

# Changes in the composition of marine and sea-ice diatoms derived from sedimentary ancient DNA of the eastern Fram Strait over the past 30,000 years

5 Heike Hildegard Zimmermann<sup>1</sup>, Kathleen Rosmarie Stoof-Leichsenring<sup>1</sup>, Stefan Kruse<sup>1</sup>, Juliane Müller<sup>2,3,4</sup>, Ruediger Stein<sup>2,3,4</sup>, Ralf Tiedemann<sup>2</sup>, Ulrike Herzsuh<sup>1,5,6</sup>

<sup>1</sup>Polar Terrestrial Environmental Systems, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, 14473, Germany

<sup>2</sup>Marine Geology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, 27568 Bremerhaven, Germany

<sup>3</sup>MARUM, University of Bremen, 28359 Bremen, Germany

10 <sup>4</sup>Faculty of Geosciences, University of Bremen, 28334 Bremen, Germany

<sup>5</sup>Institute of Biochemistry and Biology, University of Potsdam, 14476 Potsdam, Germany

<sup>6</sup>Institute of Environmental Sciences and Geography, University of Potsdam, 14476 Potsdam, Germany

*Correspondence to:* Heike H. Zimmermann (heike.zimmermann@awi.de), Ulrike Herzsuh (ulrike.herzsuh@awi.de)

**Abstract.** The Fram Strait is an area with a relatively low and irregular distribution of diatom microfossils in surface sediments, and thus microfossil records are scarce, rarely exceed the Holocene and contain sparse information about past richness and taxonomic composition. These attributes make the Fram Strait an ideal study site to test the utility of sedimentary ancient DNA (*seDaDNA*) metabarcoding. Amplifying a short, partial *rbcL* marker from samples of sediment core MSM05/5-712-2, resulted in 95.7 % of our sequences being assigned to diatoms across 18 different families with 38.6 % of them being resolved to species and 25.8 % to genus level. Independent replicates show high similarity of PCR products, especially in the oldest samples. Diatom *seDaDNA* richness is highest in the Late Weichselian and lowest in Mid- and Late-Holocene samples. Taxonomic composition is dominated by cold-water and sea-ice associated diatoms and suggests several re-organizations – after the Last Glacial Maximum, after the Younger Dryas, after the Early and after the Mid-Holocene. Different sequences assigned to, amongst others, *Chaetoceros socialis* indicate the detectability of intra-specific diversity using *seDaDNA*. We detect no clear pattern between our diatom *seDaDNA* record and the previously published IP<sub>25</sub> record of this core, although proportions of pennate diatoms increase with higher IP<sub>25</sub> concentrations and proportions of *Nitzschia cf. frigida* exceeding 2 % of the assemblage point towards past sea-ice presence.

## 1 Introduction

The marine environment is a complex ecosystem in which the distribution of organisms is controlled significantly by abiotic constraints such as sea-surface temperatures (SSTs), salinity, nutrient distribution, light conditions, and sea-ice cover (Cherkasheva et al., 2014; Ibarbalz et al., 2019; Nöthig et al., 2015; Pierella Karlusich et al., 2020). Over the past 30,000 years the subarctic North Atlantic Ocean was subject to frequent sea-ice expansions and contractions (Müller et al., 2009; Müller

and Stein, 2014; Syring et al., 2020; Werner et al., 2013), which are expected to have affected the composition of the regional species pool. Diatoms (Bacillariophyta) are unicellular, siliceous organisms that are photoautotrophic and thrive in the euphotic zone of the ocean. Owing to their sensitive responses to environmental conditions, diatoms are frequently used as indicators for paleoenvironmental reconstructions, to assess, for example, changes in surface water temperatures (Birks and Koç, 2002; Krawczyk et al., 2017; Miettinen et al., 2015), paleoproductivity (Fahl and Stein, 1997; Limoges et al., 2018), and sea-ice distribution (Smirnova et al., 2015; Weckström et al., 2013). Next to microfossil-based reconstructions, the diatom produced sea-ice proxy IP<sub>25</sub> (a highly branched isoprenoid alkene with 25 carbon atoms; Belt et al., 2007) combined with phytoplankton biomarkers (e.g. brassicasterol, dinosterol; Volkman, 1986) permit semi-quantitative reconstructions of past sea-ice distribution (Belt, 2018; Belt and Müller, 2013; Müller et al., 2009; Müller and Stein, 2014; Stein et al., 2012, 2017). However, diatoms in northern high-latitude regions are less silicified and more prone to silica dissolution (Kohly, 1998; Stabell, 1986) compared to diatoms of the southern polar oceans (Harrison and Cota, 1991). In the Fram Strait – an important area of heat exchange between Arctic and North Atlantic water masses (Untersteiner, 1988) – particularly low and irregular preservation of diatom microfossils prevails in surface sediments (Karpuz and Schrader, 1990; Stabell, 1987). The diatom records are generally underrepresented, contain sparse information about past diversity and taxonomic composition, and rarely exceed the Holocene (Jessen et al., 2010; Koç et al., 2002; Stabell, 1986). This makes the Fram Strait an excellent site to test ancient DNA metabarcoding on a sediment core (Müller et al., 2012; Müller and Stein, 2014).

Ancient DNA is a new proxy that can exploit diatoms as indicators of past marine environmental change (Coolen et al., 2007; De Schepper et al., 2019; Kirkpatrick et al., 2016). Deep-sea sediments have been reported to be rich in DNA with up to 70–90 % of the total DNA pool being extracellular DNA (Coolen et al., 2007; Dell’Anno et al., 2002; Lejzerowicz et al., 2013; Morard et al., 2017), and traces of DNA can be detected in sediments even though microfossils are absent or highly degraded (Boere et al., 2009; Coolen et al., 2009, 2013; Lejzerowicz et al., 2013; Pawłowska et al., 2020). Therefore, analyses of sedimentary ancient DNA (*sedaDNA*) could be advantageous in areas of biased preservation due to high silica dissolution rates. Beyond morphological or biogeochemical analyses, ancient DNA can distinguish cryptic species that are morphologically similar (Stoof-Leichsenring et al., 2012), trace temporal changes of intra-specific genetic variation (Epp et al., 2018; Parducci et al., 2012; Zimmermann et al., 2017b), and identify genetic relationships and microevolution (Stoof-Leichsenring et al., 2014, 2015).

Paleogenetic analyses with an emphasis on diatoms have been successfully carried out in various limnic settings ranging from the subarctic (Epp et al., 2015; Stoof-Leichsenring et al., 2014, 2015) to the tropics (Stoof-Leichsenring et al., 2012) and Antarctica (Coolen et al., 2004). Yet, it is still a relatively underrepresented branch in the marine realm with only a few published studies targeting phytoplankton ancient DNA (Boere et al., 2009, 2011b, 2011a; Coolen et al., 2006, 2007, 2009, 2013; De Schepper et al., 2019; Giosan et al., 2012; Kirkpatrick et al., 2016) or diatoms in particular (Pawłowska et al., 2020). In this study we examine ancient diatom DNA retrieved from sediments from the eastern Fram Strait and assess whether it can be used to analyze temporal changes in the taxonomic composition of diatoms. We have three principal objectives: (1) to assess the quality and replicability of the data obtained by *sedaDNA* metabarcoding, (2) to analyze temporal changes of diatom

taxonomic composition and richness, and (3) to evaluate diatom *seadDNA* as a new proxy for sea-ice reconstruction. DNA was derived from distinct samples of the comprehensively analyzed marine sediment core MSM05/5-712-2, covering the major climatic intervals since the Late Weichselian (i.e. the last ~30 kyr BP). As previous work indicates variability in the past sea-ice cover (Falardeau et al., 2018; Müller et al., 2012; Müller and Stein, 2014; Werner et al., 2011, 2013), samples were chosen according to high, medium, and low concentrations of the diatom produced sea-ice biomarker IP<sub>25</sub> (Müller et al., 2012; Müller and Stein, 2014) and we expect associated changes in the taxonomic composition. We used *seadDNA* metabarcoding by applying the diatom-specific *rbcL\_76* marker (Stoof-Leichsenring et al., 2012) which has already proved successful in low-productivity lakes of northern Siberia (Dulias et al., 2017; Stoof-Leichsenring et al., 2014, 2015), but so far has not been tested on marine sediments. The marker amplifies a short region of the *rbcL* gene, which is located on the chloroplast that is present in diatoms in several copy numbers (Vasselon et al., 2018), thereby increasing the probability of its long-term preservation. The *rbcL* gene has an adequate sequence reference database and was tested as a potential diatom barcode marker with high resolution power (Guo et al., 2015; Kermarrec et al., 2013). Furthermore, it reduces co-amplification of non-photosynthetic bacteria or archaea that are active in subsurface sediments and thus could be preferentially amplified during PCR in comparison to the highly fragmented and damaged ancient DNA.

## 80 **2 Materials and Methods**

### **2.1 Study site and sample material**

The kastenlot core MSM05/5-712-2 (N 78.915662, E 6.767167, water depth 1487 m) was collected from the western continental slope of Svalbard during the cruise of *Maria S. Merian* (Budéus, 2007) in the eastern Fram Strait in August 2007 (Fig. 1). On board, subsections of 1 m length were placed in square plastic boxes as explained in the supplement of Gersonde et al. (2012) and stored at 4°C. This may have affected DNA preservation. The Fram Strait is located between Greenland and Svalbard and connects the Arctic Ocean with the Atlantic Ocean. The study area is influenced by temperate, saline water masses that are transported northwards via the West Spitzbergen Current, which is a continuation of the North Atlantic Current (Aagaard, 1982). Furthermore, the site is located downslope from Kongsfjorden and is thus influenced by one of the major outlets of western Svalbard meltwater (Werner et al., 2013). Today, the site is located south of the winter and summer sea-ice margin and is ice-free year round (Fig. 1). Age-depth modeling suggests a maximum age of about 30 cal kyr BP for the lowermost core interval (Müller and Stein, 2014). The sampling procedure for ancient DNA analyses followed the protocol for non-frozen sediment cores explained in Epp et al. (2019) for the 12 samples taken along the core, of which the depth at 8.85 m was sampled twice to check whether different samples from the same level show similar taxonomic composition and richness, particularly the oldest sample (Table 1).

95

## 2.2 DNA extraction, PCR, and sequencing

The DNA extractions and PCR setups were prepared in a dedicated laboratory for ancient DNA at the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (Potsdam, Germany). Total DNA was extracted from 12 samples of approximately 2 g (wet weight) sediment using the same method as described in Zimmermann et al. (2017a). Each extraction batch included one negative control. The DNA concentrations were measured with the Qubit dsDNA BR Assay Kit (Invitrogen, Carlsbad, CA, USA) on a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA). As DNA concentrations were below the detection limit, we concentrated 600  $\mu\text{L}$  of each sample with the GeneJET PCR Purification KIT (Thermo Scientific, Carlsbad, CA, USA) according to the manufacturer's protocol and eluted twice with 15  $\mu\text{L}$  elution buffer. All DNA extracts were stored at  $-20^\circ\text{C}$ .

We amplified the marker *rbcl\_76* (Stoof-Leichsenring et al., 2012), a 76 bp long fragment of the plastid *rbcl* gene using tagged primers *Diat\_rbcL\_705F* (AACAGGTGAAGTTAAAGGTTTCATAYTT) and *Diat\_rbcL\_808R* (TGTAACCCATAACTAAATCGATCAT) as described in Dulias et al. (2017). The PCR reactions were set up in small batches, each including up to 9 samples and the corresponding negative control from the DNA extraction as well as a PCR no template control (NTC). For each sample, extraction negative control, and NTC, we performed 3 PCRs with different primer-tag combinations on different days. The PCR reaction mixes and conditions were prepared following the adjusted protocol for tagged *Diat\_rbcL\_705F* and *Diat\_rbcL\_808R* primers as described in Dulias et al. (2017) with the exception that 3  $\mu\text{l}$  DNA (DNA concentration 3  $\text{ng } \mu\text{L}^{-1}$ ) was used as a template. PCRs were performed with the following settings: 5 minutes at  $94^\circ\text{C}$  (initial denaturation), then 50 cycles at  $94^\circ\text{C}$  (denaturation),  $49^\circ\text{C}$  (annealing), and  $68^\circ\text{C}$  (elongation), and a final elongation step at  $72^\circ\text{C}$  for 5 minutes. Subsequently the PCR success was checked with gel-electrophoresis. All PCR products were purified with the MinElute purification Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. Elution was carried out twice in 10  $\mu\text{L}$  elution buffer. The purified PCR products were mixed in equal concentrations and sent to Fasteris SA sequencing service (Switzerland) who carried out library preparation and sequencing. The sequencing library was prepared with the Mid Output kit v. 2 according to the Fasteris Metafast protocol for low complexity amplicon sequencing and checked by qPCR. The library was sequenced (2 x 150 bp, paired-end) on the Illumina NextSeq 500 (Illumina Inc., San Diego, CA, USA). The sequence data are deposited at the European Nucleotide Archive (BioProject: PRJEB39075).

## 2.3 Bioinformatic processing

The sequences were processed, filtered, and assigned a taxonomic name according to the NCBI taxonomy using the OBITools package (Boyer et al., 2016) with the same bioinformatics parameter settings as described in Dulias et al. (2017). We refrained from clustering sequences into operational taxonomic units and used amplicon sequence variants (ASVs) as recommended by Callahan et al. (2017). The OBITools program *obiclean* can identify ASVs that have likely arisen due to PCR or sequencing errors. It uses the information of sequence counts and sequence similarities to classify whether a sequence is a variant ("internal") of a more abundant ("head") ASV (Boyer et al., 2016). To generate the reference database for the taxonomic

assignment of the sequences we downloaded the EMBL release 138 (released November 2018) and used *ecopcr* (Ficetola et al., 2010) according to the descriptions of Dulas et al. (2017) and Stoof-Leichsenring et al. (2012) containing 2,320 reference sequences. A total of 7,536,449 sequence counts were assigned to samples, 235 sequence counts to extraction negative controls and 237 counts to PCR negative controls. Of the 204 different sequence variants detected in extraction and PCR negative controls, 83% of their occurrences were singletons (Suppl. Table 1) and most likely artefacts from tag-jumping during library preparation (Schnell et al., 2015). Using R v. 3.5.0 (R Core Team, 2018), we kept only those ASVs (Suppl. Table 2) that (1) were assigned a taxonomic name based on 90–100 % similarity to an entry in the reference database, (2) were represented with at least 10 read counts in a replicate, (3) were present at least 3 times among all sequenced PCR products, (4) showed taxonomic resolution below the phylum level “Bacillariophyta”, and (5) were tagged as “internal” by *obiclean* in less than 50 % of all sequenced PCR products to reduce PCR and sequencing artefacts. Filtering with R reduced the number of read counts from 7,536,449 to 6,199,984.

## 2.4 Reproducibility of PCR replicates

Alpha-diversity estimates, such as richness, depend highly on the sequencing depth. As more taxa, especially rare ones, can be detected with increasing sequencing depth, alpha-diversity is only comparable between samples when it is estimated based on the same number of sequences. Despite efforts to reduce such differences by equimolar pooling of PCR products, the number of sequences generally varies among PCR replicates. Therefore, we analyzed the dissimilarities of PCR replicates as well as one sample replicate (8.85 m depth). We resampled the dataset 100 times to the minimum number of sequences available (25,601 counts), then, for each replicate, we calculated the mean number of sequence counts for each ASV across the 100 resampling steps (code available at: [https://github.com/StefanKruse/R\\_Rarefaction](https://github.com/StefanKruse/R_Rarefaction); Kruse, 2019). This dataset was used to calculate the proportions of each ASV per replicate. The proportional data were used for the Multiple Response Permutation Procedure (MRPP) using the R function *mrpp* on Bray Curtis Dissimilarities to test if dissimilarities within replicates of the same sample are significant. Furthermore, we applied non-metric multidimensional scaling (NMDS) using *metaMDS* to assess which replicates show high and low replicability.

## 2.5 Taxonomic composition and richness

As a measure of alpha-diversity we calculated richness of (1) ASVs (number of amplicon sequence variants) and (2) unique taxonomic names (number of grouped ASVs that were assigned to the same taxonomic name). For taxonomic composition and richness calculations, we combined the PCR replicates of the corresponding sample. This resulted in a new minimum number of sequence counts (300,415 counts) that was used for resampling the dataset 100 times according to descriptions in the preceding paragraph. The resampled dataset was used to calculate the relative abundances of each ASV per sample. Finally, ASVs, which were assigned to the same taxonomic name but had different similarities to an entry in the reference database, were summed up to one entry using the R *aggregate* function. Stratigraphic diagrams showing temporal changes in taxonomic composition were generated with *strat.plot*. All statistical analyses and visualizations were prepared with R v. 3.5.0 using the

160 packages “vegan” (Oksanen et al., 2011) and “rioja” (Juggins, 2012). For correlation analysis we interpolated IP<sub>25</sub> values using  
the methods described in Reschke et al. (2019). We transformed the IP<sub>25</sub> data using the function *zoo* from the “zoo” package  
(Zeileis and Grothendieck, 2005) and used in the function *CorIrregTimser* using the package “corit”  
(<https://github.com/EarthSystemDiagnostics/corit>) (Reschke et al., 2019). The correlation between Chaetocerotaceae and  
Thalassiosiraceae as well as between IP<sub>25</sub> and all ASVs was tested for significance using the R function *rcorr* (method Pearson)  
165 from the package “Hmisc” (Hollander and Wolfe, 1975; Press et al., 1988).

### 3 Marker specificity, taxonomic resolution, and taxonomic coverage

The amplification of the short *rbcL\_76* marker permitted the retrieval of diatom DNA from all samples and was highly specific  
for diatom sequence variants (95.7 %). Only 4.3 % of all assigned ASVs were assigned to other organismal groups: 4 %  
Bolidophyceae (a closely related sister clade of diatoms with unicellular, siliceous, flagellated algae), 0.2 % Phaeophyceae  
170 (brown algae), and 0.1 % Eustigmatophyceae (photosynthetic heterokonts). This exceeds the marker specificity found in  
previous studies focusing on lake sediment cores (84 % (Stoof-Leichsenring et al., 2012), 88 % (Stoof-Leichsenring et al.,  
2014), 90 % (Stoof-Leichsenring et al., 2015)) as well as specificity of a previous *in silico* PCR (90.4 % (Dulias et al., 2017)).  
Possibly owing to a much shorter marker size, the *rbcL\_76* marker also surpassed the usually amplified 18S rDNA markers  
with regard to specificity for diatoms (Coolen et al., 2007; De Schepper et al., 2019; Kirkpatrick et al., 2016).

175 The initial dataset contained 1,398 ASVs that were systematically filtered to a final diatom dataset containing 360 ASVs  
(6,199,984 counts) which were further grouped into 75 unique taxonomic names. The majority of diatom sequences are  
assigned to polar centric Mediophyceae (79.7 %), followed by pennate Bacillariophyceae (14.5 %), while radial centric  
Conscinodiscophyceae (1.1 %) make up only a small proportion of the dataset. The majority of ASVs are assigned as low as  
species level (38.6 %) or genus level (25.8 %) (Fig. 2a), yet the taxonomic resolution (i.e. the taxonomic assignment of ASVs)  
180 is limited by the incompleteness of the sequence reference database as indicated by the proportion of ASVs (4.8%) for which  
assignment is restricted to phylum level (Bacillariophyta).

In total, 18 different diatom families are represented in the final dataset which is dominated by ASVs assigned to the families  
Thalassiosiraceae (35 %), Bacillariaceae (15.8 %), and Chaetoceraceae (9.2 %) (Fig. 2b). Particularly dominant ASVs are  
assigned to *Thalassiosira* (21.2 %) and *Porosira* (10.9 %) (resolution only possible to genus level) and to *Chaetoceros* cf.  
185 *contortus* 1SEH-2013 (6.6 %) – all centric diatoms belong to the class Mediophyceae. Their dominance is likely the result of  
high paleoproductivity or of differential overrepresentation caused by preservation and/or technical biases. Indeed, in the Fram  
Strait, *Chaetoceros* and *Thalassiosira* species especially have high productivity in different hydrographical regimes (Gradinger  
and Baumann, 1991; Lalande et al., 2013) as well as at Kongsfjorden, a major outlet of western Svalbard potentially influencing  
the coring site (Hodal et al., 2012), and even in micropaleontological studies they are often dominant (Birks and Koç, 2002;  
190 Bylinskaya et al., 2016; Oksman et al., 2017, 2019). Next to high paleoproductivity, the preservation of *seda*DNA could be  
biased by differential degradation. Heavily silicified and/or spore forming diatoms such as some *Thalassiosira* and

*Chaetoceros* species may be less sensitive to dissolution, which might also improve the preservation of DNA over long time periods. Furthermore, the enrichment of centric diatoms in the *seadaDNA* record could be the result of copy number variation of the *rbcL* gene between different species and cell biovolume (Bedoshvili et al., 2009; Round et al., 1990; Vasselon et al., 2018). Technical biases often arise during PCR due to mismatches between primer sequences and primer binding sites (Nichols et al., 2018) and the high number of PCR cycles – which is needed to increase the chance of retrieving rare sequences – leads to an over-amplification of already abundant template molecules in comparison to rare ones. While the reduction of cycles could reduce this effect, fewer PCR cycles would reduce replicability (Krehenwinkel et al., 2017; Nichols et al., 2018). Despite the filtering, several distinct ASVs, (e.g. *Chaetoceros socialis*) are assigned the same taxonomic name (Suppl. Table S1, Suppl. Fig. S1). We believe that these represent either different lineages or closely related species so far not included in the database, although we cannot rule out that some ASVs in our filtered dataset might still represent PCR or sequencing artefacts.

Despite the current limitations, *seadaDNA* shows much promise. New reference sequences are added to GenBank on a daily basis, due to numerous phylogenetic studies and barcoding projects aiming to improve our systematic knowledge about taxonomic relationships and to archive the molecular inventory of global biodiversity (Degerlund et al., 2012; Li et al., 2015; Luddington et al., 2016). Taxonomic coverage of the reference database can be increased by subjecting diatom strains to single cell sequencing (Luddington et al., 2016; Sieracki et al., 2019). Most importantly, *seadaDNA* allows tracing ASVs through time and eventually relating them to, for example, environmental change, without relying on the state of taxonomic coverage of the reference database.

#### 210 **4 Quality and replicability of the data obtained by *seadaDNA* metabarcoding**

A crucial requirement for the interpretation of ancient DNA records is that PCR replicates show similar signals in richness and taxonomic composition. The recovery of taxa by *seadaDNA* metabarcoding is prone to false presences or absences. As our study lacks a morphological diatom record, false absences cannot be assessed and true presences cannot be confirmed, which means that absence in our record does not necessarily translate to a true physical absence in the past. Hence, it was important to use independent PCR replicates for each sediment sample and stringent criteria to filter the dataset to remove artefacts introduced by ancient DNA damage, PCR, and sequencing. The PCR replicates (different PCR products from the same DNA extract) of each sample show some variations (Fig. 3) in the presence and abundance of ASVs, especially for ASVs amounting to less than 1 % per sample (Suppl. Table 2, Suppl. Figure S2). We tested whether these differences are significant using MRPP ( $p = 0.001$ , number of permutations 999, observed delta = 0.3683, expected delta = 0.6548), which suggests that PCR replicates of the same sample share significantly lower dissimilarities (average 39.3 %) compared to replicates between different samples (67.5 %). The PCR replicates are highly similar in the oldest samples up to 5.8 m depth and in the youngest sample. For the oldest sample at 8.85 m depth we additionally processed a sample replicate. The PCR replicates of both sediment samples at 8.85 m depth are highly similar and cluster tightly together in the NMDS plot (Fig. 3). Although a higher

number of replicates would improve the robustness of our analysis, this indicates that reasons other than sample age and associated DNA degradation controlled replicability in this study. It is possible that higher dissimilarities between PCR replicates are the result of low amounts of template molecules.

## 5 Temporal change of taxonomic richness and composition

We used *seada*DNA metabarcoding on samples that, due to the low sample size, represent temporally restricted snapshots of the different climatic intervals since the Late Weichselian. The dataset reveals diatom taxa that have mostly been reported from the Fram Strait in modern surveys (Karpuz and Schrader, 1990; Oksman et al., 2019; von Quillfeldt, 2000) and micropaleontological records (Oksman et al., 2017; Stabell, 1986). We detect typical ice-associated (sympagic) and/or cold-water diatoms (e.g. *Nitzschia* cf. *frigida*, *Cylindrotheca closterium*, *Thalassiosira nordenskiöldii*, *T. gravida*, *T. antarctica*, (Hasle, 1976; von Quillfeldt, 1997; von Quillfeldt et al., 2003), the epiphytic *Attheya septentrionalis* (Poulin et al., 2011) and *Pseudo-nitzschia granii* (Lovejoy et al., 2002)) alongside the temperate to warm-water species *Detonula pumila* (Hasle, 1976) and *Thalassiosira angulata* (Krawczyk et al., 2013; Luddington et al., 2016; Weckström et al., 2014), and some cosmopolitans (*Minidiscus trioculatus*, *Cerataulina pelagica* (Hasle, 1976)). Beyond marine diatoms, sequences are also assigned to species preferring fresh to brackish water (e.g. *Skeletonema subsalsum* (Hasle and Evensen, 1975), *Nitzschia palea*, *N. cf. paleacea* (Husted, 1930)). Taxonomic composition and richness changes with core depth and fits well into the framework reconstructed by other proxy data (biomarkers (Müller et al., 2012; Müller and Stein, 2014); foraminifers (Werner et al., 2011, 2013, 2016; Zamelczyk et al., 2014); dinoflagellates (Falardeau et al., 2018)).

Generally, the richness of both ASVs and unique taxonomic names (ASVs grouped based on identically assigned taxonomic names) is higher in samples dated to the last glacial in comparison to those dated to the Holocene (Fig 4). A shift of diatom *seada*DNA composition is captured with some ASVs being predominantly abundant in Late Glacial samples (e.g. those assigned to *Thalassiosira gravida*, *Minidiscus trioculatus*, and *Nitzschia* cf. *paleacea*) whereas others are mainly present in Holocene samples (e.g. *Chaetoceros* cf. *contortus* 1 SHE-2013, *C. cf. pseudobrevis* 1 SHE-2013) (Fig. 5). This shift is also strongly reflected at the family level (Fig. 6). A general trend that can also be observed at the family level is an inverse relationship ( $r = -0.61$ ,  $p = 0.046$ ) of the dominant families Thalassiosiraceae and Chaetoceraceae (Fig 6).

Samples dated to the Late Weichselian and the last glacial maximum (LGM) are characterized by highest overall richness with regard to ASVs (Fig.4). The samples contain high proportions of sympagic taxa (*Thalassiosira gravida* (4.8–28.1 %), *T. antarctica* (4.9–17.9 %), *T. delicata* (3.1–7.5 %), *Chaetoceros socialis* (0.3–7 %), *Nitzschia* cf. *frigida* (0.9–3.3%), *Porosira* (6–10.4 %)) as well as littoral (*Haslea avium* (1.3–3.4 %)), oceanic (*Minidiscus trioculatus* (1.4–4.4 %)), and brackish to freshwater taxa (*Nitzschia* cf. *paleacea* (0.3–5.9%), *Skeletonema subsalsum* (1–2.6 %)) (Fig. 5). Overall, ASVs assigned to the families Thalassiosiraceae, Bacillariaceae, and Naviculaceae (Fig. 6) dominate these samples. The relatively high proportions of sympagic taxa, especially of *Nitzschia* cf. *frigida*, in samples dated to the Late Weichselian and LGM are in accordance with previously reconstructed cold sea-surface conditions based on high proportions of the polar planktic



foraminifer *Neogloboquadrina pachyderma* (Zamelczyk et al., 2014), low dinocyst concentrations with a dominance of phototrophic taxa (Falardeau et al., 2018), and moderate concentrations of both the sea-ice proxy IP<sub>25</sub> and the phytoplankton biomarker brassicasterol (Müller and Stein, 2014). It is conceivable that a heterogeneous and dynamic environment produced by winter sea-ice cover with ice-free conditions during summer could allow diverse diatom communities to develop in the different habitats and over the seasons, which is suggested by the highest overall numbers of ASVs in samples dated to this time span. *Thalassiosira* and *Chaetoceros* are also abundant in the LGM microfossil assemblage identified further south at the Knipovich Ridge (eastern Fram Strait), although species of the genera *Rhizosolenia* and *Coscinodiscus* are strongly represented there as well (Bylinskaya et al., 2016).

Samples taken from Heinrich Stadial 1 and Bølling/Allerød phases are characterized by high proportions of ASVs assigned to *Thalassiosira* (35.1–65.3 %), *T. antarctica* (6.8–9.8 %), and *T. delicata* (4.1–5.3 %), and abundant *Skeletonema subsalsum* (2.1–5.6 %), whereas proportions of *Thalassiosira gravida* (3.4–11.3 %), *Nitzschia cf. frigida* (0.6–0.7 %), *Haslea avium* (0.2–0.4 %), and *Minidiscus trioculatus* (0–0.7%) are lower compared to the LGM. The taxonomic composition of the sample dated to Heinrich Stadial 1 suggests a partial re-organization of the diatom *sedaDNA* composition, which took place either gradually or abruptly sometime between 20.5 and 15.1 cal kyr BP. The low proportions of *Nitzschia cf. frigida* and other cold-water and sea-ice diatoms in samples corresponding to Heinrich Stadial 1 and Bølling/Allerød correspond well to low IP<sub>25</sub> and moderate brassicasterol concentrations in these samples, reflecting a reduction of the sea-ice cover (Müller and Stein, 2014). In the Bølling/Allerød sample the relatively high abundance of sequences assigned to the brackish–freshwater preferring diatom *Skeletonema subsalsum* (Hasle and Evensen, 1975) might be explained by elevated meltwater discharge from Svalbard. Low surface-water salinity was inferred previously from the dinocyst record of this core about a century earlier, possibly resulting from melting of the Barents Sea ice-sheet (Falardeau et al., 2018). Higher sedimentation rates and thus higher temporal resolution during the Bølling/Allerød could have affected the *sedaDNA* signal, yet the distinct shift in taxonomic composition suggests that this is rather an effect of the changing environmental conditions during this phase.

The Younger Dryas sample exhibits moderate proportions of *Thalassiosira antarctica* (11.8 %) and *T. delicata* (2.3 %) and is marked by the presence of the sympagic diatom *Cylindrotheca closterium* (0.3 %) and relative increases in *Porosira* (9.1 %), *Haslea avium* (5.7 %), and *Nitzschia cf. frigida* (1.3 %). Higher proportions of sympagic diatoms in this sample point towards colder conditions in comparison to the Bølling/Allerød sample and the presence of sea ice. Severe and extended sea-ice cover at the coring site are indicated by heterotrophic species in the dinocyst record (Falardeau et al., 2018), peak concentrations of IP<sub>25</sub>, and very low brassicasterol concentrations (Müller and Stein, 2014). A diatom microfossil record from Hinlopen Strait northwest of Spitzbergen detects first diatom occurrence during the Younger Dryas (10.8 <sup>14</sup>C kyr BP) with more than 30 % sea-ice associated species yet with a different taxonomic composition in comparison to our record (Koç et al., 2002).

The Early Holocene sample, which is marked by peak proportions of several families such as Skeletonemaceae (19.6 %) and Bacillariaceae (18.3%) and considerable increases of Chaetocerotaceae from 6.5 to 32.4 % combined with a strong decrease of Thalassiosiraceae from 57.8 to 44.4 % and Naviculaceae from 5.7 to 0.2 %, points towards a second partial re-organization of the taxonomic composition between 12.8 and 10.1 cal kyr BP. In particular, the sample is dominated by sequences assigned

290 to *Skeletonema* (16.8 %), *Thalassiosira* (14.5 %), and *Pseudo-nitzschia granii* (11.4 %). *Thalassiosira angulata*, a species associated with low sea-ice concentrations (Oksman et al., 2019) and temperate water masses (Weckström et al., 2014), displays peak proportions (1.7 %) in this sample. Lower proportions of sympagic diatoms in this sample are in accordance with the sea-ice retreat reconstructed from low IP<sub>25</sub> and high brassicasterol concentrations (Müller and Stein, 2014) and high proportions of the subpolar planktic foraminifers *Turborotalita quinqueloba* in another core further northwest (Werner et al., 295 2016). The diatom composition recorded by *seDNA* is also quite different to what has been found in microfossil records from the Fram Strait and the Greenland, Iceland, and Norwegian Seas, where first diatom microfossils are recorded between 13.4 and 9 cal kyr BP (Koç et al., 1993; Schröder-Ritzrau et al., 2001). Southwest of Svalbard, diatom-rich sediments dated to 10.1 and 9.8 cal kyr BP are attributed to the inflow of Atlantic surface water and a retreat of the Polar Front (Jessen et al., 2010; Stabell, 1986) and are composed mostly of *Coscinodiscus* spp., *Rhizosolenia*, and *Paralia sulcata* (Jessen et al., 2010; 300 Rigual-Hernández et al., 2017; Stabell, 1986). In concordance with this, sequences assigned to *Paralia* have highest proportions in the Early Holocene sample. *Coscinodiscus* is likely resolved to class level as sequences (Coscinodiscophyceae), and thus are present in our record albeit only sparsely. *Rhizosolenia* was not detected in this sample, which could be explained by poor DNA preservation or regional differences of the past diatom communities between the two coring sites.

The sample dated to the Mid-Holocene is marked by low diatom richness and by peak proportions of the families 305 Chaetoceraeae (32.4 %) and Anomoeneidae (2.1 %). It is dominated by sequences assigned to *Chaetoceros* cf. *contortus* 1SEH-2013 (27 %), *Thalassiosira* (23.3 %), and *T. antarctica* (9 %), accompanied by abundant *Porosira* (7.9 %) and *Skeletonema* (3.9 %). The relatively low richness suggests a loss of diversity between the Early and Mid-Holocene. The Mid-Holocene sample is characterized by a diatom composition that cannot be clearly related to sea ice, yet the low richness is supported by low diatom concentrations in a sediment core from Mohn Ridge (Koç et al., 1993) and low phytoplankton 310 productivity (Müller et al., 2012). The near absence of *Nitzschia* cf. *frigida*, however, does not match with the reconstructed strong sea-surface cooling and sea-ice growth from lower proportions of *Turborotalita quinqueloba* (Werner et al., 2011, 2013), low concentrations of CaCO<sub>3</sub>, and high concentrations of IP<sub>25</sub> and ice-rafted detritus (Müller et al., 2012).

The samples dated to the Roman Warm Period and the Little Ice Age (Late Holocene) have a similar richness to that of the Mid-Holocene sample. This phase is characterized by cold-water and ice-associated taxa with peak proportions of *Porosira* 315 (15.1–22.9 %) accompanied by significant proportions of *Chaetoceros* cf. *contortus* 1SEH-2013 (3.6–14.9 %), *Attheya* (2.3–6.7 %), *Thalassiosira antarctica* (1.4–4.7%), *T. delicata* (4.2–6.3 %), and *Chaetoceros socialis* (0.9–3.6 %). Richness is lowest in the youngest investigated sample which was dated to the Little Ice Age. The sample is characterized by a peak of *Chaetoceros* cf. *pseudobrevis* 1SEH-2013 (3.9 %), *Attheya* (6.7 %), and the sympagic diatoms *Porosira* (15.1 %), *Thalassiosira nordenskiöldii* (1.1 %), and *Nitzschia* cf. *frigida* (5.3 %). The Late-Holocene samples are characterized by 320 elevated proportions of sympagic taxa in comparison to the Mid-Holocene sample which is in agreement with Neoglacial cooling (increasing IP<sub>25</sub> concentrations and moderate to high concentrations of brassicasterol (Müller et al., 2012), increases in ice-rafted detritus and the dominance of polar planktic foraminifers (Werner et al., 2011)). Diatom microfossil records located southwest of Svalbard display an increase in diatom abundance since approximately 1.5 cal kyr BP (Rigual-Hernández

et al., 2017; Stabell, 1986). The microfossil record of Rigual-Hernández et al. (2017) is mostly composed of *Chaetoceros* resting spores which matches the increase of Chaetocerotaceae in our *sedaDNA* data. The composition of the record published by Stabell (1986) is more similar to Rigual-Hernández et al.'s (2017) Early-Holocene diatom maximum with *Coscinodiscus* spp., *Rhizosolenia*, and *Paralia sulcata*. Differences between composition and diversity of our samples and the records of Rigual-Hernández et al. (2017) and Stabell (1986) suggest either regional differences or differential preservation of microfossils and *sedaDNA* between the sites.

## 330 **6 Potential of diatom *sedaDNA* as a proxy for sea-ice distribution**

The *sedaDNA* record generally contains a high proportion of sequences assigned to cold-water and sea-ice associated diatoms, such as *Nitzschia cf. frigida*, *Thalassiosira antarctica*, and *T. nordenskiöldii* (Hasle, 1976; Poulin et al., 2011; von Quillfeldt, 1997; von Quillfeldt et al., 2003). Furthermore, pennate diatoms, which often dominate bottom ice layers (Van Leeuwe et al., 2018), display higher proportions in samples dated to the LGM, the Younger Dryas, and the Late Holocene with moderate IP<sub>25</sub> concentrations (Fig. 6). Among ASVs, *Nitzschia cf. frigida* (ASV 709), *Attheya* (ASV 28), and Bacillariophyta (ASV 154) show increased proportions with IP<sub>25</sub> exceeding 0.8 µg g<sup>-1</sup> organic carbon (Suppl. Fig. S3). In contrast, Bacillariophyta (ASV 245), *Gomphonema* (ASV 586), and *Thalassiosira* (ASV 1017) are detected only in samples with IP<sub>25</sub> of less than 0.5 µg g<sup>-1</sup> sediment.

The sea-ice proxy IP<sub>25</sub> is produced by only a few known ice-associated diatoms: *Haslea kjellmanii* (Cleve) Simonsen, *H. spicula* (Hickie) Lange-Bertalot, and *Pleurosigma stuxbergii* var. *rhomboides* (Cleve in Cleve and Grunow) Cleve (Brown et al., 2014; Limoges et al., 2018). We detected none of the known producers due to the incompleteness of the reference database (for the species of *Haslea*) and/or because the DNA of these species was not preserved in sufficient quantities as these species are low in abundance (Brown et al., 2014). Only for *Pleurosigma stuxbergii* is there a publicly available reference containing our marker region and is included in our database, whereas 15 references for non-IP<sub>25</sub> producing *Haslea* species are available in our database. Hence, an absence in our record does not mean an absence in the past communities. Sequences assigned to *Haslea avium* in our *sedaDNA* record do not show a linear relationship with IP<sub>25</sub>, but are nevertheless present in most samples. Our data, in combination with other proxy data from this core, suggest that sequences assigned to *Nitzschia cf. frigida* could be a useful indicator of past sea-ice distribution in *sedaDNA* records. Samples with low concentrations of the sea-ice proxy IP<sub>25</sub> (< 0.4 µg g<sup>-1</sup> total organic carbon) have less than 2 % of sequences assigned to *Nitzschia cf. frigida*, whereas samples characterized by moderate to high (> 0.8 µg g<sup>-1</sup> total organic carbon) concentrations of IP<sub>25</sub> display highly variable proportions without a clear relationship. *Nitzschia frigida* (Grunow) forms arborescent colonies (Medlin and Hasle, 1990) and is often abundant from late winter in the bottom layer of nearshore first-year ice and in multi-year ice in the Arctic pack-ice zones (Krawczyk et al., 2017; Melnikov et al., 2002; Olsen et al., 2017; Poulin et al., 2011; von Quillfeldt et al., 2003), but it can also be found in the water column during vernal under-ice or ice-edge blooms (Hasle and Heimdal, 1998; Olsen et al., 2017). Recently, De Schepper et al. (2019) used *sedaDNA* metabarcoding on a core from the East Greenland Sea and identified the

Mediophyceae OTU\_5051 which was significantly correlated to the IP<sub>25</sub> concentrations. The class Mediophyceae is the most diverse and dominant group in our dataset and contains several sea-ice associated taxa, such as *Thalassiosira antarctica*, *T. nordenskiöldii*, and *Porosira glacialis*. Yet their tendency to prevail in both sea ice and open water might be responsible for the non-linear relationship with IP<sub>25</sub> observed in our record. This is supported by Weckström et al. (2013) who find no specific response of sea-ice diatom microfossil composition to either IP<sub>25</sub> concentrations or observational sea-ice data in the Labrador Sea.

## 7 Conclusions

For the first time in a marine environment, our study targets high-resolution, diatom-specific sedimentary ancient DNA using a DNA metabarcoding approach. We show that diatom DNA is preserved with substantial taxonomic richness in the eastern Fram Strait over the past 30,000 years even though diatom microfossils are recorded in the Svalbard region only since the Younger Dryas. This highlights the advantage of our approach for paleoenvironmental reconstructions aiming to identify drivers of community-level taxonomic composition and diversity, especially in regions known for their irregular and poor diatom microfossil preservation, such as the Fram Strait. The *rbcL\_76* marker is highly diatom specific and provides detailed taxonomic resolution, mostly at genus and species level. The shortness of this marker is a strong advantage that leads to adequate replicability and high quality as diversity patterns do not show conspicuous signs of bias by age-associated DNA fragmentation. The *seadNA* record captures substantial temporal change of diatom taxonomic composition and richness with four compositional re-organizations: the first between 20.51 and 15.1 cal kyr BP, the second between 12.8 and 10.1 cal kyr BP, the third between 10.1 and 5 cal kyr BP, and the fourth between 5 and 1.6 cal kyr BP. Our record extends diatom compositional and diversity information back to the Late Weichselian as microfossil records in the Fram Strait are rare and extend only as far as the Younger Dryas due to poor preservation. Increasing proportions of pennate diatoms are associated with increased IP<sub>25</sub> concentrations, and sympagic diatoms are present, but with no clear pattern with regard to biomarker signals. Recommendations for future work with sedimentary ancient DNA in the context of sea-ice reconstructions include the preparation of reference genomes and a more targeted enrichment, for example of genes that help species to adapt to sea ice and allow them to cope with rapidly changing environmental conditions.

## 380 Supplement

### Code/Data availability

The raw sequencing data are deposited at Dryad (Zimmermann et al., 2020). The rarefaction script is available at [https://github.com/StefanKruse/R\\_Rarefaction](https://github.com/StefanKruse/R_Rarefaction) (Kruse, 2019).

## Author contributions

385 H.H.Z., K.R.S.L., and U.H. conceived and designed the study; H.H.Z. performed experiments and data analysis; K.S.L. sampled the core and provided laboratory equipment; S.K. wrote resampling R script; R.S., R.T., J.M. and U.H. provided the framework for the study (samples, funding); H.H.Z. wrote the paper that all co-authors commented on.

## Competing interests

The authors declare that they have no conflict of interest.

## 390 Acknowledgements

We thank Sarah Olischläger and Iris Eder for support with the laboratory work and the captain and crew of the RV Maria S. Merian. This research was funded by the Initiative and Networking Fund of the Helmholtz Association. U.H. is financed by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 772852, GlacialLegacy Project). J.M. received financial support from a Helmholtz Research Grant VH-NG-395 1101. We thank Cathy Jenks for English correction.

## References

- Aagaard, K.: Inflow from the Atlantic Ocean to the Polar Basin, in *The Arctic Ocean*, pp. 69–81, Palgrave Macmillan, London., 1982.
- 400 Bedoshvili, Ye. D., Popkova, T. P. and Likhoshway, Ye. V.: Chloroplast structure of diatoms of different classes, *Cell Tiss. Biol.*, 3(3), 297–310, doi:10.1134/S1990519X09030122, 2009.
- Belt, S. T.: Source-specific biomarkers as proxies for Arctic and Antarctic sea ice, *Organic Geochemistry*, 125, 277–298, doi:10.1016/j.orggeochem.2018.10.002, 2018.
- 405 Belt, S. T. and Müller, J.: The Arctic sea ice biomarker IP<sub>25</sub>: a review of current understanding, recommendations for future research and applications in palaeo sea ice reconstructions, *Quaternary Science Reviews*, 79, 9–25, doi:10.1016/j.quascirev.2012.12.001, 2013.
- Belt, S. T., Massé, G., Rowland, S. J., Poulin, M., Michel, C. and LeBlanc, B.: A novel chemical fossil of palaeo sea ice: IP<sub>25</sub>, *Organic Geochemistry*, 38(1), 16–27, doi:10.1016/j.orggeochem.2006.09.013, 2007.
- Birks, C. J. A. and Koç, N.: A high-resolution diatom record of late-Quaternary sea-surface temperatures and oceanographic conditions from the eastern Norwegian Sea, *Boreas*, 31(4), 323–344, doi:10.1111/j.1502-3885.2002.tb01077.x, 2002.
- 410 Boere, A. C., Abbas, B., Rijpstra, W. I. C., Versteegh, G. J. M., Volkman, J. K., Sinninghe Damsté, J. S. and Coolen, M. J. L.: Late-Holocene succession of dinoflagellates in an Antarctic fjord using a multi-proxy approach: paleoenvironmental genomics, lipid biomarkers and palynomorphs, *Geobiology*, 7(3), 265–281, doi:10.1111/j.1472-4669.2009.00202.x, 2009.

- Boere, A. C., Rijpstra, W. I. C., De Lange, G. J., Sinninghe Damsté, J. S. and Coolen, M. J. L.: Preservation potential of ancient plankton DNA in Pleistocene marine sediments, *Geobiology*, 9(5), 377–393, doi:10.1111/j.1472-4669.2011.00290.x, 2011a.
- 415 Boere, A. C., Sinninghe Damsté, J. S., Rijpstra, W. I. C., Volkman, J. K. and Coolen, M. J. L.: Source-specific variability in post-depositional DNA preservation with potential implications for DNA based paleoecological records, *Organic Geochemistry*, 42(10), 1216–1225, doi:10.1016/j.orggeochem.2011.08.005, 2011b.
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P. and Coissac, E.: OBITools: a Unix-inspired software package for DNA metabarcoding, *Molecular Ecology Resources*, 16, 176–182, doi:10.1111/1755-0998.12428, 2016.
- 420 Brown, T. A., Belt, S. T., Tatarek, A. and Mundy, C. J.: Source identification of the Arctic sea ice proxy IP<sub>25</sub>, *Nature Communications*, 5, 4197, doi:10.1038/ncomms5197, 2014.
- Budéus, G.: Short Cruise Report RV Maria S. Merian Cruise MSM05/5, University of Hamburg, Institute of Oceanography. [online] Available from: <https://www.lfd.uni-hamburg.de/en/merian/wochenberichte/wochenberichte-merian/msm05/msm05-5-scr.pdf>, 2007.
- 425 Bylinskaya, M. E., Golovina, L. A., Radionova, E. P., Pokrovskii, B. G. and Lavrushin, V. Yu.: Paleoenvironments in the Fram Strait during Marine Isotope Stages 2–6 based on planktonic paleobiological and stable-isotope proxies and ice-rafted debris, *Quaternary International*, 420, 272–279, doi:10.1016/j.quaint.2015.09.086, 2016.
- Callahan, B. J., McMurdie, P. J. and Holmes, S. P.: Exact sequence variants should replace operational taxonomic units in marker-gene data analysis, *ISME J*, 11(12), 2639–2643, doi:10.1038/ismej.2017.119, 2017.
- 430 Cherkasheva, A., Bracher, A., Melsheimer, C., Köberle, C., Gerdes, R., Nöthig, E.-M., Bauerfeind, E. and Boetius, A.: Influence of the physical environment on polar phytoplankton blooms: A case study in the Fram Strait, *Journal of Marine Systems*, 132, 196–207, doi:10.1016/j.jmarsys.2013.11.008, 2014.
- Coolen, M. J. L., Muyzer, G., Rijpstra, W. I. C., Schouten, S., Volkman, J. K. and Sinninghe Damsté, J. S.: Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake, *Earth and Planetary Science Letters*, 223(1), 225–239, doi:10.1016/j.epsl.2004.04.014, 2004.
- 435 Coolen, M. J. L., Boere, A., Abbas, B., Baas, M., Wakeham, S. G. and Sinninghe Damsté, J. S.: Ancient DNA derived from alkenone-biosynthesizing haptophytes and other algae in Holocene sediments from the Black Sea, *Paleoceanography*, 21(1), doi:10.1029/2005PA001188, 2006.
- Coolen, M. J. L., Volkman, J. K., Abbas, B., Muyzer, G., Schouten, S. and Damsté, J. S. S.: Identification of organic matter sources in sulfidic late Holocene Antarctic fjord sediments from fossil rDNA sequence analysis, *Paleoceanography*, 22(2), doi:10.1029/2006PA001309, 2007.
- 440 Coolen, M. J. L., Saenz, J. P., Giosan, L., Trowbridge, N. Y., Dimitrov, P., Dimitrov, D. and Eglinton, T. I.: DNA and lipid molecular stratigraphic records of haptophyte succession in the Black Sea during the Holocene, *Earth and Planetary Science Letters*, 284(3), 610–621, doi:10.1016/j.epsl.2009.05.029, 2009.
- 445 Coolen, M. J. L., Orsi, W. D., Balkema, C., Quince, C., Harris, K., Sylva, S. P., Filipova-Marinova, M. and Giosan, L.: Evolution of the plankton paleome in the Black Sea from the Deglacial to Anthropocene, *PNAS*, 110(21), 8609–8614, doi:10.1073/pnas.1219283110, 2013.
- De Schepper, S., Ray, J. L., Skaar, K. S., Sadatzki, H., Ijaz, U. Z., Stein, R. and Larsen, A.: The potential of sedimentary ancient DNA for reconstructing past sea ice evolution, *ISME J*, doi:10.1038/s41396-019-0457-1, 2019.

- 450 Degerlund, M., Huseby, S., Zingone, A., Sarno, D. and Landfald, B.: Functional diversity in cryptic species of *Chaetoceros socialis* Lauder (Bacillariophyceae), *J Plankton Res*, 34(5), 416–431, doi:10.1093/plankt/fbs004, 2012.
- Dell’Anno, A., Stefano, B. and Danovaro, R.: Quantification, base composition, and fate of extracellular DNA in marine sediments, *Limnology and Oceanography*, 47(3), 899–905, doi:10.4319/lo.2002.47.3.0899, 2002.
- 455 Dulias, K., Stoof-Leichsenring, K. R., Pestryakova, L. A. and Herzsuh, U.: Sedimentary DNA versus morphology in the analysis of diatom-environment relationships, *J Paleolimnol*, 57(1), 51–66, doi:10.1007/s10933-016-9926-y, 2017.
- Epp, L. S., Gussarova, G., Boessenkool, S., Olsen, J., Haile, J., Schröder-Nielsen, A., Ludikova, A., Hassel, K., Stenøien, H. K., Funder, S., Willerslev, E., Kjær, K. and Brochmann, C.: Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes, *Quaternary Science Reviews*, 117(0318), 152–163, doi:10.1016/j.quascirev.2015.03.027, 2015.
- 460 Epp, L. S., Kruse, S., Kath, N. J., Stoof-Leichsenring, K. R., Tiedemann, R., Pestryakova, L. A. and Herzsuh, U.: Temporal and spatial patterns of mitochondrial haplotype and species distributions in Siberian larches inferred from ancient environmental DNA and modeling, *Scientific Reports*, 8(1), doi:10.1038/s41598-018-35550-w, 2018.
- Epp, L. S., Zimmermann, H. H. and Stoof-Leichsenring, K. R.: Sampling and Extraction of Ancient DNA from Sediments, in *Ancient DNA: Methods and Protocols*, edited by B. Shapiro, A. Barlow, P. D. Heintzman, M. Hofreiter, J. L. A. Paijmans, and A. E. R. Soares, pp. 31–44, Springer New York, New York, NY., 2019.
- 465 Fahl, K. and Stein, R.: Modern organic carbon deposition in the Laptev Sea and the adjacent continental slope: surface water productivity vs. terrigenous input, *Organic Geochemistry*, 26(5), 379–390, doi:10.1016/S0146-6380(97)00007-7, 1997.
- Falardeau, J., de Vernal, A. and Spielhagen, R. F.: Paleoceanography of northeastern Fram Strait since the last glacial maximum: Palynological evidence of large amplitude changes, *Quaternary Science Reviews*, 195, 133–152, doi:10.1016/j.quascirev.2018.06.030, 2018.
- 470 Fetterer, F., Knowles, K., Meier, W. N., Savoie, M. and Windnagel, A. K.: Sea Ice Index, Version 3. Monthly and daily GIS compatible shapefiles of median ice extent, , doi:10.7265/n5k072f8, 2017.
- Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P. and Pompanon, F.: An In silico approach for the evaluation of DNA barcodes, *BMC Genomics*, 11(1), 434, doi:10.1186/1471-2164-11-434, 2010.
- 475 Gersonde, R.: The expedition of the research vessel “Sonne” to the subpolar North Pacific and the Bering Sea in 2009 (SO202-INOPEX), *Berichte zur Polar- und Meeresforschung = Reports on polar and marine research*, 643, doi:https://doi.org/10.2312, 2012.
- Giosan, L., Coolen, M. J. L., Kaplan, J. O., Constantinescu, S., Filip, F., Filipova-Marinova, M., Kettner, A. J. and Thom, N.: Early anthropogenic transformation of the Danube-Black Sea system, *Scientific Reports*, 2(1), doi:10.1038/srep00582, 2012.
- 480 Gradinger, R. R. and Baumann, M. E. M.: Distribution of phytoplankton communities in relation to the large-scale hydrographical regime in the Fram Strait, *Mar. Biol.*, 111(2), 311–321, doi:10.1007/BF01319714, 1991.
- Guo, L., Sui, Z., Zhang, S., Ren, Y. and Liu, Y.: 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, *Int. J. Syst. Evol. Microbiol.*, 65(Pt 4), 1369–1380, doi:10.1099/ijs.0.000076, 2015.

- 485 Harrison, W. G. and Cota, G. F.: Primary production in polar waters: relation to nutrient availability, 1, 10(1), 87–104, doi:10.3402/polar.v10i1.6730, 1991.
- Hasle, G. R.: The biogeography of some marine planktonic diatoms, *Deep Sea Research and Oceanographic Abstracts*, 23(4), 319-IN6, doi:10.1016/0011-7471(76)90873-1, 1976.
- 490 Hasle, G. R. and Evensen, D. L.: Brackish-water and fresh-water species of the diatom genus *Skeletonema* Grev. I. *Skeletonema subsalsum* (A. Cleve) Bethge, *Phycologia*, 14(4), 283–297, doi:10.2216/i0031-8884-14-4-283.1, 1975.
- Hasle, G. R. and Heimdal, B. R.: The net phytoplankton in Kongsfjorden, Svalbard, July 1988, with general remarks on species composition of Arctic phytoplankton, *Polar Research*, 17(1), 31–52, doi:10.3402/polar.v17i1.6605, 1998.
- 495 Hodal, H., Falk-Petersen, S., Hop, H., Kristiansen, S. and Reigstad, M.: Spring bloom dynamics in Kongsfjorden, Svalbard: nutrients, phytoplankton, protozoans and primary production, *Polar Biol*, 35(2), 191–203, doi:10.1007/s00300-011-1053-7, 2012.
- Hollander, M. and Wolfe, D. A.: Nonparametric statistical methods., *Biometrische Zeitschrift*, 17(8), 526–526, doi:10.1002/bimj.19750170808, 1975.
- Husted, F.: Bacillariophyta (Diatomeae), in *Die Subwasser Flora Mitteleuropas*, vol. 10, edited by A. Pascher, p. 466, Gustav Fischer, Jena., 1930.
- 500 Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P., Endo, H., Gasol, J. M., Gregory, A. C., Mahé, F., Rigonato, J., Royo-Llonch, M., Salazar, G., Sanz-Sáez, I., Scalco, E., Soviadan, D., Zayed, A. A., Zingone, A., Labadie, K., Ferland, J., Marec, C., Kandels, S., Picheral, M., Dimier, C., Poullain, J., Pisarev, S., Carmichael, M., Pesant, S., Babin, M., Boss, E., Iudicone, D., Jaillon, O., Acinas, S. G., Ogata, H., Pelletier, E., Stemmann, L., Sullivan, M. B., Sunagawa, S., Bopp, L., de Vargas, C., Karp-Boss, L., Wincker, P., Lombard, F., Bowler, C., Zinger, L., Acinas, S. G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., de Vargas, C., Follows, M., Gorsky, G., Grimsley, N., Guidi, L., Hingamp, P., Iudicone, D., Jaillon, O., Kandels, S., Karp-Boss, L., Karsenti, E., Not, F., Ogata, H., Pesant, S., Poulton, N., Raes, J., Sardet, C., Speich, S., Stemmann, L., Sullivan, M. B., Sunagawa, S. and Wincker, P.: Global Trends in Marine Plankton Diversity across Kingdoms of Life, *Cell*, 179(5), 1084-1097.e21, doi:10.1016/j.cell.2019.10.008, 2019.
- 510 Jessen, S. P., Rasmussen, T. L., Nielsen, T. and Solheim, A.: A new Late Weichselian and Holocene marine chronology for the western Svalbard slope 30,000–0 cal years BP, *Quaternary Science Reviews*, 29(9), 1301–1312, doi:10.1016/j.quascirev.2010.02.020, 2010.
- Juggins, S.: rioja: Analysis of Quaternary science data, R package version 0.7-3., 2012.
- Karpuz, N. K. and Schrader, H.: Surface sediment diatom distribution and Holocene paleotemperature variations in the Greenland, Iceland and Norwegian Sea, *Paleoceanography*, 5(4), 557–580, doi:10.1029/PA005i004p00557, 1990.
- 515 Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Humbert, J. F. and Bouchez, A.: Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms, *Molecular Ecology Resources*, 13(4), 607–619, doi:10.1111/1755-0998.12105, 2013.
- Kirkpatrick, J. B., Walsh, E. A. and D’Hondt, S.: Fossil DNA persistence and decay in marine sediment over hundred-thousand-year to million-year time scales, *Geology*, 44(8), 615–618, doi:10.1130/G37933.1, 2016.



- 520 Koç, N., Jansen, E. and Haflidason, H.: Paleoceanographic reconstructions of surface ocean conditions in the Greenland, Iceland and Norwegian seas through the last 14 ka based on diatoms, *Quaternary Science Reviews*, 12(2), 115–140, doi:10.1016/0277-3791(93)90012-B, 1993.
- Koç, N., Klitgaard-Kristensen, D., Hasle, K., Forsberg, C. F. and Solheim, A.: Late glacial palaeoceanography of Hinlopen Strait, northern Svalbard, 1, 21(2), 307–314, doi:10.3402/polar.v21i2.6492, 2002.
- 525 Kohly, A.: Diatom flux and species composition in the Greenland Sea and the Norwegian Sea in 1991–1992, *Marine Geology*, 145(3–4), 293–312, doi:10.1016/S0025-3227(97)00115-1, 1998.
- Krawczyk, D. W., Witkowski, A., Lloyd, J., Moros, M., Harff, J. and Kuijpers, A.: Late-Holocene diatom derived seasonal variability in hydrological conditions off Disko Bay, West Greenland, *Quaternary Science Reviews*, 67, 93–104, doi:10.1016/j.quascirev.2013.01.025, 2013.
- 530 Krawczyk, D. W., Witkowski, A., Moros, M., Lloyd, J. M., Høyer, J. L., Miettinen, A. and Kuijpers, A.: Quantitative reconstruction of Holocene sea ice and sea surface temperature off West Greenland from the first regional diatom data set, *Paleoceanography*, 32(1), 18–40, doi:10.1002/2016PA003003, 2017.
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B. and Gillespie, R. G.: Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding, *Scientific Reports*, 7(1), 17668, doi:10.1038/s41598-017-17333-x, 2017.
- 535 Kruse, S.: R code for resampling and thus normalizing of count data to the minimum number of counts across a set of samples (e.g. sedaDNA sequence/pollen taxa counts per sample along a sediment core), R. [online] Available from: [https://github.com/StefanKruse/R\\_Rarefaction](https://github.com/StefanKruse/R_Rarefaction) (Accessed 1 August 2019), 2019.
- Lalande, C., Bauerfeind, E., Nöthig, E.-M. and Beszczynska-Möller, A.: Impact of a warm anomaly on export fluxes of biogenic matter in the eastern Fram Strait, *Progress in Oceanography*, 109, 70–77, doi:10.1016/j.pocan.2012.09.006, 2013.
- Lejzerowicz, F., Esling, P., Majewski, W., Szczuciński, W., Decelle, J., Obadia, C., Arbizu, P. M. and Pawlowski, J.: Ancient DNA complements microfossil record in deep-sea subsurface sediments, *Biology Letters*, 9(4), 20130283, doi:10.1098/rsbl.2013.0283, 2013.
- 545 Li, C. L., Ashworth, M. P., Witkowski, A., Dąbek, P., Medlin, L. K., Kooistra, W. H. C. F., Sato, S., Zgłobicka, I., Kurzydłowski, K. J., Theriot, E. C., Sabir, J. S. M., Khiyami, M. A., Mutwakil, M. H. Z., Sabir, M. J., Alharbi, N. S., Hajarrah, N. H., Qing, S. and Jansen, R. K.: New insights into Plagiogrammaceae (Bacillariophyta) based on multigene phylogenies and morphological characteristics with the description of a new genus and three new species, *PLOS ONE*, 10(10), e0139300, doi:10.1371/journal.pone.0139300, 2015.
- 550 Limoges, A., Ribeiro, S., Weckström, K., Heikkilä, M., Zamelczyk, K., Andersen, T. J., Tallberg, P., Massé, G., Rysgaard, S., Nørgaard-Pedersen, N. and Seidenkrantz, M.-S.: Linking the modern distribution of biogenic proxies in high Arctic Greenland shelf sediments to sea ice, primary production, and Arctic-Atlantic inflow, *Journal of Geophysical Research: Biogeosciences*, 123(3), 760–786, doi:10.1002/2017JG003840, 2018.
- 555 Lovejoy, C., Legendre, L., Martineau, M.-J., Bâcle, J. and von Quillfeldt, C. H.: Distribution of phytoplankton and other protists in the North Water, *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(22), 5027–5047, doi:10.1016/S0967-0645(02)00176-5, 2002.
- Luddington, I. A., Lovejoy, C. and Kaczmarek, I.: Species-rich meta-communities of the diatom order Thalassiosirales in the Arctic and northern Atlantic Ocean, *J Plankton Res*, 38(4), 781–797, doi:10.1093/plankt/fbw030, 2016.

- Medlin, L. K. and Hasle, G. R.: Some *Nitzschia* and related diatom species from fast ice samples in the Arctic and Antarctic, *Polar Biology*, 10(6), doi:10.1007/BF00233693, 1990.
- 560 Melnikov, I. A., Kolosova, E. G., Welch, H. E. and Zhitina, L. S.: Sea ice biological communities and nutrient dynamics in the Canada Basin of the Arctic Ocean, *Deep Sea Research Part I: Oceanographic Research Papers*, 49(9), 1623–1649, doi:10.1016/S0967-0637(02)00042-0, 2002.
- Miettinen, A., Divine, D. V., Husum, K., Koç, N. and Jennings, A.: Exceptional ocean surface conditions on the SE Greenland shelf during the Medieval Climate Anomaly, *Paleoceanography*, 30(12), 1657–1674, doi:10.1002/2015PA002849, 2015.
- 565 Morard, R., Lejzerowicz, F., Darling, K. F., Lecroq-Bennet, B., Winther Pedersen, M., Orlando, L., Pawlowski, J., Mülitz, S., Vargas, C. de and Kucera, M.: Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroecology, *Biogeosciences*, 14(11), 2741–2754, doi:https://doi.org/10.5194/bg-14-2741-2017, 2017.
- Müller, J. and Stein, R.: High-resolution record of late glacial and deglacial sea ice changes in Fram Strait corroborates ice–ocean interactions during abrupt climate shifts, *Earth and Planetary Science Letters*, 403, 446–455, doi:10.1016/j.epsl.2014.07.016, 2014.
- 570 Müller, J., Massé, G., Stein, R. and Belt, S. T.: Variability of sea-ice conditions in the Fram Strait over the past 30,000 years, *Nature Geoscience*, 2(11), 772–776, doi:10.1038/ngeo665, 2009.
- Müller, J., Werner, K., Stein, R., Fahl, K., Moros, M. and Jansen, E.: Holocene cooling culminates in sea ice oscillations in Fram Strait, *Quaternary Science Reviews*, 47, 1–14, doi:10.1016/j.quascirev.2012.04.024, 2012.
- 575 Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., Green, R. E. and Shapiro, B.: Minimizing polymerase biases in metabarcoding, *Molecular Ecology Resources*, 18(5), 927–939, doi:10.1111/1755-0998.12895, 2018.
- Nöthig, E.-M., Bracher, A., Engel, A., Metfies, K., Niehoff, B., Peeken, I., Bauerfeind, E., Cherkasheva, A., Gäbler-Schwarz, S., Hardge, K., Kiliyas, E., Kraft, A., Kidane, Y. M., Lalande, C., Piontek, J., Thomisch, K. and Wurst, M.: Summertime plankton ecology in Fram Strait—a compilation of long- and short-term observations, *Polar Research*, 34(1), 23349, doi:10.3402/polar.v34.23349, 2015.
- 580 Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O’Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H. and Wagner, H.: *vegan: Community Ecology Package*. R package version 2.0-2., 2011.
- 585 Oksman, M., Weckström, K., Miettinen, A., Ojala, A. E. K. and Salonen, V.-P.: Late Holocene shift towards enhanced oceanic variability in a high-Arctic Svalbard fjord (79°N) at 2500 cal. yr BP, *Arktos*, 3(1), 4, doi:10.1007/s41063-017-0032-9, 2017.
- Oksman, M., Juggins, S., Miettinen, A., Witkowski, A. and Weckström, K.: The biogeography and ecology of common diatom species in the northern North Atlantic, and their implications for paleoceanographic reconstructions, *Marine Micropaleontology*, 148, 1–28, doi:10.1016/j.marmicro.2019.02.002, 2019.
- 590 Olsen, L. M., Laney, S. R., Duarte, P., Kauko, H. M., Fernández-Méndez, M., Mundy, C. J., Rösel, A., Meyer, A., Itkin, P., Cohen, L., Peeken, I., Tatarek, A., Róžańska-Pluta, M., Wiktor, J., Taskjelle, T., Pavlov, A. K., Hudson, S. R., Granskog, M. A., Hop, H. and Assmy, P.: The seeding of ice algal blooms in Arctic pack ice: The multiyear ice seed repository hypothesis, *Journal of Geophysical Research: Biogeosciences*, 122(7), 1529–1548, doi:10.1002/2016JG003668, 2017.

- 595 Parducci, L., Jørgensen, T., Tollefsrud, M. M., Elverland, E., Alm, T., Fontana, S. L., Bennett, K. D., Haile, J., Matetovici, I.,  
Suyama, Y., Edwards, M. E., Andersen, K., Rasmussen, M., Boessenkool, S., Coissac, E., Brochmann, C., Taberlet, P.,  
Houmark-Nielsen, M., Larsen, N. K., Orlando, L., Gilbert, M. T. P., Kjaer, K. H., Alsos, I. G. and Willerslev, E.: Glacial  
survival of boreal trees in northern Scandinavia, *Science*, 335(6072), 1083–1086, doi:10.1126/science.1216043, 2012.
- 600 Pawłowska, J., Łącka, M., Kucharska, M., Pawłowski, J. and Zajączkowski, M.: Multiproxy evidence of the Neoglacial  
expansion of Atlantic Water to eastern Svalbard, *Climate of the Past*, 16(2), 487–501, doi:https://doi.org/10.5194/cp-16-487-  
2020, 2020.
- Pierella Karlusich, J. J., Ibarbalz, F. M. and Bowler, C.: Phytoplankton in the Tara Ocean, *Annual Review of Marine Science*,  
12(1), 233–265, doi:10.1146/annurev-marine-010419-010706, 2020.
- 605 Poulin, M., Daugbjerg, N., Gradinger, R., Ilyash, L., Ratkova, T. and Quillfeldt, C. von: The pan-Arctic biodiversity of marine  
pelagic and sea-ice unicellular eukaryotes: a first-attempt assessment, *Mar Biodiv*, 41(1), 13–28, doi:10.1007/s12526-010-  
0058-8, 2011.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., Eds.: *Numerical recipes in C: the art of scientific  
computing*, Cambridge University Press, Cambridge [Cambridgeshire]; New York., 1988.
- von Quillfeldt, C. H.: Distribution of diatoms in the Northeast Water Polynya, Greenland, *Journal of Marine Systems*, 10(1),  
211–240, doi:10.1016/S0924-7963(96)00056-5, 1997.
- 610 von Quillfeldt, C. H.: Common diatom species in Arctic spring blooms: their distribution and abundance, *Botanica Marina*,  
43(6), doi:10.1515/BOT.2000.050, 2000.
- von Quillfeldt, C. H., Ambrose, W. G. and Clough, L. M.: High number of diatom species in first-year ice from the Chukchi  
Sea, *Polar Biology*, 26(12), 806–818, doi:10.1007/s00300-003-0549-1, 2003.
- 615 R Core Team: *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna,  
Austria. [online] Available from: <https://www.R-project.org/>, 2018.
- Reschke, M., Kunz, T. and Laepple, T.: Comparing methods for analysing time scale dependent correlations in irregularly  
sampled time series data, *Computers & Geosciences*, 123, 65–72, doi:10.1016/j.cageo.2018.11.009, 2019.
- 620 Rigual-Hernández, A. S., Colmenero-Hidalgo, E., Martrat, B., Bárcena, M. A., de Vernal, A., Sierro, F. J., Flores, J. A.,  
Grimalt, J. O., Henry, M. and Lucchi, R. G.: Svalbard ice-sheet decay after the Last Glacial Maximum: New insights from  
micropalaeontological and organic biomarker paleoceanographical reconstructions, *Palaeogeography, Palaeoclimatology,  
Palaeoecology*, 465, 225–236, doi:10.1016/j.palaeo.2016.10.034, 2017.
- Round, F. E., Crawford, R. M. and Mann, D. G.: *Diatoms: Biology and morphology of the genera*, Cambridge University  
Press., 1990.
- 625 Schlitzer, R.: Interactive analysis and visualization of geoscience data with Ocean Data View, <https://odv.awi.de>, *Computers  
& Geosciences*, 28(10), 1211–1218, doi:10.1016/S0098-3004(02)00040-7, 2002.
- Schnell, I. B., Bohmann, K. and Gilbert, M. T. P.: Tag jumps illuminated – reducing sequence-to-sample misidentifications in  
metabarcoding studies, *Mol Ecol Resour*, 15(6), 1289–1303, doi:10.1111/1755-0998.12402, 2015.

- 630 Schröder-Ritzrau, A., Andrleit, H., Jensen, S., Samtleben, C., Schäfer, P., Matthiessen, J., Hass, H. C., Kohly, A. and Thiede, J.: Distribution, export and alteration of fossilizable plankton in the Nordic Seas, in *The northern North Atlantic: a changing environment*, edited by P. Schäfer, W. Ritzrau, M. Schlüter, and J. Thiede, pp. 81–104, Springer, Berlin, Heidelberg., 2001.
- Sieracki, M. E., Poulton, N. J., Jaillon, O., Wincker, P., Vargas, C. de, Rubinat-Ripoll, L., Stepanauskas, R., Logares, R. and Massana, R.: Single cell genomics yields a wide diversity of small planktonic protists across major ocean ecosystems, *Scientific Reports*, 9(1), 6025, doi:10.1038/s41598-019-42487-1, 2019.
- 635 Smirnova, M. A., Kazarina, G. Kh., Matul, A. G. and Max, L.: Diatom evidence for paleoclimate changes in the northwestern Pacific during the last 20000 years, *Oceanology*, 55(3), 383–389, doi:10.1134/S0001437015030157, 2015.
- Stabell, B.: A diatom maximum horizon in upper Quaternary deposits, *Geologische Rundschau*, 75(1), 175–184, doi:10.1007/BF01770186, 1986.
- Stabell, B.: The diatom flora in the Fram Strait. A comparison with the Pleistocene/Holocene boundary flora in the Norwegian Sea, *Polar Research*, 5(3), 321–322, doi:10.3402/polar.v5i3.6902, 1987.
- 640 Stein, R., Fahl, K. and Müller, J.: Proxy reconstruction of Cenozoic Arctic Ocean sea-ice history - from IRD to IP<sub>25</sub>, *Polarforschung*, 82(1), 37–71, doi:10.2312/polarforschung.82.1.37, 2012.
- Stein, R., Fahl, K., Gierz, P., Niessen, F. and Lohmann, G.: Arctic Ocean sea ice cover during the penultimate glacial and the last interglacial, *Nat Commun*, 8(1), 1–13, doi:10.1038/s41467-017-00552-1, 2017.
- 645 Stoof-Leichsenring, K. R., Epp, L. S., Trauth, M. H. and Tiedemann, R.: Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation, *Molecular Ecology*, 21(8), 1918–1930, doi:10.1111/j.1365-294X.2011.05412.x, 2012.
- Stoof-Leichsenring, K. R., Bernhardt, N., Pestryakova, L. A., Epp, L. S., Herzsuh, U. and Tiedemann, R.: A combined paleolimnological/genetic analysis of diatoms reveals divergent evolutionary lineages of *Staurosira* and *Staurosirella* (Bacillariophyta) in Siberian lake sediments along a latitudinal transect, *J Paleolimnol*, 52(1–2), 77–93, doi:10.1007/s10933-014-9779-1, 2014.
- 650 Stoof-Leichsenring, K. R., Herzsuh, U., Pestryakova, L. A., Klemm, J., Epp, L. S. and Tiedemann, R.: Genetic data from algae sedimentary DNA reflect the influence of environment over geography, *Scientific Reports*, 5, 12924, doi:10.1038/srep12924, 2015.
- 655 Syring, N., Stein, R., Fahl, K., Vahlenkamp, M., Zehlich, M., Spielhagen, R. F. and Niessen, F.: Holocene changes in sea-ice cover and polynya formation along the eastern North Greenland shelf: New insights from biomarker records, *Quaternary Science Reviews*, 231, 106173, doi:10.1016/j.quascirev.2020.106173, 2020.
- Untersteiner, N.: On the ice and heat balance in Fram Strait, *J. Geophys. Res.*, 93(C1), 527, doi:10.1029/JC093iC01p00527, 1988.
- 660 Van Leeuwe, M., Tedesco, L., Arrigo, K. R., Assmy, P., Campbell, K., Meiners, K. M., Rintala, J.-M., Selz, V., Thomas, D. N. and Stefels, J.: Microalgal community structure and primary production in Arctic and Antarctic sea ice: A synthesis, *Elem Sci Anth*, 6(1), 4, doi:10.1525/elementa.267, 2018.
- Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., Tapolczai, K. and Domaizon, I.: Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring, *Methods in Ecology and Evolution*, 9(4), 1060–1069, doi:10.1111/2041-210X.12960, 2018.

- 665 Volkman, J. K.: A review of sterol markers for marine and terrigenous organic matter, *Organic Geochemistry*, 9(2), 83–99, doi:10.1016/0146-6380(86)90089-6, 1986.
- Weckström, K., Massé, G., Collins, L. G., Hanhijärvi, S., Bouloubassi, I., Sicre, M.-A., Seidenkrantz, M.-S., Schmidt, S., Andersen, T. J., Andersen, M. L., Hill, B. and Kuijpers, A.: Evaluation of the sea ice proxy IP<sub>25</sub> against observational and diatom proxy data in the SW Labrador Sea, *Quaternary Science Reviews*, 79, 53–62, doi:10.1016/j.quascirev.2013.02.012, 670 2013.
- Weckström, K., Miettinen, A., Caissie, B., Pearce, C., Ellegaard, M., Krawczyk, D. and Witkowski, A.: Sea surface temperatures in Disko Bay during the Little Ice Age – caution needs to be exercised before assigning *Thalassiosira kushirensis* resting spore as a warm-water indicator in palaeoceanographic studies, *Quaternary Science Reviews*, 101, 234–237, doi:10.1016/j.quascirev.2014.07.015, 2014.
- 675 Werner, K., Spielhagen, R. F., Bauch, D., Hass, H. C., Kandiano, E. and Zamelczyk, K.: Atlantic Water advection to the eastern Fram Strait — Multiproxy evidence for late Holocene variability, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 308(3), 264–276, doi:10.1016/j.palaeo.2011.05.030, 2011.
- Werner, K., Spielhagen, R. F., Bauch, D., Hass, H. C. and Kandiano, E.: Atlantic Water advection versus sea-ice advances in the eastern Fram Strait during the last 9 ka: Multiproxy evidence for a two-phase Holocene, *Paleoceanography*, 28(2), 283–680 295, doi:10.1002/palo.20028, 2013.
- Werner, K., Müller, J., Husum, K., Spielhagen, R. F., Kandiano, E. S. and Polyak, L.: Holocene sea subsurface and surface water masses in the Fram Strait – Comparisons of temperature and sea-ice reconstructions, *Quaternary Science Reviews*, 147, 194–209, doi:10.1016/j.quascirev.2015.09.007, 2016.
- 685 Zamelczyk, K., Rasmussen, T. L., Husum, K., Godtlieb, F. and Hald, M.: Surface water conditions and calcium carbonate preservation in the Fram Strait during marine isotope stage 2, 28.8–15.4 kyr, *Paleoceanography*, 29, 1–12, doi:10.1002/2012PA002448, 2014.
- Zeileis, A. and Grothendieck, G.: zoo: S3 Infrastructure for Regular and Irregular Time Series, *Journal of Statistical Software*, 14(1), 1–27, doi:10.18637/jss.v014.i06, 2005.
- 690 Zimmermann, H. H., Raschke, E., Epp, L. S., Stoof-Leichsenring, K. R., Schwamborn, G., Schirrmeister, L., Overduin, P. P. and Herzschuh, U.: Sedimentary ancient DNA and pollen reveal the composition of plant organic matter in Late Quaternary permafrost sediments of the Buor Khaya Peninsula (north-eastern Siberia), *Biogeosciences*, 14(3), 575–596, doi:https://doi.org/10.5194/bg-14-575-2017, 2017a.
- Zimmermann, H. H., Raschke, E., Epp, L. S., Stoof-Leichsenring, K. R., Schirrmeister, L., Schwamborn, G. and Herzschuh, U.: The history of tree and shrub taxa on Bol'shoi Lyakhovsky Island (New Siberian Archipelago) since the last interglacial 695 uncovered by sedimentary ancient DNA and pollen data, *Genes*, 8(10), 273, doi:10.3390/genes8100273, 2017b.
- Zimmermann, H. H., Stoof-Leichsenring, K. R., Kruse, S., Müller, J., Stein, R., Tiedemann, R. and Herzschuh, U.: Diatom sedimentary ancient DNA metabarcoding from western Fram Strait and Kronotsky Peninsula, *Dryad*, Dryad dataset, doi:https://doi.org/10.5061/dryad.bnzs7h481, 2020.

700

Figure 1: Map showing the coring site of MSM05/5-712-2 with bathymetric data derived from Ocean Data View (Schlitzer, 2002) and median March sea-ice extent from 1981-2010 (white, dashed line; [https://nsidc.org/data/seaice\\_index/archives](https://nsidc.org/data/seaice_index/archives) [accessed 06.08.2019] (Fetterer et al., 2017)).

705

Figure 2: Number of diatom amplicon sequence variants (ASVs) assigned (a) to different taxonomic levels and (b) to different families.

710

Figure 3: Non-metric multidimensional scaling plot based on the filtered and resampled diatom dataset with the PCR replicates (indicated by an underscore with a number) of a sample (depth in m) linked in a polygon of sample-specific colour: light green = 0.12m, green = 0.3m, light violet = 0.9 m, violet = 3.1 m, light blue = 3.71 m, dark blue = 5.8 m, turquoise = 6.68 m, grey = 7.27 m, red = 7.85 m, brown = 8.56 m, orange = 8.85 m (2 sample replicates with 3 PCR-replicates each).

715

Figure 4: Barplots showing the rarefied (a) number of amplicon sequence variants (ASVs) per sample and (b) grouped ASVs assigned to the same taxonomic name for each sample with depth (m) of the sediment core MSM05/5-712-2.

720

Figure 5: Stratigraphic diagrams with sequences assigned to (a) pennate diatoms (blue) and (b) centric diatoms (green) and higher level sequences assigned between family and phylum level (gray) The taxonomic composition with relative proportions (%) of the 360 detected sequence variants is grouped into 75 unique taxonomic names based on identically assigned taxonomic names of sediment core MSM05/5-712-2.

725

Figure 6: Proportions of sequences assigned to diatoms grouped on family level and down-core proportions of centric (blue) and pennate diatoms (green) as well as concentrations of the sea-ice biomarker IP<sub>25</sub> (Müller et al., 2012; Müller and Stein, 2014) and the phytoplankton biomarker brassicasterol (Müller et al., 2012; Müller and Stein, 2014). Higher level (grey) contains sequences assigned between family and phylum level. TOC = total organic carbon.