Improving the paleoceanographic proxy tool kit – On the biogeography and ecology of the sea ice-associated species *Fragilariopsis oceanica*, *Fragilariopsis reginae-jahniae* and *Fossula arctica* in the northern North Atlantic

Kaarina Weckström, Benjamin Redmond Roche, Arto Miettinen, Diana Krawczyk, Audrey Limoges, Steve Juggins, Sofia Ribeiro, Maija Heikkilä

A long-term perspective is essential for understanding environmental change. To be able to access the past, environmental archives such as marine and lake sediments that store information in the form of diverse proxy records are used. Whilst many analytical techniques exist to extract the information stored in these proxy records, the critical assessment and refinement of current methods in addition to developing new methods is crucial to improving our understanding. This study aims to improve our knowledge on diatom species used for reconstructing ocean surface conditions, especially temperature and sea ice variability over time. We define the distribution and the relationship to sea surface temperature (SST) and sea ice concentrations (SIC) of the species *Fragilariopsis oceanica*, *Fragilariopsis reginae-jahniae* and *Fossula arctica* using diatom training sets from the northern North Atlantic. We further assess the effect of separating these species compared to grouping them under *F. oceanica*, as has been done in the past. Our results suggest that while these three species share similarities such as the preference for stratified waters induced by sea ice or glacier meltwater, they also exhibit heterogeneous distributions across the northern North Atlantic, with individual optima for SST and SIC. This also affects quantitative reconstructions based on our data, resulting in lower SST and higher SIC estimates when the species are separated in the surface sediment and down-core diatom assemblages.

1. Introduction

To better understand our present environment, and to be able to project future changes, firm knowledge is essential on how the present ecosystems have evolved and what their long-term natural variability is. To achieve this, we need to use environmental archives, such as marine sediments, which preserve information of the past in the form of physical, chemical and biological proxy records. Over the years, a large number of analytical techniques have been developed, allowing us to access these records. While many proxies and analytical techniques to extract them are part of the routine tool kit of paleo-scientists, proxy development and refinement are still required. This is especially true for relatively new biogeochemical proxies and sedaDNA (Belt, 2018; Armbrrecht et al., 2019; Park et al., 2019), but also for microfossils, where individual species’ ecologies are still often poorly known (Van der Zwaan et al., 1999; Eynaud, 2011; Heikkilä et al., 2014; Pieńkowski et al., 2017; Oksman et al., 2019). It is also frequently forgotten, that species’ responses to their environment are complex, and can usually not be reduced to a single variable.

In the North Atlantic region — our study area — a prime focus is the reconstruction of past climate and ocean variability (e.g., Andersen et al., 2004; Hald et al., 2004; Berner et al., 2011; Spielhagen et al., 2011; Müller et al., 2012; Miettinen et al., 2012; Pearce et al., 2013; Krawczyk et al., 2017; Falardeau et al., 2018). At higher latitudes, due to the rapid decline in the extent and thickness of Arctic sea ice and the resulting transition of the Arctic into a seasonally sea-ice free region...
(Stroeve and Notz, 2018), much effort has recently gone into reconstructing sea-ice changes over longer time scales to better constrain its past natural variability (Massé et al., 2008; Müller et al., 2009; Navarro-Rodriguez et al., 2013; Sha et al., 2014; Cabedo-Sanz et al., 2016). As sea ice is an important climate modulator, such information is crucial for building robust climate models (Notz et al., 2016). For this purpose, a range of physical (e.g., IRD), biomarker (e.g., IP25) and microfossil (e.g., dinoflagellate cysts, foraminifers, diatoms) proxies are available. While each of these proxies has its merits, it also has shortcomings (de Vernal et al. 2013). Diatoms are widely used both for qualitative and quantitative reconstructions of past sea ice conditions. Despite their wide use, the ecologies of most common marine diatom species are still insufficiently known. This is also true for sea ice-associated species, which this study focuses on.

*Fragilariopsis oceanica* and *Fragilariopsis cylindrus* were described already in the 19th century (Cleve, 1873, 1883), and are often cited as sea-ice indicator species in the Northern Hemisphere (e.g., Jiang et al., 2001; Justwan and Koc, 2008; Sha et al., 2014; Miettinen et al., 2015; Krawczyk et al., 2017). However, the latter has been shown to have a far more complex response to sea ice than previously assumed (von Quillfeldt, 2004; Oksman et al., 2019). These two species have recently also been used to reconstruct past behavior of tidewater glaciers via their freshwater inputs (Oksman et al., 2017). *Fragilariopsis reginae-jahniae* and *Fossula arctica* are relatively recently described species (Witkowski et al., 2000; Hasle et al., 1996), which may be the reason for the limited information available on their ecological preferences, although they generally are considered to be associated with sea ice (von Quillfeldt, 2000, 2001; Witkowski et al., 2000). Both *Fragilariopsis reginae-jahniae* and *Fossula arctica* have likely been included in the species count data of *Fragilariopsis oceanica*, particularly in older studies, owing to the similar morphologies of the species (von Quillfeldt, 2000). Given the growing interest to reconstruct longer-term sea ice variability (e.g., Kinnard et al., 2011), we need to both develop new and refine our currently existing proxies, which motivates our present work. By studying the biogeography and ecology of these potential sea-ice indicator species, we aim to refine our understanding of their responses to different ice and cold-water environments, to ultimately make better inferences as to what their presence and relative abundances are suggesting in down-core records.

For this purpose, modern marine diatom training sets are highly useful. These data sets consist of both modern species assemblages analyzed from surface sediments and environmental data collected at the sea surface from the same locations, and they usually cover relatively large geographic areas (e.g., Caisse 2012; Sha et al., 2014; Ren et al. 2014; Miettinen et al., 2015; Krawczyk et al., 2017). Our study focuses on the North Atlantic, which hosts to date the largest and geographically most extensive training set in the Northern Hemisphere (the North Atlantic training set; Koc et al. 1993; Andersen et al., 2004; Miettinen et al., 2015), covering the main areas of the region between 42°N and 79°N. This training set was largely collected before the description of *Fragilariopsis reginae-jahniae* and *Fossula arctica*, hence these morphologically similar species are likely included in the count data of

Fragilariopsis oceanica. In addition to this training set, we have included available sites from the West Greenland (Baffin Bay) training set described in Krawczyk et al. (2017).

Our specific aims are to 1) improve the existing ecological knowledge on this group of species, 2) assess their value as sea-ice indicators, and 3) determine the importance of including these species separately (rather than grouping them under Fragilariopsis oceanica) for both qualitative and quantitative down-core reconstructions.

2. Materials and methods

2.1. Training sets & study area

The large North Atlantic training set (most recent version presented in Miettinen et al., 2015) encompasses altogether 183 sites. Of these, 46 sites — where material/microscopy slides were available — were reanalyzed for this study. Half of the re-analyzed sites are located to the west of Greenland in Baffin Bay and the Labrador Sea, and the other half to the east of Greenland in the Nordic Seas (Fig. 1). The West Greenland training set (Krawczyk et al., 2017) consists of 35 $^{210}$Pb- and $^{137}$Cs-dated surface sediment samples that were available for this study.

Within the area encompassed by these two training sets, sea ice extent varies markedly: To the west of Greenland, maximum sea-ice extent reaches far southwards into the Labrador Sea, with Baffin and Hudson Bays generally freezing over every year (Tang et al., 2004; Fetterer et al., 2017). An exception is the North Water Polynya (NOW, in northern Baffin Bay), where even during maximum sea-ice extent ~10% is kept open by local winds, upwelling of the West Greenland Current (WGC) waters and a well delineated ice bridge at the southern margin of the Kane Basin (Tang et al., 2004). To the east of Greenland, maximum sea-ice extent is limited to the waters north and east of Svalbard and along the east Greenland margin (Fetterer et al., 2017).

During the summer sea-ice minimum in the eastern Fram Strait, north of Svalbard and the Barents Sea, the median ice edge (1981–2010) is located at 81°N. Sea ice extends northwards from 73°N on the Greenland coastline, progressively reaching further seawards along the Greenland shelf (Fetterer et al., 2017). Baffin Bay, however, is ice free during the summer sea-ice minimum, apart from thin strips of sea ice along the coasts, particularly along Baffin Island (Tang et al., 2004). Sea ice is limited to the west of Devon Island and 78°N in the Kane Basin (Fetterer et al., 2017). This dichotomous sea-ice extent at parallel latitudes is fundamentally attributed to the transfer of heat by the North Atlantic Current and the strength of the North Atlantic Oscillation (NAO, Taylor and Stephens, 1998).

2.2. Re-analysis of the 46-site subset diatom samples

The counting was undertaken with a ZEISS Axio Imager A2 upright light microscope with immersion oil at 1000× magnification, and photomicrographs were taken using the ZEN Blue software (Fig. 2). A total of 100 diatom valves of the species targeted in our study were counted per sample, with a maximum scope of 10 transects for slides with low abundances. The count sums of the targeted species largely encompass the 100 diatom valves of the species targeted in our study were with low abundances. The count sums of the targeted species largely (compared to the lanceolate valve shape and striae thickness of the otherwise very similar F. oceanica); and Fossula arctica from the characteristic capitale ends (although this is not well developed in smaller valves) and thin discontinuous striae (Table 1, see detailed descriptions in Hasle et al., 1996; Lundholm and Hasle, 2010). The counts of the separated third diatom species Frugiariopsis oceanica, Fragilariopsis reginae-jahniae and Fossula arctica were converted into percentages relative to what the grouped F. oceanica originally constituted in the assemblage of each sample (Miettinen et al., 2015).

2.3. Statistical analysis

To define the relationship of the three species with their environment, we used August sea-surface temperature (aSST) and April sea-ice concentration (aSIC) for the numerical analyses. These have been found to be statistically significant variables (and months, i.e., August for SST and April for sea ice) in explaining diatom assemblage distribution in the North Atlantic (Berner et al., 2008; Miettinen et al., 2015). In addition, unlike in the case of SST and SIC, other environmental data such as nutrient concentrations are not available for many marine regions. Thus aSST and aSIC are the most commonly used parameters in quantitative diatom-based paleoceanographic reconstructions (Berner et al., 2008, 2011; Sha et al., 2014, 2015; Miettinen et al., 2015; Krawczyk et al., 2017).

The type and significance of the response of each taxon to both aSST and aSIC was assessed by fitting a series of Huisman-Offf-Fresco (HOF) hierarchical response models (Huisman et al., 1993) following Jansen and Oksanen (2013). This was done using both training sets. The procedure fits seven models of increasing complexity, from a null model or flat response (i.e., no relationship, model I), through monotone sigmoid (II), monotone sigmoid with plateau (III), unimodal symmetric (IV), unimodal skewed (V), bimodal with equal peaks (VI) and bimodal with unequal peaks (VII). It then selects the most parsimonious model using Akaake information criterion corrected for small sample size (AICc) and a bootstrap approach (500 permutations) to ensure model stability. Taxa are deemed to have a statistically significant relationship to either aSST or aSIC if the selected response yields statistically significant improvement in fit over a null or flat model.

For multivariate analyses and the creation of transfer functions, only the North Atlantic training set samples were used, due to potential differences in taxonomy between the two training sets regarding some other species than the Fragilariopsis/Fossula species group. Based on the gradient length of the species data in detrended correspondence analysis (DCA; gradient length 1.8 SD units of species turnover), redundancy analysis (RDA) with Monte Carlo permutation (999 randomizations) was run and presented as a biplot of species and environmental variables. The “grouped” Fragilariopsis oceanica (such as it is in the North Atlantic training set) was included as a supplementary variable (Fig. 3). RDA was undertaken with CANOCO version 5.01 (Ter Braak and Šmilauer, 2012).

aSST and aSIC were reconstructed using a weighted averaging (WA) transfer function (Birks, 1995) based on a modern calibration dataset consisting of the 46 re-analyzed surface sediment samples from the North Atlantic training set. The prediction error was estimated using h-block cross-validation (Burman et al., 1994), in which samples closer than a cut-off distance (h) from a target sample were excluded from contributing to the prediction of that sample. h-block cross validation was selected to allow for spatial dependency in the calibration data, which can lead to underestimation of the prediction error because of pseudoreplication. The cut-off distance (h) was estimated by assessing the spatial structure of the residuals of the surface sample predictions (Trachsel and Telford, 2016). All numerical analyses (with the exception of RDA) and Figs. 4–6 were produced using the R 3.4.1 software for statistical computing (R Core Team, 2017).

3. Results and discussion

3.1. Species distribution and environmental responses

3.1.1. Fragilariopsis reginae-jahniae Witkowski, Lange-Bertalot & Metzeltin 2000

Distribution in the northern North Atlantic...
Fig. 2. Photomicrographs of (A) *Fragilariopsis oceanica*; (B) *Fragilariopsis reginae-jahniae*; and (C) *Fossula arctica*. Images captured with a ZEISS Axio Imager A2 at 1000× magnification. Images from surface sediment sample HU2008–029–036, North Water Polynya.

Table 1
Summary of the morphology, ecology and distribution of the studied species. The current table caption should be the header of a separate section of the text, i.e. 3.2
The effect of the species separation on qualitative and quantitative paleoenvironmental interpretations

<table>
<thead>
<tr>
<th>Species</th>
<th>Morphology</th>
<th>Ecology &amp; distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Fragilariopsis oceanica</em></td>
<td>Ellipsoidal to lanceolate shape with rounded apices. Continuous striae and eccentric raphe with a central nodule. Apical axis ~25–45 μm, transapical axis ~4.5–6.5 μm.</td>
<td>Cold-water species abundant below 7 °C, found at a variety of sea-ice concentrations (highest abundances at both &lt; 25% and &gt; 75% aSIC), preference for stratified waters (cold, fresh surface layer). MIZ species, related to generally high aSIC (highest abundances &gt; 75%), and cold aSST (highest abundances around 3 °C), preference for stratified waters (cold, fresh surface layer). Sea-ice species, related to high aSIC (highest abundances &gt; 75%) and cold aSST (optimum 2.6 °C), preference for stratified waters (cold, fresh surface layer). Potential polynya species (affinity for higher nutrient availability).</td>
</tr>
<tr>
<td><em>Fragilariopsis reginae- jahniae</em></td>
<td>Elongate, linear valve shape with continuous striae (slightly thicker compared to <em>F. oceanica</em>). Apical axis ~40–60 μm, transapical axis ~5–6 μm with a distinctive nodule on the eccentric raphe.</td>
<td></td>
</tr>
<tr>
<td><em>Fossula arctica</em></td>
<td>Characteristic capitate ends and thin discontinuous striae. Larger valves are more elongate with clear capitate apices, whilst smaller valves are elliptical with near-rounded apices. Apical axis ~8–50 μm, transapical axis ~5.5–7.5 μm.</td>
<td></td>
</tr>
</tbody>
</table>
The diatom *Fragilariopsis reginae-jahniae* has previously been described as an Arctic sea ice-associated species (Witkowski et al., 2000; Jensen, 2003; Lundholm and Hasle, 2010; Krawczyk et al., 2014; Krawczyk et al., 2017). While it has low to moderate abundances across the northern North Atlantic data sets, ranging from 0 to 14.3% of the total assemblages, our results largely confirm this observation.

East of Greenland, the Greenland Sea is characterized by the di-chotomous temperature gradient of the warm Atlantic Water dominated Norwegian Atlantic Current and West Spitsbergen Current to the east with relatively high average aSST, and the cold Arctic Water-dominated and ice-loaded East Greenland Current (EGC) to the west with low aSST (Woodgate et al., 1999) (Fig. 4). In the Greenland Sea, highest abundances of the species (up to 7.2%) occur northwards of 78°N, in the central Fram Strait and on the West Spitsbergen Slope close to the winter sea-ice edge. These waters are cold (aSST < 4 °C) and the spring sea-ice concentrations vary between low to moderate (aSIC 0.5–34%). The North Atlantic Current-derived warmer waters in the eastern Greenland Sea (aSIC < 2.4%, aSST ca. 5–6 °C) generally have low abundances of *F. reginae-jahniae* (< ca. 3%).

South of 75°N in the Greenland Sea, relative abundances of *F. reginae-jahniae* are very low (< 1.2%). These sites are located within the winter sea-ice zone, and exhibit moderate to high aSIC (ca. 40–60%) and very low aSSTs (< 2 °C). The nearshore Scoresby Sound Shelf site (which is the southernmost site east of Greenland, where the species is present) is influenced both by the cold and fresher EGC and by the outflow of fresh glacier meltwater from the 38,000 km² system of fjords, and yields a higher abundance (5.5%) at very high levels of aSIC (80%). East of Greenland the species appears to be absent south of 69°N.

To the west of Greenland, the WGC transports both cold polar water from the EGC and temperate Atlantic Water from the Irminger Current (IC) northwards, whereas the Baffin-Labrador Current brings cold fresher Arctic water to the south along the west side of Baffin Bay (Ribergaard, 2014; Oksman et al., 2017). The general trend of increasing abundances northwards is repeated, apart from one clear exception: the highest abundance in the whole data set (14.3%) is reached at the site off Cumberland Peninsula, which is strongly influenced by the cold Baffin Current (BC) and glacial meltwater runoff from land (Cook et al., 2019) at very high aSIC (86.9%). This abundance peak is analogous to the peak at the Scoresby Sound Shelf east of Greenland.

In the northern Labrador Sea, the abundances of *F. reginae-jahniae* are generally very low (< 1% apart from two sites at ca. 2.5 and 4%, and with the exception of the Cumberland Peninsula site discussed above). The Labrador Sea sites have very variable aSIC consisting of drift ice (1–80%, mean = 38% ± 28%) and they are clearly influenced by warm and saline Atlantic Water (aSSTs 4–6.5 °C). In the southern and central Baffin Bay the abundances of *F. reginae-jahniae* are clearly higher than in the Labrador Sea (up to ca. 5%). Lowest abundances can be found in the middle, where the upwelling of the warmer, saline IC-derived waters and less meltwater-induced stratification may explain the lower values (Boertmann et al., 2013; Krawczyk et al., 2017; Oksman et al., 2017). The highest abundances occur along the West Greenland coast northwards of Disko Bay. Here the sea-ice edge prevails until June (Seidenkrantz, 2013), and its proximity may explain the elevated abundances. The two sites located nearshore Cape Adair and Home Bay in western Baffin Bay, which also show higher abundances compared to the middle parts of the bay, are influenced by meltwater from marine-terminating glaciers and by the Arctic water carried by the BC (Cook et al., 2019).

Northern Baffin Bay is influenced by cold Arctic Water from the Arctic Ocean transported via the Nares Strait and the Canadian Arctic Archipelago (Tang et al., 2004), with very high aSIC (78.7–97.8%) and low average aSST (2.8 °C). The combination of currents, winds and physical barrier of the ice bridge result in the NOW remaining partially ice-free throughout the year (Dunbar, 1969; Tang et al., 2004). In northern Baffin Bay, the abundances of *F. reginae-jahniae* are consistently relatively high (4.1–9.6%). The highest relative abundances occur in the NOW, likely due to the very high aSIC concentrations (> 80%) that lead to water column stratification during the early spring melt typical for NOW, possibly enhanced by the nutrient-rich waters of the polynyas (Barber and Massom, 2007).

**Species – environment relationship**

Based on the RDA analysis (Fig. 3), *F. reginae-jahniae* is related to relatively high aSIC and low aSSTs, similarly to other sea-ice associated species such as *Porosira glacialis*, *Actinocyclus curvatulus* and *Coscinodiscus occlus-iris* (Oksman et al., 2019). Moreover, the ecological response curves (Fig. 5) verify that *F. reginae-jahniae* has a statistically significant relationship to both aSST with highest abundances around 3 °C, and to aSIC with highest abundances at > 75%, although the relationship to sea ice shows some scatter, and similar abundances can be found at aSIC < 50%.

*Fragilariopsis reginae-jahniae* appears to be common but not very abundant at higher latitudes, in agreement with previous literature (Jensen, 2003; Krawczyk et al., 2014; Krawczyk et al., 2017). The highest average abundances in the NOW and near Lancaster and Jones Sounds, along the June sea-ice edge in southeastern Baffin Bay, as well as at the West Fram Strait winter/spring sea-ice edge are indicative of *F. reginae-jahniae* being primarily a marginal ice zone (MIZ) species (as also suggested by Limoges et al., 2018), with a preference for cold, fresher waters with moderate to high aSIC. In agreement with this interpretation, the highest abundances (14.2%) occurred in cold aSSTs (2.7 °C) with glacier meltwater influence and very high aSIC (86.9%) at the lower latitude site off Cumberland Peninsula. Thus, high abundances of *F. reginae-jahniae* could be interpreted as indicative of a MIZ...
environment in an Arctic water–dominated area that experiences generally high sea-ice concentrations. Additionally, relative abundances can also be comparatively high in areas where the meltwater input from tidewater glaciers results in a lid of colder and fresher water, similar to MIZ conditions (Table 1).

3.1.2. *Fragilariopsis oceanica* (Cleve) Hasle 1965
(Basionym. Fragilaria oceanica (Cleve), Synonym. Fragilaria arctica Grunow in Cleve & Grunow, Nitzschia grunowii Hasle)

**Distribution in the northern North Atlantic**

The separated *Fragilariopsis oceanica* is a common and relatively abundant (0.5–30%) species in the North Atlantic, averaging in excess of 10% (and up to 19%) of the total diatom assemblages above 75°N in the northern Greenland Sea and in excess of 14% (and up to 23%) north of 74°N in Baffin Bay (Fig. 4).

Around Svalbard the species shows relatively similar abundances — ca. 10–16% of the total assemblage — at a range of sea-ice conditions from the low aSIC in the path of the West Spitsbergen Current (0 – ca. 10%), to the moderate aSIC concentrations of up to 33% in the path of the cold and fresh East Spitsbergen Current closer to Svalbard.

Along the East Greenland Margin, the cold EGC accommodates sea-ice formation much further to the south than in the eastern Fram Strait. Slightly higher abundances of *F. oceanica* (up to 18.6%) were observed at the central Fram Strait winter sea-ice edge (aSIC: 15.2–33.8%). South of 75°N in the Greenland Sea, low abundances of *F. oceanica* (0.7–3.4%) were comparable to *F. reginae-jahniae* in the moderate to high aSIC and very low aSST conditions along the winter sea-ice zone of the East Greenland Margin. The highest abundance of *F. oceanica* (29.4%)
recorded for the 46 re-analyzed samples of the training set occurred nearshore located on the Scoresby Sound Shelf. The influence of fresh meltwater from the marine-terminating glaciers of the Scoresby Sound fjord system is likely playing a significant role for this anomalously high abundance, in addition to the cold and low-density EGC. A recurring polynya forms at the entrance of the Sound (Sorensen, 2012), which, as an open-water more productive system (Smith Jr. and Barber, 2007) may also play a role for the elevated abundances. East of Greenland, the species is largely absent south of 69°N.

In the northern Labrador Sea, the abundances of *F. oceanica* are low to moderate (ca. 3.5% on average, excluding one anomalous site off Cumberland Peninsula). In addition to this anomalous site, two sites (included in the average) show abundances of ca. 10%. The Labrador Sea sites have very variable aSIC (1–80%, mean = 38% σ = 28%) and are clearly influenced by warm Atlantic Water. Interestingly, these two sites are at opposite ends regarding their aSIC concentrations (1 and 63%). Comparable to *F. reginae-jahniae*, the highest abundance of *F. oceanica* (19.3%) occurred at the nearshore cold-water ice-laden site off Cumberland Peninsula. It is likely that glacier runoff, possibly coupled with iceberg-derived meltwater, were causal for this high abundance.

Unlike *F. reginae-jahniae*, the abundances of *F. oceanica* were relatively similar between the Labrador Sea and the southern and central Baffin Bay, where aSIC is generally high (at all sites > 89%) and the average aSST is relatively low (around 3 °C). Additionally, the abundances of *F. oceanica* along the West Greenland coast north of Disko Bay are lower than those of *F. reginae-jahniae*, otherwise the overall patterns for these two species are similar.

In the northern Baffin Bay, at the entrance of Lancaster Sound and

---

**Fig. 5.** Ecological response curves to aSST and aSIC for *Fragilariopsis reginae-jahniae*, *F. oceanica*, *Fossula arctica* and the grouped *F. oceanica*. 
in the NOW, abundances are moderate to high (11.6–22.5%). The stratified water column during the spring bloom (possibly coupled with the elevated nutrient concentrations of the polynya) is likely the reason for these elevated abundances. Only one site just south of the NOW had a low abundance of *F. oceanica* (5.5%) at very high aSIC (94.4%).

### Species-environment relationships

RDA analysis places *F. oceanica* close to the cold-water species *Rhizosolenia hebetata* f. *hebetata*, which can be found at a variety of sea-ice conditions, and clearly separated from a larger group of species (such as *Thalassiosira angulata*, *Shionodiscus oestrupii* and *Thalassionema nitzschioideae*) found higher up along the aSST gradient. *Fragilariopsis oceanica* appears less strongly related to higher aSIC compared to e.g., *oceanica*. *Coscinodiscus oculus-iridis* is a cold-water species related to cold aSST/high aSIC conditions revealed by *F. reginae-jahniae*, and it can be found at a variety of sea-ice concentrations with nearly equal abundances at < 25% and > 75% aSIC. Previous studies have suggested that *F. oceanica* is associated with pack ice and the water column spring bloom coeval to initial sea-ice melt (Von Quillfeldt, 2000; von Quillfeldt, 2001; Miettinen et al., 2015 (grouped *F. oceanica*; Oksman et al., 2019). This tendency for *F. oceanica* is a cold-water species with highest abundances at high aSIC (both relationships are statistically significant), it does not have a well-defined low temperature optimum like *F. reginae-jahniae*, but is overall abundant below ca. 7 °C, and it can be found at a variety of sea-ice concentrations with nearly equal abundances at < 25% and > 75% aSIC.

Based on the relatively few existing studies, *Fossula arctica* seems to be related to Arctic waters that are influenced by sea ice, appearing early in the spring bloom, before *F. oceanica* and other *Fragilariopsis* species (Hasle et al., 1996; von Quillfeldt, 2000, 2001; Onodera et al., 2015). Similarly to both *F. oceanica* and *F. reginae-jahniae*, it is considered to be associated to the cold low salinity conditions of the MIZ (Limoges et al., 2018).

The abundance of *F. arctica* in the northern North Atlantic is very distinctive, with low abundances in the Greenland Sea (0–4.3%) and the northern Labrador Sea (0–1.1%), a near-absence along the WGC – dominated West Coast of Greenland, slightly higher abundances in the southern and central Baffin Bay (up to 8.9%), and a significant increase to maximum abundances in the NOW (13.7–26.4%) (Fig. 4). Exceptions to this clear general pattern were the few coastal sites – Scoresby Sound Shelf, off Cumberland Peninsula and Home Bay, where elevated abundances of also the other two species were observed.

### Species-environment relationships

Based on the RDA (Fig. 3), *Fossula arctica* is a sea-ice species related to high aSIC and cold aSSTs, similarly to *Coscinodiscus oculus-iridis*, *Porosira glacialis* and *Actinocyclus curvatus*. The affinity for high aSIC appears to be stronger than that of *F. reginae-jahniae* and *F. oceanica*. Moreover, the environmental response curves (Fig. 5) verify that *F. arctica* has a statistically significant relationship to aSST with an optimum at 2.6 °C and highest abundances at > 75% aSIC. The exceptionally high abundances of this species in the NOW significantly contribute to its affinity for cold aSST/high aSIC conditions revealed by the ecological response curves, however, there may be other environmental factors that partially drive these high abundances in the NOW, compared to other sites with similar temperature, salinity (fresher surface layer) and sea-ice conditions. Indeed, von Quillfeldt (1997) found that *F. arctica* constituted up to 77.7% of total assemblages from...
another large Arctic polynya, the Northeast Water Polynya (described as ‘Fragilariopsis sp. 1’). The high abundances of *F. arctica* in both the NOW and the Northeast Water Polynya (von Quillfeldt, 1997) suggest that *F. arctica* could be considered as a characteristic polynya species. Analysis of diatom assemblages from other polynya environments could further verify this finding. In addition to high aSIC and low aSST conditions, the upwelling of the nutrient-rich WGC in the NOW (Melling et al., 2001), coupled with the early season productivity that the ice-free area affords (Christie and Sommerkorn, 2012; Meltote, 2013), may be conducive to the enhanced NOW abundances of *F. arctica* in our data set.

Warm and saline Atlantic Water appears to be unaccommodating for *F. arctica* as it is near absent along the pathway of the West Greenland Current. Comparable to the other two species, *F. arctica* seems to thrive in stratified waters, where a colder, fresher surface layer is formed above ambient sea-water temperatures and salinities, such as in the MIZ or close to glacier meltwater influence, agreeing with previous studies (e.g. von Quillfeldt, 2000; Limoges et al., 2018). A noteworthy observation regarding the species’ distribution is the significantly higher relative abundance in the NOW compared to other areas with low aSST/high aSIC and stratified waters, which suggests that additional environmental factors typical of a polynya setting (such as higher nutrient concentrations) play a role (Table 1). *Fragilariopsis reginae-jahniae* and particularly *F. oceanica* displayed higher abundances at the Scoresby Sound shelf, where a recurring polynya, the “Scoresby Sound Water” form. *Fossula arctica*, however, is not showing elevated abundances at this site, which may, based on our hypothesis that this is a “polynya species”, indicate that the site on the Scoresby Sound shelf is not affected by the polynya, but rather by meltwater from the large fjord system.

3.2. The effect of the species separation on qualitative and quantitative paleoenvironmental interpretations

Based on our study on the 46-site subset of the North Atlantic data set, the grouped *F. oceanica*, presented lately in Oksman et al. (2019), included significant abundances of both *F. arctica* and *F. reginae-jahniae*. The separated diatom species analyzed here exhibit heterogeneous distributions across the northern North Atlantic, with individual maximum abundances occurring in different regions with different aSST and aSIC (Figs. 4 and 5). This indicates that separating the species will increase the accuracy of future ecological interpretations. The grouped *F. oceanica* was present at a temperature gradient ranging between 1.2 and 6.6 °C, and had a statistically significant relationship to both aSST with highest abundances at 2.5 °C, and to aSIC with highest abundances at > 75% (Fig. 5). The separated *F. oceanica*, on the other hand, had clearly less well-defined relationships with the two environmental gradients, showing a wide scatter of high abundances at aSSTs between 1.4 and 7 °C and nearly equally high abundances at aSIC < 25% and > 75%. It appears that the grouped *F. oceanica* showed a stronger affinity to low aSST and high aSIC due to the inclusion of *F. arctica* and *F. reginae-jahniae*.

To explore if the separation of the three species (compared to lumping them together under *Fragilariopsis oceanica*) affects quantitative SST and sea ice reconstructions, we created transfer functions for aSST and aSIC (under “grouped” and “separated”) using the 46-site subset of the large North Atlantic training set. We underline that these two transfer functions were created only for the purpose of testing the differences, and are not meant to be used for future paleoenographic reconstructions. We found that the WA-model performance differed only slightly between the “grouped” vs. “separated” aSST and aSIC transfer functions: $r^2 = 0.9515$ and RMSEP = 1.0226 for the “grouped” and $r^2 = 0.9513$ and RMSEP = 1.0227 for the “separated” aSST transfer functions, and again $r^2 = 0.5583$ and RMSEP = 22.8286 for the “grouped” and $r^2 = 0.5724$ and RMSEP = 22.4532 for the “separated” aSIC transfer functions. The aSST transfer functions are clearly more robust compared to the aSIC models, likely due to a smaller number of species having a well-defined relationship with sea-ice concentrations in this training set (e.g., Oksman et al., 2019). However, the separation of the species had an effect on the down-core reconstruction of both environmental variables in core AMD15-CASQ1-BC (77°16.746′ N, 74°21.428′ W, 702 m water depth), especially in the case of aSIC (Fig. 6). While the difference between the aSST reconstructions was only marginal using this training set (the “separated” transfer function & core assemblages resulting in lower aSST estimates), the aSIC reconstruction showed 5–10% higher sea-ice concentrations when the species were separated. It should be noted that some previous aSST reconstructions using the full North Atlantic training set where the species are currently not separated have resulted in absolute aSST values higher than what is to be expected (e.g., Oksman et al., 2019). Core AMD15-CASQ1-BC was chosen as it had relatively high percentage abundances of *Fragilariopsis reginae-jahniae* and *Fossula arctica*, at average values of 14 and 22%, respectively. Based on this exercise it appears that in areas where all three species are abundant, quantitative reconstructions of especially past sea-ice concentrations should be carried out using training sets where these species have been separated. Further, given the distinct differences in their biogeography and relationships with aSST and aSIC (Figs. 5 and 6), it is advised to separate these species also in qualitative down-core analyses, as this will allow a much more nuanced interpretation of past oceanographic conditions in the northern North Atlantic.

4. Conclusions

- The three diatom species analyzed here — *Fragilariopsis reginae-jahniae*, *Fragilariopsis oceanica* and *Fossula arctica* — exhibit heterogeneous distributions across the northern North Atlantic, with individual maximum abundances occurring in different regions with different aSST and aSIC.
- All three species appear to thrive in stratified waters caused by sea ice and/or glacier melt (cold, fresher surface water resembling MIZ conditions) and are overall most abundant north of ca. 70°N.
- Abundances of *F. arctica* are suggested to be enhanced by polynya conditions (availability of upwelling nutrients) and this species could be a potential polynya indicator in paleoceanographic studies.
- Quantitative aSST and aSIC reconstructions result in lower temperatures and higher sea ice concentrations when the species are identified separately rather than combined under *F. oceanica* (which has often been done in the past).
- We advise to separate these species in future studies, as this will allow a much more nuanced interpretation of the past oceanography in the northern North Atlantic.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to acknowledge the following funders: Academy of Finland (Project 307282); Grenlands Forskningsråd and Greenland Institute of Natural Resources for funding the project ‘Sea ice during Holocene’, which supported the collection of the West Greenland surface sediment dataset during the annual survey with M/T Paamiut.

The data will be available after publication at PANGAEA, or via request from author (West Greenland training set).