FFRs supported by the tree topologies were considered as putative cryptic species. If two distinct well-supported clades shared only one allele of one gene, they were still considered as two putative cryptic species. If a ml-FFR consisted of a single individual and it was closely related to another ml-FFR in a phylogeny, it was united with it into one phylogroup.

In addition, we employed two automatic species delimitation algorithms. First, multirate Poisson tree processes (MPTP) approach was employed. The analyses were run on MPTP 0.2.0 web server (Kapli et al., 2017) with default settings, using the three-gene ML phylogeny and SNP ML phylogeny as inputs. Second, assemble species by automatic partitioning (ASAP) analyses were run on a web server with default settings using the three-gene and SNP alignments as inputs (Puillandre et al., 2021).

2.7 | Sequence clustering

To evaluate the taxonomical resolution of SSU-based DNA barcoding at the intramorphospecies level, SSU sequences of *Ph. albescens* were clustered with two similarity thresholds that were applied for species determination or clustering of the operational taxonomic units (OTUs) in previous studies: (1) 99.1% (Borg Dahl et al., 2019; Borg Dahl, Brejnrod, et al., 2018; Borg Dahl, Shchepin, et al., 2018; Shchepin et al., 2017, 2019) and (2) 98% (Clissmann et al., 2015; Gao et al., 2019; Kamono et al., 2013; Shchepin et al., 2019). SSU sequences were trimmed to match ca. 350 bp fragment covered by the primers S3bF/S31R used in DNA metabarcoding studies (columns 98–457 in Alignment S2). The --cluster_size command implemented in VSEARCH was used for clustering.

2.8 | Pairwise genetic distances

Genetic distances were calculated for all pairs of sequences from the three-gene data set and the SSU data set. Additionally, within-phylogroup and between-phylogroup pairwise genetic distances were computed for nonsingleton phylogroups. The analyses were conducted in MEGA X (Kumar et al., 2018). Genetic distance was defined as the number of base differences between two sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). Standard error estimates were obtained by a bootstrap procedure (100 replicates).

2.9 | Diversity estimates

Phylogroup and SSU sequence variant diversity estimates (richness, sample completeness, Shannon and Simpson diversity) together with individual-based rarefaction and extrapolation sampling curves for phylogroup richness of *Ph. albescens* were obtained using the R package INEXT2.0.20 (Hsieh et al., 2016). In addition, the same analyses were performed for six other myxomycete species based on data from published phylogeographic studies: *Badhamia melanospora*, *Didymium nivicola*, *Hemitrichia serpula*, *Lepidoderma chailletii*, *Meriderma atrosporum*, *Trichia varia* (Aguilar et al., 2014; Dagamac et al., 2017; Feng et al., 2016; Feng & Schnittler, 2015; Janik et al., 2020; Shchepin et al., 2016).

2.10 | Computer simulation

To test for the presence of genetic isolation between members of different phylogroups of *Physarum albescens*, a python script was written (Shchepin, 2021b). It takes a table of observed combinations

of genetic variants of three independently inherited genes and creates 1000 times the first offspring generation, consisting of a specified number of individuals (equal to the number of studied specimens in the input table) with randomly combined multilocus genotypes.

To produce a new simulated individual, the algorithm randomly samples with replacement two "parent" individuals and randomly chooses one SSU variant (from parent 1 or parent 2, non-Mendelian inheritance), one COI variant (from parent 1 or parent 2, non-Mendelian inheritance), and two EF1A variants (one of the two alleles from parent 1 and one of the two alleles from parent 2, Mendelian inheritance). The simulation does not consider that in myxomycetes some variants of SSU and COI are preferentially preserved during the homogenization process (Ferris et al., 1983; Silliker & Collins, 1988; Silliker et al., 2002). To test if this factor would influence the results, an alternative simulation was created: here, a weight value (0, 1, or 2) was randomly assigned to each SSU and each COI variant. Whenever an individual was produced, the weights of the parental SSU and COI variants were compared, and the variant with a larger weight was inherited. If the weights were identical, the variant to be inherited was chosen randomly.

To minimize the possible influence of geographic isolation, the input data included only specimens collected in 2–3 adjacent valleys within a given mountain range. Five input data sets were compiled: (1) Forty-one specimens from Khibiny Mts. (phylogroups Ha, Hb, Hc), (2) Thirty-six specimens from Spanish Sierra Nevada (phylogroups B, D, Fa, G, M), (3) Sixteen specimens from Khibiny Mts. (phylogroup Hb only), (4) Twenty-two specimens from Khibiny Mts. (phylogroup Hc only), and (5) Eleven specimens from Spanish Sierra Nevada (phylogroup D only).

For each data set, the percentage of individuals with genetic variants coming from different phylogroups ("mixed" genotypes) in the observed and simulated data was calculated. Additionally, the average number of unique multilocus genotypes over 1000 simulated generations was calculated.

The Sierra Nevada data set was additionally tested with different values for generation size (from 100 to 5000, with an increment of 100) to determine the influence of generation size on the number of unique genotypes in the simulated offspring.

3 | RESULTS

3.1 | Genetic diversity of *Ph. albescens*

The 368 specimens of *Physarum albescens* collected throughout the Northern Hemisphere revealed 44 unique variants of partial SSU sequences. The monophyly of this morphospecies was confirmed by a two-gene phylogeny of Physaraceae (Figure S1). A phylogeny based on the SSU sequences (Figure S2) separated *Ph. albescens* into 13 clades of closely related specimens (phylogroups), with four additional phylogroups consisting of one specimen each (singletons). For a subset of specimens from each phylogroup, partial sequences of a mitochondrial gene COI and nuclear gene EF1A were obtained.

Among 204 COI sequences, 46 unique variants (haplotypes) were found. The highest number of genetic variants was found for the exons of the EF1A gene: 58 unique heterozygous variants (corresponding to 87 allelic variants) among 194 sequences. The three-gene genotypes' total diversity equalled 85, with 61% of them occurring in one specimen only (Figure 2a). In all studied specimens of *Ph. albescens*, the amplified fragment of EF1A contained a spliceosomal intron 59–110 bp in length near the end. This intron was absent in all other species of Physaraceae sequenced to date except for one accession of *Badhamia nitens* (MA-Fungi 57896) that was sister to *Ph. albescens* in the two-gene phylogeny (Alignment S1).

A robust phylogeny based on the three-gene data set separated *Ph. albescens* into multiple phylogroups with high statistical support. This pattern was reproduced in the phylogenetic network based on the same data (Figure 3a). The analysis of combinations between the sequence variants of three independently inherited gene markers (SSU, COI, and EF1A) revealed 22 ml FFRs, six of them consisting of a single specimen (Figure 2c). Based on these ml-FFRs and the topology of the three-gene tree, 18 phylogroups representing putative cryptic species and two singleton phylogroups were delimited according to the rules described in methods section (Figure 2b). The presence of ml-FFRs matching the clades in the three-gene phylogeny suggested the absence of recombination between the members of different phylogroups, that is, their reproductive isolation (Figure 2c). Only two exceptions were found: SSU variant 21 appeared in specimens from two closely related phylogroups (Ha and Hb), and the EF1A variant 32 was shared by sister phylogroups Ha and He.

Two species delimitation algorithms applied to the three-gene data set, mPTP and ASAP, suggested the presence of 17 and 19 species (including singletons), respectively (Table 1). Both algorithms united several phylogroups together (Fa+Fb, Ha+He, Hb+Hc). In addition, ASAP split the phylogroup I into three species.

For the three-gene data set, the pairwise genetic distance between specimens of *Ph. albescens* averaged 48.34 (SE 0.21, SD 27.47, min 0.00, max 102.00) base differences between two sequences. The average within-phylogroup distance was as low as 3.01 (SE 1.10, SD 4.69, min 0.00, max 19.67), while the average between-phylogroup distance was 62.48 (SE 1.32, SD 16.35, min 5.83, max 87.85). For SSU, the average between-phylogroup distance was 62.48 (SE 0.31, SD 4.05, min 0.896, max 22.56). Matrices of genetic distances and their descriptive statistics for SSU, COI, EF1A, and the three-gene data can be found in Table S2.

3.2 | GBS results

Raw Illumina sequencing data for 98 specimens representing seven phylogroups of *Ph. albescens* contained 87.6 million paired-end reads. After quality filtering, clustering withing samples, and filtering clusters by sequencing depth, 1,220,930 consensus reads were produced. Consensus reads were assembled into 76,192 contigs. Of them, 4521 contigs that covered at least 10 samples were retained (length: min 35, max 304, N50 200). Only 664 contigs were present
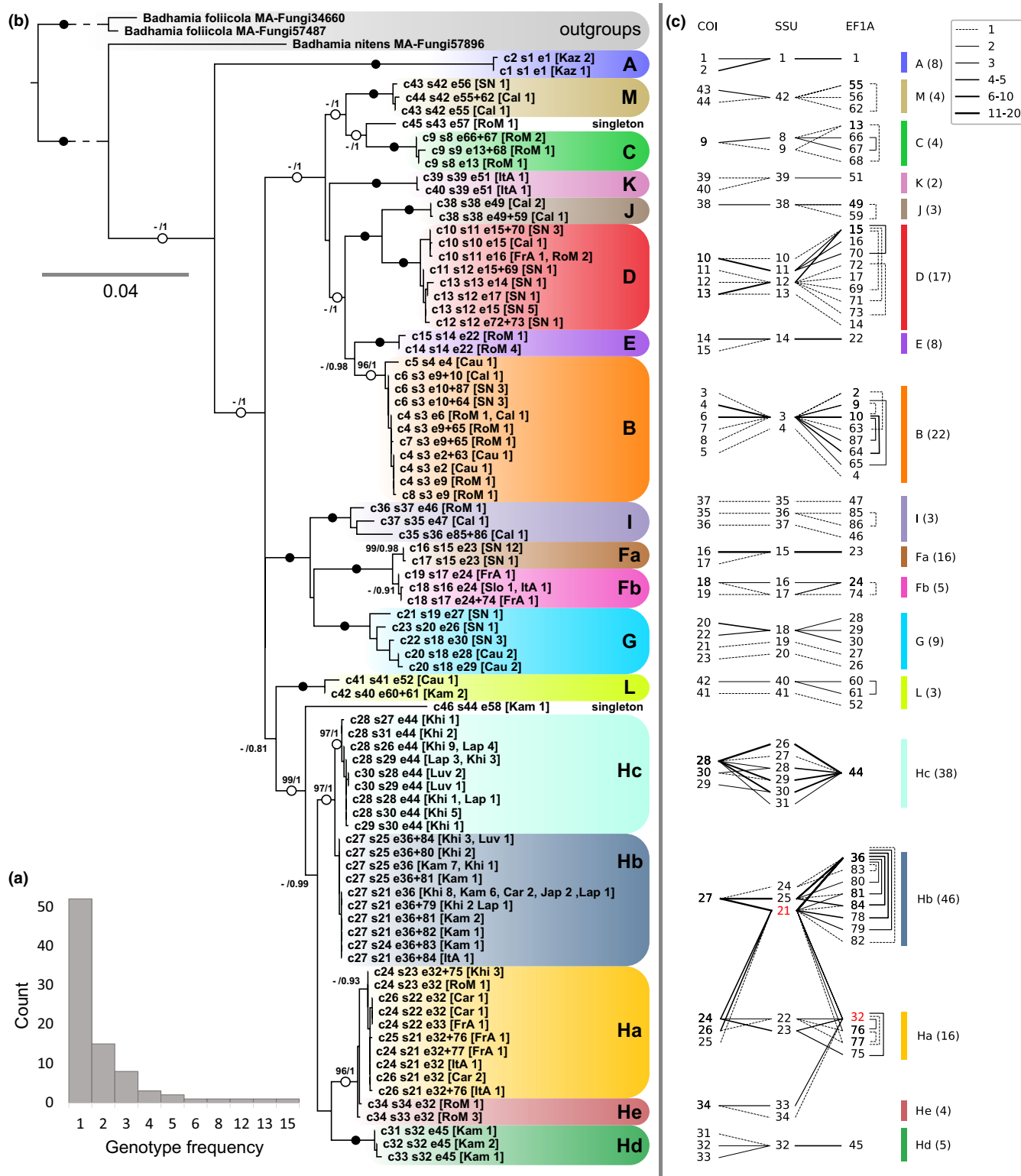
in at least half of the samples. Each sample was on average represented by 11,305 reads (min 1537, max 38,800) and 1180 contigs (min 111, max 1833).

SNP calling resulted in an alignment of 25,589 SNPs with 71.63% missing sites. A phylogenetic network (Figure 3b) that was produced using this alignment separated the specimens into the same seven phylogroups as the three-gene phylogeny (Figure 2b).

The SNP-based ML analysis produced a phylogeny with maximum bootstrap support for almost all major branches (Figure 4). Its topology again confirmed the separation of the investigated specimens into the same seven phylogroups as the SSU and the three-gene phylogenies.

Two species delimitation algorithms, MPTP and ASAP, produced different results when applied to the SNP data set (Table 1). MPTP
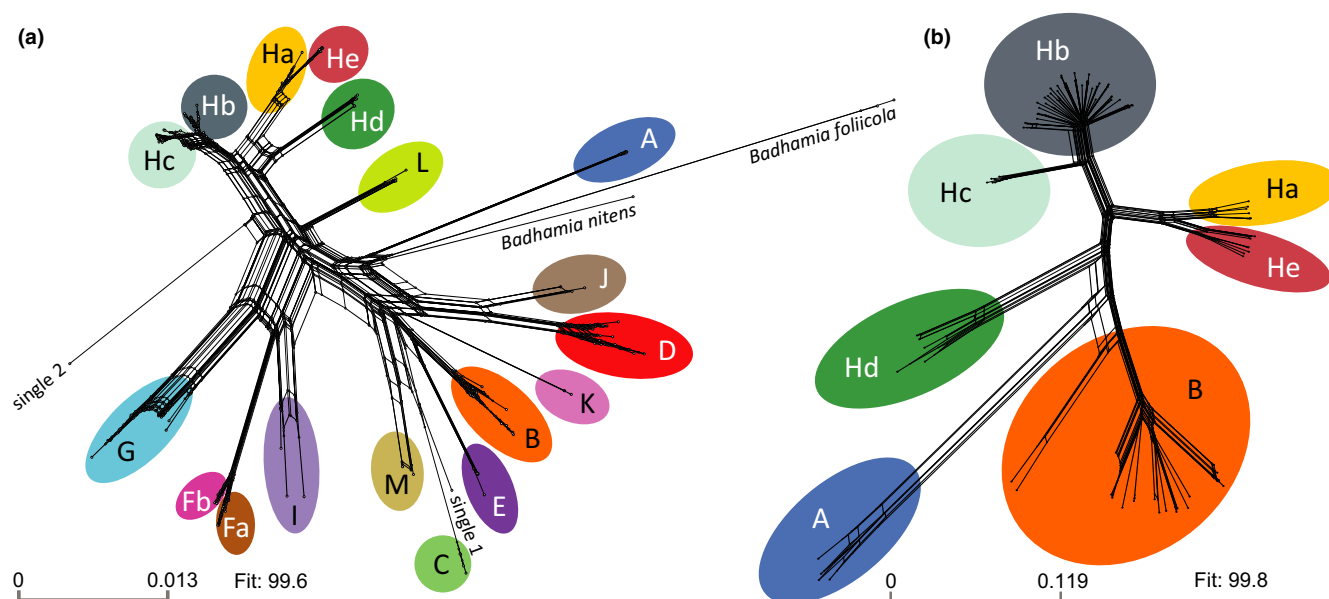
FIGURE 3 Phylogenetic networks produced for *Ph. albescens* in SplitsTree5 using the three-gene data set (a) and the SNP data set for seven phylogroups investigated with GBS (b). Phylogroups are highlighted with the same colours as in Figure 2. Scale bar indicates the number of substitutions per site

supported delimitation of the phylogroups Ha, Hb, Hc, Hd, and He, but suggested splitting of B into six species. ASAP supported delimitation of A and Hd, but sugested splitting B into two species and uniting Hb with Hc and Ha with He.

ABBA-BABA tests were conducted for ten trios of closely related phylogroups Ha–He to detect possible introgression events. Four trios showed *D* statistic significantly different from zero ($Z > 3$, *p*-adjusted <.00135) and large f_4-ratio values (0.18–0.23) (Table S3). A sliding-window analysis demonstrated that SNPs showing ABBA-BABA patterns for each trio were distributed over 25–29 different contigs (Table S3). However, a single gene-flow event can lead to multiple elevated *D* and f_4-ratio results. When a branch-specific $f_b(C)$ was calculated that partially disentangles correlated f_4-ratio results, only two of the eight branches in the phylogeny showed significant excess sharing of derived alleles

with at least one other phylogroup (Figure 5). Specifically, an internal branch representing the common ancestor of phylogroups Ha and He showed an excess sharing of derived alleles with phylogroups Hb and Hc. One more $f_b(C)$ signal involves two of the same four phylogroups: Hc and Ha. These results are consistent with the observation of an SSU variant shared by Ha and Hb and an EF1A variant shared by Ha and He (Figure 2c), which, however, could be also explained by incomplete lineage sorting. It should also be taken into account that a single introgression event can still lead to significant $f_b(C)$ values across multiple related phylogroups (Malinsky et al., 2018).

Pairwise genetic Hamming distances between technical replicates ranged between 0.0011 and 0.0097, while distances between unique specimens averaged 0.2446 (min 0.0012, max 0.5292). Among the phylogroups A, B, Ha, Hb, Hd, and He, some pairs of

TABLE 1 Results of delimitation of cryptic species within *Physarum albescens* obtained using different methods

Phylogroup	SSU barcode		Three-gene data		GBS SNPs	
	98% similarity	99.1% similarity	MPTP	ASAP	MPTP	ASAP
A	5	6	17	10	1	4
B	3*	4	6	9	2^, 3^, 4^, 5^, 6^, 7^	1^, 5^
C	8^, 9^	15	10	18		
D	6*	7	4	2		
E	3*	13	7	14		
Fa	2*	3*	1*	1*		
Fb	2*	3*	1*	1*		
G	4*	5	3	3		
Ha	1*	1*	13*	6*	10	3*
Hb	1*	1*	12*	5*	9	2*
Hc	1*	1*	12*	5*	8	2*
Hd	1*	19	14	15	12	6
He	1*	14^, 17^	13*	6*	11	3*
I	4^, 7^, 8^	8^, 10^	2	8^, 13^, 17^		
J	6*	9	5	4		
K	8*	11	8	12		
L	1*	2	16	11		
M	3*	12	9	16		
Single1	8*	16	11	19		
Single2	10	18	15	7		

Note: "Phylogroup" – species delimited manually based on ml-FFRs and topologies of the three-gene and SNP phylogenies. "SSU barcode" – species delimited based on two different SSU barcoding similarity thresholds. MPTP and ASAP delimitations are presented separately for the three-gene data and for the GBS SNP data. GBS SNP data are available only for seven phylogroups. Predicted species that contain members of more than one phylogroup are marked with *Predicted species that split one phylogroup into several species are marked with ^.

specimens showed a genetic distance comparable to the distance between technical replicates (0.0012–0.0094). Due to this, 18 groups of specimens that showed pairwise genetic distances not exceeding 0.01 were denoted as putative clones (Figure 4, Table S3). In total, 57 of 98 specimens were included into putatively clonal groups, with groups consisting of 2–8 specimens. Within phylogroup Hb, nine putatively clonal groups were found.

In most pairs of putatively clonal specimens, both specimens were collected from the same locality. However, in 10 of 18 putatively clonal groups, some specimens were collected from distant localities (240–44,610 m), including different valleys of the same mountain range and different mountain ranges in a region (Figure 4, Table S3). Two groups included putatively clonal specimens that were collected in two different years and at large distances from each other (44,610 and 3540 m).

3.3 | Observed and simulated number of unique multilocus genotypes

Five data sets containing information on genetic variants of three independently inherited genes (SSU, EF1A, COI) in specimens of *Ph. albescens* were used to test for the presence of reproductive isolation between different phylogroups. For populations of *Ph. albescens* from the Khibiny Mts. (data set 1, phylogroups Ha, Hb, and Hc) and Spanish Sierra Nevada (data set 2, phylogroups B, D, Fa, G, and M), the observed number of unique multilocus genotypes was outside the distance of three standard deviations (3σ) from the mean number of unique multilocus genotypes for 1000 simulated offspring generations consisting of the same number of individuals as the parent generation (Table S4, Figure S3). In other words, the observed number of combinations of three genetic variants is

FIGURE 4 ML tree based on 25,589 single-nucleotide polymorphisms from the GBS analysis for 108 samples (98 specimens plus technical replicates) of *Ph. albescens*. Phylogroups are highlighted with the same colours as in Figure 2. Labels show specimen numbers with sequence variants of the three analysed genes (same format as in Figure 2b; "-" stands for missing data), followed by figures for the estimated heterozygosity and error rate. Labels of technical replicates are shown in red font. Groups of putatively clonal specimens are marked with grey vertical lines, and geographic distances (minimal and maximal) between the localities where members of a group were collected are shown, with "0 m" denoting specimens collected from the same locality. Putatively clonal groups that include specimens collected in different years are marked with an asterisk. Bootstrap values above 70 are indicated. Fully supported branches (100) are marked with a solid circle
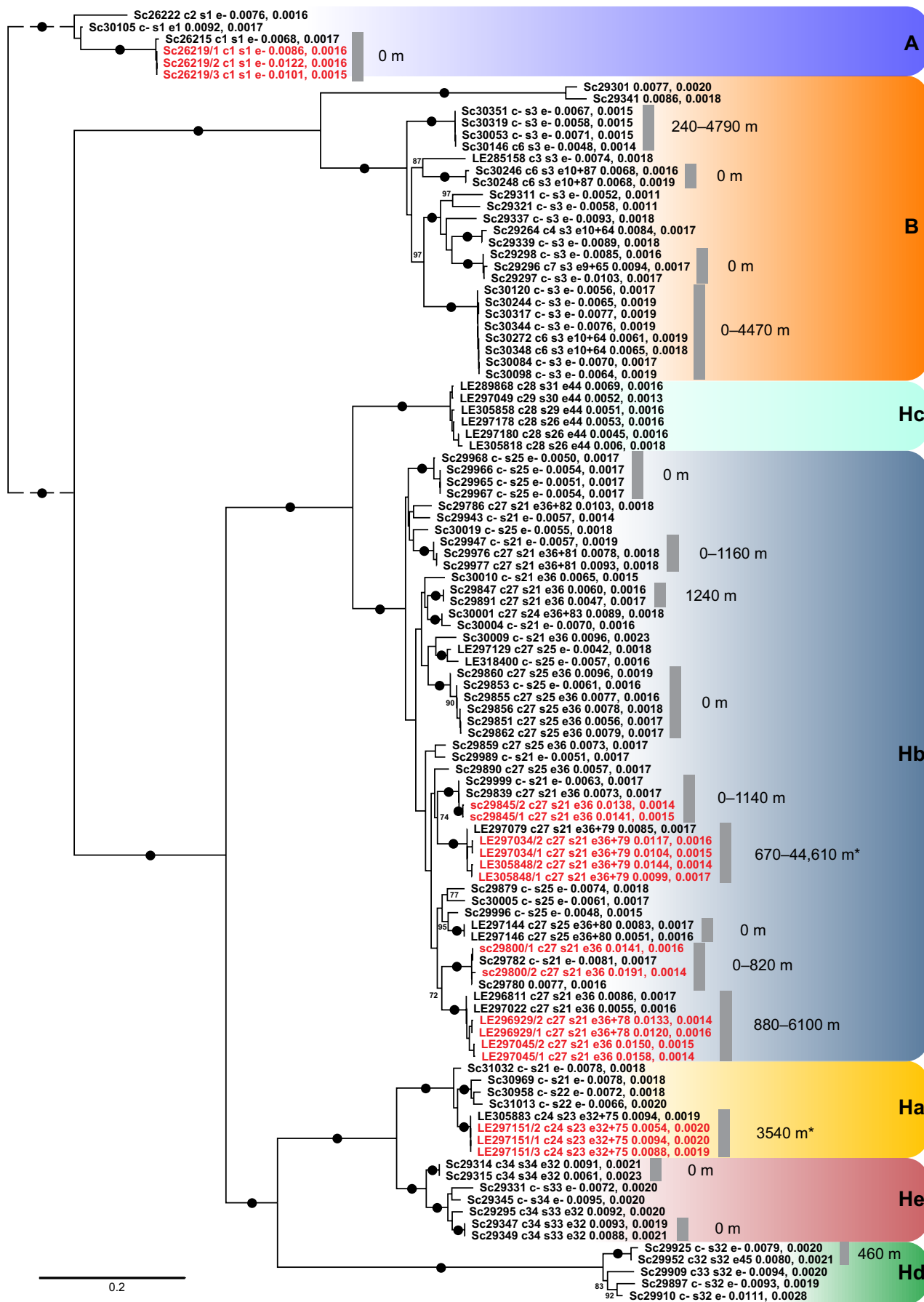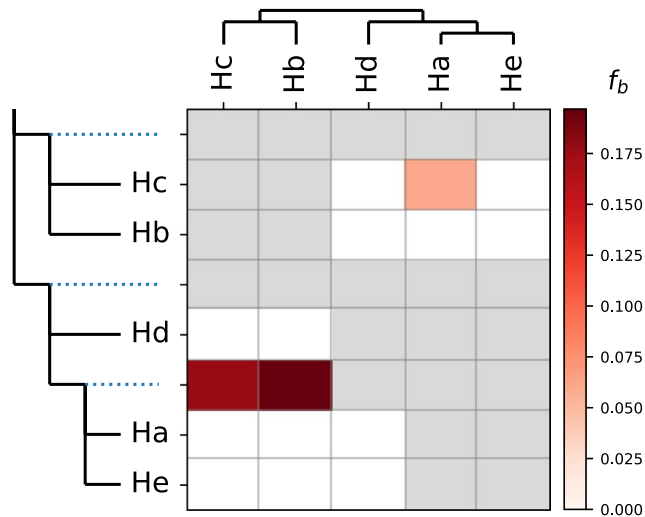
FIGURE 5 Results of *f*-branch inference for five phylogroups of *Ph. albescens* performed in Dsuite. The branch-specific statistic $f_b(C)$ identifies excess sharing of derived alleles between the tree branch on the *y*-axis and the phylogroup *C* on the *x*-axis. Grey data points in the matrix correspond to tests that are not applicable to the provided phylogeny

significantly lower than what would be expected in a randomly sampled panmictic population.

In the case of the data sets 3–5, the simulation was run separately for three phylogroups that were represented by more than 10 specimens and more than three unique multilocus genotypes within a region, assuming that they were reproductively isolated from other phylogroups (phylogroups Hb and Hc from Khibiny Mts., and D from Sierra Nevada). For these data sets, the observed number of combinations of three genetic variants did not differ significantly from the simulation results (Table S4).

When the influence of generation size on the mean genotype number over 1000 simulated generations was tested for the data set from the Sierra Nevada, a strong positive correlation was found (*R* 0.89, *p* < .01) (Figure S4a). However, when generations of different sizes were sub-sampled to the size of the parent generation, this correlation disappeared (*R* −0.02, *p* < .88), and the mean genotype number showed a negligible variation (mean 28.69, SD 0.07) (Figure S4b).

3.4 | Recombination patterns in simulated populations

Each of the 1000 offspring generations simulated under the assumption of panmictic populations for data sets 1 and 2 contained 34%–83% or 50%–94% (for Khibiny Mts. and Sierra Nevada, respectively) of individuals with "mixed" genotypes, consisting of genetic variants from different phylogroups (Table S4). Such "mixed" multi-locus genotypes were completely absent in the input data on *Ph. albescens*.

The observed patterns of combinations of genetic variants of three genes clearly differ from the combinations simulated under the assumption that individuals from different phylogroups recombine freely. As seen from the examples shown in Figure 6, phylogroups cannot be separated in the simulated generations, since COI, SSU, and EF1A variants from different phylogroups combine freely with each other, while this never happens among the combinations observed in the input data.

3.5 | Species-level similarity thresholds for SSU

Clustering SSU sequences at 98% and 99.1% similarity threshold resulted in 10 and 19 clusters, respectively (Table 1). None of the two thresholds allowed to differentiate all phylogroups. Clustering at 98% similarity formed separate clusters only for the phylogroup A and one of the singletons. Six clusters contained sequences from more than one phylogroup. Clustering at 99.1% similarity produced separate clusters for 13 phylogroups and both singletons; only two clusters included more than one phylogroup. In both cases, they united phylogroups that appeared closely related in phylogenies (Fa and Fb; Ha, Hb, and Hc).

3.6 | Regional phylogroup diversity

Diversity estimates for the 20 phylogroups (two represented by singletons only) in the investigated regions revealed a sample coverage of 99.44% (the probability that the next sampled specimen would belong to one of the previously found phylogroups; Chao & Jost, 2012) (Figure 7a, Table S1). The asymptotic estimate of the total phylogroup richness was 21.99 ± 3.73 (for the formula, see Chao et al., 2014). Among the five regions that were sampled most intensively, the sample coverage varied from 94.41% in the Rocky Mts. to 100% in the mountains of the Kola Peninsula (Figure 7b, Table S1). The same two regions showed the highest and the lowest phylogroup richness estimates: for the Rocky Mts., the observed and estimated phylogroup richness was eight and 12.41 ± 7.06, respectively; for the Kola Pen., these numbers were 3 and 3.00 ± 0.02 (Figure 8).

Some phylogroups showed a very restricted distribution in the Northern Hemisphere. For example, the phylogroups C, E, He, I, and J occurred only in the mountains of western North America, Fa – only in Spanish Sierra Nevada, Fb – in European Alps and the Pyrenees, Hc – mountains of the Kola Pen., and Hd – in the mountains of the Kamchatka Pen. (Figure 8).

3.7 | Diversity comparison with other species

Ph. albescens showed the highest observed phylogroup and SSU sequence diversity compared to the other six species that were extensively studied with molecular markers (Figure S6, Table S5). When extrapolated diversity values were compared, *Ph. albescens* still showed the highest expected phylogroup diversity, while
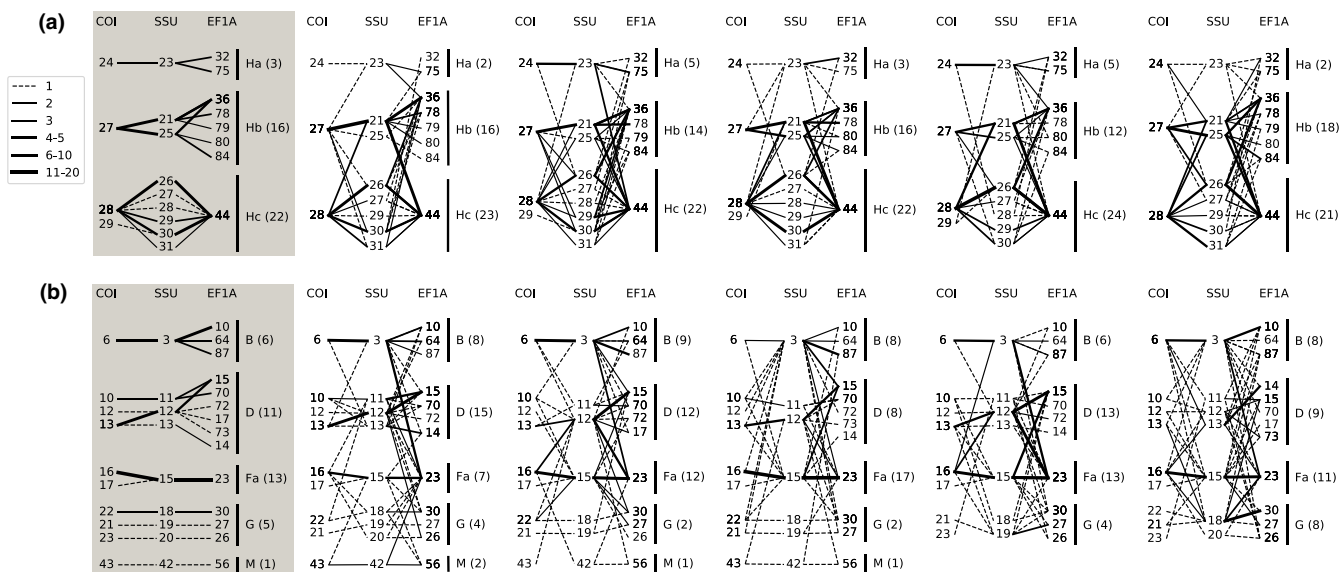
FIGURE 6 Observed and simulated combinations of the variants of three independently inherited genes for two data sets: a – Khibiny Mts., b – Sierra Nevada. The first graph on the left represents combinations observed in the input data (parent generation). The other graphs depict five generations randomly chosen from 1000 simulated offspring generations. In the graphs with simulated data, phylogroups are assigned according to the respective SSU variant. For more explanations, see caption for Figure 2c
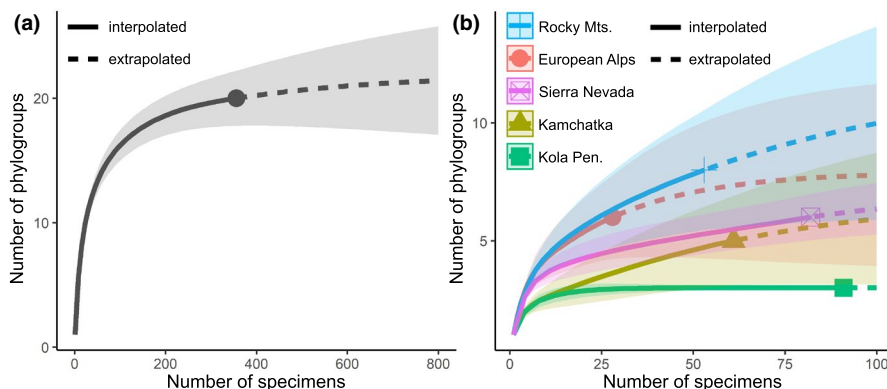

FIGURE 7 Individual-based interpolation (rarefaction) and extrapolation sampling curves for phylogroup richness of the *Ph. albescens*. This graph shows the number of detected (including singletons) and expected phylogroups within the investigated regions. The 95% confidence interval is shown as a transparent area. (a) Curve for the total phylogroup richness. (b) Curves for phylogroup richness in five regions that were sampled most intensively

SSU sequence diversity was expected to be higher in *Hemitrichia serpula*.

3.8 | Small-scale phylogroup distribution

Among the 164 localities of approx. 50 m diam., in 60 localities more than one specimen of *Ph. albescens* was collected. In many cases, specimens from different phylogroups occurred in the same locality: of the 60 localities, 24 (40%) harboured members of two or more different phylogroups (mean 1.53, SD 0.79) (Figure S5). The highest number of phylogroups per locality was found in the Rocky Mts.: up to five in one locality.

4 | DISCUSSION

4.1 | Phylogroups in *Ph. albescens*

We investigated the genetic structure of populations of the nivicolous myxomycete *Physarum albescens* in different mountain ranges of the Northern Hemisphere using three independent genetic markers (SSU, EF1A, COI). With 324 studied specimens, *Ph. albescens* becomes the most extensively sequenced species of myxomycetes to date. The results of our study show an extraordinarily high intraspecific phylogroup diversity that largely exceeds the observed and expected phylogroup diversity in other studied myxomycete species (Figure S6, Table S5). Our main finding is that *Ph. albescens*
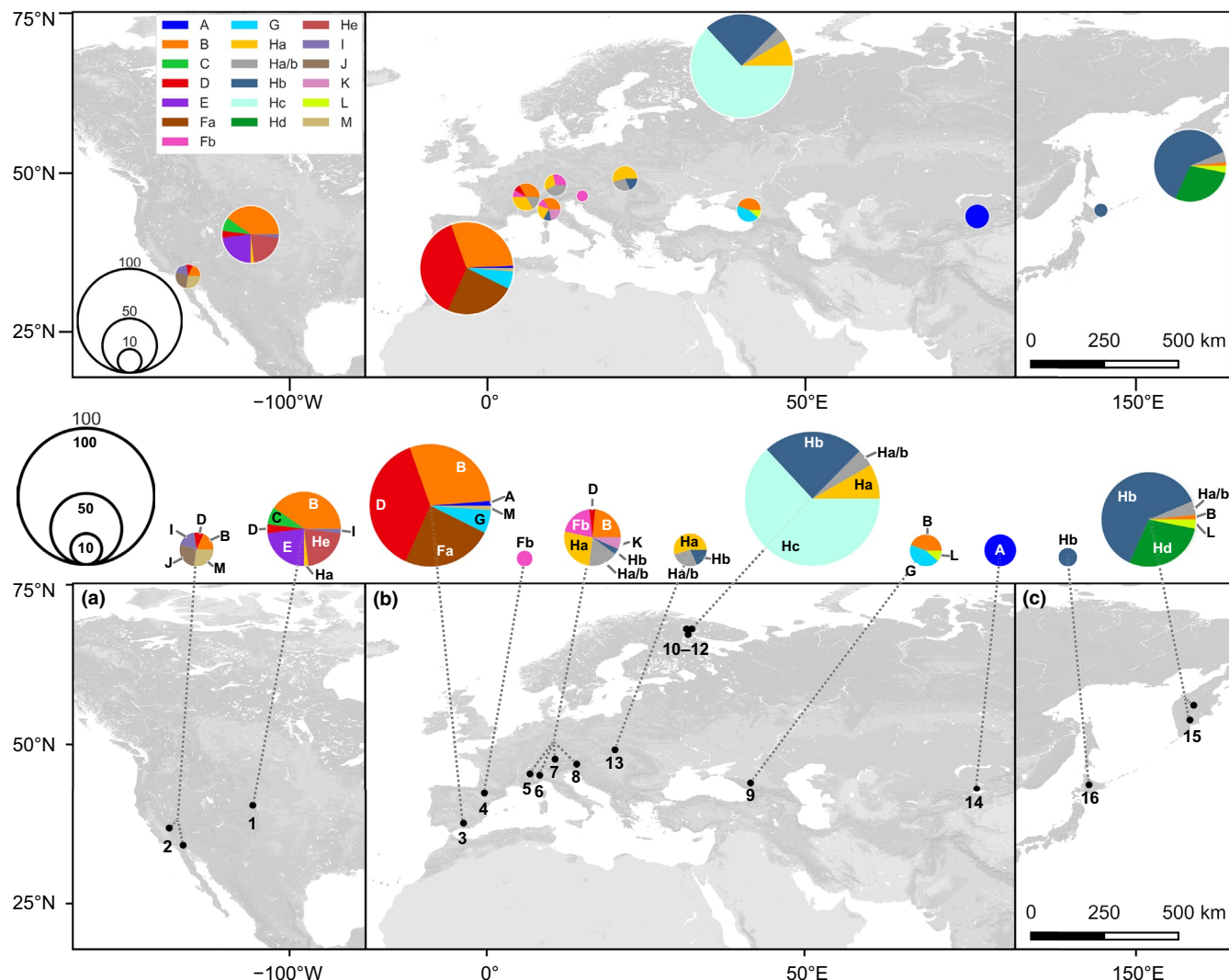
FIGURE 8 Regions sampled for *Ph. albescens*. The circle diameter reflects the number of studied specimens collected from a region. The area of the coloured sectors shows proportions of phylogenetic groups among the specimens collected from a region (see details in the Results section). Specimens sequenced only for SSU and showing sequence variant s21 shared by Ha and Hb are shown in grey. (a) North America (1, Rocky Mts., 2, Angeles National Forest, San Bernardino National Forest, San Jasinto Mt., and Sequoia National Park). (b) Europe (3, Sierra Nevada Mts; 4, French Pyrenees; 5, French Alps; 6, Italian Alps; 7, German Alps; 8, Slovenian Alps; 9, Northern Caucasus; 10–12, Kola Peninsula: Khibiny Mts., Chunatundra Mts., Luvenga Tundra Mts., 13, Western Carpathians) and Central Asia (14, Ile Alatau Mts.). (c) East Asia (15, Kamchatka; 16, Hokkaido). Maps were produced with QGIS 3.12.1 and Mapbox online service

represents not a single species, but a species complex, since this morphologically defined species splits into at least 18 phylogroups that can be considered as cryptic species. According to the extrapolated diversity estimates, even more phylogroups could be found with an extended sampling. Previous studies reported two to seven large phylogroups (putative cryptic species) per morphospecies in other complexes of myxomycete species (Aguilar et al., 2014; Dagamac et al., 2017; Feng et al., 2016; Feng & Schnittler, 2015; Janik et al., 2020; Novozhilov et al., 2013; Shchepin et al., 2016).

The three-gene data set used here included 302 parsimony-informative variable sites, which allowed to construct a robust phylogeny and a phylogenetic network that, in combination with the analysis of ml-FFRs, showed the separation into 18 phylogroups

(Figures 2b and 3a). Two automatic species delimitation algorithms applied to the three-gene data set (mPTP and ASAP) agreed in delimitation of 13 phylogroups and suggested uniting some of the closely related phylogroups together (Table 1), partly supporting our manual phylogroup delimitation. High-resolution GBS data that were obtained for a subset of samples from seven phylogroups further supported the separation revealed from the three-gene data. Judging by the topology of the phylogenetic tree and network (Figures 3b and 4) and mPTP results for the GBS data (Table 1), it is possible that some of these phylogroups split into even more cryptic species (for example, phylogroup B). We applied here a conservative approach and tried to keep the number of recognized phylogroups to a minimum.

4.2 | Reproductive isolation

In all intensively sampled mountain ranges, members of several different phylogroups of *Ph. albescens* occur together, and even within one locality, two or more phylogroups were often found. However, even though members of different phylogroups may share habitats (Figure S5), they do not exchange genetic information, as seen from the observed combinations of genetic variants in 214 sampled colonies of sporocarps (Figure 2c). This observation supports the hypothesis of reproductive isolation between phylogroups. We consider incomplete lineage sorting or hybridization to be the most probable explanation for the two exceptions (one SSU variant shared by Ha and Hb and one EF1A variant shared by Ha and He). There are three reasons for this: (1) in both cases, these variants are shared by two closely related phylogroups separated by short branches, which is indicative of a recent speciation event, (2) the separation of the phylogroups Ha, Hb, and He is strongly supported by the SNP phylogenies (Figures 3b and 4), and (3) ABBA-BABA tests and *f*-branch statistics indicate a possible gene flow event(s) between the discussed phylogroups. The fact that only two of the eight branches in the input tree for Fbranch showed an excess of shared derived alleles with other branches supports the idea that gene flow is generally restricted even between the most closely related phylogroups.

The computer simulation results showed that the observed numbers of combinations of three genetic variants in populations of *Ph. albescens* from neighbouring valleys in the Khibiny Mts. and the Sierra Nevada were significantly lower than what would be expected for a single randomly sampled panmictic population. Moreover, among the individuals sampled from each of the 1000 replicates of the simulated panmictic populations, 34%–94% of individuals showed "mixed" genotypes consisting of genetic variants from different phylogroups, while such genotypes did not occur in the individuals sampled from the real populations, although the sample size was the same.

Based on the observations listed above and the simulation results, we conclude that the revealed phylogroups of *Ph. albescens* are reproductively isolated from each other, although rare hybridization events between recently diverged phylogroups cannot be excluded. As such, they can be seen as biological species in Mayr's strict sense (1942). In this case, another biological feature, the monophyly of the respective phylogroups, supports this (compare de Queiroz, 2005), but we do not expect to find consistent morphological differences between all phylogroups. The frequent coexistence of different phylogroups (cryptic species) in one region can be explained by sympatric speciation caused by mutations in the mating-type systems or polyploidisation (see review in Clark & Haskins, 2010). Alternatively, the observed phylogroup distribution could have resulted from an allopatric or parapatric speciation followed by range expansion.

4.3 | Mode of reproduction

The reproduction mode of myxomycetes in natural populations has long been in question. Amoebozoa are principally capable of sexual reproduction (Lahr et al., 2011), but amoebae reproduce as well asexually by cell fission. Although myxomycetes are considered to be predominantly heterothallic sexual organisms, in vitro experiments with several easily cultivable species have shown that at least in culture exceptions do often occur (see discussion in Feng et al., 2016; Walker & Stephenson, 2016).

From the observed pattern of combinations of genetic variants within phylogroups in the current study and the high number of three-marker genotypes found, we can assume that in most lineages from time to time sexual recombination does occur. However, the GBS data show a considerable percentage of specimens of putatively clonal origin, forming 18 putatively clonal groups within six of the seven genotyped phylogroups (Figure 4). The most obvious explanation for this could be the fragmentation of a single plasmodium, with different parts of it moving in different directions and forming genetically identical colonies of sporocarps. However, in 10 of 18 groups, some specimens showing nearly identical genotypes (genetic distance within the range of technical errors) were found on a large distance from the others (240–44,610 m) (Figure 4, Table S3). Moreover, in two cases, members of a putatively clonal group were collected in two different years (2013 and 2015). In such cases, the genotypic identity of different specimens cannot be explained by fragmentation of a single plasmodium, as it can be done for the specimens found in one season in a few metres from each other. In the Sierra Nevada and Khibiny Mts., members of some putatively clonal groups were found in different valleys, so their distribution cannot be explained by a passive drift of pieces of plasmodia with streams of water running down from melting snow patches. This could be a sign of asexual lineages dispersing via airborne spores, as suggested for *Didymium nivicola* (Janik et al., 2020). A second possible explanation (as well speculative) could be that these genetically nearly identical specimens result from repeated mating of a few dominating myxamoebal strains that are widely distributed within a large area and persist from year to year.

Our results demonstrate that sexual and asexual reproduction coexist in natural populations of *Ph. albescens*. Despite these intriguing results, the proportions of sexual and asexual reproduction remain unclear. Crossing experiments or studies of ploidy levels at different life cycle stages could shed light on this question. However, these methods require in vitro cultivation of *Ph. albescens* at least from spore to plasmodium, which we failed to achieve (Shchepin et al., 2014).

4.4 | Geographical distribution

The cryptic species in the *Ph. albescens* complex show well-pronounced phylogeographical patterns (Figure 8). Most of the 18 cryptic species have a limited geographical distribution, and none was found in all of the studied regions. Moreover, some phylogroups were found to be abundant in only one region and totally absent in all other studied regions and thus may represent local endemics. Nevertheless, in most studied regions, several phylogroups coexist

and are often found in one habitat. The observed geographical distribution may be at least partially due to the niche differentiation among the members of the *Ph. albescens* complex (hypothesis not tested herein). The presence of niche differentiation between closely related cryptic species was demonstrated previously for *Badhamia melanospora* and *Hemitrichia serpula* using environmental niche modeling (Aguilar et al., 2014; Dagamac et al., 2017).

While reports of *Ph. albescens* from the Southern Hemisphere are rare, the largest observed and predicted phylogroup diversity in the Northern Hemisphere is recorded from the Rocky Mts., which may represent the centre of diversity for this species complex. However, since other mountains in North America are sampled poorly or not sampled at all, the real centre of diversity can be somewhere else on this continent.

4.5 | Resolution of SSU barcoding

Species determination based on the similarity of short nucleotide sequences to references (DNA barcoding) is becoming a widely used method in myxomycete research. Here, with only one exception (SSU variant 21 shared by Ha and Hb), a partial SSU sequence is sufficient to unambiguously classify a herbarium specimen of *Ph. albescens* as belonging to a particular phylogroup. However, in environmental DNA studies, sequences are usually clustered prior to taxonomic assignment to reduce sequencing noise. Several different similarity thresholds were applied previously for the determination of myxomycete species in SSU-based DNA barcoding studies and clustering OTUs in metabarcoding studies: (1) 99.1% (Borg Dahl et al., 2019; Borg Dahl, Brejnrod, et al., 2018; Borg Dahl, Shchepin, et al., 2018; Shchepin et al., 2017, 2019), (2) 98% (Clissmann et al., 2015; Gao et al., 2019; Kamono et al., 2013; Shchepin et al., 2019), (3) 97% (Fiore-Donno et al., 2016), and (4) 96% (Kamono et al., 2013). We have evaluated the ability of the two more popular similarity thresholds (98% and 99.1%) to differentiate between the cryptic species within *Ph. albescens*.

With a 98% similarity threshold, the species complex split into ten clusters, six of them containing sequences of more than one phylogroup. The 99.1% similarity threshold showed better results: 13 cryptic species and both singleton phylogroups could be separated, and only two clusters included more than one phylogroup (in both cases, they united closely related phylogroups within clades F and H).

However, none of the two thresholds can differentiate all 18 cryptic species. This can present a serious limitation for the barcoding-based identification of environmental sequences due to two reasons: (1) in DNA metabarcoding studies, usually only one genetic marker is sequenced. Thus, the taxonomic resolution cannot be increased by the analysis of additional genes; even if two or more different gene markers are sequenced from the same environmental sample, it can often be impossible to show that two sequences of different genes originate from the same species; (2) environmental sequencing data are inherently noisy due to PCR artifacts and high

error rates of the high-throughput sequencing platforms, thus clustering at thresholds higher than 98% may lead to a high number of erroneous OTUs. In comparison, lower thresholds would lump even more species together into one OTU. In a DNA metabarcoding study based on OTUs, the phylogroup Hc from the Kola Peninsula would not differ from Hb distributed over Eurasia and from Ha, which occurs in Europe and North America; Fa from Sierra Nevada would form one OTU with Fb from European Alps and Pyrenees. A possible solution could be the use of amplicon sequence variants (ASVs) instead of OTUs. ASVs represent every unique biological sequence in the analysed sample (Callahan et al., 2017). This approach could increase the taxonomic resolution of the amplicon-based eDNA studies, but very stringent data filtering is required to reduce sequencing noise. However, even the use of ASVs in some cases would not help to differentiate between SSU sequences of the phylogroups Ha and Hb due to the presence of at least one shared SSU sequence variant. In contrast, COI sequences are unique for each of the 18 phylogroups and more genetically diverse, and could potentially serve as DNA barcodes with better resolution than SSU sequences. COI has been promoted as a barcode for metazoans since the very appearance of the concept of DNA barcoding (Hebert et al., 2003). Later this gene was shown to be a more suitable barcode than SSU for different groups of protists, including some members of Amoebozoa (Nassonova et al., 2010; Zlatogursky et al., 2016) and SAR (Heger et al., 2011; Zhao et al., 2016). Taxonomic resolution and convenience of COI as a DNA barcode for myxomycetes needs to be tested on a broader taxon sample.

5 | CONCLUSIONS

Although the nivicolous myxomycete *Ph. albescens* seems to have a cosmopolitan distribution over the mountains of the Northern Hemisphere, our results demonstrate that it represents a complex of at least 18 cryptic species, each of them showing a much narrower distribution than the morphospecies as a whole. This again challenges the ubiquity hypothesis for myxomycetes and supports the moderate endemicity hypothesis. Despite the fact that members of different phylogroups co-occur in most of the studied regions and many microhabitats, our data suggest that introgression events between phylogroups, if they happen, should be rare and occur mostly between recently diverged phylogroups.

On the one hand, such patterns of multiple cryptic species within morphospecies seem to be common in myxomycetes and may represent a general mode of evolution in this group: all hitherto investigated morphospecies turned out to be complexes of several cryptic species (Aguilar et al., 2014; Dagamac et al., 2017; Feng & Schnittler, 2015; Janik et al., 2020; Leontyev et al., 2015; Shchepin et al., 2016). This means that we should expect the real number of myxomycete species to be several times bigger than c. 1050 currently recognized species (Lado, 2005–2021). Further efforts should be made to identify cryptic species within other myxomycete species and their geographical distribution, to find new morphological

traits that could help to differentiate between them, and to test for the presence of niche differentiation between members of cryptic species complexes. This data would allow us to better estimate the total expected myxomycete species diversity and to re-evaluate our knowledge on ecology and distribution of some presumably ubiquitous morphospecies.

On the other hand, the presence of groups of presumably clonal specimens together with a high genotypic diversity within phylogroups suggests that sexual and asexual reproduction coexist in natural populations of *Ph. albescens*. This conclusion is consistent with the results of crossing experiments conducted for other species of myxomycetes, which revealed a frequent occurrence of apomictic strains, including facultatively apomictic ones (Clark & Haskins, 2013). If the presence of apomictic strains among sexual strains is usual in myxomycetes, this makes species delimitation and interpretatin of species-level ecological data even more complicated.

In addition, it should be pointed out that SSU barcode proved to lack resolution in some cases to differentiate some closely related phylogroups in *Ph. albescens*. Since COI showed a better resolution in currrent study, it should be evaluated on a broader sample of taxa as a possibly better DNA barcode for myxomycetes.

The large cryptic diversity and biogeographical patterns revealed for *Ph. albescens* are consistent with the presence of multiple cryptic species with restricted distribution found for other soil-inhabiting protists as different as diatoms, streptophyte algae, and testate amoebae (Pinseel et al., 2020; Ryšánek et al., 2015; Singer et al., 2019). It seems possible that these patterns of diversity and distribution can be universal for soil-inhabiting protists. Further research on soil protist diversity and biogeography is essential to estimate the number of endemic protist species and their ability to persist in changing environments.

AUTHOR CONTRIBUTIONS

OS, YN and MS designed the study. OS, YN, JW, IP, NF, VG, MBD, NHAD, YY and MS collected samples. OS, MBD, NHAD, MS, IP and NF acquired sequencing data. OS, MB and NHAD prepared samples for GBS. OS performed all bioinformatic analyses. OS, YN and MS wrote the manuscript. All authors edited and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The GBS data: NCBI Sequence Read Archive BioProject PRJNA706537 (Shchepin et al., 2021). DNA sequences: GenBank accessions MW691477–MW691847, MW692988–MW693025, MW701443–MW701877. Sequence alignments are included in the supplementary materials (Additional files 1–4). Information on studied samples and collection sites is included in the supplementary materials (Additional file 6). Source code for the program visualizing recombination patterns: https://github.com/Gurdhhu/recombination_graph. Source code for the computer simulation: https://github.com/Gurdhhu/recombination_simulation.

ORCID

Oleg Shchepin https://orcid.org/0000-0001-9327-7655

REFERENCES

Aguilar, M., Fiore-Donno, A. M., Lado, C., & Cavalier-Smith, T. (2014). Using environmental niche models to test the 'everything is everywhere' hypothesis for *Badhamia*. *The ISME Journal*, *8*, 737–745. https://doi.org/10.1038/ismej.2013.183

Borg Dahl, M., Brejnrod, A. D., Russel, J., Sørensen, S. J., & Schnittler, M. (2019). Different degrees of niche differentiation for bacteria, fungi, and myxomycetes within an elevational transect in the German Alps. *Microbial Ecology*, *78*(3), 764–780. https://doi.org/10.1007/s00248-019-01347-1

Borg Dahl, M., Brejnrod, A. D., Unterseher, M., Hoppe, T., Feng, Y., Novozhilov, Y., Sørensen, S. J., & Schnittler, M. (2018). Genetic barcoding of dark-spored myxomycetes (Amoebozoa) – Identification, evaluation and application of a sequence similarity threshold for species differentiation in NGS studies. *Molecular Ecology Resources*, *18*(2), 306–318. https://doi.org/10.1111/1755-0998.12725

Borg Dahl, M., Shchepin, O. N., Schunk, C., Menzel, A., Novozhilov, Y. K., & Schnittler, M. (2018). A four year survey reveals a coherent pattern between occurrence of fruit bodies and soil amoebae populations for nivicolous myxomycetes. *Scientific Reports*, *8*, 11662. https://doi.org/10.1038/s41598-018-30131-3

Bryant, D., & Moulton, V. (2004). Neighbor-net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, *21*(2), 255–265. https://doi.org/10.1093/molbev/msh018

Callahan, B., McMurdie, P., & Holmes, S. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*, 2639–2643. https://doi.org/10.1038/ismej.2017.119

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species

diversity studies. *Ecological Monographs*, *84*(1), 45–67. https://doi.org/10.1890/13-0133.1

Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology*, *93*(12), 2533–2547. https://doi.org/10.1890/11-1952.1

Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/sysbio/syw037

Clark, J., & Haskins, E. F. (2010). Reproductive systems in the myxomycetes: A review. *Mycosphere*, *1*, 337–353.

Clark, J., & Haskins, E. F. (2013). The nuclear reproductive cycle in the myxomycetes: A review. *Mycosphere*, *4*, 233–248. https://doi.org/10.5943/mycosphere/4/2/6

Clissmann, F., Fiore-Donno, A. M., Hoppe, B., Krüger, D., Kahl, T., Unterseher, M., & Schnittler, M. (2015). First insight into dead wood protistean diversity: A molecular sampling of bright-spored Myxomycetes (Amoebozoa, slime moulds) in decaying beech logs. *FEMS Microbiology Ecology*, *91*(6), fiv050.

Dagamac, N. H. A., Rojas, C., Novozhilov, Y. K., Moreno, G. H., Schlueter, R., & Schnittler, M. (2017). Speciation in progress? A phylogeographic study among populations of *Hemitrichia serpula* (Myxomycetes). *PLoS One*, *12*(4), e0174825. https://doi.org/10.1371/journal.pone.0174825

De Queiroz, K. (2005). Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences*, *102*, 6600–6607. https://doi.org/10.1073/pnas.0502030102

Doyle, J. J. (1995). The irrelevance of allele tree topologies for species delimitation, and a non-topological alternative. *Systematic Botany*, *20*, 574–588. https://doi.org/10.2307/2419811

Dress, A. W. M., & Huson, D. H. (2004). Constructing splits graphs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *1*(3), 109–115. https://doi.org/10.1109/TCBB.2004.27

Eaton, D., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics (Oxford, England)*, *36*(8), 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Feng, Y., Klahr, A., Janik, P., Ronikier, A., Hoppe, T., Novozhilov, Y. K., & Schnittler, M. (2016). What an intron may tell: Several sexual biospecies coexist in *Meriderma* spp. (Myxomycetes). *Protist*, *167*(3), 234–253. https://doi.org/10.1016/j.protis.2016.03.003

Feng, Y., & Schnittler, M. (2015). Sex or no sex? Independent marker genes and group I introns reveal the existence of three sexual but reproductively isolated biospecies in *Trichia varia* (Myxomycetes). *Organisms Diversity & Evolution*, *15*, 631–650.

Ferris, P. J., Vogt, V. M., & Truitt, C. L. (1983). Inheritance of extrachromosomal rDNA in *Physarum polycephalum*. *Molecular and Cellular Biology*, *3*(4), 635–642. https://doi.org/10.1128/mcb.3.4.635

Fiore-Donno, A. M., Kamono, A., Meyer, M., Schnittler, M., Fukui, M., & Cavalier-Smith, T. (2012). 18S rDNA phylogeny of *Lamproderma* and allied genera (Stemonitidales, Myxomycetes, Amoebozoa). *PLoS One*, *7*(4), e35359.

Fiore-Donno, A. M., Meyer, M., Baldauf, S. L., & Pawlowski, J. (2008). Evolution of dark-spored Myxomycetes (slime molds): Molecules versus morphology. *Molecular Phylogenetics and Evolution*, *46*, 878–889. https://doi.org/10.1016/j.ympev.2007.12.011

Fiore-Donno, A. M., Weinert, J., Wubet, T., & Bonkowski, M. (2016). Metacommunity analysis of amoeboid protists in grassland soils. *Scientific Reports*, *11*(6), 19068. https://doi.org/10.1038/srep19068

Flot, J. F., Couloux, A., & Tillier, S. (2010). Haplowebs as a graphical tool for delimiting species: A revival of Doyle's "field for recombination" approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evolutionary Biology*, *10*, 372. https://doi.org/10.1186/1471-2148-10-372

Foissner, W. (2008). Protist diversity and distribution: Some basic considerations. *Biodiversity and Conservation*, *17*, 235–242. https://doi.org/10.1007/s10531-007-9248-5

Gao, Y., Zhang, X., He, G., Shchepin, O. N., Yan, S., & Chen, S. (2019). Influence of forest type on dark-spored myxomycete community in subtropical forest soil, China. *Soil Biology and Biochemistry*, *139*, 107606. https://doi.org/10.1016/j.soilbio.2019.107606

Geisen, S., Lara, E., Mitchell, E. A., Völcker, E., & Krashevska, V. (2020). Soil protist life matters! *Soil Organisms*, *92*(3), 189–196. https://doi.org/10.25674/so92iss3pp189

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.

Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Technical Journal*, *29*(2), 147–160. https://doi.org/10.1002/j.1538-7305.1950.tb00463.x

Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *270*(1512), 313–321.

Heger, T. J., Pawlowski, J., Lara, E., Leander, B. S., Todorov, M., Golemansky, V., & Mitchell, E. A. (2011). Comparing potential COI and SSU rDNA barcodes for assessing the diversity and phylogenetic relationships of cyphoderiid testate amoebae (Rhizaria: Euglyphida). *Protist*, *162*(1), 131–141. https://doi.org/10.1016/j.protis.2010.05.002

Hoppe, T., & Schnittler, M. (2015). Characterization of myxomycetes in two different soils by TRFLP analysis of partial 18S rRNA gene sequences. *Mycosphere*, *6*(2), 216–227. https://doi.org/10.5943/mycosphere/6/2/11

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*, 1451–1456.

Huelsenbeck, J. P., & Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*(8), 754–755. https://doi.org/10.1093/bioinformatics/17.8.754

Huson, D. H. (1998). SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics (Oxford, England)*, *14*(1), 68–73. https://doi.org/10.1093/bioinformatics/14.1.68

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Janik, P., Lado, C., & Ronikier, A. (2020). Range-wide phylogeography of a nivicolous protist *Didymium nivicola* Meyl. (Myxomycetes, Amoebozoa): Striking contrasts between the Northern and the Southern Hemisphere. *Protist*, *171*, 125771.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Kamono, A., Kojima, H., Matsumoto, J., Kawamura, K., & Fukui, M. (2009). Airborne myxomycete spores: Detection using molecular techniques. *Naturwissenschaften*, *96*, 147–151. https://doi.org/10.1007/s00114-008-0454-0

Kamono, A., Meyer, M., Cavalier-Smith, T., Fukui, M., & Fiore-Donno, A. M. (2013). Exploring slime mould diversity in high-altitude forests and grasslands by environmental RNA analysis. *FEMS Microbiology Ecology*, *94*, 98–109. https://doi.org/10.1111/1574-6941.12042

Kapli, T., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, *33*(11), 1630–1638. https://doi.org/10.1093/bioinformatics/btx025

Katoh, K., Kuma, K., Toh, H., & Miyata, T. (2005). MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, *33*, 511–518. https://doi.org/10.1093/nar/gki198

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and

visualization. *Briefings in Bioinformatics*, 20(4), 1160–1166. https://doi.org/10.1093/bib/bbx108

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547–1549. https://doi.org/10.1093/molbev/msy096

Lado, C. (2005–2021). *NomenMyx: An on line nomenclatural information system of Eumycetozoa* [Electronic database]. Retrieved from http://www.nomen.eumycetozoa.com

Lahr, D. J. G., Parfrey, L. W., Mitchell, E. A., Katz, L. A., & Lara, E. (2011). The chastity of amoebae: Re-evaluating evidence for sex in amoeboid organisms. *Proceedings of the Royal Society B*, 278, 2081–2090. https://doi.org/10.1098/rspb.2011.0289

Leontyev, D., Schnittler, M., & Stephenson, S. L. (2015). A critical revision of the *Tubifera ferruginosa*-complex. *Mycologia*, 107, 959–985.

Liu, Q. S., Yan, S. Z., & Chen, S. L. (2015). Further resolving the phylogeny of Myxogastria (slime molds) based on COI and SSU rRNA genes. *Russian Journal of Genetics*, 51, 46–53. https://doi.org/10.1134/S1022795414110076

Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, 21(2), 584–595. https://doi.org/10.1111/1755-0998.13265

Malinsky, M., Svardal, H., Tyers, A. M., Miska, E. A., Genner, M. J., Turner, G. F., & Durbin, R. (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecology & Evolution*, 2(12), 1940–1955. https://doi.org/10.1038/s41559-018-0717-x

Mann, D. G., & Vanormelingen, P. (2013). An inordinate fondness? The number, distributions, and origins of diatom species. *The Journal of Eukaryotic Microbiology*, 60(4), 414–420. https://doi.org/10.1111/jeu.12047

Mayr, E. (1942). *Systematics and the origin of species*. Colombia University Press.

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

Minh, B. Q., Nguyen, M. A., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 30(5), 1188–1195. https://doi.org/10.1093/molbev/mst024

Nassonova, E., Smirnov, A., Fahrni, J., & Pawlowski, J. (2010). Barcoding amoebae: Comparison of SSU, ITS and COI genes as tools for molecular identification of naked lobose amoebae. *Protist*, 161(1), 102–115. https://doi.org/10.1016/j.protis.2009.07.003

Novozhilov, Y. K., Mitchell, D. W., Okun, M. V., & Shchepin, O. N. (2014). New species of *Diderma* from Vietnam. *Mycosphere*, 5, 554–564. https://doi.org/10.5943/mycosphere/5/4/8

Novozhilov, Y. K., Okun, M. V., Erastova, D. A., Shchepin, O. N., Zemlyanskaya, I. V., García-Carvajal, E., & Schnittler, M. (2013). Description, culture and phylogenetic position of a new xerotolerant species of *Physarum*. *Mycologia*, 105(6), 1535–1546. https://doi.org/10.3852/12-284

Novozhilov, Y. K., Prikhodko, I. S., & Shchepin, O. N. (2019). A new species of *Diderma* from Bidoup Nui Ba national park (southern Vietnam). *Protistology*, 13, 126–132. https://doi.org/10.21685/1680-0826-2019-13-3-2

Pinseel, E., Janssens, S. B., Verleyen, E., Vanormelingen, P., Kohler, T. J., Biersma, E. M., Sabbe, K., Van de Vijver, B., & Vyverman, W. (2020). Global radiation in a rare biosphere soil diatom. *Nature Communications*, 11(1), 2382. https://doi.org/10.1038/s41467-020-16181-0

Puillandre, N., Brouillet, S., & Achaz, G. (2021). ASAP: Assemble species by automatic partitioning. *Molecular Ecology Resources*, 21(2), 609–620. https://doi.org/10.1111/1755-0998.13281

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584. https://doi.org/10.7717/peerj.2584

Ronikier, A., & Ronikier, M. (2009). How 'alpine' are nivicolous myxomycetes? A worldwide assessment of altitudinal distribution. *Mycologia*, 101(1), 1–16. https://doi.org/10.3852/08-090

Ryšánek, D., Hrčková, K., & Škaloud, P. (2015). Global ubiquity and local endemism of free-living terrestrial protists: Phylogeographic assessment of the streptophyte alga *Klebsormidium*. *Environmental Microbiology*, 17(3), 689–698.

Schnittler, M., Dagamac, N. H. A., Woyzichovski, J., & Novozhilov, Y. K. (2021). Biogeographical patterns in myxomycetes. In S. L. Stephenson, & C. Rojas (Eds.), *Myxomycetes. Biology, systematics, biogeography and ecology* (2nd ed., pp. 377–416). Elsevier, Academic Press.

Schnittler, M., Erastova, D. A., Shchepin, O. N., Heinrich, E., & Novozhilov, Y. K. (2015). Four years in the Caucasus – Observations on the ecology of nivicolous myxomycetes. *Fungal Ecology*, 14, 105–115. https://doi.org/10.1016/j.funeco.2015.01.003

Schnittler, M., Shchepin, O. N., Dagamac, N. H. A., Borg Dahl, M., & Novozhilov, Y. K. (2017). Barcoding myxomycetes with molecular markers: Challenges and opportunities. *Nova Hedwigia*, 104, 323–341. https://doi.org/10.1127/nova_hedwigia/2017/0397

Shchepin, O. N. (2021a, March 21). Visualization of recombination patterns [Source code]. https://github.com/Gurdhhu/recombination_graph

Shchepin, O. N. (2021b, March 21). Recombination simulation [Source code]. https://github.com/Gurdhhu/recombination_simulation

Shchepin, O. N., Dagamac, N. H., Sanchez, O. M., Novozhilov, Y. K., Schnittler, M., & Zemlyanskaya, I. V. (2017). DNA barcoding as a tool for identification of plasmodia and sclerotia of myxomycetes (Myxogastria) appearing in moist chamber cultures. *Mycosphere*, 8(10), 1904–1913. https://doi.org/10.5943/mycosphere/8/10/13

Shchepin, O. N., Novozhilov, Y. K., & Schnittler, M. (2014). Nivicolous myxomycetes in agar culture: Some results and open problems. *Protistology*, 8(2), 53–61.

Shchepin, O. N., Novozhilov, Y. K., & Schnittler, M. (2016). Disentangling the taxonomic structure of the *Lepidoderma chailletii-carestianum* species complex (Myxogastria, Amoebozoa): Genetic and morphological aspects. *Protistology*, 10(4), 117–129. https://doi.org/10.21685/1680-0826-2016-10-4-1

Shchepin, O. N., Novozhilov, Y. K., Woyzichovski, J., Bog, M., Prikhodko, I. S., Fedorova, N. A., & Schnittler, M. (2021). GBS data for herbarium specimens of a myxomycete *Physarum albescens*. *NCBI Sequence Read Archive. BioProject* PRJNA706537.

Shchepin, O. N., Schnittler, M., Erastova, D. A., Prikhodko, I. S., Borg Dahl, M., Azarov, D. V., Chernyaeva, E. N., & Novozhilov, Y. K. (2019). Community of dark-spored myxomycetes in ground litter and soil of taiga forest (Nizhne-Svirskiy Reserve, Russia) revealed by DNA metabarcoding. *Fungal Ecology*, 39, 80–93. https://doi.org/10.1016/j.funeco.2018.11.006

Silliker, M. E., & Collins, O. R. (1988). Non-mendelian inheritance of mitochondrial DNA and ribosomal DNA in the myxomycete, *Didymium iridis*. *Molecular & General Genetics: MGG*, 213(2–3), 370–378. https://doi.org/10.1007/BF00339605

Silliker, M. A., Liles, J. L., & Monroe, J. A. (2002). Patterns of mitochondrial inheritance in the myxogastrid, *Didymium iridis*. *Mycologia*, 94, 939–946. https://doi.org/10.2307/3761862

Singer, D., Mitchell, E. A. D., Payne, R. J., Blandenier, Q., Duckert, C., Fernández, L. D., Fournier, B., Hernández, C. E., Granath, G., Rydin, H., Bragazza, L., Koronatova, N. G., Goia, I., Harris, L. I., Kajukało, K., Kosakyan, A., Lamentowicz, M., Kosykh, N. P., Vellak, K., & Lara, E. (2019). Dispersal limitations and historical factors determine the biogeography of specialized terrestrial

protists. *Molecular Ecology*, *28*(12), 3089–3100. https://doi.org/10.1111/mec.15117

Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, *44*(W1), W232–W235. https://doi.org/10.1093/nar/gkw256

Walker, L. M., & Stephenson, S. L. (2016). The species problem in myxomycetes revisited. *Protist*, *167*(4), 319–338. https://doi.org/10.1016/j.protis.2016.05.003

Warren, D. L., Geneva, A. J., & Lanfear, R. (2017). RWTY (R We There Yet): An R package for examining convergence of Bayesian phylogenetic analyses. *Molecular Biology and Evolution*, *34*(4), 1016–1020. https://doi.org/10.1093/molbev/msw279

Zhao, Y., Yi, Z., Gentekaki, E., Zhan, A., Al-Farraj, S. A., & Song, W. B. (2016). Utility of combining morphological characters, nuclear and mitochondrial genes: An attempt to resolve the conflicts of species identification for ciliated protists. *Molecular Phylogenetics and Evolution*, *94*(Pt B), 718–729. https://doi.org/10.1016/j.ympev.2015.10.017

Zlatogursky, V. V., Kudryavtsev, A., Udalov, I. A., Bondarenko, N., Pawlowski, J., & Smirnov, A. (2016). Genetic structure of a morphological species within the amoeba genus *Korotnevella* (Amoebozoa: Discosea), revealed by the analysis of two genes. *European Journal of Protistology*, *56*, 102–111. https://doi.org/10.1016/j.ejop.2016.08.001

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Shchepin, O., Novozhilov, Y., Woyzichovski, J., Bog, M., Prikhodko, I., Fedorova, N., Gmoshinskiy, V., Borg Dahl, M., Dagamac, N. H. A., Yajima, Y., & Schnittler, M. (2022). Genetic structure of the protist *Physarum albescens* (Amoebozoa) revealed by multiple markers and genotyping by sequencing. *Molecular Ecology*, 31, 372–390. https://doi.org/10.1111/mec.16239