



RESEARCH PAPER

Journal of
Biogeography

WILEY

Phylogenetic diversity and environment form assembly rules for Arctic diatom genera—A study on recent and ancient sedimentary DNA

Kathleen R. Stoof-Leichsenring¹ | Luidmila A. Pestryakova² | Laura S. Epp¹ |
Ulrike Herzschuh^{1,3,4}

¹Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Polar Terrestrial Environmental Systems, Potsdam, Germany

²Department for Geography and Biology, North-Eastern Federal University of Yakutsk, Yakutsk, Russia

³Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

⁴Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

Correspondence

Kathleen R. Stoof-Leichsenring, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Polar Terrestrial Environmental Systems, Telegrafenberg A45, 14473 Potsdam, Germany.
Email: kathleen.stoof-leichsenring@awi.de

Present address

Laura S. Epp, University of Konstanz, Limnological Institute, Mainaustrasse 252, 78464, Konstanz, Germany

Funding information

Federal Ministry of Education and Research, Grant/Award Number: 5.2711.2017/4.6; Russian Foundation for Basic Research, Grant/Award Number: 18-45-140053; North-Eastern Federal University

Handling Editor: Jani Heino

Abstract

Aim: This study investigates taxonomic and phylogenetic diversity in diatom genera to evaluate assembly rules for eukaryotic microbes across the Siberian tree line. We first analysed how phylogenetic distance relates to taxonomic richness and turnover. Second, we used relatedness indices to evaluate if environmental filtering or competition influences the assemblies in space and through time. Third, we used distance-based ordination to test which environmental variables shape diatom turnover.

Location: Yakutia and Taymyria, Russia: we sampled 78 surface sediments and a sediment core, extending to 7,000 years before present, to capture the forest-tundra transition in space and time respectively.

Taxon: Arctic freshwater diatoms.

Methods: We applied metabarcoding to retrieve diatom diversity from surface and core sedimentary DNA. The taxonomic assignment binned sequence types (lineages) into genera and created taxonomic (abundance of lineages within different genera) and phylogenetic datasets (phylogenetic distances of lineages within different genera).

Results: Contrary to our expectations, we find a unimodal relationship between phylogenetic distance and richness in diatom genera. We discern a positive relationship between phylogenetic distance and taxonomic turnover in spatially and temporally distributed diatom genera. Furthermore, we reveal positive relatedness indices in diatom genera across the spatial environmental gradient and predominantly in time slices at a single location, with very few exceptions assuming effects of competition. Distance-based ordination of taxonomic and phylogenetic turnover indicates that lake environment variables, like HCO_3^- and water depth, largely explain diatom turnover.

Main conclusion: Phylogenetic and abiotic assembly rules are important in understanding the regional assembly of diatom genera across lakes in the Siberian tree line ecotone. Using a space-time approach we are able to exclude the influence of geography and elucidate that lake environmental variables primarily shape the assemblies. We conclude that some diatom genera have greater capabilities to adapt to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Journal of Biogeography published by John Wiley & Sons Ltd

environmental changes, whereas others will be putatively replaced or lost due to the displacement of the Arctic tundra biome under recent global warming.

KEYWORDS

ancient sedimentary DNA, Arctic lakes, assembly rules, climate change, diatoms, environmental filtering, phylogenetic diversity, Siberian tree line

1 | INTRODUCTION

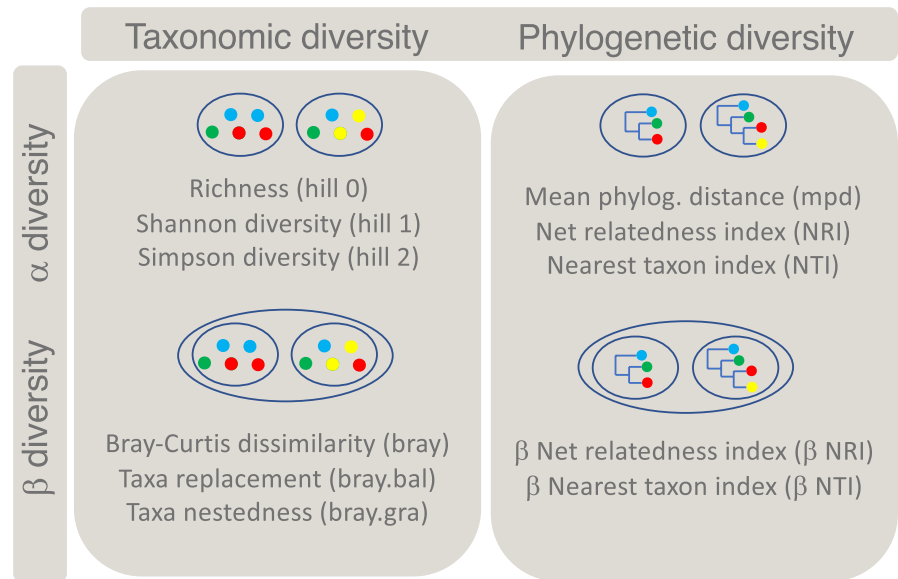
Assembly rules constrain the formation of community composition out of a pool of available species. Exploring which and how intensely assembly rules affect species composition helps to understand past, recent and prospective community composition. Particularly in times of recent global warming it is important to understand how community composition will change under human-induced, rapid environmental change compared to natural environmental variability (Bellard, Bertelsmeier, Leadley, Thuiller, & Courchamp, 2012; Davey & Blaxter, 2010; Thakur et al., 2017). In environmentally sensitive Arctic areas, the magnitude of biotic change and its consequences on ecosystem services (Frainer et al., 2017; Vincent et al., 2012) require special attention. Globally, too little awareness has been paid to the effects of global warming on microbial life (Cavicchioli et al., 2019; Dutta & Dutta, 2016; Heger et al., 2014), which presents an urgent need to better understand the rules by which microbial assemblies are influenced and sustained. The term 'assembly rules' was introduced by Diamond in 1975 and different terminologies have developed since (Götzenberger et al., 2012). In our study we categorize assembly rules into (a) phylogenetic factors, which are determined by historical diversification and speciation of taxa, (b) abiotic (environmental characteristics of the habitat) and biotic (interactions between co-occurring taxa) factors and (c) spatial factors (geographic distance). These influences of the assembly rules on taxa composition can be elucidated by looking at the diversity patterns of communities. Alpha- and beta-diversity measures can be obtained from taxonomic diversity (based on the number of taxa) as well as from phylogenetic diversity (based on the phylogenetic distance between taxa; Figure 1). Assembly rules in minute life-forms such as diatoms are less widely explored than in plants (Echeverría-Londoño et al., 2018; Kraft et al., 2015; Swenson, 2011) or animals (Arnan, Cerdá, & Retana, 2017) and only a few empirical studies exist (Bennett, Cumming, Ginn, & Smol, 2010; Bottin, Soininen, Alard, & Rosebery, 2016; Rimet et al., 2016; Stooft-Leichsenring et al., 2015). It is known that diatoms are unicellular eukaryotic minute life-forms and environmentally highly sensitive (Alverson, 2008). Thus, most diatom studies find environment to be the major factor determining diatom communities at different spatial scales by decomposing the variation in taxonomic community to spatial- versus environmental-related variation (Bennett et al., 2010). Although there is the expectation that diatoms can freely disperse (Soininen, 2007), other studies on global diatom distribution find biogeographic (geographic and evolutionary) factors strongly contribute to diatom community

structure (Vyverman et al., 2007). In general, we have to assume that similar assembly rules known from plants act on unicellular microbes such as diatoms (Heino & Soininen, 2005). Because it is known that morphologically similar diatoms can bear hidden genetic diversity (Pinseel et al., 2019), adding phylogenetic to taxonomic diversity will enable us to discern assembly rules in a more robust way (Van den Wyngaert, Most, Freimann, Ibelings, & Spaak, 2015).

In general, the relative importance of assembly rules to community structure can be assessed by null model approaches (deviation of observed from random patterns). These are used to examine if the observed co-occurrence of species in a community associate with the spatial, environmental or biotic similarity of the investigated locations. To evaluate the relative impact of environment in comparison to the role of geography, it would be advantageous to analyse diversity patterns along ecological gradients that provide different ecological niches over a geographical space and at a fixed place back in time with different ecological periods at one geographical site. Under the assumption that evolutionary relatedness is positively correlated with similar ecological requirements of species, analyses of phylogenetic diversity allow us to infer the relative importance of environmental factors and biotic interactions by comparing the observed diversity with the total diversity in the present assemblies (Webb, Ackerly, & Kembel, 2008; Webb, Ackerly, McPeck, & Donoghue, 2002). A phylogenetic clustering, meaning more closely related taxa in a community than expected by random null communities, is probably related to environmental filtering, whereas phylogenetically even or over-dispersed communities can be a result of taxa competition for habitats leading to the exclusion of phylogenetically similar taxa (Webb et al., 2002). Because the composition (e.g. species richness and phylogenetic diversity) of intrageneric diversity is generally characterized by species that share ecologically similar requirements due to their close phylogenetic relatedness, the analyses of generic rather than total community diversity is more suitable to elucidate the impact of either competitive exclusion or environmental filtering (Voskamp, Baker, Stephens, Valdes, & Willis, 2017; Webb et al., 2002).

The Siberian tree line presents a huge ecological gradient from tundra to boreal vegetation with embedded Arctic lake systems and sediment archives of tree line lakes preserve environmental changes and diatom communities of the past. Morphological investigations have attributed Holocene diatom variations to vegetation changes in the lake catchments (Herzschuh et al., 2013; Pestryakova, Herzschuh, Wetterich, & Ulrich, 2012). A few genetic studies from the area find that the occurrence of intrageneric diatom lineages is

FIGURE 1 Overview of taxonomic and phylogenetic diversity metrics used in the study based on Bray–Curtis dissimilarity partitioning (phylog. = phylogenetic)



linked to environmental conditions (i.e. vegetation in the catchment) rather than geographical constraints (Stoof-Leichsenring et al., 2014, 2015). Because species of a genus are generally composed of ecologically similar taxa, addressing diversity within genera will presumably result in stronger competition between taxa of the same genus than taxa of different genera. Adding a community's phylogenetic information would then allow us to speculate on the influence of competition over environmental filtering in intrageneric lineages.

In this case, metabarcoding approaches on sedimentary DNA using phylogenetically informative marker genes facilitate the analyses of phylogenetic diversity in natural communities (Nanjappa, Audic, Romac, Kooistra, & Zingone, 2014), although ancient sedimentary DNA in particular allows the retrieval of only very short marker genes and thus limited phylogenetic content (Drummond et al., 2015; Stoof-Leichsenring et al., 2014). The large subunit of the Ribulose-1,5-bisphosphate carboxylase/oxygenase (rbcL) has been used frequently for phylogenetic analyses (Guo, Sui, Zhang, Ren, & Liu, 2015; Theriot, 2010), but is most powerful in determining lower taxonomic levels and thus has been used to identify intrageneric and intraspecific genetic diversity and diversification (Abarca, Jahn, Zimmermann, & Enke, 2014; Evans, Wortley, & Mann, 2007; Hamsher, Evans, Mann, Poullickova, & Saunders, 2011; Kermarrec, Bouchez, Rimet, & Humbert, 2013). Moreover, short rbcL metabarcodes have been successfully applied for diatom diversity assessment (Dulias, Stoof-Leichsenring, Pestryakova, & Herzsuh, 2017; Rimet, Vasselon, A.-Keszte, & Bouchez, 2018). Combining taxonomic with phylogenetic diversity retrieved from sedimentary DNA archives is a fairly new approach. It offers the opportunity to directly analyse taxonomic and phylogenetic diversity in space as well as through time and allows the analysis of community shifts at one geographic place under changing environmental conditions. It can help evaluate the effect of competition between lineages depending on their phylogenetic relatedness.

This study investigates assembly rules for diatom genera across the vegetation gradient of the Siberian tree line and via Holocene

time slices of a lake sediment core, which experienced vegetation changes through time. We combine taxonomic diversity (binning lineages/sequence types into taxonomically defined genera) and phylogenetic diversity obtained from a short phylogenetically informative chloroplast marker specified for metabarcoding using lake sediments. With our study we aim to evaluate the assembly rules that structure the community composition of diatoms at the intrageneric level. First, we investigate how phylogenetic factors (phylogenetic distance in diatom genera) affect taxonomic richness and turnover in diatom genera. Second, we attempt to discern if abiotic or biotic factors (environmental filtering or intrageneric competition) or geography primarily influence the assembly by comparing taxonomic and phylogenetic community signals along a similar ecological gradient in space as well as through time. Third, we investigate the extent to which environmental variables and geographic distance explain taxonomic and phylogenetic turnover.

2 | MATERIAL AND METHODS

2.1 | Material and sampling locations

Sediment and water samples were collected during joint field expeditions of the Alfred Wegner Institute (AWI) and North-Eastern Federal University of Yakutsk (NEFU) from 78 lakes along a latitudinal transect in the northern lowlands of Yakutia and Taymyria, Russia. The localities range from the edge of the Arctic tundra over forest tundra to boreal forest (Figure 2). The lakes investigated were formed by thermokarst processes and are characterized by their small size and catchment area and water chemistry parameters (Table S1 in Appendix S1). Surface sediments were sampled using a bottom sampler and subsamples (mainly the first centimetre) were taken with a sterile spatula and stored in sterile plastic tubes until further analyses. The sediment core 11-CH-12A was sampled during an expedition to Chatanga in 2011 and details about sampling and

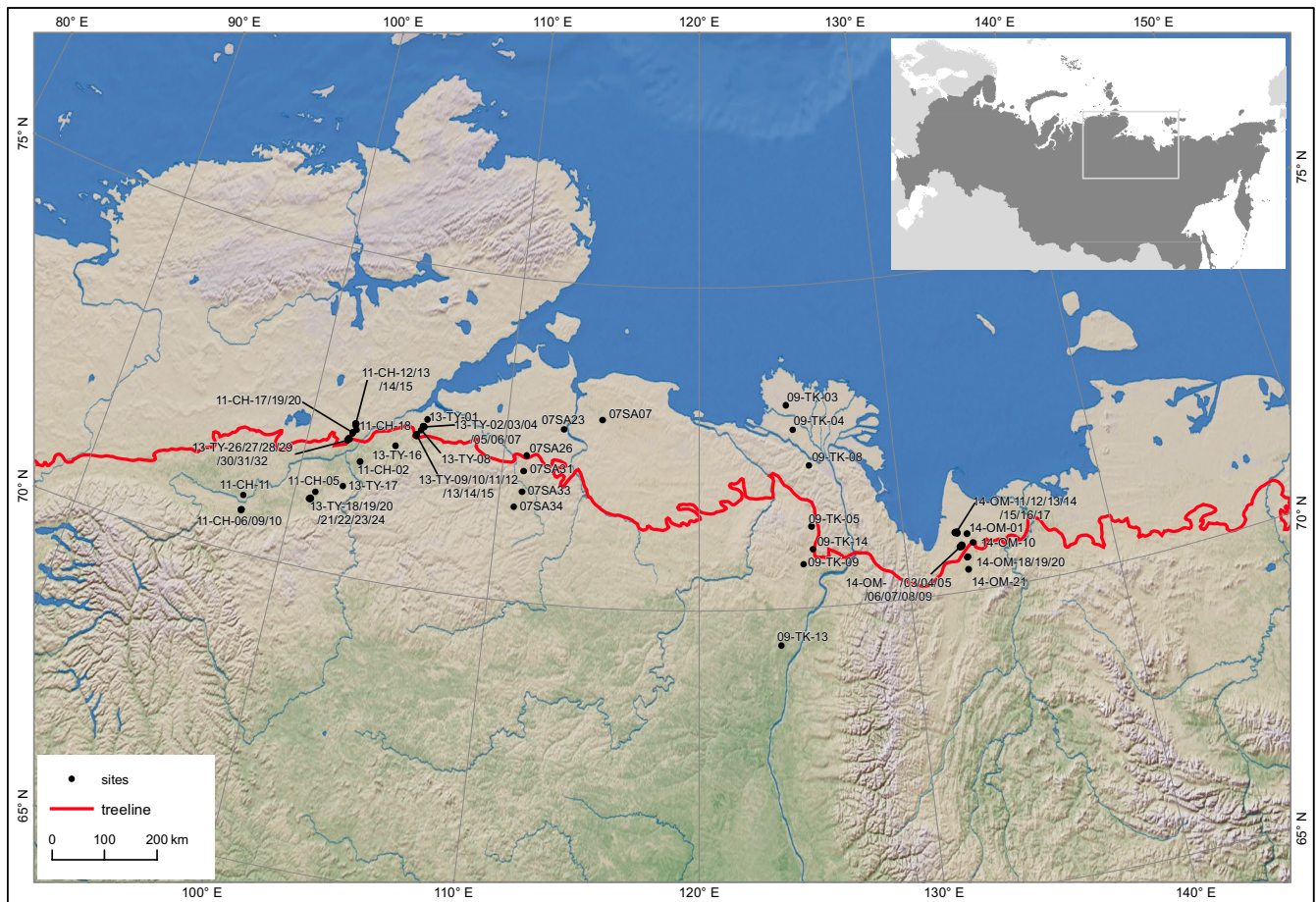


FIGURE 2 Map of the 78 lake sampling sites across the tree line ecotone in northern Siberia. The map was created with Esri Arc GIS Version 10.2 and Natural Earth Background Map

core chronology are described in Stooft-Leichsenring et al. (2015). Based on pollen analyses, we infer vegetation turnover from the warmer and forested mid-Holocene period to arctic tundra at the present day (Klemm, Herzschuh, & Pestryakova, 2016).

2.2 | Genetic assessment of diversity in diatom genera

Surface sediments were isolated according to the descriptions in Stooft-Leichsenring et al. (2015), except for the extraction of surface samples from Taymyr (TY) and Omoloy (OM) lakes, which were isolated with the same method but in the genetic laboratories at Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research in Potsdam. Core sediment samples (11-CH-12A) were handled in the dedicated ancient DNA laboratory at Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research in Potsdam, specially equipped for ancient DNA work. Precise descriptions of the DNA extraction, PCR amplification of the short *rbcL* 76bp metabarcode and Illumina Next-Generation-Sequencing are given in the (Appendix S1). For bioinformatic analyses of obtained sequence reads, we used the OBITools pipeline, which includes a taxonomic assignment of sequence types with the *ecotag* algorithm (Boyer et

al., 2016). The complete sequencing data were obtained from two independent Illumina sequencing runs and were joined after the OBITools pipeline and further manually filtered (for details see Appendix S1). After the filtering steps, we kept only sequence types which were taxonomically identified to diatom genus level and those genera which occurred in at least 20 of 78 lakes. Similar guidelines were used for the core data, but genera were kept if they occurred in 7 of 10 core samples. Then, we split up the entire dataset into single genus datasets. The final dataset for surface samples of the lakes comprises 19 different diatom genera, meaning 19 subdatasets, whereas the core dataset contains 9 diatom genera resulting in 9 subdatasets.

2.3 | Taxonomic and phylogenetic diversity metrics

The taxonomic diversity dataset consists of sequence types that are binned into taxonomically defined genera, whereas the phylogenetic diversity datasets are based on the phylogenetic distances of lineages within the identified genera. We tested the suitability of the 76bp metabarcode to assess proper phylogenetic distances within genera by comparing pairwise distances of reference sequences covering the short (76bp) and three longer *rbcL* fragments



(1398bp, 939bp and 365bp). The 1398bp rbcL fragment have been successfully used to recover phylogenetic relationships in the genus *Sellaphora* (Evans et al., 2007; Vanormelingen, Evans, Chepurnov, Vyverman, & Mann, 2013) and a slightly shorter fragment was used to demonstrate phylogenetic relationships in *Aulacoseira* (Edgar & Theriot, 2004). Restricted by the available reference sequences, we could test intrageneric distances in 4 genera (*Sellaphora*, *Nitzschia*, *Navicula* and *Cyclotella*) by comparing the 1398bp and the 76bp rbcL barcode. We used 11 genera for comparing the 939bp with the 76bp rbcL barcode and 16 genera in the comparison between the barcodes 365bp and 76bp (Table S2 in Appendix S1). The comparison of pairwise phylogenetic distances between reference sequences of the short (76bp) and three longer overlapping rbcL fragments revealed significant correlations for all the tested datasets (Table S2), except datasets *Asterionella* (76bp vs. 365bp, $N = 6$) and *Melosira* (76bp vs. 939bp, $N = 3$) showed no significant correlation, which is most likely due to the low number of available reference sequences (Table S2 in Appendix S1).

We applied taxonomic and phylogenetic alpha- and beta-diversity metrics (Figure 1) on the generic datasets. We calculated Simpson's diversity (hill 2) as a measure of taxonomic alpha diversity within genera using the R package 'vegetarian' (Charney & Record, 2012). The Simpson diversity gives an estimate of the effective number of species (with down-weighting of rare taxa), which, in our study, is the number of different lineages within a genus. Prior to the analyses of the effective number of taxa in each genus, the datasets were rarefied to the minimum number of counts per genus (surface samples = 1,010, core samples = 391). To obtain the taxonomic beta diversity we used R package 'betapart' and calculated the Bray-Curtis dissimilarity to quantify the total taxonomic turnover (bray) for each genus and the proportion of replacement of taxa (bray.bal), as well as abundance gradient components between two sites (Baselga, 2012) and calculated the overall mean of each metric per genus. Prior to analyses of taxonomic beta diversity, we rarefied surface and core data to the minimum sample count (surface samples = 2,899, core samples = 94,899) and resampled the data to create rarefied datasets which were used for the subsequent statistical analyses. For rarefaction analyses we used the R script (https://github.com/StefanKruze/R_Rarefaction).

For phylogenetic alpha diversity within the genera we calculated the uncorrected phylogenetic distance (pdist) using Geneious® 11.1.5 and the mean phylogenetic distance (mpd) and mean nearest taxon distance (mntd) based on abundance weighted and presence/absence data respectively. The calculation of the phylogenetic distances between lineages within one genus was conducted using the R package 'phangorn' (Schliep, 2011) by estimating the tree branch lengths obtained from a maximum likelihood tree (using *pml* function with GTR substitution model) derived from phylogenetic distances (function *dist.ml*) between the sequence types, which were saved in fasta format. Phylogenetic distances were then combined with the abundance data or presence/absence data of the sequence types within each genus by using the function *match.phylo.comm* and mpd and mntd values were calculated with the functions *ses.mpd* and

ses.mntd respectively. Resulting *ses.mpd* and *ses.mntd* values were converted into net relatedness index (NRI) and nearest taxon index (NTI), respectively, by multiplying by -1 (using R package 'picante'; Kembel et al., 2010). NRI and NTI values in single locations and core slices were calculated for the five most abundant genera in surface and core samples. NRI measures whether the sum of the pairwise phylogenetic distances among all pairs of taxa (considering differences in higher taxonomic levels) in the community is more or less than expected from randomly generated null communities, assuming that positive NRI values indicate environmental clustering, while negative values indicate even or over-dispersed communities, which could be attributed to exclusive competition between co-occurring taxa. NTI measures whether most closely related co-occurring taxa (mostly within species/genus relationships) in a community are more or less closely related than expected and follows the same interpretation as NRI. Furthermore, we used the *comdist* function in the R package 'picante' (Kembel et al., 2010) to calculate phylogenetic beta diversity (betaNRI and betaNTI), that is, the phylogenetic distances between pairs of samples weighted for the abundance of sequence types and also with presence/absence data. Correlation analyses and Loess regression were used to evaluate relationships between the diversity indices and were done in the R package 'ggplot2' (Wickham, 2016). Distance-based redundancy analyses (dbRDA) were run to estimate the influence of geography and environment on the distance matrix derived from the beta-diversity calculations for the taxonomic (total turnover and bray) and phylogenetic datasets. Hence, co-ordinates of the sampling localities were transformed to rectangular data suitable for ordination methods using principal co-ordinates of neighbourhood matrix (PCNM), with the function *pcnm*. Environmental variables were $\log(x + 1)$ transformed using the *log1p* function and were, together with the principal coordinates, used as constraints for the dbRDA with the dissimilarity matrices originating from the taxonomic and phylogenetic datasets using the function *capscale* and a subsequent stepwise forward model selection with the function *ordistep*. The selected constraints were then applied to dbRDA and their significance assessed with the function *permutest*. The calculations regarding dbRDA were performed using the R package 'vegan' (Oksanen et al., 2019) and all computations were conducted with RSTUDIO version 1.1.456.

3 | RESULTS

3.1 | Taxonomic and phylogenetic composition

The joined dataset of the two Illumina sequencing runs and subsequent filter steps resulted in 19 verified genera each containing between 6 and 156 sequence types for the 78 lake surface samples and 9 verified genera containing between 5 and 87 sequence types for the 10 core samples. The sequence types within a genus consist of intraspecific to intrageneric sequence types, as indicated by their sequence match to reference sequences, which varied between 100% and 90%.

Our results demonstrate strong differences in abundance, richness, turnover and phylogenetic distance in diatom genera. The total modern surface dataset consists of 7,534,170 sequence reads, whereof 95% are assigned to *Staurosira*, *Aulacoseira*, *Pinnularia*, *Melosira*, *Sellaphora* and *Stauroneis*; the remaining 13 genera represent <5% of total reads. We identified 986 unique sequence types and highest richness occurs in *Stauroneis*, *Staurosira*, *Aulacoseira* and *Sellaphora*, whereas lowest richness is present in *Melosira*, *Caloneis*, *Ellerbeckia* and *Tabellaria*. The taxonomic beta diversity (full replacement of taxa; bray.bal) is highest *Nitzschia* and *Neidium*, whereas *Tabellaria* and *Ellerbeckia* show least turnover. In terms of phylogenetic distances, *Ellerbeckia*, *Cyclotella* and *Staurosira* consist of phylogenetically closely related lineages (uncorrected mean pdist varies

between 1.8 and 3.9 nucleotide differences), whereas lineages in *Nitzschia*, *Stauroneis* and *Caloneis* are most distantly related (uncorrected mean pdist varies between 4.5 and 12 nucleotide differences).

The total core dataset consist of 3,613,450 sequence reads and 234 unique sequence types. About 76% of sequence reads are attributed to *Staurosira* and *Cyclotella*, whereas the richness is highest in *Staurosira*, *Aulacoseira* and *Cyclotella*. *Sellaphora* and *Stauroneis* show highest turnover, whereas *Tabellaria* and *Asterionella* represent genera with low turnover throughout the core slices. Phylogenetically closest lineages are present in *Tabellaria*, *Cyclotella* and *Staurosira* (pdist varies between 2.8 and 3.8), whereas *Stauroneis*, *Sellaphora* and *Pinnularia* comprise most distantly related lineages. The detailed results for all genera are presented in Table S3 in Appendix S1.

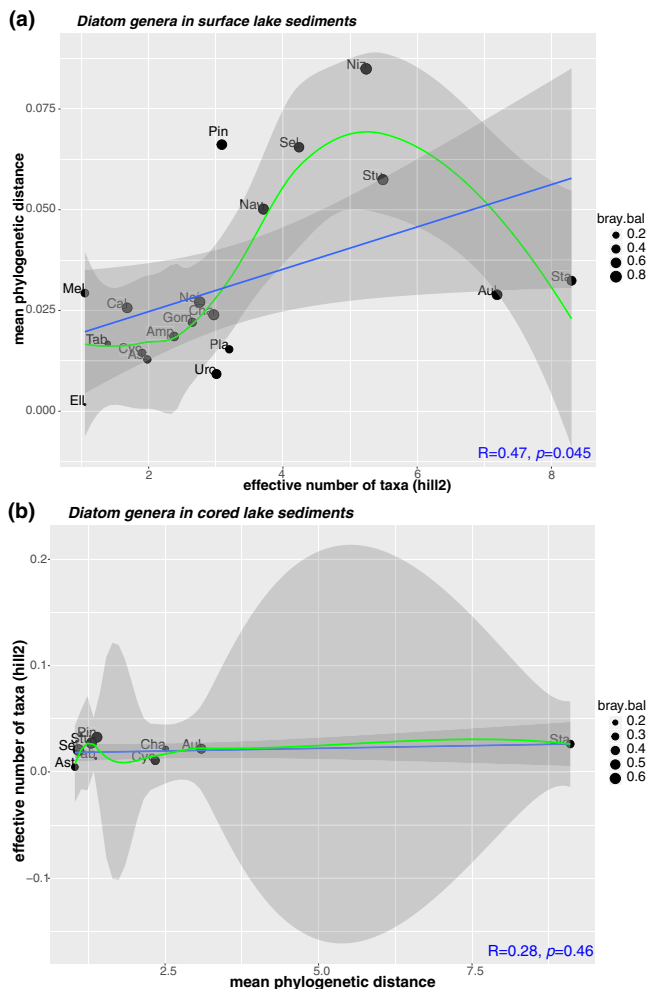


FIGURE 3 Relationship between phylogenetic (mean phylogenetic distance (mpd, abundance.weighted = TRUE) and taxonomic alpha diversity (Simpson diversity, effective number of taxa = hill 2) of (a) 19 diatom genera obtained from 78 lake surface samples and (b) 9 diatom genera from 10 slices of the lake sediment core 11-CH-12A. For Loess regression (green curve) we used a 0.5 smoothing factor. Amp, Amphora; Ast, Asterionella; Aul, Aulacoseira; Cal, Caloneis; Cha, Chaetoceros; Ell, Ellerbeckia; Gom, Gomphonema; Mel, Melosira; Nav, Navicula; Nei, Neidium; Niz, Nitzschia; Pin, Pinnularia; Pla, Planothidium; Sel, Sellaphora; Sta, Staurosira; Stu, Stauroneis; Tab, Tabellaria; Uro, Urosolenia

3.2 | Phylogenetic distance and taxonomic richness and turnover in diatom genera

Our results demonstrate that the variation in diatom genera with regard to taxonomic richness and turnover is influenced by the mean phylogenetic distance (mpd) of genera. We show that genera with more distantly related lineages (high mpd) have high richness (=high effective number of lineages, hill 2) and greatest turnover (bray.bal) across the investigated ecotone compared with genera consisting of closely related lineages, which have fewer effective taxa and show a lower turnover. Two genera (*Aulacoseira* and *Staurosira*) consist of closely related lineages (moderate mpd) with a high number of effective lineages but a lower turnover across the sampled area (Figure 3a). This unimodal relationship between the phylogenetic distance and richness of diatom genera indicates the effect of counteracting factors between phylogenetic diversity and richness. No relationship could be identified between mpd and the effective number of lineages derived from core samples (Figure 3b). Only the genus *Staurosira* is notable as it is characterized by a moderate mpd but shows a much higher effective number of lineages than the other genera.

A clear positive relationship is revealed between the phylogenetic distance (mpd) and the turnover (bray.bal) of diatom genera. This means that genera with closely related sequence types show less turnover than genera consisting of more distantly related sequence types. This holds true for either abundance-weighted or presence/absence data (Table 1, Figure 4). The patterns discerned from the spatial data are partly mirrored in the core data, although the signal is weaker, probably due to the smaller number of genera and samples and because the mean phylogenetic distance is generally lower than from surface samples, as only a selection of sequence types occurs in the core locality compared with the larger spatial gradient. However, the positive relationship between mpd and bray.bal detected from surface data is also found for the core data and is most significant when using mpd values based on presence/absence rather than abundance-weighted data. Because the positive relationship between mpd and bray.bal is valid for both surface (spatial) and core (temporal) assemblies of diatom genera—that is, the turnover of taxa through time at a fixed place

TABLE 1 Results of Pearson correlation tests used to analyse the relationships between diatom's alpha- and beta-diversity metrics obtained from lake surface and core sediment samples

Diversity metrics	Lakes surface samples ($N_{\text{Genera}} = 19$)		Lake core samples ($N_{\text{Genera}} = 9$)	
	R	<i>p</i>	R	<i>p</i>
Alpha phylogenetic versus taxonomic diversity				
mpd.aw versus hill2	0.466	.045*	0.284	.458
mpd.pa versus hill2	0.437	.061*	-0.156	.688
mntd.aw versus hill2	-0.008	.973	-0.402	.284
mntd.pa versus hill2	0.041	.868	-0.372	.325
Alpha taxonomic versus beta taxonomic diversity				
hill2 versus bray.bal	0.523	.021*	-0.157	.687
Alpha phylogenetic versus beta taxonomic diversity				
mpd.aw versus bray.bal	0.662	.002**	0.545	.125
mpd.pa versus bray.bal	0.638	.003**	0.725	.027*
mntd.aw versus bray.bal	0.424	.070*	0.648	.059*
mntd.pa versus bray.bal	0.483	.036*	0.653	.056*

Note: *p* values are coded in ***0.001, **0.01, *0.05, •0.1; mpd.aw, mean phylogenetic distance (method = abundance weighted); mpd.pa, mean phylogenetic distance (method = presence/absence); mntd.aw, mean nearest taxon distance (method = abundance weighted); mntd.pa, mean nearest taxon distance (method = presence/absence); hill2, Simpson diversity index (effective number of taxa); bray.bal, proportion of taxa replacement from total taxonomic turnover.

is similar to turnover in space—we presume that geography has less influence on the regional community assembly than environmental or biotic assembly rules (Figure 4).

3.3 | Relatedness indices in diatom genera from surface and core samples

To evaluate the impact of environmental and biotic assembly rules (e.g. competition) within the five most abundant diatom genera (*Stauroneis*, *Aulacoseira*, *Pinnularia*, *Sellaphora* and *Stauroneis*) we calculated the alpha phylogenetic metrics NRI and NTI for all surface and core-slice samples (Figure 5). In most of the samples from the genera datasets we find positive NRI and NTI values indicating clustering of closely related sequences, although only some exceed a significant deviation from null communities after Bonferroni correction (Figure 5). The genera *Stauroneis*, *Aulacoseira* and *Stauroneis* show a strong significant clustering signal in many investigated localities, whereas for *Pinnularia* and *Sellaphora* only

very few samples show this signal. There are no significant negative NRI or NTI values (after correction), suggesting that few distantly related lineages co-occur in the samples analysed. Only in the genera *Sellaphora* and *Stauroneis* do NTI values in two sample sites tend to show over-dispersed communities, supported by weak *p*-values.

Most of the genera obtained from core samples indicate a shift in NRI and NTI values across the different time slices attributed to vegetation changes over the last 7,000 years, but with only a few samples exceeding the uncorrected/corrected significance thresholds, indicating neither a strong clustering nor a signal of over-dispersed communities in the core samples. Exclusively, *Stauroneis* displays strongly significant positive values indicating a clustering of closely related sequences in most of the core samples.

3.4 | Taxonomic and phylogenetic turnover in diatoms

We related total taxonomic and phylogenetic beta diversity of surface samples to environmental variables and geographic distance. We used distance-based redundancy analysis to test if the explanatory variables (a combination of three lake environment variables: water depth, pH and hydrogen carbonate) and geographic components significantly contribute to the variation of taxonomic and phylogenetic turnover of the five selected genera obtained from the lake sediment-surface samples (Table 2). Taxonomic and phylogenetic turnover of the selected genera were always significantly correlated (see Mantel test results Table 2) and both metrics are partly explained by similar variables. However, phylogenetic turnover in *Stauroneis* is exclusively explained by HCO_3^- and water depth, whereas the other genera are explained by a combination of different variables. For nearly all analysed datasets, HCO_3^- and water depth are the predominant environmental variables which contribute to the turnover of selected diatom genera across the Siberian tree line ecotone, mostly coupled with spatial variables. Except the genus *Aulacoseira* is largely explained by geography and water depth.

4 | DISCUSSION

4.1 | The effect of phylogenetic distance on richness and turnover in diatom genera

Our study reveals a strong phylogenetic influence on the richness patterns of diatom genera. We identify one group of diatom genera (*Asterionella*, *Amphora*, *Caloneis*, *Cyclotella*, *Chaetoceros*, *Gomphonema*, *Melosira*, *Neidium*, *Planorhynchium*, *Tabellaria* and *Urosolenia*) characterized by low mean phylogenetic distance (mpd) and low richness and turnover across the Siberian tree line. In contrast, a second group of genera (*Navicula*, *Nitzschia*, *Pinnularia*, *Stauroneis* and *Sellaphora*) is characterized by high mpd, but with high richness and turnover. A

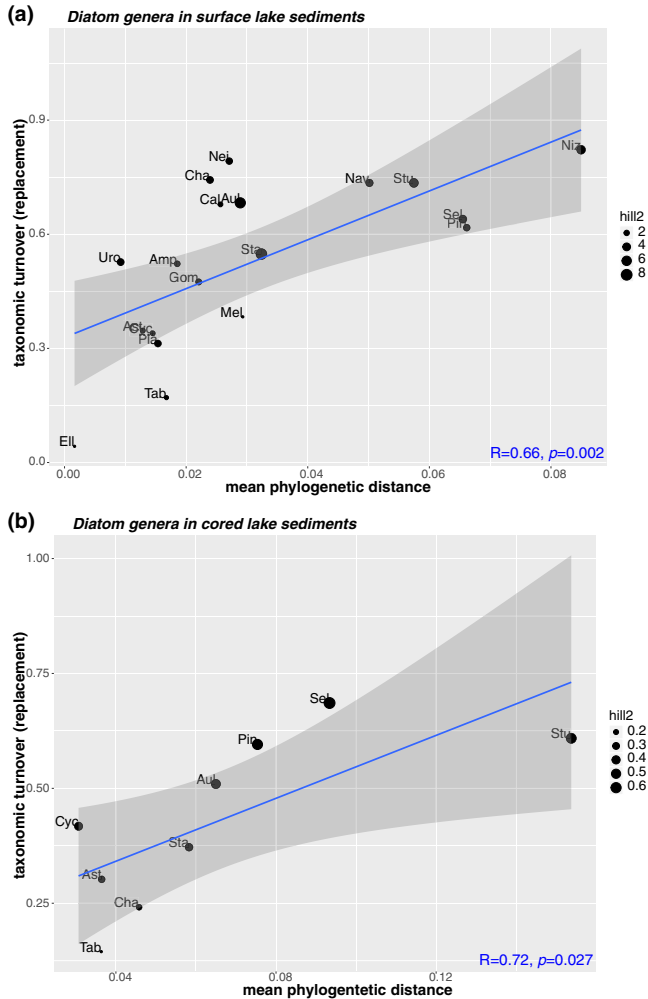


FIGURE 4 Relationship between the mean phylogenetic distance (method= abundance-weighted) and the taxonomic turnover (beta diversity) in terms of taxa replacements (bray.bal) of (a) 19 diatom genera obtained from 78 lake surface samples and (b) 9 diatom genera from 10 slices of the lake sediment core 11-CH-12A (method= presence/absence). Size of points indicates the Simpson diversity (effective number of taxa = hill2). Amp–Amphora; Ast, Asterionella; Aul, Aulacoseira; Cal, Caloneis; Cha, Chaetoceros; Ell, Ellerbeckia; Gom, Gomphonema; Mel, Melosira; Nav, Navicula; Nei, Neidium; Niz, Nitzschia; Pin, Pinnularia; Pla, Planothidium; Sel, Sellaphora; Sta, Staurosira; Stu, Stauroneis; Tab, Tabellaria; Uro, Urosolenia

third group is formed by the most abundant genera *Staurosira* and *Aulacoseira*, which are characterized by moderate mpd, but have the highest richness compared to the other two groups. Our data suggest that increasing mpd does not lead to a further increase in richness, which we hypothesize could be a result of ongoing diversification in the genus *Staurosira* and *Aulacoseira*. A similar antagonistic relationship has also been identified in microbes, in which increasing genetic richness in the absence of an increasing dissimilarity of genotypes in, for example, homogeneous habitats, has been attributed to reduced ecosystem functioning (Jousset, Schmid, Scheu, & Eisenhauer, 2011). The mpd of genera identified in the different time slices of the core samples do not show a relationship with richness, which could be a result of the reduced number of genera identified. This may be

caused by technical issues such as degradation of sedimentary DNA (Parducci et al., 2017) that might bias the retrieved genetic diversity, or dispersal limitation, which could have restricted the local genetic richness of diatoms in the past (Verleyen et al., 2009).

The mpd of diatom genera is significantly positively related to taxonomic turnover (full replacement of taxa) across the tree line ecotone (vegetation changes from tundra to forest). This relationship is also valid at a single location in one lake, which experienced climatic and related vegetation changes (as present along the spatial gradient) through time. We expected this relationship because genera with more distantly related sequence types are most probably adapted to different environments and thus would show greater turnover across the ecological gradient, as has been reported for microbial diversity (Jousset et al., 2011), whereas genera consisting of closely related sequence types tend to have similar requirements and thus show less turnover. This phenomenon could also explain the reduced turnover in *Staurosira* and *Aulacoseira* compared to other genera with more distantly related lineages, assuming that several closely related lineages complement each other and can persist in different ecological niches due to their incomplete differentiation across the tree line ecotone. Because habitat heterogeneity has been shown to affect the diversification of diatoms (Nakov, Beaulieu, & Alverson, 2018; Zorzal-Almeida, Soinin, Bini, & Bicudo, 2017), we assume similar processes affected *Staurosira* and *Aulacoseira* across the ecological gradient formed by the tree line.

Alternatively, genera with a high turnover along the spatial ecological gradient of the tree line could also result from geographically structured lineages due to historic events, for example, dispersal limitation (Verleyen et al., 2009; Vilmi, Karjalainen, & Heino, 2017). Because the relationship between mpd and turnover is also present in a sample series presenting past vegetation changes at one locality, our results support the assumption that mainly environmental factors shape the turnover in diatom genera across the ecotone, whereas the impact of geographic factors seems less important. This is also supported by the fact that geographic factors do not significantly explain the variance in, for example, *Staurosira* across the tree line. To fully rule out the impact of geography on diatom diversity across the tree line, more time series need to be investigated. However, in the core samples we identified only a significant correlation between mpd and turnover when using mpd values based on presence/absence data, which might be related to the degraded nature of ancient DNA-enhancing PCR bias, for example, and thus impacting the abundance of sequence types (Krehenwinkel et al., 2017).

4.2 | Environmental filtering shapes diatom assembly in space and through time

The phylogenetic relatedness of taxa that co-occur in a community can be used to discern whether the environment or biotic interaction is responsible for the community assemblage (Webb et al., 2002).

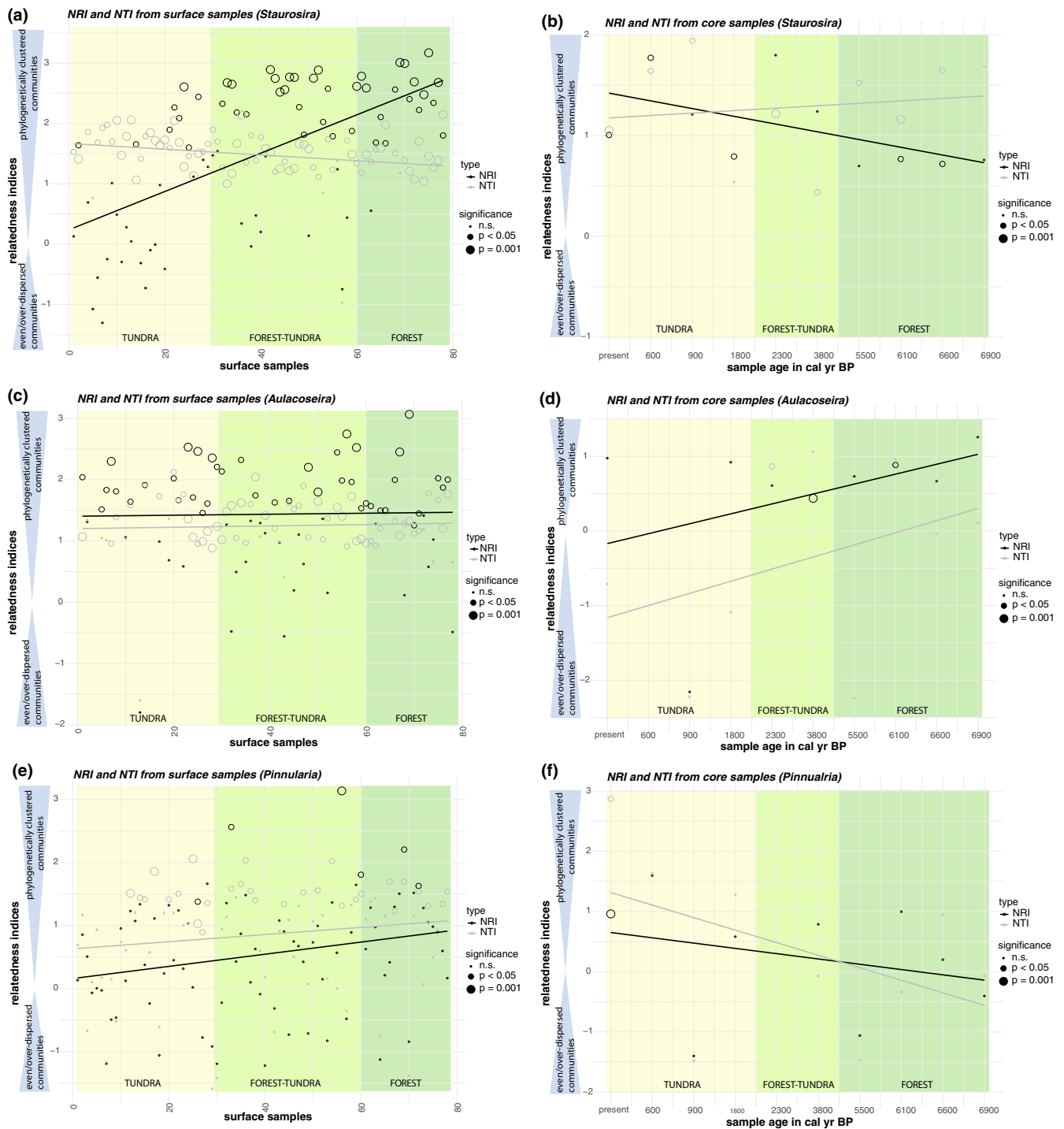


FIGURE 5 Scatter plots indicating the NRI (net relatedness index) and NTI (nearest taxon index) values for five selected genera, showing the standardized effect size of mean phylogenetic diversity (mpd) versus null communities using the method=abundance.weighted, which includes variation of sequence types abundance in the 78 lake sediment–surface samples across the Siberian tree line and in 10 core sediment samples (core 11-CH-12A). The background is coloured according to the dominant vegetation type in the vicinity of the sampled lakes (personal observation and vegetation plot analyses). Sample ages are in calibrated years Before Present (cal yr BP). Vegetation changes are reconstructed from pollen analyses (Klemm et al., 2016). Significance of relatedness indices is indicated by bubble sizes. $p = .001$ indicates significant deviation of observed mpd versus null communities after Bonferroni correction. Gaps indicate no available data. a/b – *Stausosira*, c/d – *Aulacoseira*, e/f – *Pinnularia*, g/h – *Sellaphora*, i/j – *Stauroneis*

Our study supports that the phylogenetic relatedness of lineages in five selected diatom genera (using NRI and NTI metrics) primarily indicates a clustering effect defined by the co-occurrence of closely related lineages in the investigated lake localities. This effect

is driven by environmental filtering resulting in the co-occurrence of lineages due to their similar ecological requirements. This signal tends to be strongest in the genus *Stausosira*, which could be a result of either multiple invasions of minute staurosroid diatoms into the

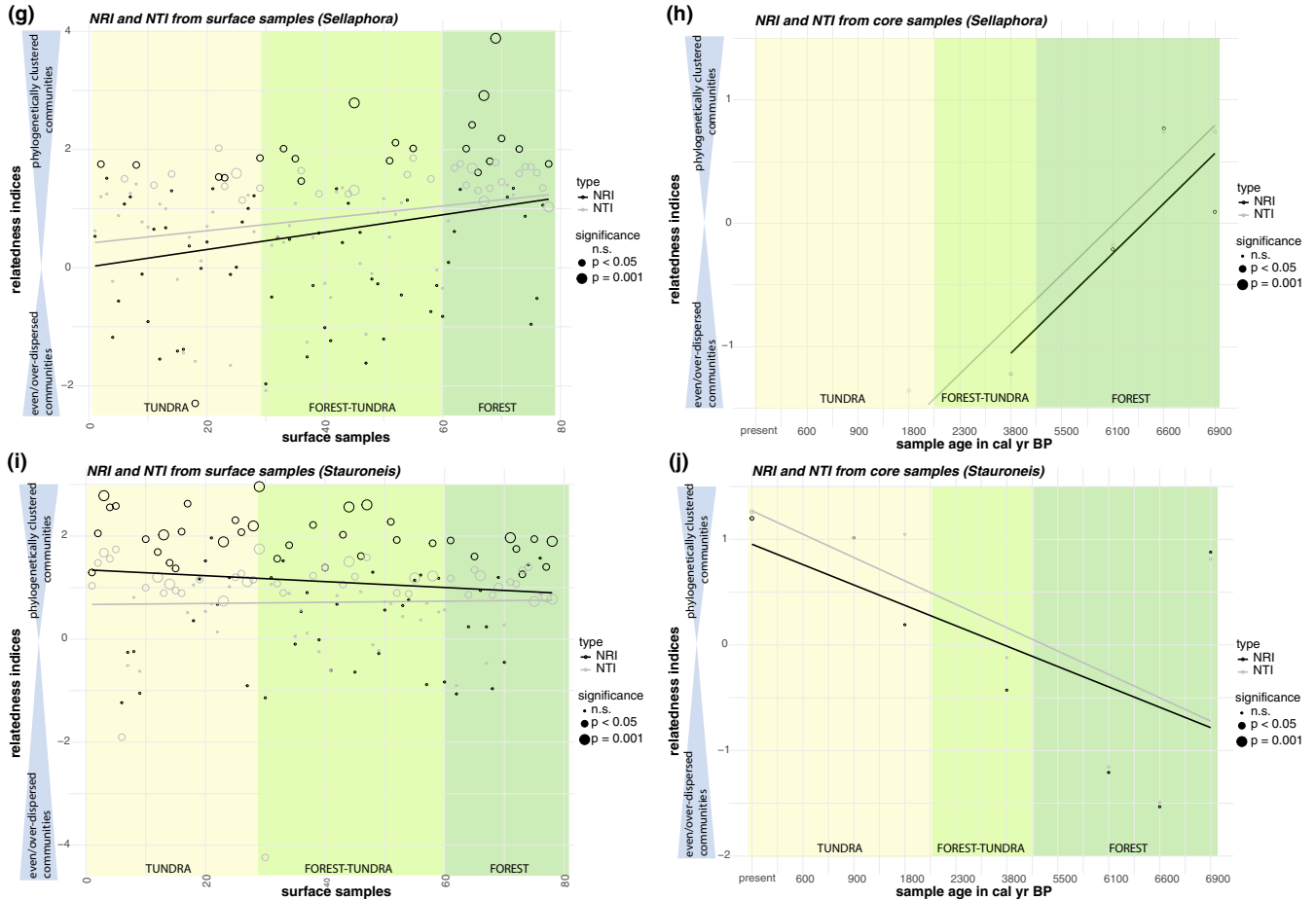


FIGURE 5 Continued

lakes with a subsequent selection of lineages adapted to the current environment or in situ radiation, which has been explored for diatoms in isolated ecosystems (Stelbrink et al., 2018). A significant (after Bonferroni correction) clustering signal in space as well as through time is also most obvious for *Stauroneis*, supporting its affinity to diversify along the ecological gradient of the Siberian tree line (Stoof-Leichsenring et al., 2015). Diversification in *Stauroneis* and related genera (*Stauroneirella* and *Pseudostauroneis*) is also detectable in the variety of present morphotypes reported from lakes of the Siberian tree line area (Stoof-Leichsenring et al., 2014) and also from other regions in the Circum-Arctic (Paull, Hamilton, Gajewski, & LeBlanc, 2008) and North America and Mongolia (Morales, Edlund, & Spaulding, 2010).

Interestingly, all other genera with more distantly related lineages present less pronounced environmental filtering and show only few significant deviations from randomized null communities, indicating mostly random community assemblages in the investigated lakes across the tree line ecotone. Core samples of different time slices and different past vegetation periods produced mostly clustering signals (predominantly in *Stauroneis*) or random patterns, and only a few samples, with weak statistical support, indicated over-dispersed or even communities.

Nearest taxon index and NRI values are known to reflect slightly different results: NTI is more suitable for detecting patterns based

on competition, whereas NRI is more efficient in detecting patterns of environmental filtering (Kraft, Cornwell, Webb, & Ackerly, 2007). Furthermore, it has been suggested that richness of sequence types in the sample compared with richness in the species pool can bias NTI and NRI values, promoting the effects of environmental filtering in a large species pool (Kraft et al., 2007). These biases could have also affected our results, as *Stauroneis*, *Aulacoseira* and *Stauroneis* show highest richness among all genera analysed and calculated indices are mostly related to environmental filtering.

4.3 | Relationship between taxonomic and phylogenetic diatom turnover and environmental change

In our study, distance-based redundancy analyses based on taxonomic and phylogenetic composition support the assembly of main diatom genera (*Aulacoseira*, *Pinnularia*, *Sellaphora*, *Stauroneis* and *Stauroneis*) being majorly explained by HCO_3^- concentrations, water depth and geography. Morphological diatom data have demonstrated that the overall diatom variance in Siberian lakes sampled from Arctic to boreal environments is mostly explained by conductivity, lake vegetation type and mean July air temperature, whereas silica, pH and water depth explain less variance (Pestryakova, Herzsich, Gorodnichev,

TABLE 2 Results of the dbRDA analyses of the 19 diatom genera datasets obtained from the lake surface samples

Genera	Taxonomic beta diversity			Phylogenetic beta diversity			Mantel test		
	Selected variables	R ² adj	p	Selected variables	R ² adj	p	R	p	
<i>Staurosira</i>	HCO ₃ ⁻ + dep + PCNM1	0.261	.001***	HCO ₃ ⁻ + dep	0.110	.001***	0.254	.001***	
<i>Aulacoseira</i>	PCNM1 + dep + PCNM7 + PCNM10 + PCNM9	0.054	.001***	PCNM1 + dep + PCNM9	0.245	.001***	0.439	.001***	
<i>Pinnularia</i>	HCO ₃ ⁻ + PCNM11 + PCNM9 + PCNM4	0.082	.002**	dep + PCNM11 + HCO ₃ ⁻	0.081	.001***	0.492	.001***	
<i>Sellaphora</i>	PCNM1 + dep + PCNM2 + HCO ₃ ⁻ + P CNM4	0.124	.001***	PCNM4 + PCNM2 + dep	0.102	.001***	0.431	.001***	
<i>Stauroneis</i>	dep + HCO ₃ ⁻	0.058	.001***	PCNM5 + pH	0.127	.004**	0.452	.001***	

Note: Variables for dbRDA were selected with the function *ordistep*, only pre-selected variables were included in the dbRDA. Mantel test compared distance matrices derived from taxonomic and phylogenetic beta diversity per genus.

p values are coded in ***0.001, **0.01, *0.05, *0.1; HCO₃⁻; Hydrogen carbonate; dep, lake water depth; pH, pH value; PCNM eigenvectors indicate spatial variables.

& Wetterich, 2018). It has been hypothesized that there is a strong interaction between lake vegetation type and conductivity, which is mainly driven by HCO₃⁻ concentrations, explaining increased HCO₃⁻ concentrations in boreal lakes compared to tundra lakes (Herzschuh et al., 2013). Likewise, our data support that HCO₃⁻ mainly explains diatom variance in the genera *Staurosira*, *Pinnularia*, *Sellaphora* and *Stauroneis*, except *Aulacoseira* is largely explained by spatial variables. Because HCO₃⁻ concentrations and spatial variables partly mirror vegetational changes across the sampled ecotone, we assume that additional environmental variables, such as vegetation type or related environmental variables that have not been exclusively considered in this study, contribute to the diatom community assembly. Geographic distance, however, seems to be of weaker influence, this has been shown in other alpine and boreal studies (Feret, Bouchez, & Rimet, 2017; Teittinen & Soininen, 2015) and has been demonstrated for the genus *Staurosira* across a similar area (Stoof-Leichsenring et al., 2015) and for this study as well.

Until recently, it is still unclear whether diatoms can generally adapt to recent global climate change (Irwin, Finkel, Müller-Karger, & Troccoli Ghinaglia, 2015) and if vegetation changes, which already resulted in the greening and browning of the tundra (Fu, Su, Wang, & Sui, 2019; Lara, Nitze, Grosse, Martin, & McGuire, 2018), will eventually cause the disappearance of specific diatoms along with the potential loss of the tundra biome in the future. Furthermore, temperature-related variables, like ice cover duration, thermal stratification (Rühland, Priesnitz, & Smol, 2003), primary production (Drake et al., 2019) and nutrient availability in Arctic lakes (Pestryakova et al., 2018) are affected by recent global warming. Lacking thermal stratification and nutrient increase (Sienkiewicz, Gąsiorowski, & Migata, 2017) as well as the lengthening of the growing season (Smol et al., 2005) resulted in drastic replacements of established diatom assemblages. Especially, the reduction in ice cover duration altered the abundance of diatoms of the genus *Fragilaria* (synonymous for *Staurosira*; Keatley, Douglas, & Smol, 2008; Smol & Douglas, 2007; Smol et al., 2005;). In contrast, growth experiments on marine diatoms suggest adaptation to rapidly increasing temperature is possible (Jin & Agustí, 2018; O'Donnell et al., 2018), whereas other studies also identify a loss of diversity in cold-adapted diatoms (Lazarus, Barron, Renaudie, Diver, & Turke, 2014). From our data, however, we can conclude that for taxonomically and phylogenetically poor genera (*Ellerbeckia*, *Urosolenia*, etc.), environmental change will putatively result quickly in the absence of lineages. In contrast, lineages of taxonomically highly diverse and phylogenetically less differentiated diatom genera (*Staurosira* and *Aulacoseira*) have greater capabilities to adapt to environmental changes by shifting their abundancies without being entirely lost. Lineages of genera, which are less taxonomically diverse but largely phylogenetically differentiated (*Nitzschia*, *Pinnularia*, *Sellaphora* and *Stauroneis*) will putatively shift in abundance, been replaced or lost. This is likely because they can poorly cope with environmental changes as they are strongly adapted to specific environmental conditions.

5 | CONCLUSION

In general, our study supports using environmental DNA from lake sediment archives to infer taxonomic and phylogenetic diversity directly through time and space, thereby allowing a deeper understanding of assembly rules by combining recent and temporal patterns of community assemblage in minute eukaryotic algae. In particular, our study reveals that lake environmental conditions and phylogenetic diversity, which is a result of evolutionary differentiation, are important factors in the assembly rules for diatom genera. The sensitivity of diatoms towards environmental change highlights the importance of understanding the effects of climate variations on arctic-boreal lake environments and community composition, especially under recent global climate change. Under the expectation that phylogenetic diversity shapes the diatom assembly by affecting richness and turnover in diatom genera, as shown for communities from Siberian tree line lakes, we conclude that some diatom genera have greater capabilities to adapt to environmental changes, whereas others will be putatively replaced or lost due to the displacement of the tundra biome and related changing environmental conditions.

ACKNOWLEDGEMENTS

We thank the Federal Ministry of Education and Research (BMBF, grant no 5.2711.2017/4.6), the Russian Foundation for Basic Research (RFBR, grant no 18-45-140053 r_a) and the Project of the North-Eastern Federal University (NEFU, Regulation SMK-P-1/2-242-17 ver. 2.0, order no. 494-OD), which partly financed the project. In addition, we would like to acknowledge Alexey Pestryakov and his family for their great support during fieldwork, Evgenia Mikhaylova (rector of NEFU) and the BIOM laboratory of NEFU.

DATA AVAILABILITY STATEMENT

Raw sequence data, description of the bioinformatic filtering of the raw sequence data and final tabular DNA sequence datasets are deposited in the Dryad Digital Repository (<https://doi.org/10.5061/dryad.h18931zg9>).

ORCID

Kathleen R. Stooft-Leichsenring  <https://orcid.org/0000-0002-6609-3217>

Laura S. Epp  <https://orcid.org/0000-0002-2230-9477>

Ulrike Herzsuh  <https://orcid.org/0000-0003-0999-1261>

REFERENCES

- Abarca, N., Jahn, R., Zimmermann, J., & Enke, N. (2014). Does the cosmopolitan diatom *Gomphonema parvulum* (Kutzing) Kutzing have a biogeography? *PLoS ONE*, 9(1), e86885. <https://doi.org/10.1371/journal.pone.0086885>
- Alverson, A. J. (2008). Molecular systematics and the diatom species. *Protist*, 159(3), 339–353. <https://doi.org/10.1016/j.protis.2008.04.001>
- Arnan, X., Cerdá, X., & Retana, J. (2017). Relationships among taxonomic, functional, and phylogenetic ant diversity across the biogeographic regions of Europe. *Ecography*, 40(3), 448–457. <https://doi.org/10.1111/ecog.01938>
- Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, 21(12), 1223–1232. <https://doi.org/10.1111/j.1466-8238.2011.00756.x>
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, 15(4), 365–377. <https://doi.org/10.1111/j.1461-0248.2011.01736.x>
- Bennett, J. R., Cumming, B. F., Ginn, B. K., & Smol, J. P. (2010). Broad-scale environmental response and niche conservatism in lacustrine diatom communities. *Global Ecology and Biogeography*, <https://doi.org/10.1111/j.1466-8238.2010.00548.x>
- Bottin, M., Soininen, J., Alard, D., & Rosebery, J. (2016). Diatom co-occurrence shows less segregation than predicted from niche modeling. *PLoS ONE*, 11(4), e0154581. <https://doi.org/10.1371/journal.pone.0154581>
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1), 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., ... Webster, N. S. (2019). Scientists' warning to humanity: Microorganisms and climate change. *Nature Reviews Microbiology*, 17(9), 569–586. <https://doi.org/10.1038/s41579-019-0222-5>
- Charney, N., & Record, S. (2012). vegetarian: Jost Diversity Measures for Community Data. R package version 1.2. <https://CRAN.R-project.org/package=vegetarian>
- Davey, J. W., & Blaxter, M. L. (2010). RADSeq: Next-generation population genetics. *Briefings in Functional Genomics*, 9(5–6), 416–423. <https://doi.org/10.1093/bfgp/elq031>
- Drake, T. W., Holmes, R. M., Zhulidov, A. V., Gurtovaya, T., Raymond, P. A., McClelland, J. W., & Spencer, R. G. M. (2019). Multidecadal climate-induced changes in Arctic tundra lake geochemistry and geomorphology: Long-term changes in Arctic lakes. *Limnology and Oceanography*, 64(S1), S179–S191. <https://doi.org/10.1002/lno.11015>
- Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C. M., ... Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *Gigascience*, 4, 46. <https://doi.org/10.1186/s13742-015-0086-1>
- Dulias, K., Stooft-Leichsenring, K. R., Pestryakova, L. A., & Herzsuh, U. (2017). Sedimentary DNA versus morphology in the analysis of diatom-environment relationships. *Journal of Paleolimnology*, 57(1), 51–66. <https://doi.org/10.1007/s10933-016-9926-y>
- Dutta, H., & Dutta, A. (2016). The microbial aspect of climate change. *Energy, Ecology and Environment*, 1(4), 209–232. <https://doi.org/10.1007/s40974-016-0034-7>
- Echeverría-Londoño, S., Enquist, B. J., Neves, D. M., Violle, C., Boyle, B., Kraft, N. J. B., ... Kerkhoff, A. J. (2018). Plant functional diversity and the biogeography of biomes in North and South America. *Frontiers in Ecology and Evolution*, 6, 219. <https://doi.org/10.3389/fevo.2018.00219>
- Edgar, S. M., & Theriot, E. C. (2004). Phylogeny of Aulacoseira (Bacillariophyta) based on molecules and morphology 1: Phylogeny of Aulacoseira. *Journal of Phycology*, 40(4), 772–788. <https://doi.org/10.1111/j.1529-8817.2004.03126.x>
- Evans, K. M., Wortley, A. H., & Mann, D. G. (2007). An assessment of potential diatom "barcode" genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in Sellaphora (Bacillariophyta). *Protist*, 158(3), 349–364. <https://doi.org/10.1016/j.protis.2007.04.001>
- Feret, L., Bouchez, A., & Rimet, F. (2017). Benthic diatom communities in high altitude lakes: A large scale study in the French Alps. *Annales de Limnologie - International Journal of Limnology*, 53, 411–423. <https://doi.org/10.1051/limn/2017025>

- Frainer, A., Primicerio, R., Kortsch, S., Aune, M., Dolgov, A. V., Fossheim, M., & Aschan, M. M. (2017). Climate-driven changes in functional biogeography of Arctic marine fish communities. *Proceedings of the National Academy of Sciences of the United States of America*, 114(46), 12202–12207. <https://doi.org/10.1073/pnas.1706080114>
- Fu, D., Su, F., Wang, J., & Sui, Y. (2019). Patterns of arctic tundra greenness based on spatially downscaled solar-induced fluorescence. *Remote Sensing*, 11(12), 1460. <https://doi.org/10.3390/rs11121460>
- Götzenberger, L., de Bello, F., Bräthen, K. A., Davison, J., Dubuis, A., Guisan, A., ... Zobel, M. (2012). Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews*, 87(1), 111–127. <https://doi.org/10.1111/j.1469-185X.2011.00187.x>
- Guo, L., Sui, Z., Zhang, S., Ren, Y., & Liu, Y. (2015). Comparison of potential diatom “barcode” genes (the 18S rRNA gene and ITS, COI, rbcL) and their effectiveness in discriminating and determining species taxonomy in the Bacillariophyta. *International Journal of Systematic Evolutionary Microbiology*, 65(Pt 4), 1369–1380. <https://doi.org/10.1099/ijs.0.000076>
- Hamsher, S. E., Evans, K. M., Mann, D. G., Poulickova, A., & Saunders, G. W. (2011). Barcoding diatoms: Exploring alternatives to COI-5P. *Protist*, 162(3), 405–422. <https://doi.org/10.1016/j.protis.2010.09.005>
- Heger, T. J., Edgcomb, V. P., Kim, E., Lukes, J., Leander, B. S., & Yubuki, N. (2014). A resurgence in field research is essential to better understand the diversity, ecology, and evolution of microbial eukaryotes. *Journal of Eukaryotic Microbiology*, 61(2), 214–223. <https://doi.org/10.1111/jeu.12095>
- Heino, J., & Soininen, J. (2005). Assembly rules and community models for unicellular organisms: Patterns in diatoms of boreal streams. *Freshwater Biology*, 50(4), 567–577. <https://doi.org/10.1111/j.1365-2427.2005.01346.x>
- Herzschuh, U., Pestyakova, L. A., Savelieva, L. A., Heinecke, L., Böhmer, T., Biskaborn, B. K., ... Birks, H. J. B. (2013). Siberian larch forests and the ion content of thaw lakes form a geochemically functional entity. *Nature Communications*, 4, 2408. <https://doi.org/10.1038/ncomms3408>
- Irwin, A. J., Finkel, Z. V., Müller-Karger, F. E., & Troccoli Ghinaglia, L. (2015). Phytoplankton adapt to changing ocean environments. *Proceedings of the National Academy of Sciences of the United States of America*, 112(18), 5762–5766. <https://doi.org/10.1073/pnas.1414752112>
- Jin, P., & Agustí, S. (2018). Fast adaptation of tropical diatoms to increased warming with trade-offs. *Scientific Reports*, 8(1), 17771. <https://doi.org/10.1038/s41598-018-36091-y>
- Jousset, A., Schmid, B., Scheu, S., & Eisenhauer, N. (2011). Genotypic richness and dissimilarity opposingly affect ecosystem functioning: Genotypic diversity and ecosystem functioning. *Ecology Letters*, 14(6), 537–545. <https://doi.org/10.1111/j.1461-0248.2011.01613.x>
- Keatley, B. E., Douglas, M. S. V., & Smol, J. P. (2008). Prolonged ice cover dampens diatom community responses to recent climatic change in high arctic lakes. *Arctic, Antarctic, and Alpine Research*, 40(2), 364–372. [https://doi.org/10.1657/1523-0430\(06-068\)\[keatley\]2.0.Co;2](https://doi.org/10.1657/1523-0430(06-068)[keatley]2.0.Co;2)
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., ... Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464.
- Kermarrec, L., Bouchez, A., Rimet, F., & Humbert, J. F. (2013). First evidence of the existence of semi-cryptic species and of a phylogeographic structure in the *Gomphonema parvulum* (Kützing) Kützing complex (Bacillariophyta). *Protist*, 164(5), 686–705. <https://doi.org/10.1016/j.protis.2013.07.005>
- Klemm, J., Herzschuh, U., & Pestyakova, L. A. (2016). Vegetation, climate and lake changes over the last 7000 years at the boreal treeline in north-central Siberia. *Quaternary Science Reviews*, 147, 422–434. <https://doi.org/10.1016/j.quascirev.2015.08.015>
- Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., Levine, J. M., & Fox, J. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, 29(5), 592–599. <https://doi.org/10.1111/1365-2435.12345>
- Kraft, N. J. B., Cornwell, W. K., Webb, O. C., & Ackerly, D. D. (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*, 170, 271–283.
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, 7(1), 17668. <https://doi.org/10.1038/s41598-017-17333-x>
- Lara, M. J., Nitze, I., Grosse, G., Martin, P., & McGuire, A. D. (2018). Reduced arctic tundra productivity linked with landform and climate change interactions. *Scientific Reports*, 8(1), 2345. <https://doi.org/10.1038/s41598-018-20692-8>
- Lazarus, D., Barron, J., Renaudie, J., Diver, P., & Turke, A. (2014). Cenozoic planktonic marine diatom diversity and correlation to climate change. *PLoS ONE*, 9(1), e84857. <https://doi.org/10.1371/journal.pone.0084857>
- Morales, E. A., Edlund, M. B., & Spaulding, S. A. (2010). Description and ultrastructure of araphid diatom species (Bacillariophyceae) morphologically similar to *Pseudostaurosira elliptica* (Schumann). *Phycological Research*, 58(2), 97–107. <https://doi.org/10.1111/j.1440-1835.2010.00567.x>
- Nakov, T., Beaulieu, J. M., & Alverson, A. J. (2018). Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (Diatoms, Bacillariophyta). *New Phytologist*, 219(1), 462–473. <https://doi.org/10.1111/nph.15137>
- Nanjappa, D., Audic, S., Romac, S., Kooistra, W. H., & Zingone, A. (2014). Assessment of species diversity and distribution of an ancient diatom lineage using a DNA metabarcoding approach. *PLoS ONE*, 9(8), e103810. <https://doi.org/10.1371/journal.pone.0103810>
- O'Donnell, D. R., Hamman, C. R., Johnson, E. C., Kremer, C. T., Klausmeier, C. A., & Litchman, E. (2018). Rapid thermal adaptation in a marine diatom reveals constraints and trade-offs. *Global Change Biology*, 24(10), 4554–4565. <https://doi.org/10.1111/gcb.14360>
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., ... Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>
- Parducci, L., Bennett, K. D., Ficetola, G. F., Alsos, I. G., Suyama, Y., Wood, J. R., & Pedersen, M. W. (2017). Ancient plant DNA in lake sediments. *New Phytologist*, 214(3), 924–942. <https://doi.org/10.1111/nph.14470>
- Paull, T. M., Hamilton, P. B., Gajewski, K., & LeBlanc, M. (2008). Numerical analysis of small Arctic diatoms (Bacillariophyceae) representing the *Staurosira* and *Staurosirella* species complexes. *Phycologia*, 47(2), 213–224. <https://doi.org/10.2216/07-17.1>
- Pestyakova, L. A., Herzschuh, U., Gorodnichev, R., & Wetterich, S. (2018). The sensitivity of diatom taxa from Yakutian lakes (north-eastern Siberia) to electrical conductivity and other environmental variables. *Polar Research*, 37(1), 1485625. <https://doi.org/10.1080/17518369.2018.1485625>
- Pestyakova, L. A., Herzschuh, U., Wetterich, S., & Ulrich, M. (2012). Present-day variability and Holocene dynamics of permafrost-affected lakes in central Yakutia (Eastern Siberia) inferred from diatom records. *Quaternary Science Reviews*, 51, 56–70. <https://doi.org/10.1016/j.quascirev.2012.06.020>
- Pinseel, E., Kulichová, J., Scharfen, V., Urbánková, P., Van de Vijver, B., & Vyverman, W. (2019). Extensive cryptic diversity in the terrestrial diatom *pinnularia borealis* (Bacillariophyceae). *Protist*, 170(2), 121–140. <https://doi.org/10.1016/j.protis.2018.10.001>
- Rimet, F., Chaumeil, P., Keck, F., Kermarrec, L., Vasselon, V., Kahlert, M., ... Bouchez, A. (2016). R-Syst:Diatom: An open-access and curated barcode database for diatoms and freshwater monitoring. *Database (Oxford)*, 2016, baw016. <https://doi.org/10.1093/database/baw016>
- Rimet, F., Vasselon, V., A.-Keszte, B., & Bouchez, A. (2018). Do we similarly assess diversity with microscopy and high-throughput sequencing? Case of microalgae in lakes. *Organisms Diversity & Evolution*, 18(1), 51–62. <https://doi.org/10.1007/s13127-018-0359-5>
- Rühland, K., Priesnitz, A., & Smol, J. P. (2003). Paleolimnological evidence from diatoms for recent environmental changes in 50 lakes across

- Canadian Arctic Treeline. *Arctic, Antarctic, and Alpine Research*, 35(1), 110–123. [https://doi.org/10.1657/1523-0430\(2003\)035\[0110:PEFDR\]2.0.CO;2](https://doi.org/10.1657/1523-0430(2003)035[0110:PEFDR]2.0.CO;2)
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, 27(4), 592–593.
- Sienkiewicz, E., Gąsiorowski, M., & Migąła, K. (2017). Unusual reaction of diatom assemblage on climate changes during the last millennium: A record from Spitsbergen lake. *Journal of Paleolimnology*, 58(1), 73–87. <https://doi.org/10.1007/s10933-017-9962-2>
- Smol, J. P., & Douglas, M. S. V. (2007). From controversy to consensus: Making the case for recent climate change in the Arctic using lake sediments. *Frontiers in Ecology and the Environment*, 5(9), 466–474. <https://doi.org/10.1890/060162>
- Smol, J. P., Wolfe, A. P., Birks, H. J. B., Douglas, M. S. V., Jones, V. J., Korhola, A., ... Weckstrom, J. (2005). Climate-driven regime shifts in the biological communities of arctic lakes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12), 4397–4402. <https://doi.org/10.1073/pnas.0500245102>
- Soininen, J. (2007). Environmental and spatial control of freshwater diatoms—A review. *Diatom Research*, 22(2), 473–490. <https://doi.org/10.1080/0269249x.2007.9705724>
- Stelbrink, B., Jovanovska, E., Levkov, Z., Ognjanova-Rumenova, N., Wilke, T., & Albrecht, C. (2018). Diatoms do radiate: Evidence for a freshwater species flock. *Journal of Evolutionary Biology*, 31(12), 1969–1975. <https://doi.org/10.1111/jeb.13368>
- Stoof-Leichsenring, K. R., Bernhardt, N., Pestryakova, L. A., Epp, L. S., Herzschuh, U., & Tiedemann, R. (2014). A combined paleolimnological/genetic analysis of diatoms reveals divergent evolutionary lineages of *Staurosira* and *Staurosirella* (Bacillariophyta) in Siberian lake sediments along a latitudinal transect. *Journal of Paleolimnology*, 52(1–2), 77–93. <https://doi.org/10.1007/s10933-014-9779-1>
- Stoof-Leichsenring, K. R., Herzschuh, U., Pestryakova, L. A., Klemm, J., Epp, L. S., & Tiedemann, R. (2015). Genetic data from algae sedimentary DNA reflect the influence of environment over geography. *Scientific Reports*, 5(1), 12924. <https://doi.org/10.1038/srep12924>
- Swenson, N. G. (2011). Phylogenetic beta diversity metrics, trait evolution and inferring the functional beta diversity of communities. *PLoS ONE*, 6(6), e21264. <https://doi.org/10.1371/journal.pone.0021264>
- Teittinen, A., & Soininen, J. (2015). Testing the theory of island biogeography for microorganisms – patterns for spring diatoms. *Aquatic Microbial Ecology*, 75(3), 239–250. <https://doi.org/10.3354/ame01759>
- Thakur, M. P., Tilman, D., Purschke, O., Ciobanu, M., Cowles, J., Isbell, F., ... Eisenhauer, N. (2017). Climate warming promotes species diversity, but with greater taxonomic redundancy, in complex environments. *Science Advances*, 3(7), e1700866. <https://doi.org/10.1126/sciadv.1700866>
- Theriot, E. C., Ashworth, M., Ruck, E., Nakov, T., & Jansen, R. K. (2010). A preliminary multigene phylogeny of the diatoms (Bacillariophyta): Challenges for future research. *Plant Ecology and Evolution*, 143(3), 278–296. <https://doi.org/10.5091/plecevo.2010.418>
- Van den Wyngaert, S., Most, M., Freimann, R., Ibelings, B. W., & Spaak, P. (2015). Hidden diversity in the freshwater planktonic diatom *Asterionella formosa*. *Molecular Ecology*, 24(12), 2955–2972. <https://doi.org/10.1111/mec.13218>
- Vanormelingen, P., Evans, K. M., Chepurinov, V. A., Vyverman, W., & Mann, D. G. (2013). Molecular species discovery in the diatom *Sellaphora* and its congruence with mating trials. *Fottea*, 13(2), 133–148. <https://doi.org/10.5507/fot.2013.012>
- Verleyen, E., Vyverman, W., Sterken, M., Hodgson, D. A., De Wever, A., Juggins, S., ... Sabbe, K. (2009). The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. *Oikos*, 118(8), 1239–1249. <https://doi.org/10.1111/j.1600-0706.2009.17575.x>
- Vilmi, A., Karjalainen, S. M., & Heino, J. (2017). Ecological uniqueness of stream and lake diatom communities shows different macroecological patterns. *Diversity and Distributions*, 23(9), 1042–1053. <https://doi.org/10.1111/ddi.12594>
- Vincent, W. F., Callaghan, T. V., Dahl-Jensen, D., Johansson, M., Kovacs, K. M., Michel, C., ... Sharp, M. (2012). Ecological implications of changes in the Arctic cryosphere. *Ambio*, 40(5), 87–99. <https://doi.org/10.1007/s13280-011-0218-5>
- Voskamp, A., Baker, D. J., Stephens, P. A., Valdes, P. J., & Willis, S. G. (2017). Global patterns in the divergence between phylogenetic diversity and species richness in terrestrial birds. *Journal of Biogeography*, 44(4), 709–721. <https://doi.org/10.1111/jbi.12916>
- Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D. A., ... Wever, A. D. (2007). Historical processes constrain patterns in global diatom diversity. *Ecology*, 88(8), 1924–1931. <https://doi.org/10.1890/06-1564.1>
- Webb, C. O., Ackerly, D. D., & Kembel, S. W. (2008). Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, 24(18), 2098–2100. <https://doi.org/10.1093/bioinformatics/btn358>
- Webb, C. O., Ackerly, D. D., McPeck, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33(1), 475–505. <https://doi.org/10.1146/annurev.ecolsys.33.010802.150448>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag.
- Zorzal-Almeida, S., Soininen, J., Bini, L. M., & Bicudo, D. C. (2017). Local environment and connectivity are the main drivers of diatom species composition and trait variation in a set of tropical reservoirs. *Freshwater Biology*, 62(9), 1551–1563. <https://doi.org/10.1111/fwb.12966>

BIOSKETCH

Kathleen Stoof-Leichsenring is a group and laboratory leader of the subresearch group of High Latitude Biodiversity in the section Environmental Terrestrial Systems at the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Biology (link to the webpage: <https://www.awi.de/en/science/geosciences/polar-terrestrial-environmental-systems/research-foci/high-latitude-biodiversity.html>). Her research interests are in palaeogenetic and palaeoenvironmental research with a focus on diatom diversity and their evolutionary patterns in space and through time.

Author contributions: K.S.-L. and U.H. designed research; K.S.-L., U.H and L.P. performed field sampling; K.S.-L. and U.H performed research, while L.E. performed core sediment subsampling and K.S.-L. performed genetic laboratory work; K.S.-L and U.H analysed data; K.S.-L. wrote the article that all co-authors commented on.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Stoof-Leichsenring KR, Pestryakova LA, Epp LS, Herzschuh U. Phylogenetic diversity and environment form assembly rules for Arctic diatom genera—A study on recent and ancient sedimentary DNA. *J Biogeogr.* 2020;00:1–14. <https://doi.org/10.1111/jbi.13786>