DR. PAVEL SKALOUD (Orcid ID : 0000-0003-1201-3290)

Article type : Original Article

Speciation in protists: Spatial and ecological divergence processes cause rapid species diversification in a freshwater chrysophyte

Pavel Škaloud[1] | Magda Škaloudová[1] | Pavla Doskočilová[1] | Jong Im Kim[2] | Woonghi Shin[2] | Petr Dvořák[3]

[1]Department of Botany, Faculty of Science, Charles University, Praha, Czech Republic

[2]Department of Biology, Chungnam National University, Daejeon, Korea

[3]Department of Botany, Palacký University, Olomouc, Czech Republic

Correspondence:

Pavel Škaloud, Department of Botany, Faculty of Science, Charles University, CZ-12800 Praha, Czech Republic.

e-mail: skaloud@natur.cuni.cz

Running title: Speciation in protists

Abstract

Though eukaryotic microorganisms are extremely numerous, diverse, and essential to global ecosystem functioning, they are largely understudied by evolutionary biologists compared to multicellular macroscopic organisms. In particular, very little is known about the speciation mechanisms which may give rise to the diversity of microscopic eukaryotes. It was postulated that enormous population sizes and ubiquitous distribution of these organisms could lead to a lack of population differentiation and therefore very low speciation rates. However, such assumptions have traditionally been based on morphospecies which may not accurately reflect the true diversity, missing cryptic taxa. In this study, we aim to articulate the major diversification mechanisms leading to the contemporary molecular diversity by employing a colonial freshwater flagellate, *Synura sphagnicola*, as an example. Phylogenetic analysis of five sequenced loci showed that *S. sphagnicola* differentiated into two morphologically distinct lineages approximately 15.4 Mya, which further diverged into several evolutionarily recent haplotypes during the late Pleistocene. The most recent haplotypes are ecologically and biogeographically much more differentiated than the old lineages, presumably because of their persistent differentiation after the allopatric speciation events. Our study shows that in microbial eukaryotes, species diversification by colonising new geographic regions or ecological resources occurs much more readily than was previously thought. Consequently, divergence times of microorganisms may be in some lineages equivalent to the estimated times of speciation in plants and animals.

KEYWORDS: biogeography, Chrysophyceae, Pleistocene, protists, speciation, taxonomy

1 | INTRODUCTION

While eukaryotic microorganisms (protists) are extremely numerous, diverse, and essential to global ecosystem functioning, due to their small size and difficulty getting into culture, our knowledge of their diversity and evolutionary processes is considerably limited. Advances in molecular techniques have brought new insights into global protist diversity and revealed a high degree of cryptic

diversification (Boenigk, Ereshefsky, Hoef-Emden, Mallet, & Bass, 2012). Molecular markers have been employed in species delineation, and revealed that classical, morphological approaches drastically underestimated diversity (Boenigk et al., 2012). Several large-scale studies based on numerous isolates from a wide range of habitats have revealed that actual protist diversity is much higher than previously assumed (Howe, Bass, Vickerman, Chao, & Cavalier-Smith, 2009; Kooistra et al., 2008; Sáez et al., 2003). More recently, the growing number of Next Generation Sequencing (NGS) studies employing amplicon sequencing of microbial eukaryotes has confirmed the vast genetic diversity and abundance of protists (Behnke et al., 2011; Gimmler, Korn, De Vargas, Audic, & Stoeck, 2016; Pawlowski et al., 2012). Moreover, several studies have shown that both terrestrial and freshwater protist communities are strongly shaped by the combination of geographic distance and environmental factors (Bates et al., 2013; Boenigk et al., 2018; Ragon, Fontaine, Moreira, & López-García, 2012).

However, little is known about mechanisms generating and maintaining the protist diversity. Uncertainty of these processes, coupled with fluid or inconsistent species concepts, however, has led some authors to the contrary opinion that protist diversity is actually quite low, perhaps less so than even that of metazoans (Fenchel & Finlay, 2006; Mora, Tittensor, Adl, Simpson, & Worm, 2011). According to these authors, the enormous population sizes and unlimited dispersal characteristics of protists should lead to the lack of population differentiation with no barriers to gene flow, leading to very low speciation rates (Fenchel & Finlay, 2006). They further postulate that the use of genetic data has brought confusion into the estimations of real diversity in protists, because the variation in molecular markers reflected the accumulation of neutral mutations over historical times rather than the existence of large numbers of cryptic taxa. Therefore, it is essential to understand the nature of genetic variability in protists, in particular, to disentangle the drivers of population differentiation that lead to speciation processes.

Several examples of population differentiation in protists have been described so far, along with the geographical, ecological, or temporal patterns. Geographic speciation seems to represent the major isolating force, especially in diatoms (Casteleyn et al., 2010; Evans et al., 2009; Souffreau,

Vanormelingen, Verleyen, Sabbe, & Vyverman, 2010). Population differentiation has been reported among both distantly and closely related localities, indicating that gene flow may be limited to a few kilometres (Rengefors, Logares, & Laybourn-Parry, 2012). A remarkable example of recent ecological speciation has been reported by Logares et al. (2007), who identified the post-glacial marine-freshwater transition of a dinoflagellate connected with the speciation event. Host specialization has been shown to be the major diversification force in symbiotic algae (LaJeunesse & Thornhill, 2011; Thornhill, Lewis, Wham, & Lajeunesse, 2014). Finally, temporal differentiation of phytoplankton has been reported in the marine diatom *Ditylum brightwellii* (Rynearson, Newton, & Armbrust, 2006) and freshwater raphidophyte *Gonyostomum semen* (Lebret, Kritzberg, Figueroa, & Rengefors, 2012).

Silica-scaled chrysophytes are an interesting group to study speciation mechanisms in protists because of their potential to improve our knowledge gathered primarily from diatoms and dinoflagellates. First, recently published phylogenetic investigations demonstrate the existence of evolutionarily recent radiations within several traditionally characterized species (Jo, Kim, Škaloud, Siver, & Shin, 2016; Scoble & Cavalier-Smith, 2014; Škaloud, Škaloudová, Procházková, & Němcová, 2014). Second, the biogenesis of species-specific siliceous scales allows us to morphologically differentiate most of these extant taxa (Siver et al., 2015). Finally, deposition of silica scales in sediments of ponds and lakes leaves a recognizable fossil record, which enables the calibration of divergence times (Siver et al., 2015; Siver, Lott, & Wolfe, 2013).

In this study, we aim to articulate the major diversification mechanisms that generated the contemporary molecular diversity within the silica-scaled chrysophyte *Synura sphagnicola* (Chrysophyceae, Stramenopiles), a colonial, cosmopolitan freshwater flagellate that often dominates in slightly acidic waters. To study both molecular and morphological diversity in detail, we cultured more than 70 newly established strains isolated from water bodies differing in climatic and chemical characteristics. We used a time calibrated phylogenetic analysis of five molecular loci to infer patterns of diversity and to estimate divergence times among the lineages. Then, we assessed the relative contributions of geography, climate, and habitat in promoting the diversification of lineages at

different times of their evolutionary history. We recovered rapid diversification that occurred during the late Pleistocene, with the most recent haplotypes being ecologically and biogeographically much more differentiated than the old, morphologically distinguishable lineages which diverged in Miocene. Accordingly, the colonisation of new geographic regions or ecological resources seems to represent an important diversification force in protists, despite their enormous population sizes and cosmopolitan dispersal.

2 | MATERIALS AND METHODS

2.1 | Sampling, isolation, and cultivation

The strains used in this study were obtained during an extensive investigation of the genus *Synura* in Europe, Newfoundland, and Korea (2007–2017). The material was sampled using a plankton net with 20-μm mesh. Abiotic factors including water pH, temperature and specific conductivity were measured using a combined pH/conductometer (WTW 340i; WTW GmbH, Weilheim, Germany). Collected samples were kept in a polystyrene box equipped by a cooling gel pad for a few hours until they were processed at the research base. To establish unialgal cultures, the individual colonies were isolated from the natural samples by micropipetting, and transferred into separate wells of a 96-well polypropylene plate filled with the MES-buffered DY IV liquid medium (pH ≈ 6; Andersen, Berges, Harirson, & Watanabe, 2005). All cultures were then grown in the 50 ml Erlenmeyer flasks at 15°C, under the permanent illumination of 7-40 μmol m^{-2} s^{-1} (various light sources, including fluorescent and LED lamps). From a total of 2,450 isolated cultures, 71 strains morphologically well corresponded with the description of *Synura sphagnicola*. These strains were further analysed in this study (Supporting Information Table S1).

2.2 | DNA sequencing and phylogenetic analyses

DNA was isolated as described in Škaloud et al. (2014). Five molecular loci were amplified by PCR: the hyper variable nuclear ITS rDNA, plastid *psa*A, and mitochondrial *cox*I were obtained from the 69, 45 and 39 strains, respectively. Two additional molecular markers (nuclear SSU rDNA and plastid *rbc*L) were obtained from the five selected strains to assess the phylogenetic position and estimate the ages of *S. sphagnicola* lineages. The amplification of ITS rDNA was performed as described in Kynčlová, Škaloud, & Škaloudová (2010). Amplification of the *psa*A and *cox*I markers were performed according to Boo et al. (2010), with the newly designed forward *cox* primer Synura-cox1-F1 (5'-TTT GAT CCT GCA GGN GGD GG-3'). Amplification of the SSU rDNA and *rbc*L was performed as described in Škaloud, Kristiansen, & Škaloudová (2013). The PCR products were purified and sequenced at Macrogen Inc. in Seoul, Korea, using ABI 3730 DNA analyzer (Applied Biosystems, USA).

Multiple alignments of the ITS rDNA, *psa*A, *cox*I, *rbc*L and SSU rDNA loci sequences were either manually built in MEGA6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013) or constructed using MAFFT v6, applying the Q-INS-i strategy (Katoh, Misawa, Kuma, & Miyata, 2002). The positions with deletions prevailing in a majority of sequences were removed from the alignment. The resulting datasets comprised of 447 nucleotide sites of the ITS rDNA, 958 sites of the *psa*A, 958 sites of the *psa*A, 542 sites of the *cox*I, 1,663 sites of the SSU rDNA, and 1,236 sites of the *rbc*L gene.

The Bayesian evolutionary analysis was performed to infer the phylogenetic position and estimate the split age of *S. sphagnicola* lineages, using the program BEAST v1.8.2. The analysis was performed on the concatenated and partitioned SSU rDNA and *rbc*L dataset (Supporting Information Table S2), using the lognormal relaxed clock model. The most appropriate partition-specific substitution models were recognized by the BIC-based model in jModelTest 2.1.4 as follows: (i) TIM2 + I + Γ for SSU rDNA, (ii) TIM1 + I + Γ for the 1st *rbc*L codon partition, (iii) TPM3 + I + Γ for the 2nd *rbc*L codon partition, and (iv) TIM3 + I + Γ for the 3rd *rbc*L codon partition. However, the GTR + I + Γ substitution model was applied to all partitions, as the most similar applicable option. A birth-death diversification process was used as a prior on the distribution of node heights. Eleven

secondary calibration constraints were used, following the inferred molecular divergence times published by Siver et al. (2015). These includes the *Synura* calibration points 3, 4, 5, 6, 7, 9, 14, and 15, the ages of genera *Mallomonas* (124 Ma, Stdev 15) and *Synura* (100 Ma, Stdev 15), and the estimated divergence between *Mallomonas* and *Synura* (130 Ma, Stdev 15). Five MCMC analyses were run for 100 million generations (burn-in 20 million generations). The convergence diagnostics was performed in Tracer 1.6. To check for the reliability of using secondary calibration constraints, we further performed the molecular clock analysis based on fewer but primary calibrations. We analysed the concatenated SSU rDNA and *rbc*L dataset of Chrysophyceae published by Škaloud et al. (2013), applying seven calibration points as follows: (1) the split between *Mallomonas cratis* and *M. pseudocratis*, (2) the node including all species from the section Mallomonas, (3) the node including *M. insignis* and *M. bangladeshica*, (4) the node including *M. matvienkoae*, *M. caudata*, *M. oviformis* and *M. heterospina*, (5) the parent node of *M. insignis*, (6) the parent node of *Synura uvella*, and (7) the node including *S. spinosa*, *S. curtispina* and *S. mollispina*. For all calibration points, we used an offset of 40 Ma, a mean of 15, and a standard deviation of 0.6, based on the observation of above-mentioned *Mallomonas* and *Synura* scales in the laminated sediments of the Giraffe Pipe core (Siver, Lott, & Wolfe, 2009; Siver & Wolfe, 2005), estimated to have a minimum age of approximately 40 Ma (Siver & Wolfe, 2009). All the other parameters were the same as described above. The estimated node ages were highly congruent in both analyses (e.g., the split of *S. sphagnicola* was estimated to 15.4 ± 8.4 and 16.2 ± 10.3 Mya ago, respectively).

Genealogical relationships of ITS rDNA, *psa*A, and *cox*I loci were investigated by constructing the maximum parsimony haplotype networks using the Haplotype Viewer (Ewing, available at www.cibiv.at/~greg/haploviewer). Since the loci were incongruent in haplotype composition and their relationships, the species tree was inferred by the StarBEAST analysis using the two-step approach. First, the concatenated alignment of SSU rDNA, *rbc*L, ITS rDNA, *psa*A and *cox*I sequences of 14 Petersenianae taxa (Supporting Information Table S3) was built to estimate the substitution rates of ITS rDNA, *psa*A, and *cox*I molecular loci. The following partition-specific substitution models were selected according to the BIC-based model in jModelTest 2.1.4: (i) GTR + I + Γ for SSU rDNA, (ii)

GTR + I + Γ for the *rbc*L partition, (iii) JC for the ITS rDNA partition, (iv) HKY + I for the *cox*I partition, and (iv) HKY + I for the *psa*A partition. The alignment was analysed by BEAST v1.8.2 (Suchard et al., 2018), with substitution and clock models unlinked among the partitions. Lognormal relaxed clock models were selected for the partitions, and a birth-death diversification process was used as a prior on the distribution of node heights. Four temporal constraints were used to calibrate the phylogeny, following the inferred molecular divergence times published by Siver et al. (2015). These constrains include nodes 5, 9, 14 and 15 in Siver et al. (2015). Five Markov Chain Monte Carlo (MCMC) analyses were run for 50 million generations, sampling every 10 000 generation. After the diagnosis for convergence using Tracer 1.6, the log files were merged using the burn-in set to 10 million generations. Accordingly, the mean rates and coefficients of variation were obtained for ITS rDNA (9.569E-3, 1.414), *psa*A (2.571E-3, 0.417), and *cox*I (4.962E-2, 0.222) loci, respectively. The estimated rates were then applied to infer the phylogenetic position and node ages of *S. sphagnicola* lineages, using the StarBEAST as implemented in BEAST2 v2.4.7 (Bouckaert et al., 2014). The analysis was performed on the concatenated and partitioned ITS rDNA, *psa*A and *cox*I dataset, using only those strains for which all three partitions were successfully sequenced (Supporting Information Table S1). A total of 7 haplotypes were defined in StarBEAST Taxon sets (1A, 1B, 1C, 1D, 2A, 2B, 2C), according to the results of haplotype networks. MCMC analyses were run for 100 million generations, using the same setting as specified above. The convergence diagnostics was performed in Tracer 1.6.

2.3 | Statistical analyses

Even though we measured the temperature, pH and conductivity for the most of studied sites, these values are likely influenced by short-term variations occurring as a result of biological activity and climatic deviations. Therefore, more biologically meaningful variables were calculated for each sampling site, using several pre-processed environmental data layers. First, altitude and 19 bioclimatic variables were obtained from WorldClim database (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) at resolution of 2.5 arc minutes to characterize the general climatic conditions. Since the studied

organisms naturally exhibit temporal changes in their composition and abundance, we additionally retrieved average, minimum and maximum temperatures for the period of one month preceding the sampling date, using the timeseries datasets in Chelsa database (Karger et al., 2017). Finally, seven physical and chemical soil properties were obtained from the SoilGrids database (soilgrids.org): pH, cation exchange capacity, clay, silt, and sand content, and organic carbon stock and content. At every sampling site, climatic and habitat data were obtained by applying either a 1-km or 5-km buffer to limit the effects of spatial bias. No sampling site was located on a hillside, to prevent inaccurate data buffering. Geographical distance (latitude and longitude) was transformed to the principal coordinates of neighbour matrices (PCNM; Dray, Legendre, & Peres-Neto, 2006) by distGPS and pcnm functions in BoSSA (Lefeuvre, 2018) and vegan (Oksanen et al., 2016) packages in R (ver. 3.3.3; R Development Core Team), respectively.

The relative effects of climate, habitat and geographical distance on the variance in *Synura* diversity were analysed using variation partitioning in redundancy analyses, using the varpart function in vegan package in R. The effects of environmental and spatial factors were analysed at three levels of diversity, using the affiliation of samples into i) two main *S. sphagnicola* lineages, ii) four sp1 haplotypes 1A-D, and iii) three sp2 haplotypes 2A-C, as a response variable. The analyses were run on the subset of 48 samples for which the affiliation into one of the seven sp1 and sp2 haplotypes was recovered (Supporting Information Table S1). To identify significant explanatory variables associated with the tested levels of *Synura* diversity, forward selection by redundancy analysis was used prior to variation partitioning, using the ordiR2step function in vegan package in R.

The morphological analyses were based on fifteen cultures randomly selected in both the lineage sp1 (seven cultures) and sp2 (eight cultures), respectively. The cultures were inoculated into fresh media, and after 4-5 weeks of cultivation they were examined with a JEOL 1011 transmission electron microscope. For each of these cultures, the five characters were measured in at least 40 randomly scales, as follows: (1) scale length, (2) scale width, (3) spine length, (4) spine width, and (5) rim width. The morphological differentiation of *S. sphagnicola* lineages was evaluated using the principal component analyses (PCA). To test for the morphological distinction of the lineages, we

performed the linear discriminant analysis (LDA). We further evaluated the linear discriminant analysis by randomly selecting ¾ of scales, estimating the parameters on the training data, and classifying the remaining ¼ scales of the holdout sample. We repeated this 100 times. All the analyses were performed in R, using the packages scales (Wickham, 2017), vegan, MASS (Ripley, Venables, Hornik, & Ripley, 2015), and DiscriMiner (Gaston & Sanchez, 2013).

3 | RESULTS

3.1 | Phylogenetic relationships and timing of divergence

A total of 71 strains were isolated from various water bodies in Newfoundland, Korea, the Czech Republic, Norway, and British Isles (Figure 1). All strains were inferred to be closely related, forming a single, monophyletic group (Figure 2a). The clade was split into two distinct lineages at about 15.4 (± 8.4) Mya ago, here referred to as sp1 and sp2. These lineages frequently co-occurred in European areas, but were not found together in Newfoundland and Korea (Figure 1). Additional molecular diversity was revealed within both of these lineages, by sequencing fast evolving ITS rDNA, *psa*A, and *cox*I loci (Figures 2b,c). Four haplotypes were identified within the lineage sp1. The haplotypes 1A and 1B were closely related, differentiated only by *psa*A locus. Whereas 1A occurred only in the Czech Republic, the latter haplotype was isolated from the Czech Republic, Sweden, Scotland, and Ireland (Figure 3). The haplotypes 1C and 1D were recognized by ITS rDNA and *psa*A loci, and were both endemic to Korea. Molecular diversity within the lineage sp2 was recovered only by the *cox*I locus, identifying three haplotypes. The haplotypes 2A and 2B were closely related, the former sampled in the Czech Republic, Scotland, and England, and the latter distributed in the Czech Republic, Norway, and Newfoundland (Figure 3). The haplotype 2C, occurring in Northern Europe (Scotland, Ireland, Sweden, and Norway), was rather distantly related to the other haplotypes.

Since ITS rDNA, *psa*A, and *cox*I loci fragment the molecular diversity within the sp1 and sp2 lineages differently, we inferred the putative, time-calibrated species tree of *S. sphagnicola* haplotypes based on all haplotype data. The resulting species tree strongly supported the differentiation of all haplotypes, providing evidence for their recent radiation in the Pleistocene (Figure 2c). Using the

haplotype-specific mean rates, we estimate sp1 and sp2 diverged 17.7 (± 12.1) Mya, which corresponds closely to the 15.4 (± 8.4) Mya estimate based on the fossil-calibrated phylogeny of Synurales (Figure 2a).

3.2 | Explaining patterns in molecular diversity

In order to assess the role of environmental and spatial processes on the pattern of observed genetic diversity, we performed variation partitioning analyses on three levels of diversity (Figure 4). At the level of lineages sp1 and sp2, only the net effect of climate made a significant contribution to explaining the overall variation (P = 0.029; Figure 4a). Season precipitation (BIO15) and wettest month precipitation (BIO13) were the main factors explaining the divergence of the lineages. The spatial and habitat differences were nested within the effect of the climate. Most of the variation (64%) remained unexplained, suggesting that the lineages are highly correlated to their environmental preferences and spatial distribution.

In contrast, when analysing the diversity within the recently diverging haplotypes there was much less unexplained variance (29% and 22%, respectively). The haplotypes were well differentiated by their geography and climate (Figures 3 and 4b,c). Interestingly, the patterns of variation partitioning were very similar, irrespective of explaining the variability within the lineage sp1 or sp2. The biggest part of variation was explained by the spatial pattern (sp1 haplotypes: 32% net effect, P = 0.001; sp2 haplotypes: 41% net effect, P = 0.010), followed by the climatic factors (sp1 haplotypes: 16% net effect, P = 0.004; sp2 haplotypes: 18% net effect, P = 0.043). Mean diurnal temperature range (BIO2) and wettest month precipitation (BIO13) were the most informative bioclimatic variables in the model.

3.3 | Morphological analyses

All strains morphologically well fit the circumscription of *S. sphagnicola*: they formed multicellular colonies composed of two-flagellated cells up to 11(-16) μm in diameter, often filled with bright red granules (Figure 5a). The cells were covered by a layer of imbricate scales composed of a perforated basal plate with an upturned rim and an apical spine (Figure 5b). The principal component analysis (PCA) clearly separated the sp1 and sp2 lineages, based on the morphometric characteristics of silica scales analysed in 15 cultures (Figure 5c). The linear discriminant analysis (LDA) showed significant morphological differentiation of the lineages, indicating that all but one morphological feature (spine width) was significant for lineage recognition ($P < 0.00001$), with spine length being the most discriminating character (Wilk's $\lambda = 0.43$, $F = 1063.4$). Indeed, the average spine length was 1.58 ± 0.31 in sp1 and 2.64 ± 0.56 in sp2, respectively (Figure 5d). The predictive discrimination of individual scales on the basis of their morphology reached 94.1%. In addition to the distinction between the sp1 and sp2 lineages, we also detected apparent differences among the cultures belonging to the same lineages (Figure 5c). This variation cannot be explained by the origin of analysed strains, since the morphologically most similar strains were isolated from geographically distantly related and ecologically distinct localities. It seems that the morphological variation could be rather attributed to the distinction between the recently diverging haplotypes, however, additional measurements based on much larger samples are required to support this explanation.

4 | DISCUSSION

4.1 | Protist speciation

In general, speciation models assume that species may diverge either with or without gene flow, or by secondary contact between the diverging species (Richards, Servedio, & Martin, 2018). Gene flow restriction is the most commonly documented mechanism (Sobel, Chen, Watt, & Schemske, 2010), occurring in both spatially structured (allopatry) and well-mixed, initially randomly mating populations (classical sympatry). However, species may also emerge in sympatry when gene flow is present (Richards et al., 2018). Interestingly, genomic studies based primarily on animal models show

that high levels of genome-wide divergence can be established along with ongoing gene flow (Martin et al., 2013; Ravinet et al., 2018). Unfortunately, no such genomic studies have been performed in protists so far, with the exception of investigations based on yeasts (Leducq et al., 2016). From a genetic point of view, isolating barriers may be based on the divergent ecological selection of alleles (e.g. habitat and temporal isolation), cytological divergence (e.g. polyploidisation, chromosome rearrangements), or species hybridization (Futuyma, 2013).

Herein, we had an excellent opportunity to document the patterns in evolutionary divergence in eukaryotic microorganisms (protists), based on our investigations of spatial and ecological structuring of genetic diversity in *S. sphagnicola*. We were able to overcome several obstacles that have limited research in this field so far. First, it is usually very hard to trace recently derived protist lineages in nature due to their small size and paucity of differentiating morphological features. Second, detailed comparative research requires labour-intensive cultivation practices. We were able to circumvent these issues due to the pre-existing, well characterized phylogeny of the genus *Synura* (Jo et al., 2016; Škaloud et al., 2014) with a concurrent well-established fossil record (Siver, Wolfe, Rohlf, Shin, & Jo, 2013; Siver et al., 2015). This enabled us to set up reliable calibration points on several diversification events. Although logistical constraints prevent investigation of hundreds of strains, the observed differentiation of lineages in organisms studied herein is so strong that it cannot be considered the result of random processes.

The results of our variation partitioning analyses clearly show that the most recently diverged haplotypes differ significantly by their distribution and ecological preferences. Since the distribution patterns of recently derived haplotypes most likely reflects the possible causes of speciation (Barraclough & Nee, 2001), allopatric and divergent ecological speciation seems to represent a major speciation force in *S. sphagnicola* despite the assumed enormous population sizes and unlimited dispersal of this protist taxon. Specifically, we showed that obvious restriction in gene flow was maintained between geographically differentiated populations, despite only recently diverging. However, how can populations be resistant to gene flow under the continuous influx of immigrant cells? One possible explanation could lie in the unfavourable conditions to which the protists are

exposed during their long-distance transport. Drought, freezing, and UV radiation may be lethal for protist propagules, making it impossible for them to colonize geographically remoted new habitats. Indeed, freshwater diatoms were recognized to be extremely sensitive to desiccation stress (Souffreau et al., 2010). On the other hand, approximately 200 viable airborne protist taxa were recorded around the world (Sharma, Rai, Singh, & Brown, 2007), indicating that under some microclimatic and physico-chemical conditions microorganisms may be able to successfully colonize remote habitats.

However, high migration rates do not necessarily prevent speciation processes. In plants and animals, several examples of species divergence have been reported for populations occurring over local distance scales, even under a high migration frequency. For example, sexual selection (Elmer, Lehtonen, & Meyer, 2009; Higashi, Takimoto, & Yamamura, 1999) or disjunction in flowering time (Savolainen et al., 2006), which are usually tightly coupled with ecological specialization (Feder et al., 1994; Mallet, Meyer, Nosil, & Feder, 2009), are among the best-documented examples of speciation maintained in the face of gene flow. In protists, which are characterized by rather rare sexual reproduction, population differentiation may be related to mechanisms that effectively prevent immigrant genotypes from being incorporated into local populations, such as the founder effect of resident species that have monopolized the resources (De Meester, Gómez, Okamura, & Schwenk, 2002). Indeed, numerous studies determined the importance of the founder effect in structuring bacterial (Svoboda, Lindström, Ahmed Osman, & Langenheder, 2018), fungal (James & Vilgalys, 2001), and even protist (Weisse et al., 2011) communities. Protists are generally characterized not only by high dispersal rates but also by rapid growth rates and short generation times. Rapid population growth combined with the formation of seed banks may result in a strong numerical advantage for the first colonisers. In addition, rapid adaptation of populations to local environments may strongly enhance the priority effect that hinders immigration (Urban & De Meester, 2009). The combination of frequent clonal and rare sexual reproduction, typical for the majority of protist species, represents a powerful tool for promoting local adaptation (De Meester et al., 2002).

In addition to the dominant spatial pattern of recent *S. sphagnicola* haplotypes, a significant part of the overall genetic variability has been explained by the net effect of climatic and habitat

factors. This may represent one of the consequences of speciation following founder effect; in other words, the ecological differentiation of closely related species to local conditions. However, this pattern may also relate to the divergent ecological selection that occurs during sympatry. Indeed, closely related *S. sphagnicola* genotypes were sometimes found to occur in sympatry (Figure 3c). In protists, the above-mentioned combination of frequent clonal and rare sexual reproduction may primarily facilitate a specific mechanism of ecological speciation referred to as "periodic selection" (Atwood, Schneider, & Ryan, 1951; Cohan, 2001; Maynard Smith & Haigh, 2008). This is characterized by periods of relaxed selection followed by bouts of pronounced selection, yielding a vast reduction of genetic diversity (Cohan, 2001). However, if a new mutant differs in its ecological niche, it could escape the diversity-purging effects of periodic selection, giving rise to the new, ecologically defined species (Škaloud & Rindi, 2013).

4.2 | Revisiting species concept of protists

The proper delimitation of species is an essential requirement for both biodiversity assessments and a correct understanding of their ecology, biogeography, and evolutionary history. As is the case for almost every group of organisms, the species concept in protists has been discussed in many publications yet is still highly controversial (Boenigk et al., 2012; Leliaert et al., 2014; Malavasi et al., 2016). The great majority of protist species are still delimited morphologically, although the morphology alone is generally insufficient to distinguish natural species units (Hoef-Emden, 2007; Škaloud & Rindi, 2013). However, this concept was frequently applied in proposing and testing hypotheses of distribution patterns and speciation mechanisms in protists (Fenchel, 2005; Finlay & Fenchel, 1999), which can lead to confusing inferences.

Importantly, the lack of clear morphological delineation between recently derived taxa of protists may obscure evolutionary relationships and tempos of change. This bias, based on the morphological perception of species, is well demonstrated when comparing macro- and microorganisms. Estimated divergence times of plant or animal sister species generally range from tens of thousands to several million years, with a considerable portion of species originating during

the Pleistocene glaciations (Johnson & Cicero, 2004; Wallis, Waters, Upton, & Craw, 2016). For example, estimated speciation times span 0.03-4.1 Mya in birds (Johnson & Cicero, 2004), 0.002-4.6 Mya in cichlids (McCune, 2001; Genner et al., 2007), 0.5-1.5 Mya in sea stars (Foltz, Nguyen, Kiger, & Mah, 2008), about 1.5 Mya in frogs (Sasa, Chippindale, & Johnson, 1998), and less than 1 Mya in endemic cyprinodontid and semionotid fishes, fruit flies and *Tetragnatha* spiders (McCune, 2001). In addition, extremely fast speciation times, evolving within thousands of generations, have been observed in numerous plant and animal taxa (Hendry, Nosil, & Rieseberg, 2007).

In contrast to macroorganisms, the estimated divergence times of sister protist species are usually much longer. Dinoflagellate species divergence varies from 5.5 to 32 Mya (Leaw et al., 2016; Lowe, Martin, Montagnes, & Watts, 2012), diatom divergence times then generally span 2.9-11.7 Mya (Brown & Sorhannus, 2010; Sörhannus, Fenster, Hoffman, & Burckle, 1991; Whittaker, Rignanese, Olson, & Rynearson, 2012). Noteworthy, more recent divergence times were estimated for those protist groups producing a variety of beautifully ornamented shells or scales, such as Radiolaria (about 1.7 Mya; Ishitani, Ujiié, de Vargas, Not, & Takahashi, 2012) and Coccolithophores (0.3-5.6 Mya; Sáez et al., 2003). Obviously, the scarcity of morphological features in protists lead to ignoring recently diverged lineages and recognizing rather evolutionary old lineages as species units, often artificially lumping together ecologically well differentiated young lineages. In addition, since increasing phylogenetic resolution well correlates with detecting geographically restricted lineages (Bass & Boenigk, 2011), such taxonomic bias can fundamentally influence our estimates of principal speciation mechanisms towards the higher probability of allopatric speciation.

In addition, since morphological change often lags behind ecological change, morphologically defined taxa may mask cryptic, ecologically distinct lineages, leading to erroneous estimations of their ecological niches (Malavasi et al., 2016; Škaloud & Rindi, 2013). Considering our model, *Synura sphagnicola* is currently regarded as a broadly distributed species occurring in a wide variety of mainly acidic water bodies (Siver, 1989). However, our data indicates it may consist of several both geographically and ecologically distinct entities. Though we are fully aware that additional sampling is necessary to support our observations, the majority of haplotypes exhibit non-random distribution

patterns (Figure 3c). For example, it seems that haplotype 1A is restricted to central Europe, whereas the haplotypes 1C and 1D are distributed in more humid and warmer Korean environments. Similarly, the haplotypes 2A and 2C appear to be restricted in their distribution to colder regions of Europe.

We feel that our perception of protist species should be revised considerably in terms of distinguishing between much more recent and morphologically similar, but genetically, spatially, and ecologically isolated lineages. It is highly probable that numerous protist species originated during the Quaternary period, as a consequence of a fragmentation of previously contiguous populations by the advances of glaciers and associated sea level variations. Indeed, some striking examples of Pleistocene protist speciation were already documented, e.g. in marine foraminifer *Globigerinoides ruber* (Aurahs, Grimm, Hemleben, Hemleben, & Kucera, 2009), marine diatom *Pseudo-nitzschia pungens* (Casteleyn et al., 2010), freshwater diatom *Stephanodiscus yellowstonensis* (Theriot, Fritz, Whitlock, & Conley, 2006) and brackish-freshwater dinoflagellates *Peridinium aciculiferum* and *Scrippsiella hangoei* (Logares et al., 2007).

5 | CONCLUSIONS

To summarize, we report rapid diversification of a freshwater microbial eukaryote that occurred during the late Pleistocene, resulting in several geographically and ecologically diverse populations. Our study shows that ecological and allopatric speciation seem to occur much more readily in microbial eukaryotes than was previously thought. The colonisation of new geographic regions or ecological resources seems to promote protist speciation, despite their enormous population sizes and cosmopolitan dispersal. Though it is generally difficult to compare speciation times of microorganisms and macroorganisms due to their different generation times, these findings imply they may be equivalent, contrary to hypotheses that suggest the slow evolution of protist species (Fenchel & Finlay, 2003, 2006). Future studies employing detailed, genome-wide population analyses, such as restriction-site associated DNA sequencing (RAD-seq), will help us to better understand the ecology of evolutionarily young haplotypes and their roles in ecosystem functioning, as well as to establish a solid taxonomy based on the identification of ecological and evolutionary meaningful units.

ACKNOWLEDGEMENTS

This research was supported by the Czech Science Foundation (No. 17-13254S) to PŠ and National Research Foundation of Korea (No. NRF-2015R1A2A2A01003192) to WS. We are extremely grateful to Dale A. Casamatta for English revision.

DATA ACCESSIBILITY

DNA sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession numbers MK322768–MK322930. DNA alignments are freely available on Mendeley Data: http://dx.doi.org/10.17632/83dcg6ypp4.1.

AUTHOR CONTRIBUTIONS

Conceptualization, P.Š. and P.Dv.; Methodology, P.Š.; Formal Analysis, P.Š. and P.Do; Investigation, M.Š., P.Do., J.I.K. and W.S.; Resources, P.Š., M.Š., P.Do., J.I.K. and W.S.; Writing – Original draft, P.Š.; Writing – Review & Editing, P.Š., M.Š. and P.Dv.; Visualization, P.Š.; Funding Acquisition, P.Š. and W.S.

REFERENCES

Andersen, R., Berges, J., Harirson, P., & Watanabe, M. (2005). Appendix A—Recipes for Freshwater and Seawater Media. In Andersen R.A. (Ed.), *Algal Culturing Techniques* (pp. 429–538). Amsterdam: Elsevier.

Atwood, K. C., Schneider, L. K., & Ryan, F. J. (1951). Periodic selection in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, *37*, 146–155. doi:10.1073/pnas.37.3.146

Aurahs, R., Grimm, G. W., Hemleben, V., Hemleben, C., & Kucera, M. (2009). Geographical distribution of cryptic genetic types in the planktonic foraminifer *Globigerinoides ruber*. *Molecular Ecology*, *18*(8), 1692–1706. doi:10.1111/j.1365-294X.2009.04136.x

Barraclough, T. G., & Nee, S. (2001). Phylogenetics and speciation. *Trends in Ecology and Evolution*, *16*(7), 391–399. doi:10.1016/S0169-5347(01)02161-9

Bass, D., & Boenigk, J. (2011). Everything is everywhere: A twenty-first century de-/reconstruction with respect to protists. In Fontaneto D. (Ed.), *Biogeography of Microscopic Organisms is Everything Small Everywhere?* (pp. 88–110). Cambridge University Press. doi:10.1017/CBO9780511974878.007

Bates, S. T., Clemente, J. C., Flores, G. E., Walters, W. A., Parfrey, L. W., Knight, R., & Fierer, N. (2013). Global biogeography of highly diverse protistan communities in soil. *ISME Journal*, *7*(3), 652–659. doi:10.1038/ismej.2012.147

Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R. R., & Stoeck, T. (2011). Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environmental Microbiology*, *13*, 340–349. doi:10.1111/j.1462-2920.2010.02332.x

Boenigk, J., Ereshefsky, M., Hoef-Emden, K., Mallet, J., & Bass, D. (2012). Concepts in protistology: Species definitions and boundaries. *European Journal of Protistology*, *48*, 96–102. doi:10.1016/j.ejop.2011.11.004

Boenigk, J., Wodniok, S., Bock, C., Beisser, D., Hempel, C., Grossmann, L., … Jensen, M. (2018). Geographic distance and mountain ranges structure freshwater protist communities on a European scale. *Metabarcoding and Metagenomics*, *2*, e21519. doi:10.3897/mbmg.2.21519

Boo, S. M., Kim, H. S., Shin, W., Boo, G. H., Cho, S. M., Jo, B. Y., … Yoon, H. S. (2010). Complex phylogeographic patterns in the freshwater alga *Synura* provide new insights into ubiquity vs. endemism in microbial eukaryotes. *Molecular Ecology*, *19*(19), 4328–4338. doi:10.1111/j.1365-294X.2010.04813.x

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*, e1003537. doi:10.1371/journal.pcbi.1003537

Brown, J. W., & Sorhannus, U. (2010). A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): Substantive underestimation of putative fossil ages. *PLoS ONE*, *5*(9), 1–11. doi:10.1371/journal.pone.0012759

Casteleyn, G., Leliaert, F., Backeljau, T., Debeer, A.-E., Kotaki, Y., Rhodes, L., … Vyverman, W. (2010). Limits to gene flow in a cosmopolitan marine planktonic diatom. *Proceedings of the National Academy of Sciences*, *107*(29), 12952–12957. doi:10.1073/pnas.1001380107

Cohan, F. M. (2001). Bacterial species and speciation. *Systematic Biology*, *50*(4), 513–524. doi:10.1080/10635150118398

De Meester, L., Gómez, A., Okamura, B., & Schwenk, K. (2002). The Monopolization Hypothesis and the dispersal-gene flow paradox in aquatic organisms. *Acta Oecologica*, *23*(3), 121–135. doi:10.1016/S1146-609X(02)01145-1

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*. doi:10.1016/j.ecolmodel.2006.02.015

Elmer, K. R., Lehtonen, T. K., & Meyer, A. (2009). Color assortative mating contributes to sympatric divergence of neotropical cichlid fish. *Evolution*. doi:10.1111/j.1558-5646.2009.00736.x

Evans, K. M., Chepurnov, V. A., Sluiman, H. J., Thomas, S. J., Spears, B. M., & Mann, D. G. (2009).

Highly differentiated populations of the freshwater diatom *Sellaphora capitata* suggest limited dispersal and opportunities for allopatric speciation. *Protist*, *160*(3), 386–396. doi:10.1016/j.protis.2009.02.001

Feder, J. L., Opp, S. B., Wlazlo, B., Reynolds, K., Go, W., & Spisak, S. (1994). Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.91.17.7990

Fenchel, T. (2005). Cosmopolitan microbes and their 'cryptic'species. *Aquatic Microbial Ecology*, *41*, 49–54. doi:10.3354/ame041049

Fenchel, T., & Finlay, B. J. (2003). Is microbial diversity fundamentally different from biodiversity of larger animals and plants? *European Journal of Protistology*, *39*(4), 486–490. doi:10.1078/0932-4739-00025

Fenchel, T., & Finlay, B. J. (2006). The diversity of microbes: Resurgence of the phenotype. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1475), 1965–1973. doi:10.1098/rstb.2006.1924

Finlay, B. J., & Fenchel, T. (1999). Divergent perspectives on protist species richness. *Protist*, *150*(3), 229–233. doi:10.1016/S1434-4610(99)70025-8

Foltz, D. W., Nguyen, A. T., Kiger, J. R., & Mah, C. L. (2008). Pleistocene speciation of sister taxa in a North Pacific clade of brooding sea stars (Leptasterias). *Marine Biology*, *154*, 593–602. doi:10.1007/s00227-008-0952-9

Futuyma, D. J. (2013). *Evolution* (Third Edition). Sunderland, MA: Sinauer Associates, Inc.

Gaston, A., & Sanchez, M. G. (2013). DiscriMiner : Tools of the Trade for Discriminant Analysis. *R Package Version 0.1-29*.

Genner, M. J., Seehausen, O., Lunt, D. H., Joyce, D. A., Shaw, P. W., Carvalho, G. R., & Turner, G. F. (2007). Age of cichlids: New dates for ancient lake fish radiations. *Molecular Biology and Evolution*, *24*, 1269–1282. doi:10.1093/molbev/msm050

Gimmler, A., Korn, R., De Vargas, C., Audic, S., & Stoeck, T. (2016). The Tara Oceans voyage reveals global diversity and distribution patterns of marine planktonic ciliates. *Scientific Reports*, *6*, 33555. doi:10.1038/srep33555

Hendry, A. P., Nosil, P., & Rieseberg, L. H. (2007). The speed of ecological speciation. *Functional Ecology*, *21*, 455–464. doi:10.1111/j.1365-2435.2007.01240.x

Higashi, M., Takimoto, G., & Yamamura, N. (1999). Sympatric speciation by sexual selection. *Nature*. doi:10.1038/990087

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978. doi:10.1002/joc.1276

Hoef-Emden, K. (2007). Revision of the Genus *Cryptomonas* (Cryptophyceae) II: Incongruences between the classical morphospecies concept and molecular phylogeny in smaller pyrenoid-less cells. *Phycologia*, *46*(4), 402–428. doi:10.2216/06-83.1

Howe, A. T., Bass, D., Vickerman, K., Chao, E. E., & Cavalier-Smith, T. (2009). Phylogeny, taxonomy, and astounding genetic diversity of Glissomonadida ord. nov., the dominant gliding zooflagellates in soil (Protozoa: Cercozoa). *Protist*, *160*(2), 159–189. doi:10.1016/j.protis.2008.11.007

Ishitani, Y., Ujiié, Y., de Vargas, C., Not, F., & Takahashi, K. (2012). Phylogenetic relationships and evolutionary patterns of the order Collodaria (Radiolaria). *PLoS ONE*, *7*, e35775. doi:10.1371/journal.pone.0035775

James, T. Y., & Vilgalys, R. (2001). Abundance and diversity of *Schizophyllum commune* spore clouds in the Caribbean detected by selective sampling. *Molecular Ecology*, *10*, 471–479. doi:10.1046/j.1365-294X.2001.01224.x

Jo, B. Y., Kim, J. I., Škaloud, P., Siver, P. A., & Shin, W. (2016). Multigene phylogeny of *Synura* (Synurophyceae) and descriptions of four new species based on morphological and DNA evidence. *European Journal of Phycology*, *51*(4), 413–430. doi:10.1080/09670262.2016.1201700

Johnson, N. K., & Cicero, C. (2004). New mitochondrial DNA data affirm the importance of pleistocene speciation in North American birds. *Evolution*, *58*, 1122–1130. doi:10.1111/j.0014-3820.2004.tb00445.x

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., … Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*, 170122. doi:10.1038/sdata.2017.122

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*, 3059–3066. doi:10.1093/nar/gkf436

Kooistra, W. H. C. F., Sarno, D., Balzano, S., Gu, H., Andersen, R. A., & Zingone, A. (2008). Global diversity and biogeography of *Skeletonema* species (Bacillariophyta). *Protist*, *159*(2), 177–193. doi:10.1016/j.protis.2007.09.004

Kynčlová, A., Škaloud, P., & Škaloudova, M. (2010). Unveiling hidden diversity in the *Synura petersenii* species complex (Synurophyceae, Heterokontophyta). *Nova Hedwigia Beiheft*, *136*, 283–298. doi:10.1127/1438-9134/2010/0136-0283

LaJeunesse, T. C., & Thornhill, D. J. (2011). Improved resolution of reef-coral endosymbiont (Symbiodinium) species diversity, ecology, and evolution through psbA non-coding region genotyping. *PLoS ONE*, *6*(12). doi:10.1371/journal.pone.0029013

Leaw, C. P., Tan, T. H., Lim, H. C., Teng, S. T., Yong, H. L., Smith, K. F., … Lim, P. T. (2016). New scenario for speciation in the benthic dinoflagellate genus *Coolia* (Dinophyceae). *Harmful Algae*, *55*(March), 137–149. doi:10.1016/j.hal.2016.02.010

Lebret, K., Kritzberg, E. S., Figueroa, R., & Rengefors, K. (2012). Genetic diversity within and genetic differentiation between blooms of a microalgal species. *Environmental Microbiology*, *14*(9), 2395–2404. doi:10.1111/j.1462-2920.2012.02769.x

Leducq, J. B., Nielly-Thibault, L., Charron, G., Eberlein, C., Verta, J. P., Samani, P., … Landry, C. R. (2016). Speciation driven by hybridization and chromosomal plasticity in a wild yeast. *Nature Microbiology*. doi:10.1038/nmicrobiol.2015.3

Lefeuvre, P. (2018). BoSSA: A Bunch of Structure and Sequence Analysis. *URL Https://Cran.r-Project.Org*.

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J. M., Zuccarello, G. C., & De Clerck, O. (2014). DNA-based species delimitation in algae. *European Journal of Phycology*. doi:10.1080/09670262.2014.904524

Logares, R., Rengefors, K., Kremp, A., Shalchian-Tabrizi, K., Boltovskoy, A., Tengs, T., … Klaveness, D. (2007). Phenotypically different microalgal morphospecies with identical ribosomal DNA: A case of rapid adaptive evolution? *Microbial Ecology*, *53*(4), 549–561. doi:10.1007/s00248-006-9088-y

Lowe, C. D., Martin, L. E., Montagnes, D. J. S., & Watts, P. C. (2012). A legacy of contrasting spatial genetic structure on either side of the Atlantic–Mediterranean transition zone in a marine protist. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(51), 20998–21003. doi:10.1073/pnas.1214398110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1214398110

Malavasi, V., Škaloud, P., Rindi, F., Tempesta, S., Paoletti, M., & Pasqualetti, M. (2016). DNA-based taxonomy in ecologically versatile microalgae: A re-evaluation of the species concept within the coccoid green algal genus Coccomyxa (Trebouxiophyceae, Chlorophyta). *PLoS ONE*. doi:10.1371/journal.pone.0151137

Mallet, J., Meyer, A., Nosil, P., & Feder, J. L. (2009). Space, sympatry and speciation. *Journal of Evolutionary Biology*. doi:10.1111/j.1420-9101.2009.01816.x

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., … Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, *23*(11), 1817–1828. doi:10.1101/gr.159426.113

Maynard Smith, J., & Haigh, J. (2008). The hitch-hiking effect of a favourable gene. *Genetics Research*, *23*, 23–35. doi:10.1017/S0016672308009579

McCune, A. R. (2001). Rapid speciation in species flocks. In D. E. G. Briggs & P. R. Crowther (Eds.), *Palaeobiology II* (pp. 143–146). Blackwell Scientific.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How many species are there on earth and in the ocean? *PLoS Biology*, *9*, e1001127. doi:10.1371/journal.pbio.1001127

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Wagner, H. (2016). Vegan: Community Ecology Package. *URL Https://Cran.r-Project.Org*,

Https://Github.Com/Vegandevs/Vegan. doi:10.4135/9781412971874.n145
Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., … de Vargas, C. (2012). CBOL protist working group: Barcoding Eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biology*, *10*(11), e1001419. doi:10.1371/journal.pbio.1001419
Ragon, M., Fontaine, M. C., Moreira, D., & López-García, P. (2012). Different biogeographic patterns of prokaryotes and microbial eukaryotes in epilithic biofilms. *Molecular Ecology*, *21*(15), 3852–3868. doi:10.1111/j.1365-294X.2012.05659.x
Ravinet, M., Yoshida, K., Shigenobu, S., Toyoda, A., Fujiyama, A., & Kitano, J. (2018). The genomic landscape at a late stage of stickleback speciation: High genomic divergence interspersed by small localized regions of introgression. *PLoS Genetics*, *14*(5), e1007358. doi:10.1371/journal.pgen.1007358
Rengefors, K., Logares, R., & Laybourn-Parry, J. (2012). Polar lakes may act as ecological islands to aquatic protists. *Molecular Ecology*, *21*(13), 3200–3209. doi:10.1111/j.1365-294X.2012.05596.x
Richards, E., Servedio, M., & Martin, C. (2018). New criteria for sympatric speciation in the genomic era. *BioRxiv*. doi:10.1101/367623
Ripley, B., Venables, B., Hornik, K., & Ripley, M. B. (2015). Package "MASS." *Tutorial*. doi:ISBN 0-387-95457-0
Rynearson, T. A., Newton, J. A., & Armbrust, E. V. (2006). Spring bloom development, genetic variation, and population succession in the planktonic diatom *Ditylum brightwellii*. *Limnology and Oceanography*, *51*(3), 1249–1261. doi:10.4319/lo.2006.51.3.1249
Sáez, A. G., Probert, I., Geisen, M., Quinn, P., Young, J. R., & Medlin, L. K. (2003). Pseudo-cryptic speciation in coccolithophores. *Proceedings of the National Academy of Sciences*, *100*, 7163–7168. doi:10.1073/pnas.1132069100
Sasa, M. M., Chippindale, P. T., & Johnson, N. A. (1998). Patterns of postzygotic isolation in frogs. *Evolution*, *52*, 1811–1820. doi:10.2307/2411351
Savolainen, V., Anstett, M. C., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V., … Baker, W. J. (2006). Sympatric speciation in palms on an oceanic island. *Nature*. doi:10.1038/nature04566
Scoble, J. M., & Cavalier-Smith, T. (2014). Scale evolution in paraphysomonadida (Chrysophyceae): Sequence phylogeny and revised taxonomy of paraphysomonas, new genus clathromonas, and 25 new species. *European Journal of Protistology*, *50*(5), 551–592. doi:10.1016/j.ejop.2014.08.001
Sharma, N. K., Rai, A. K., Singh, S., & Brown, R. M. (2007). Airborne algae: Their present status and relevance. *Journal of Phycology*, *43*(4), 615–627. doi:10.1111/j.1529-8817.2007.00373.x
Siver, P. A. (1989). The distribution of scaled chrysophytes along a pH gradient. *Canadian Journal of Botany*. doi:10.1139/b89-268
Siver, P. A., Jo, B. Y., Kim, J. I., Shin, W., Lott, A. M., & Wolfe, A. P. (2015). Assessing the evolutionary history of the class Synurophyceae (Heterokonta) using molecular, morphometric, and paleobiological approaches. *American Journal of Botany*, *102*(6), 921–941. doi:10.3732/ajb.1500004
Siver, P. A., Lott, A. M., & Wolfe, A. P. (2009). Taxonomic significance of asymmetrical helmet and lance bristles in the genus Mallomonas (Synurophyceae) and their discovery in Eocene lake sediments. *European Journal of Phycology*, *44*(4), 447–460. doi:10.1080/09670260902903061
Siver, P. A., Lott, A. M., & Wolfe, A. P. (2013). A summary of Synura taxa in early Cenozoic deposits from Northern Canada. *Nova Hedwigia*, (September), 181–190.
Siver, P. A., & Wolfe, A. P. (2005). Eocene Scaled Chrysophytes with Pronounced Modern Affinities. *International Journal of Plant Sciences*, *166*(3), 533–536. doi:10.1086/428702
Siver, P. A., & Wolfe, A. P. (2009). Tropical Ochrophyte Algae From the Eocene of Northern Canada: a Biogeographic Response To Past Global Warming. *Palaios*, *24*(3), 192–198. doi:10.2110/palo.2007-p07-077r
Siver, P. A., Wolfe, A. P., Rohlf, F. J., Shin, W., & Jo, B. Y. (2013). Combining geometric morphometrics, molecular phylogeny, and micropaleontology to assess evolutionary patterns in Mallomonas (Synurophyceae: Heterokontophyta). *Geobiology*, *11*(2), 127–138. doi:10.1111/gbi.12023
Škaloud, P., Kristiansen, J., & Škaloudová, M. (2013). Developments in the taxonomy of silica-scaled

chrysophytes - From morphological and ultrastructural to molecular approaches. *Nordic Journal of Botany*, *31*(4), 385–402. doi:10.1111/j.1756-1051.2013.00119.x

Škaloud, P., & Rindi, F. (2013). Ecological differentiation of cryptic species within an asexual protist morphospecies: A case study of filamentous green alga *Klebsormidium* (Streptophyta). *Journal of Eukaryotic Microbiology*, *60*, 350–362. doi:10.1111/jeu.12040

Škaloud, P., Škaloudová, M., Procházková, A., & Němcová, Y. (2014). Morphological delineation and distribution patterns of four newly described species within the *Synura petersenii* species complex (Chrysophyceae, Stramenopiles). *European Journal of Phycology*, *49*(2), 213–229. doi:10.1080/09670262.2014.905710

Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation. *Evolution*, *64*, 295–315. doi:10.1111/j.1558-5646.2009.00877.x

Sörhannus, U., Fenster, E. J., Hoffman, A., & Burckle, L. (1991). Iterative evolution in the diatom genus Rhizosolenia Ehrenberg. *Lethaia*. doi:10.1111/j.1502-3931.1991.tb01178.x

Souffreau, C., Vanormelingen, P., Verleyen, E., Sabbe, K., & Vyverman, W. (2010). Tolerance of benthic diatoms from temperate aquatic and terrestrial habitats to experimental desiccation and temperature stress. *Phycologia*, *49*(4), 309–324. doi:10.2216/09-30.1

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*, vey016. doi:10.1093/ve/vey016

Svoboda, P., Lindström, E. S., Ahmed Osman, O., & Langenheder, S. (2018). Dispersal timing determines the importance of priority effects in bacterial communities. *ISME Journal*, *12*, 644–646. doi:10.1038/ismej.2017.180

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729. doi:10.1093/molbev/mst197

Theriot, E. C., Fritz, S. C., Whitlock, C., & Conley, D. J. (2006). Late Quaternary rapid morphological evolution of an endemic diatom in Yellowstone Lake, Wyoming. *Paleobiology*. doi:10.1666/02075.1

Thornhill, D. J., Lewis, A. M., Wham, D. C., & Lajeunesse, T. C. (2014). Host-specialist lineages dominate the adaptive radiation of reef coral endosymbionts. *Evolution*, *68*(2), 352–367. doi:10.1111/evo.12270

Urban, M. C., & De Meester, L. (2009). Community monopolization: Local adaptation enhances priority effects in an evolving metacommunity. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 4129–4138. doi:10.1098/rspb.2009.1382

Wallis, G. P., Waters, J. M., Upton, P., & Craw, D. (2016). Transverse Alpine Speciation driven by glaciation. *Trends in Ecology and Evolution*, *31*, 916–926. doi:10.1016/j.tree.2016.08.009

Weisse, T., Berendonk, T., Kamjunke, N., Moser, M., Scheffel, U., Stadler, P., & Weithoff, G. (2011). Significant habitat effects influence protist fitness: evidence for local adaptation from acidic mining lakes. *Ecosphere*, *2*(12), 134. doi:10.1890/ES11-00157.1

Whittaker, K. A., Rignanese, D. R., Olson, R. J., & Rynearson, T. A. (2012). Molecular subdivision of the marine diatom *Thalassiosira rotula* in relation to geographic distribution, genome size, and physiology. *BMC Evolutionary Biology*, *12*, 209. doi:10.1186/1471-2148-12-209

Wickham, H. (2017). scales: Scale Functions for Visualization. *R Package Version 0.5.0*. doi:10.1371/journal.pone.0091650

Figure captions:

FIGURE 1 Distribution of two main *S. sphagnicola* lineages. Five regions are enlarged in inset figures for better clarity.

FIGURE 2 Phylogenetic relationships and timing of divergence. (a) Time-calibrated phylogeny for the Synurales based on concatenated SSU rDNA and *rbc*L sequences. The lineages sp1 and sp2 are given in blue and green, respectively. (b) Statistical parsimony networks of four sp1 and three sp2 haplotypes based on the analysis of three loci sequences. Circle sizes are proportional to haplotype frequency. Lines between haplotypes are single mutational steps; small dots indicate missing haplotypes. (c) Time-calibrated coalescence species tree with the visualisation of the most probable topologies. In (a) and (c), mean divergence times are given for selected nodes, along with 95% highest posterior density (HPD) values in square brackets. HPD values are illustrated for each node by blue bars. Time axis is Mya, along with chronological dating of geologic intervals.

FIGURE 3 Ecology and distribution of seven sp1 and sp2 haplotypes. Differences in the distribution of seven *S. sphagnicola* haplotypes along the gradient of (a) mean diurnal temperature range (BIO2) and (b) precipitation of wettest month (BIO13). (c) Distribution patterns in five investigated regions. Multiple samples at the same location are presented by arrows.

FIGURE 4 Venn's diagrams showing the relative effects of climate, habitat and geographical distance on the variance in *Synura* diversity. Areas not covered by coloured ellipses represent the portion of unexplained variability. (a) Differences between the sp1 and sp2 lineages. (b) Differences among the sp1 haplotypes. (c) Differences among the sp2 haplotypes. Significant net effects are indicated by asterisks (* $P < 0.05$, ** $P < 0.01$). Values lower than 1% are not shown.

FIGURE 5 Morphological analyses. (a) A colony of *S. sphagnicola* shown in light microscope. Note the layer of scales encircling the cells. (b) A silica scale as shown in transmission electron microscope. (c) Principal Component Analysis (PCA) ordination of 810 silica scales showing the morphological distinction of the lineages sp1 and sp2. The scales are clustered based on their affiliation to 15 investigated strains, and their haplotype identity (if known) and geographical origin is provided by symbols (see Figure 4) and country acronyms (CZ – Czech Republic, KR – Korea, NO – Norway, SW – Sweden, UK – United Kingdom). (d) The distribution of spine lengths in sp1 and sp2 lineages.