The 16S rDNA microbiome of the Arctic foraminifera *Neogloboquadrina pachyderma* is comprised of hydrocarbon-degrading bacteria and a diatom chloroplast store.

Clare Bird\(^1\), Kate Darling\(^{1,2}\), Rabecca Thiessen\(^3\) and Anna J. Pieńkowski\(^{4,5}\)

\(^1\)Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK.
\(^2\)School of Geosciences, Grant Institute, King’s Buildings, University of Edinburgh, Edinburgh, UK.
\(^3\)Department of Physical Sciences, MacEwan University, Edmonton, T5J 4S2, AB, Canada.
\(^4\)Geohazards Research Unit, Institute of Geology, Adam Mickiewicz University, Poznań, Poland
\(^5\)Department of Arctic Geology, UNIS (University Centre in Svalbard), Longyearbyen, Svalbard, Norway

Correspondence to: Clare Bird (clare.bird2@stir.ac.uk)

**Abstract.** *Neogloboquadrina pachyderma* is the only true polar species of planktonic foraminifera. It therefore plays a crucial role in the calcite flux, and in reconstructions and modelling of seasonality and environmental change within the high latitudes. The rapidly changing environment of the polar regions of the North Atlantic and Arctic Oceans poses challenging conditions for this (sub)polar species in terms of temperature, sea-ice melt, calcite saturation, ocean pH and contraction of the polar ecosystem. To model the potential future for this important high latitude species, it is vital to investigate the modern ocean community structure throughout the annual cycle of the Arctic to understand the inter-dependencies of *N. pachyderma*. We use 16S rDNA metabarcoding and TEM to identify the microbial interactions of *N. pachyderma* during the summer ice-free conditions in Baffin Bay. We demonstrate that the *N. pachyderma* diet consists of diatoms and bacteria. The core microbiome is defined as the 16S rDNA amplicon sequencing variants (ASVs) found in 80 % of individuals investigated. This core microbiome consists of two diatom chloroplast ASVs and seven bacterial ASVs and accounts for, on average, 50 % of the total ASVs in any individual. The bacterial ASVs represent hydrocarbon-degrading bacteria, including those found routinely in the diatom phycosphere. On average the two chloroplast ASVs compose 40 % of the core microbiome. Significantly, an average of 55.7 % of all ASVs in any individual are of chloroplast origin. TEM highlights the importance of diatoms to this species, conclusively revealing that chloroplasts remain undigested in the foraminiferal cytoplasm in very high numbers, comparable to those observed in kleptoplastic benthic foraminifera. Diatoms are the major source of kleptoplasts in benthic foraminifera and other kleptoplastic groups, but this adaptation has never been observed in a planktonic foraminifer. Further work is required to understand the association between *N. pachyderma*, diatoms and their chloroplasts in the pelagic Arctic realm. It may confer an advantage to this species for survival in this extreme habitat, but it could also become compromised by the rapidly changing climate.
Introduction

*Neogloboquadrina pachyderma* is the predominant planktonic foraminiferal morphospecies of the polar oceans (e.g. Bé and Tolderlund, 1971; Bé, 1977) and the major marine calcifier in the (sub)polar province of the northern North Atlantic (Kohfeld et al., 1996), the largest ocean carbon sink in the Northern Hemisphere (Gruber et al., 2002). Present-day temperature conditions confine *N. pachyderma* to the (sub)polar water masses in the high-latitude North Atlantic/Arctic Ocean, where summer sea surface temperature (SST) remains below 10°C (Tolderlund and Bé, 1971; Duplessy et al., 1991). Here, *N. pachyderma* exhibits strong seasonal productivity in a highly predictable pattern, as winter mixing re-supplies nutrients to surface waters, triggering the seasonal succession of maximal phytoplankton blooms and zooplankton abundance followed by more nutrient depleted summer conditions (Jonkers and Kucera, 2015).

As the major component of both the modern and Quaternary fossil (sub)polar assemblage, the calcite shells of *N. pachyderma* continue to make the major contribution to the reconstructions and modelling of seasonality and environmental change within the North Atlantic and Arctic Ocean (e.g. Simstich et al., 2003; Kretschmer et al., 2016; Altuna et al., 2018; Brummer et al., 2020; Livsey et al., 2020). However, the Arctic is now an unrelentingly warming ecosystem, with seasonal sea ice cover constantly reducing and likely to disappear within a short time frame (Serreze et al., 2009; Meier et al., 2021). The North Atlantic/Arctic *N. pachyderma* is already predicted to be particularly sensitive to the forecasted changes in seawater carbonate chemistry (Manno et al., 2012), with consequent implications for the calcite flux and the biological pump. Under ocean acidification conditions, Arctic *N. pachyderma* show reduced carbonate production moderated by ocean warming, making it difficult to predict future climate change impacts as the polar habitat of *N. pachyderma* decays. Although the North Atlantic/Arctic *N. pachyderma* population has clearly survived the extremes of Quaternary climate cyclicity in the past (Brummer et al., 2020), it is unknown whether *N. pachyderma* will find itself spatially displaced from its adaptive ecological range in the Arctic ecosystem (Jonkers et al., 2019; Greco et al., 2022), as we transition into the unknown territory of anthropogenically driven global warming.

The warming ocean is affecting all marine organisms at multiple trophic levels (Poloczanska et al., 2016; Meredith et al., 2019; Deutsch et al., 2015). It has already been demonstrated that some planktonic protist species cannot habitat track their optimal temperatures as their environment changes and may undergo extinction once local thresholds are exceeded (Trubovitz et al., 2020). Such thresholds are unknown for *N. pachyderma*, and there may soon be no true polar refugia into which to retreat. At the beginning of the 21st century within our Baffin Bay study area, the Pikialasorsuaq (the former “North Water Polynya”) was considered a region of high biological productivity (Tremblay et al., 2002; 2006). However, increasing oligotrophic conditions have been reported in the last decade, driven by meltwater from the Greenland Ice Sheet and nearby glaciers increasing stratification, and reduced mixing/upwelling causing a reduction specifically in diatom mediated net community production (Bergeron et al., 2014). Since diatoms are considered a major food source for *N. pachyderma* (Schibel and Hemleben 2017; Greco et al., 2021) this reduction in diatom primary productivity could pose an added challenge. To model the potential future
for this important high latitude species, it is vital to investigate the modern ocean community structure throughout the annual cycle of the Arctic to understand the inter-dependencies of *N. pachyderma*.

Although our understanding of Arctic (sub)polar *N. pachyderma* annual/seasonal population structure and ecological behaviour is increasing (e.g. Carstens and Wefer, 1992; Kohfeld et al, 1996; Jonkers et al., 2010; Jonkers et al., 2013; Greco et al., 2019; Meilland et al., 2022), it is far from complete. Small-subunit ribosomal RNA (SSU rRNA) genotyping indicates that only a single *N. pachyderma* genotype (Type I) occupies the whole of the North Atlantic and Arctic Ocean water mass (Darling et al., 2004, 2007). The sub-polar North Pacific *N. pachyderma* (Type VII) is not adapted to live in Arctic polar waters (Darling et al., 2007). This is good news for all the Arctic ecological investigations based on this taxon (e.g. Altuna et al., 2018; Greco et al., 2019; Meilland et al., 2022), since different genotypes of *N. pachyderma* fill distinctly different niches within the divergent ecosystems which they inhabit (Darling et al., 2017). Metabarcoding investigations of other taxa within the *Neogloboquadrina* genus already highlight the diversity and complexity of the ecological community networks and symbiont/predator/prey interactions which exist between prokaryotes and protists within the water column (Bird et al., 2018). Using the 16S rDNA metabarcoding approach together with fluorescence microscopy, or transmission electron microscopy (TEM), Bird et al. (2018) determined the taxonomic character, trophic interactions, food source and putative symbiotic associations of *N. incompta* and *N. dutertrei* in the California Current system. Results highlight their similar feeding strategy of forming feeding cysts of particulate organic matter (POM) in the water column, but that such behaviour provides no clues to their choice of prey or potential symbiotic associations. Evidently, ecological concepts of individual planktonic foraminifera must be systematically revised, as each morphospecies and potentially each genotype has most likely evolved individually distinct interactions with the marine microbial assemblage.

A recent study used single-cell metabarcoding targeting 18S rDNA to characterise the interactions of *N. pachyderma* with the local eukaryote community (Greco et al., 2021), since the majority of data on feeding behaviour in planktonic foraminifera suggest that biotic interactions are likely to be mainly with herbivorous eukaryotes (Kohfeld, Fairbanks et al., 1996; Manno and Pavlov, 2014; Pados and Spielhagen, 2014; Schiebel and Hemleben, 2017). However, since no direct investigations have been carried out on the feeding behaviour or diet of Arctic *N. pachyderma*, this remains in question. The data shown in Greco et al. (2021), indicate that the *N. pachyderma* interactome is dominated by diatoms, with Crustacea and Syndiniales (a potential parasite) also present. Here we complement this study by examining the single cell 16S rDNA metabarcodes of the *N. pachyderma* microbiome to investigate prokaryote biotic and trophic interactions and their potential symbiotic associations.

In addition, we further investigate the cellular structures within *N. pachyderma* individual specimens using TEM, to examine the cellular position of the bacterial/chloroplast sources of DNA within the *N. pachyderma* cell. Our results have direct implications for understanding trophic interactions within this at-risk habitat; interpretation of the seasonal shell geochemical record; for modelling *N. pachyderma* population dynamics and the carbonate flux under climate change; and understanding the evolutionary pressures experienced by this morphospecies.
2 Materials and methods

2.1 Sampling locality and collection methods

Sampling was undertaken in Baffin Bay in July/August 2017 aboard CCGS Amundsen as part of ArcticNet Expedition 2017 (Leg 2b; https://arcticnet.ula-val.ca/expeditions-2017) and August/Sept. 2018 aboard the CCGS Hudson 2018042 expedition. In both cases the samples were taken in open water with no sea ice cover. Details of sampling stations and collections are listed in Table 1. Sample provenance was either individual foraminifera or the water column. Samples were analysed by 16S metabarcoding of the foraminiferal microbiome or the water column bacterial assemblages from stations along the Amundsen cruise track.

Foraminifera were collected at seven stations by vertical net tow from 200 m depth to the surface. Foraminifera for genotyping and microbiome analysis were wet picked on board, rinsed in 0.2 μm filtered surface seawater and preserved in 100 μl RNALater® (Ambion™). This reagent conserves cell integrity, inhibits nucleases at ambient temperatures, and dissolves the calcite shell. Samples were stored at 4°C for 4 hours then transferred to -20°C until processing.

CTD data and water samples were collected from three stations (101, 115 and 323) in northern Baffin Bay (Fig. 1; Table 1) in 2017. Stations 101 and 115 are both situated within the biologically important Pikialasorsuaq between Greenland and Canada, which remains sea-ice-free in winter (Eegeesiak et al., 2017). Station 101 is located close to southeast Ellesmere Island in relatively shallow water (350 m water depth). The deeper Station 115 (653 m) is at a similar latitude to Station 101, but closer to Greenland. Station 323, at 789 m the deepest of the three stations, is situated farther south and outside the Pikialasorsuaq; it is in Lancaster Sound at the entrance to the Northwest Passage.

Water samples were collected from five depths; surface, 50 m, 100 m, 150 m, and 200 m. 2L of seawater was filtered from each depth at each station on to 0.2 μm polycarbonate filters. Filters were then individually placed in a 1.5ml microfuge tube and covered with RNALater® (Ambion™). Tubes were stored at 4°C for 4 hours then transferred to -20°C until processing.

Foraminifera for TEM analysis were collected on the 2018 CCGS Hudson cruise (2018042) (Table 1). They were wet picked as described above and placed directly in TEM buffer (4% glutaraldehyde, 2% paraformaldehyde in salt adjusted phosphate buffered saline), stored at room temperature for 12 hours, then kept at 4°C until further processing.
Table 1. Stations and sample information including provenance, depth, and analysis type. Specimen IDs are either WC (water column) or Fm (foraminifera). This is followed by station identification (e.g. 101), and water depth (e.g. 050 = 50 metres), and replicate ID (e.g. a or b or c etc.)

<table>
<thead>
<tr>
<th>Cruise</th>
<th>Station</th>
<th>Sampling date</th>
<th>Specimen IDs</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
<th>Water depth (m)</th>
<th>Sample depth (m)</th>
<th>Provenance</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMD20</td>
<td>101</td>
<td>JUL-24-2017</td>
<td>WC101_000a</td>
<td>76.3844</td>
<td>-77.4033</td>
<td>350</td>
<td>surface</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC101_050a</td>
<td>76.3844</td>
<td>-77.4033</td>
<td>350</td>
<td>50</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC101_100a</td>
<td>76.3844</td>
<td>-77.4033</td>
<td>350</td>
<td>100</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC101_150a</td>
<td>76.3844</td>
<td>-77.4033</td>
<td>350</td>
<td>150</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC101_200a</td>
<td>76.3844</td>
<td>-77.4033</td>
<td>350</td>
<td>200</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td>101</td>
<td>JUL-24-2017</td>
<td>WC323_000a</td>
<td>76.3419</td>
<td>-71.2192</td>
<td>653</td>
<td>surface</td>
<td>Water column</td>
<td>microbiome</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC323_050a</td>
<td>76.3419</td>
<td>-71.2192</td>
<td>653</td>
<td>50</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC323_100a</td>
<td>76.3419</td>
<td>-71.2192</td>
<td>653</td>
<td>100</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC323_150a</td>
<td>76.3419</td>
<td>-71.2192</td>
<td>653</td>
<td>150</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC323_200a</td>
<td>76.3419</td>
<td>-71.2192</td>
<td>653</td>
<td>200</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fm115a</td>
<td>76.3419</td>
<td>-71.2192</td>
<td>653</td>
<td>surface-200m</td>
<td>Foraminifera</td>
<td>microbiome, genotyping</td>
</tr>
<tr>
<td>323</td>
<td>JUL-31-2017</td>
<td>WC323_000a</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>surface</td>
<td>Water column</td>
<td>microbiome</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC323_050a</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>50</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WC323_100a</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>100</td>
<td>Water column</td>
<td>microbiome</td>
</tr>
<tr>
<td>Sample Code</td>
<td>Date</td>
<td>Location</td>
<td>Depth</td>
<td>Count</td>
<td>Description</td>
<td>Annotation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td>----------------</td>
<td>--------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC323_150a</td>
<td>JUL-31-2017</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>150 Water column microbiome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC323_150b</td>
<td>JUL-31-2017</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>150 Water column microbiome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC323_200a</td>
<td>JUL-21-2017</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>200 Water column microbiome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WC323_200ab</td>
<td>JUL-21-2017</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789</td>
<td>200 Water column microbiome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>323</td>
<td>JUL-31-2017</td>
<td>Fm323a</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-31-2017</td>
<td>Fm323b</td>
<td>74.1593</td>
<td>-80.4753</td>
<td>789 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>176</td>
<td>JUL-21-2017</td>
<td>Fm175a</td>
<td>69.6032</td>
<td>-65.3938</td>
<td>281 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-21-2017</td>
<td>Fm175b</td>
<td>69.6032</td>
<td>-65.3938</td>
<td>281 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-21-2017</td>
<td>Fm175c</td>
<td>69.6032</td>
<td>-65.3938</td>
<td>281 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-21-2017</td>
<td>Fm175d</td>
<td>69.6032</td>
<td>-65.3938</td>
<td>281 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB2</td>
<td>JUL-22-2017</td>
<td>FmBB2a</td>
<td>72.7678</td>
<td>-66.0002</td>
<td>2372 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-22-2017</td>
<td>FmBB2b</td>
<td>72.7678</td>
<td>-66.0002</td>
<td>2372 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-22-2017</td>
<td>FmBB2c</td>
<td>72.7678</td>
<td>-66.0002</td>
<td>2372 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-22-2017</td>
<td>FmBB2d</td>
<td>72.7678</td>
<td>-66.0002</td>
<td>2372 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-22-2017</td>
<td>FmBB2e</td>
<td>72.7678</td>
<td>-66.0002</td>
<td>2372 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>JUL-29-2017</td>
<td>Fm129a</td>
<td>78.3254</td>
<td>-74.1124</td>
<td>514 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JUL-29-2017</td>
<td>Fm129b</td>
<td>78.3254</td>
<td>-74.1124</td>
<td>514 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>301</td>
<td>AUG-03-2017</td>
<td>Fm301a</td>
<td>74.2778</td>
<td>-83.3641</td>
<td>716 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-03-2017</td>
<td>Fm301b</td>
<td>74.2778</td>
<td>-83.3641</td>
<td>716 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-03-2017</td>
<td>Fm301c</td>
<td>74.2778</td>
<td>-83.3641</td>
<td>716 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-03-2017</td>
<td>Fm301d</td>
<td>74.2778</td>
<td>-83.3641</td>
<td>716 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-03-2017</td>
<td>Fm301e</td>
<td>74.2778</td>
<td>-83.3641</td>
<td>716 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-03-2017</td>
<td>Fm301f</td>
<td>74.2778</td>
<td>-83.3641</td>
<td>716 surface-200m Foraminifera microbiome, genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hudson 2018042</td>
<td>AUG-23-2018</td>
<td>BB1</td>
<td>66.8605</td>
<td>-61.0668</td>
<td>337 100 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-23-2018</td>
<td>BB2</td>
<td>66.8605</td>
<td>-61.0668</td>
<td>337 100 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUG-31-2018</td>
<td>BB8</td>
<td>68.6999</td>
<td>-63.7084</td>
<td>1560 50 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>SEP-02-2018</td>
<td>BB9B</td>
<td>68.543415</td>
<td>-63.461252</td>
<td>1543 100 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>SEP-02-2018</td>
<td>BB9C</td>
<td>68.543415</td>
<td>-63.461252</td>
<td>1543 100 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>SEP-04-2018</td>
<td>BB11</td>
<td>66.1371</td>
<td>-61.3659</td>
<td>160 100 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>SEP-04-2018</td>
<td>BB12</td>
<td>66.1371</td>
<td>-61.3659</td>
<td>160 100 Foraminifera TEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Map of sampling stations (numbered red spots) and the site of Pikialasorsuaq (the former ‘North Water Polyna’ (dashed yellow line)). Stations BB2, 101, 115, 129, 301 and 323 were part of the CCGS Amundsen ArcticNet Expedition 2017 and Stations 05, 55 and 69 were part of the CCGS Hudson 2018042 Expedition 2018. Inset shows the sampling location within the wider region. The base maps were drawn in Ocean Data View v.5.6.2 (Schlitzer 2018).

2.2 DNA extractions, foraminferal 18S rDNA genotyping and 16S rDNA metabarcoding.

Downstream washing of individual cells preserved in RNALater® for genotyping was carried out to remove the shell and shell-associated external contaminants according to Bird et al. (2017). After washing, DNA was extracted from individual
foraminifer specimens using the DOC extraction method to identify the specific genotype (Holzmann and Pawlowski, 1996). PCR amplification of the foraminiferal 18S rRNA gene was performed with three rounds of PCR using a Phire Hot start DNA polymerase master mix (Thermo-Scientific) 3% DMSO and an annealing temperature of 58°C with 25 cycles. Post-RNAlater stored DNA, crushed into 40µl DOC extraction buffer was diluted 1 in 20 in PCR grade water. Primer pairs were as follows: Primary PCR: C5-sB, secondary PCR: N5-N6, tertiary PCR: 14F1-N6. PCR products between rounds were diluted 1 in 100 PCR grade water and 1 µl was used in the following round of PCR. Cloning to account for intra-individual variation was carried out according to Darling et al. (2016). DNA sequencing was carried out using the BigDye® Terminator v3.1 Cycle Sequencing Kit and an ABI 3730 DNA sequencer (both Applied Biosystems).

Filtrate from water samples was extracted for DNA using the DNeasy power water kit (Qiagen). Filters were removed from RNAlater® (Ambion™) to clean 1.5 ml microfuge tubes and centrifuged for 1 min at 10,000 xg. Excess RNAlater® (Ambion™) was removed and the filter was transferred to the bead beating tubes of the DNeasy power water kit and processed following the manufacturer’s protocol. A control DNeasy power water kit extraction on a clean filter was carried out in parallel. Alongside foraminiferal and water samples, six reagent controls were also processed across two 16S rRNA gene metabarcoding sequencing runs. On the first run there were two PCR controls containing no DNA template, an extraction control containing 2.5µl DOC buffer only and an extraction control containing 1µl of elute from a Qiagen DNA extraction of a clean 0.2 µm polycarbonate filter. On the second run a PCR control and a Qiagen extraction control were processed. The V4 region of the 16S rRNA gene was chosen for amplification using barcoded 515F (Parada et al., 2016) forward and 806R (Apprill et al., 2015) reverse primer pairs modified from the original primer pair (Caporaso et al., 2011, Walters et al., 2016). Each DNA sample and control was PCR amplified with unique dual barcoded tags that enabled demultiplexing of the samples after being pooled for sequencing. PCR reactions were performed in triplicate and contained 1 Unit Phusion DNA polymerase (ThermoScientific) 1 x Phusion HF buffer, 0.2 mM each dNTP, 0.4 µM of each primer, 0.4mM MgCl2 and 2.5 µl (foraminifera) or 1 µl (water column) of template DNA in a 50 µl volume made up with PCR grade water (Sigma). All PCR reactions were set up in a UV sterilization cabinet (GE healthcare). Reaction tubes and PCR mixtures were treated for 15 minutes with 15 W UV light (wavelength = 254 nm) to destroy contaminating DNA, prior to addition of dNTPs, DNA polymerase primers and template DNA (Padua et al., 1999). PCR products were run on a 1.2 % agarose gel and the triplicate reactions were pooled before purification with the Wizard® SV Gel and PCR Clean–Up System (Promega).

The purified amplicons were quantified using a Qubit® 2 fluorometer (ThermoFisher Scientific) prior to pooling at equimolar concentrations for DNA sequencing. DNA sequencing was performed on two separate runs at Edinburgh Genomics using an Illumina MiSeq v3 to generate 253 base pair (bp) paired-end reads.

2.3 Quality filtering paired end reads, rarefaction, taxonomic assignment and sequence filtering.

The Quantitative Insights in Microbial Ecology 2 pipeline (QIIME2, Bolyen et al., 2019) was used for initial analyses. Sequences were trimmed and denoising was carried out on each run separately using the DADA2 plugin (Callahan et al., 2016) for quality filtering, dereplication, removal of singletons, chimera identification and removal and merging paired-end reads.
This method generates amplicon sequence variants (ASVs) and a set of representative sequences. Alpha rarefaction was carried out in QIIME2 (metrics: observed OTUs; Shannon; and faith pd) to assess whether the sequencing depth was adequate to detect bacterial diversity. Samples were rarefied to the lowest sequencing depth observed across all samples and sampling depth was adequate across all samples. Taxonomy was assigned using an SKlearn classifier pre-trained on the database SILVA-132 99% OTUs from the V4 515f/806R region of the sequences (Quast et al., 2013). Prior to further analysis, abundance data were normalised to the total number of counts per sample as a relative abundance, and only ASVs present at > 0.1 % abundance averaged across all samples were retained (Prazeres et al., 2018). Unassigned taxa or those taxa assigned to mitochondria or eukaryotes (totalling 71 ASVs) were removed from the sample set leaving 59 samples, 130 ASVs and 4,620,174 reads for analysis.

2.4 Statistical analyses

Statistical analyses were performed in R v 4.0 (R Core Team, 2017). Differences between the provenances (foraminiferal microbiome and water column assemblage) were tested using the package mvabund (Wang et al., 2012). This generalized linear model approach takes into consideration the variability in the mean-variance relationships across the ASVs and absolute counts are used for the analysis. The negative-binomial model best fit the dataset, and ANOVA was performed to test for significant difference between provenances using mvabund and Vegan (Oksanen et al., 2017).

Differences in taxa associated with the different provenances, stations and water depths were assessed using the packages Phyloseq (McMurdie and Holmes, 2013) and DESeq2 (Love at al., 2014). DESeq2 was used rather than ANCOM because ANCOM assumes that <25 % of the ASVs are changing between provenances, and here this assumption does not hold true (Mandal et al., 2015). Graphical presentation of results was performed with ggplot2 (Wickham, 2009). Bray-Curtis dissimilarity between different samples was calculated and visualised using non-multidimensional scaling (NMDS) in Phyloseq. To determine if provenance, sample depth or station significantly affected the assemblages a multivariate analysis of variance (ANOVA) was carried out using the Adonis function in the Vegan R package (Oksanen et al., 2018).

Compositional differences, and specific taxa that were significantly different between provenances were identified using log2 of fold change analysis in DESeq2 by converting the phyloseq-object to a DESeq2 object. The DESeq2 analysis was run with size factor type set to “poscounts” which allows values of zero in the sample counts and accounts for the data transferal from a phyloseq-object (van den Berge et al., 2018). The significance test was set to “Wald”, and a “local” fit type for fitting of dispersions. The core microbiome, here defined as ASVs present across 80 % of the foraminifera, was identified using Microbiome (Lahti and Shetty et al., 2017).

2.5 Transmission electron microscopy

TEM was used to observe and document the structural relationships between any endobiotic micro-organisms and foraminiferal cells. After fixation in TEM fixative (see methods section 2.2) specimens were post–fixed in 1 % Osmium Tetroxide in 0.1 M Sodium Cacodylate for 45 minutes, followed by a further three 10-minute washes in distilled water.
Specimens were then set in small cubes of 1% low melting point agarose and decalcified in 0.1 M EDTA (pH 7.4) for 1 hour and 48 hours at 4 °C. Fixed cells were then dehydrated in 50%, 70% and 90% ethanol for 2 x 15 minutes followed by 100% ethanol 4 x 15 minutes. Two 10–minute changes in Propylene Oxide were carried out prior to being embedded in TAAB 812 resin. Sections, 1 μm thick, were cut on a Leica Ultracut ultramicrotome, stained with Toluidine Blue, and then viewed under a light microscope to select suitable specimen areas for investigation. Ultrathin sections, 60 nm thick, were cut from selected areas, stained in Uranyl Acetate and Lead Citrate and then viewed with a JEOL JEM–1400 Plus TEM.

3 Results

3.1 The water column

Water samples were taken from three locations: stations 101, 115 and 323 (Fig. 1, Table 1). The water at Station 115, to the east of the Pikialasorsuaq, is derived from warm Atlantic water, and the surface water temperature here is as high as 4°C, with a steep thermocline to 40 m (Fig. 2). At stations 101 and 323 in the west of the Pikialasorsuaq, and to the southwest respectively, Pacific-derived water enters via the colder Arctic Ocean (Tremblay et al., 2002; Bergeron and Tremblay, 2014) and consequently the surface waters are colder. Whilst station 323 has a surface temperature of 3°C, it very rapidly drops to -1°C by 20 m and station 101 has a surface maximum temperature of only 0.75°C. The chlorophyll maximum is closely associated with the temperature maximum at all stations.

Figure 2. CTD plots for temperature, salinity, and fluorescence at stations 101 (350m), 115 (653m) and 323 (789m) where both water and foraminifera were collected.

The general microbial assemblages in the water column across the three stations (101, 115 and 323) display a similar pattern of assemblage composition with depth (Fig. 3). Surface waters contain either no, or extremely low relative abundances of
chloroplast or archaeal ASVs. Chloroplast ASV relative abundance increases steeply with depth however, with the highest abundance found in the 50 m water samples, before numbers reduce again. This pattern agrees with our CTD data, where the chlorophyll maximum occurs between 20-40 m across all stations (Fig. 2). The chloroplast ASVs make up an average of 2.97% of the ASVs in the water column.

Archaeal ASVs (Thaumarchaeota and Euryarchaeota) are most predominant below 50 m, peaking in the 100-150 m water samples. Bray Curtis dissimilatory (Fig. A1) and multivariate analysis (Adonis; vegan) indicate systematic differences in composition with both depth and station. Depth drives 52% of the variability (Pr = 0.001) compared to Station driving 17% of the variability (Pr = 0.001).

Figure 3. The relative abundance of 16S rDNA ASVs generated from the foraminiferal specimens (Fm) and the water column (WC). Note that foraminifera were successfully processed from 7 stations (see Fig. 1) and water was collected and processed from three stations. Taxa are shown at the phylum level except for chloroplast derived 16S ASVs, which are grouped together. Black lines in bars indicate divisions of ASVs. Brackets at top indicate stations from which samples were taken, e.g. Fm101 are foraminifera samples from station 101, whereas WC101 are water samples from station 101.
3.2 Statistical comparison of water column vs foraminiferal ASVs

The combined microbial assemblages of the 31 water column samples and the 28 foraminifera specimens (see Table 1) consist of 130 identified ASVs after removal of very low abundance ASVs (< 0.1 % abundance averaged across all samples). The raw 16S metabarcoding dataset was submitted to the NCBI Sequencing Read Archive (BioProject accession PRJNA984332). By far the most prevalent ASVs in the foraminifera are those from chloroplasts (55.7 %), and particularly diatom chloroplasts (52.7 %). Proteobacteria ASVs represent the most abundant ASVs in the water column (49.8 %). (Fig. 3). Multivariate analysis determined that the composition of ASVs between the two provenances were significantly different (Bray Curtis, mvabund: LRT=574.3, P<0.001; Adonis: F.model=24.703, Pr(>F)=0.001).

The ASVs driving the statistically significant compositional differences in the water column (Fig. 4) are 14 Archaea, (Euryarchaeota, Thermoarchaeota), 34 Proteobacteria, eight Bacteroidetes and four Marinimicrobia. The ASVs driving compositional differences in the foraminifera are 13 diatom chloroplasts, 11 Proteobacteria, seven Bacteriodetes, a Planctomycetes and an Actinobacteria. The ASVs of most significance (Log2 fold Change > +10/-10) are listed in Table A1.

Figure 4. Differential abundance testing of ASVs between Provenances using DESeq2. The Log2 fold change in ASVs is the log-ratio of the ASV means in the water column and foraminifera. ASVs with positive Log2 fold change are significantly more
abundant in the water column assemblage and ASVs with negative values indicate ASVs which are significantly more abundant in the foraminiferal assemblages.

A more meaningful comparison of differences between foraminifera and water column is to remove location as a variable which drives 17% of the difference between water samples. We therefore compared all samples from station 101, which provided sufficient numbers (seven foraminifera and 12 water samples) to statistically verify that the microbiome compositional differences remained significant within a single water column, (Adonis: F.model=10.304, Pr(>F)=0.001) and compositional differences reflected those of the entire data set (Figs. A2 & A3).

3.3 Foraminiferal ASV profiles

All the foraminifera were genotyped as *N. pachyderma* Type I (NCBI GenBank Accession numbers OR137988-OR138014), consistent with it being the only genotype found in the Arctic region to date (Darling et al., 2004; 2007). Since foraminifera were sampled by vertical net tow from 200 m depth to the surface, no depth correlation data are available. However, since station accounts for 17% of the variation across water samples, we might expect to see some variation in specimen ASV profiles between foraminifera from different stations too. A Bray Curtis dissimilarity NMDS plot (Fig. A4) shows a degree of clustering of foraminifera by station (Adonis: F.model= 3.079, Pr(>F) = 0.001), with 49% of the variation in the foraminifera driven by station. Of note are two specimens, Fm176b and Fm101e, which contained only 11 and nine distinct ASVs respectively. All other foraminifera contain 22 or more ASVs. The phyla with some ASVs showing higher abundance counts in the foraminifera compared to the water column are Planctomycetes, Firmicutes, Bacteriodetes, and the group “Chloroplast” (Fig. A5).

3.3.1 Chloroplasts ASV profiles in *N. pachyderma*

ASVs are assigned to 17 distinct chloroplast-containing taxa across all water and foraminifera samples, contributing, on average, 55.7% of all ASVs in the foraminifera and only 2.97% in the water column (Fig. 3). More specifically, those ASVs assigned to diatom chloroplasts contribute to 52.7% of all ASVs in the foraminifera and only 2.92% in the water column, highlighting the major importance of diatoms (Class Bacillariophyceae) in the diet of the foraminifera compared to other phytoplankton taxa. Of the 17 chloroplast ASVs identified, 13 drive the significant difference between the foraminifera and the water column, due to greater abundance in the foraminifera. These are chloroplasts from three *Cylindrotheca* sp. (ASV18, ASV31 and ASV23), three *Fragilariopsis* sp. (ASV47, ASV84 and ASV59), *Synedra hyperborea* (ASV35), a *Melosira* sp. (ASV111), an *Attneya* sp. (ASV16), a single ASV from the Skeletonema family (ASV99), two ASVs identified only to the Bacillariophyceae class of diatoms (ASV14 and ASV15) and a *Chaetoceros* sp. (ASV102). The most statistically significant (in driving the differences between the foraminifera and the water column) are the two ASVs belonging to a *Cylindrotheca* sp.
chloroplasts (ASV18) and a *Fragilariopsis* sp. chloroplast (ASV47; Table A1). The relative abundance of chloroplast ASVs in each sample is shown in Figure 5.

![Figure 5](https://doi.org/10.5194/egusphere-2024-497)

**Figure 5.** The relative abundance of the 17 chloroplast 16S rDNA ASVs generated from the foraminiferal specimens (Fm) and the water column (WC). Taxa are shown at the genus level, except where noted in the key. Black lines in bars indicate divisions of ASVs. Brackets at top indicate stations from which samples were taken, e.g. Fm101 are foraminifera samples from station 101, whereas WC101 are water samples from station 101.

The most abundant chloroplast ASV (comprising > 90 % in some samples) is ASV15, identified only to the level of class Bacillariophyceae (Fig. 5). It has a significant Log2 fold change of -4.934 (padj = 7.35E-17). This is not as significant as for some other diatom chloroplast ASVs (Table A1), probably due to its higher relative abundance in the water column (Fig. 5). In a Blastn search, this ASV corresponds 100 % identity and coverage to the chloroplast 16S rRNA gene of the diatom *Chaetoceros gelidus* (accession NC_063631.1). ASV15 is common within Baffin Bay as it shows a >80 % relative abundance (when analysing chloroplasts only) in the water column samples from all but station 115, which comprises > 50 % ASV15.
ASV15 is therefore relatively common in both the water column and the foraminifera and is clearly a major food source for *N. pachyderma* in Baffin Bay during the summer months.

At station 101 (west side of the polynya; Fig. 1), six out of the seven foraminifera specimens contained > 45 % chloroplast ASVs, and only one specimen contained < 20 % chloroplast ASVs (Fm101a; Fig. 3). When analysing only chloroplast ASVs, six of the seven foraminifera specimens contained >80 % ASV15 (*Chaetoceros gelidus*). Specimen Fm101e, however, contained over 99 % chloroplast ASVs belonging to two streptophyta (Fig. 5). Streptophyta ASVs were only found in one other foraminifera, Fm176b at station 176, at the most southerly cruise station for which we have no water column 16S data.

In fact, despite containing a streptophyta ASV, specimen Fm176b contained a very low relative abundance of chloroplast ASVs (Fig. 3). Except for specimen Fm101e, the diatom ASVs at the cooler station 101 can be said to mirror the diatom population profile in the water column where > 90 % of water column chloroplast ASVs were ASV15 (*Chaetoceros gelidus*) (Fig. 5).

The two foraminifera processed at station 115 (where Atlantic-derived warmer water is found) contained <30 % and <10 % chloroplast ASVs (Fig. 3). Again, this reflects the water column where we find the lowest relative proportion of chloroplast ASVs across the three stations (Fig. 3). *Fragilariopsis* sp. contributed the greatest proportion of chloroplast ASVs in station 115 specimens (Fig. 5). The proportion of *Fragilariopsis* ASVs are higher in the water column at this station relative to other stations, although ASV15 (*Chaetoceros gelidus*) remains the major diatom ASV present. In addition, a non-diatom chloroplast (*Triparma laevis*) is also present in both the water column and the foraminiferal specimens. This is a relative of the diatoms, and, like most diatom species, forms siliceous plates. *Triparma laevis* ASVs are only detected in the upper water column at 50-100 m, and therefore it is highly likely that the two foraminiferal specimens were also collected from this depth.

Finally, >50 % of the ASVs in the two foraminifera from station 323 are chloroplast ASVs (Fig. 3), and of those, >90 % are ASV15 (*Chaetoceros gelidus*; Fig. 5). This reflects the high proportion in the water column at this station.

Of the other foraminiferal specimens taken from stations with no comparative water column data, the foraminifera from station 176 show the greatest diversity in chloroplast ASVs. They contained a much higher relative proportion of *Fragilariopsis*, *Melisosira*, and *Cylindrotheca* diatoms, and Streptophyta across the different specimens. This may indicate higher comparative diatom and algal (Streptophyta) diversity at this more southerly station.

### 3.3.2 Prokaryote ASV profiles in *N. pachyderma*

ASVs are assigned to 113 distinct prokaryote taxa across all water and foraminifera samples. The major classes are Gammaproteobacteria (phylum Proteobacteria) which contributes on average 19.7 % of ASVs and Bacteriodia (phylum Bacteriodetes) which contributes on average 14.1 % of ASVs. Those ASVs that drive the significant compositional differences between provenances, being significantly more abundant in the foraminiferal samples, are 11 Proteobacteria, eight Bacteriodetes, three Firmicutes, a Planctomycetes and an Actinobacterium (Fig. 4). More details on these ASVs including their finer scale taxonomy can be found in Table A2.
3.3.3 The *N. pachyderma* core microbiome

The core microbiome of *N. pachyderma* is defined here as ASVs found in 80% of the foraminiferal specimens across all stations. This core microbiome could be made up of organisms which (i) the foraminifera specifically target for food, or (ii) are routinely passively ingested due to close association with specific food sources, or (iii) are endo(symb)ionts. 16S metabarcoding indicates that there are nine core ASVs. Two are represented by diatoms: ASV15, *Chaetoceros gelidus* (27/28 specimens) and ASV59, identified in BLASTn as *Fragilariopsis cylindrica* (100% match to accession NC_045244.1, 24/28 specimens). Then seven bacterial ASVs from the Flavobacteriaceae family (ASV19, 23/28), the genus *Pseudoalteromonas* (ASV26, 25/28), the genus *Aurantivirga* (ASV27, 24/28), the genus *Paraglaciecola* (ASV34, 25/28), the family Halieaceae (ASV74, 26/28), the genus *Micrococcus* (ASV10, 25/28) and genus *Bradyrhizobium* (ASV116, 23/28; Fig. 6). Of these nine ASVs, six are also significantly more abundant in the foraminifera than the water column, driving the significant differences between the provenances. These are, in order of significance, *Micrococcus* ASV10, *Fragilariopsis cylindrica* ASV59, *Bradyrhizobium* ASV116, *Chaetoceros gelidus* ASV15, Flavobacteriaceae family ASV19, and Halieaceae ASV74. This foraminiferal core microbiome makes up, on average 49.6% of the ASVs in the *N. pachyderma* of Baffin Bay, whereas it makes up only 10.89% of ASVs in the water column.
Figure 6. The average relative abundances of the 16S rDNA ASVs found across *N. pachyderma* Type I specimens in Baffin Bay during summer 2017. The average relative abundance of the nine core ASVs (found in ≥ 80 % of specimens) are taxonomically labelled and designated the “core microbiome”. These make up 49.6 % of the ASVs in the microbiome. 50.4 % are ASVs found in fewer than 80 % of specimens and are designated “non-core”.

3.4 TEM analysis

16S metabarcoding data indicate that diatoms are the majority diet for *N. pachyderma*. TEM imaging was carried out on samples collected during the 2018 cruise (Fig. 1: Table 1) to further investigate the diet/endobionts in this genotype.

Whole diatoms, including frustules, were observed within the foraminiferal cell (Fig. 7a), although the pyrenoid-dissecting lamellae characteristic of diatom chloroplasts were not visible in our fixed samples. Empty frustules were also observed both inside and outside the foraminiferal cell. Those outside may have been ejected after the diatom organic material was digested/removed (Fig. 7b) and were likely caught in the external cytoplasm and rhizopodial network at the time of sampling and fixation, as has been reported previously (Spindler et al., 1984). These observations support the previous literature indicating that *N. pachyderma* eats diatoms.

In addition, the TEM images also show that *N. pachyderma* contains unexpectedly high numbers of chloroplasts throughout the cell from the cell periphery to the cell centre (Fig. 7c, 7d & 7e, Fig. A6). The level of preservation does not allow us to observe the number of membranes surrounding the chloroplasts, the pyrenoid-dissecting lamellae, or the degradation state of each chloroplast. Nevertheless, in several of the chloroplasts, obvious lenticular pyrenoids, and horseshoe-shaped arrays of thylakoid membranes are visible (Fig. 7d & 7e) as found in *Chaetoceros* spp. (Bedoshvili et al., 2009). There are abundant lipid droplets located amongst, and immediately adjacent to the chloroplasts at the cell periphery (Fig. 7d).
Figure 7. TEM images of three individual *N. pachyderma* specimens. (a) Intact diatoms (d) observed inside specimen BB2 (Table 1). The outer membrane is identified by the black arrows, where the pore shape can be observed. v = internal cell vacuole. ex = external to the foraminiferal cell. (b) External (ex) to specimen BB2 (Table 1) where debris, including empty diatom frustules (f), is apparent. The outer membrane is identified by black arrows, c = chloroplasts inside BB2, v = internal vacuole. Large vacuoles were observed in several specimens which may be a result of the fixation process. Scale bar of 8a and 8b 2µm. (c) Overview of thin section of BB11 (Table 1) showing chambers and identifying region of cell shown in (e). Scale bar 10µm. (d) Many chloroplasts (c) inside specimen BB1 (Table 1) close to the cell periphery. Additional chloroplasts with obvious pyrenoids are indicated with a white arrowhead. Black spots are lipid droplets (black arrow points to an example). ex = external to the foraminiferal cell. Scale bar 1µm. (e) The centre of specimen BB11 (Table 1). Image location within the foraminifera is identified by black box in (c). Chloroplasts (c) are present at the core of the cell where the chambers coalesce. Additional chloroplasts with obvious pyrenoids are identified by black arrowheads.
4 Discussion

In this study our aim has been to investigate the microbiome within the polar planktonic foraminifera *N. pachyderma* from the Arctic Baffin Bay region. We define the microbiome as the combined taxa identified by taxonomic assignment of 16S ASVs generated by metabarcoding. This will include food, any endo(symb)ionts, and chloroplast-containing eukaryotes identified by their chloroplast 16S ASVs. Shedding light on their feeding preferences as well as any microbial associations that form part of the “interactome” in the context of the changing climate may afford some clues as to the ability of *N. pachyderma* to withstand/adapt to its rapidly changing environment, and its contribution to the carbonate cycle and ocean alkalinity. For example, eco-physiological and trait-based models indicate that symbiont barren foraminifera, which *N. pachyderma* is believed to be, are predicted to experience reduced numbers and habitat decline (Roy et al., 2015), and the non-spinose species biomass is likely to be reduced by up to 11 % by 2050 (Grigoratou et al., 2022). Sound knowledge of the eco-physiology and traits of foraminifera is required for model accuracy, and to that end the genotype and 16S microbiome of the Arctic polar *N. pachyderma* has been investigated.

4.1 Divergent Feeding strategies in the Neogloboquadrinids

Our findings support previous literature stating that *N. pachyderma* feeds on diatoms (Spindler and Dieckmann, 1986; Schiebel and Hemleben, 2017, Greco et al., 2021). This study and that of Greco et al., (2021) indicate that *N. pachyderma* feeds predominantly on diatoms, class Bacillariophyceae, and occasionally other algae (e.g., streptophyta and *Triparnia laevis* - this study). However, our 16S study further reveals that *N. pachyderma* also consume bacteria, since their bacterial ASV composition is significantly different from the water column profile. This is most likely driven by selectivity in feeding behaviour, where particulate organic matter (POM) is gathered around the shell to form a feeding cyst. Once formed, the foraminifer sits within the POM microhabitat, becoming isolated from the water column. This behaviour is already observed in the Neogloboquadrinids *N. dutertrei* and *N. incompta* (Bird et al., 2018; Fehrenbacher et al., 2018), and in *Globigerinita glutinata* (Spindler et al., 1984).

Although the Neogloboquadrinids all feed within the POM microhabitat, we suggest that they are feeding on very different components within the cyst. Work carried out by Bird et al. (2018) demonstrates that of the three Neogloboquadrinids, *N. incompta* contains the highest proportion of bacterial ASVs (>99.8 %), indicating that it targets the bacteria within the POM microhabitat. The small proportion of chloroplast ASVs (<0.2 %) in the *N. incompta* study indicate that POM is not being passively phagocytosed, but that the bacteria, rather than algae are being specifically selected as food. In contrast, *N. dutertrei* contained only 2 %-4 % bacterial ASVs and instead maintains a pelagophyte algal endosymbiont population, and selectively feeds on other protists within the POM. The small proportion of intracellular bacteria in *N. dutertrei* indicates that it too does not specifically phagocytose POM itself. However, in the case of *N. pachyderma*, we find a higher proportion of bacterial ASVs than was identified in *N. dutertrei* and a higher proportion of chloroplast ASVs than found in *N. incompta*, suggesting that *N. pachyderma* may feed on the POM directly for food. This has been suggested by Greco et al. (2021) who also
demonstrated that the *N. pachyderma* 18S ASV assemblage revealed little difference in intracellular diatom ASVs between the surface dwelling and the deeper dwelling specimens living in diatom-free waters. This finding led them to suggest that *N. pachyderma* feeds on detrital diatoms, which is supportive of a POM-cyst mode of feeding. However, the 16S data and TEM images (Fig. 7, Fig. A6) in our study indicate that they also feed on living diatoms. Further, our evidence indicates that they either actively or passively consume the bacteria living in the diatom phycosphere.

### 4.2 The *N. pachyderma* core microbiome

All specimens were genotyped as *N. pachyderma* Type I, which is unsurprising given that all specimens genotyped to date from the polar waters of the Norwegian Sea and the Fram Strait have been identified as Type I (Darling et al., 2004; 2007). Two of the specimens, Fm176b and Fm101e (Fig. 3) had a much-reduced microbiome compared to all other specimens (11 and nine ASVs respectively, compared to at least 22 in all other specimens). One possible explanation for this is that these specimens were close to the end of their life span and had stopped feeding prior to producing and releasing gametes (Fehrenbacher et al., 2018). In the core microbiome analysis, Fm176b was missing seven of eight core ASVs and Fm101e was missing six of eight, highlighting their unusually sparse microbiomes compared to all other specimens (Table 2).

**Table 2.** The taxonomic assignment, Log2 fold change and abundance characteristics of the nine ASVs that make up the foraminiferal core microbiome.

<table>
<thead>
<tr>
<th>ASV</th>
<th>Log2 FC</th>
<th>Taxonomy</th>
<th>ASV Relative abundance in <em>foraminifera</em></th>
<th>Total ASV counts in <em>foraminifera</em></th>
<th>ASV Relative abundance in water</th>
<th>Total ASV counts in water</th>
<th>Forams ASV present in</th>
<th>ASV missing from</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASV15</td>
<td>-4.9343</td>
<td>Bacillariophyceae (<em>Chaetoceros gelidus</em>)</td>
<td>30.30%</td>
<td>785,765</td>
<td>3.05%</td>
<td>61,913</td>
<td>27/28</td>
<td>Fm176b</td>
</tr>
<tr>
<td>ASV59</td>
<td>-6.1951</td>
<td>Fragilariopsis (<em>Fragilariopsis cylindricus</em>)</td>
<td>9.33%</td>
<td>241,785</td>
<td>0.05%</td>
<td>918</td>
<td>24/28</td>
<td>Fm176b, Fm101a, Fm101b, Fm101e,</td>
</tr>
<tr>
<td>ASV10</td>
<td>-10.8437</td>
<td>Micrococcus</td>
<td>0.30%</td>
<td>7766</td>
<td>0.00%</td>
<td>0</td>
<td>25/28</td>
<td>FmBB2c, Fm101e, Fm301f,</td>
</tr>
<tr>
<td>ASV19</td>
<td>-3.0763</td>
<td>Flavobacteriaceae (Family)</td>
<td>2.15%</td>
<td>55,659</td>
<td>0.49%</td>
<td>9,989</td>
<td>23/28</td>
<td>Fm176b, Fm101d, Fm101e, Fm101g, Fm301a</td>
</tr>
<tr>
<td>ASV26</td>
<td>NA</td>
<td>Pseudoalteromonas</td>
<td>3.50%</td>
<td>90,829</td>
<td>4.09%</td>
<td>83,015</td>
<td>25/28</td>
<td>Fm176b, Fm101e, Fm301b</td>
</tr>
<tr>
<td>ASV27</td>
<td>NA</td>
<td>Aurantivirga</td>
<td>0.76%</td>
<td>19,758</td>
<td>1.77%</td>
<td>35,847</td>
<td>24/28</td>
<td>Fm176b, FmBB2a, FmBB2c, Fm101e,</td>
</tr>
<tr>
<td>ASV34</td>
<td>NA</td>
<td>Paraglaciecola</td>
<td>2.35%</td>
<td>60,932</td>
<td>1.22%</td>
<td>24,881</td>
<td>25/28</td>
<td>Fm176b, FMS01a, Fm301f,</td>
</tr>
<tr>
<td>ASV74</td>
<td>-2.5643</td>
<td>Halieaceae (Family)</td>
<td>1.20%</td>
<td>31,125</td>
<td>0.31%</td>
<td>6,224</td>
<td>26/28</td>
<td>Fm176b, Fm101e</td>
</tr>
<tr>
<td>ASV116</td>
<td>-4.9486</td>
<td>Bradyrhizobium</td>
<td>0.07%</td>
<td>1791</td>
<td>0.004%</td>
<td>86</td>
<td>23/28</td>
<td>Fm176b, Fm101b, Fm101c, Fm115a, Fm301d</td>
</tr>
</tbody>
</table>

https://doi.org/10.5194/egusphere-2024-497
Preprint. Discussion started: 7 March 2024
© Author(s) 2024. CC BY 4.0 License.
Whilst the core microbiome accounts for almost 50% of the total microbiome in *N. pachyderma*, the relative contributions of the nine core ASVs is highly variable, from 0.09% - 30.3% (Fig. 6). This variability may reflect the foraminiferal food preferences if they are part of the diet, or if endo(sym)bionts, may reflect potential differences in their contribution to cellular activity, their role/function in the *N. pachyderma* cell, or that their involvement could be a function of ontogenetic stage. Of the nine core ASVs, two (ASV15 and ASV59) are taxonomically assigned to diatom chloroplasts. Both contribute to the significant difference in assemblage composition between the foraminifera and the water column since there are twelve times (ASV15) and 236 times (ASV59) more ASV counts in the foraminifera than the water column (Table A3). Chloroplasts were exceptionally abundant in our TEM images (Fig. 7 and Fig. A4), which appear very similar to observations made in kleptoplastic benthic species (Jauffrais et al., 2018; Jesus et al., 2022). This raises important questions about the nature of the relationship between the chloroplasts and *N. pachyderma* Type I, which is discussed below. Empty diatom frustules were also observed in the TEM images (Fig. 7), which is highly consistent with previous reports that diatoms are a significant part of the *N. pachyderma* diet (Hemleben et al., 1989; Scheibel & Hemleben, 2017; Greco et al., 2021).

The remaining core microbiome consists of diverse bacterial genera. Of particular significance is *Micrococcus* (ASV10, 25/28 foraminifera), since its ASV counts total 7766 in the 28 foraminifera but zero in water samples (Table 2). They make up only 0.3% of the total ASV assemblage in the foraminifera, and 0% in the water column. The absence of ASV10 from the water column might suggest that this ASV could belong to a bacterial endo(sym)biont of *N. pachyderma*, albeit it occurs in low numbers within the cell. The mode of endobiont transfer could potentially be from an agamont mother cell to the gamont daughter cell, since asexual reproduction has been reported for this species (Kimoto and Tsuchiya., 2006; Davis et al., 2020; Meilland et al., 2022) and endobiont transfer has recently been observed during asexual reproduction in the foraminifer *Globigerinita uvula* (Takagi et al., 2020). Prior to these observations of asexual reproduction and symbiont transfer, all known symbionts of planktonic foraminifera were thought to be acquired directly from the water column by the gamonts due to the large symbiont size relative to the gametes (Hemleben et al., 1989; Bijma et al., 1990), water column prevalence, and lack of genetic drift between water column specimens and endobionts (Bird et al., 2017). *Micrococcus* is a hydrocarbon degrading genus (Atlas et al., 1995) that is also capable of chitin degradation (Annamalai et al., 2010) and is used in the production of the bioplastic polyhydroxybutyrate (Mohanrasu et al., 2021). There may be a role for Micrococcus within the foraminiferal host in helping to breakdown copepod-derived chitin phagocytosed in POM (Greco et al., 2021). In addition, since diatoms are a major producer of hydrocarbons (Stonik & Stonik, 2015), and are clearly eaten in large numbers, there may be a role for *Micrococcus* to degrade such hydrocarbons within the foraminiferal cell. Indeed, several other significant hydrocarbon degrading genera are found in *N. pachyderma*. For example, of the four *Colwellia* ASVs (48, 98, 119 and 125), two ASVs (48 and 119) drive the significant differences between the foraminiferal microbiome and the water column. Combining all four *Colwellia* ASVs, *Colwellia* are present in all but the two potentially gametogenic specimens, Fm176b and Fm101e. Interestingly, *Colwellia* have been shown to break down hydrocarbons most efficiently at temperatures of around 4°C.
A further ASV from a hydrocarbon degrading species is ASV40 which is taxonomically assigned to the genus Moritella. This genus of gammaproteobacteria are psychrophilic gram-negative facultative anaerobes and were found to carry out hydrocarbon degradation in the waters of the Northwest Passage in the Canadian Arctic Archipelago (Garneau et al., 2016). This ASV was identified in 15 of 28 foraminifera. Pseudoalteromonas are also known hydrocarbon degraders (e.g. Calderon et al., 2018) and a single Pseudoalteromonas ASV was identified as a core member of the microbiome. Whether the abundance of hydrocarbon degraders within the foraminifera is a function of their presence in the hydrocarbon-producing diatom’s “phycosphere” (Bell & Mitchell, 1972), or whether there is a more specific foraminiferal-endobiont interaction with any of these genera is yet to be determined.

There are three further core ASVs assigned to bacterial taxa: ASV19 from family Flavobacteriaceae; ASV74 from genus OM(NOR) of the family Halieaceae; and ASV116, a Bradyrhizobium sp. (Fig. 6). All three of these ASVs also contribute to the significant differences between foraminifera and water column samples (Fig. 4). Flavobacteriaceae are a large family of bacteria that are widely distributed in the marine environment. Tisserand et al. (2020) isolated ten species from Baffin Bay, and all were shown to grow on exudates (dissolved organic matter) from two Arctic diatoms (Fragilariopsis cylindricus and Chaetoceros neogracilis). This infers that Flavobacteriaceae may be part of the diatom phycosphere, consumed alongside diatoms by feeding N. pachyderma, leading to the higher proportion of Flavobacteriaceae ASV19 in the foraminifera compared to the water column. Additionally, Flavobacteriaceae are often found associated with detritus (as well as algae, fish and invertebrates; Gavriilidou et al., 2020). Therefore, N. pachyderma, which feeds from a detrital feeding cyst on diatoms and diatom detritus (Greco et al., 2021), could passively consume Flavobacteriaceae present in the detritus.

ASV74 is attributed to the OM(NOR) genus of the Family Halieaceae, (order Cellvibrionales, Spring et al., 2015). Total ASV74 counts for the foraminifera are 31,125 (mean relative abundance = 1.2 %) and 6,224 for the water column (mean relative abundance = 0.31 %). In total, 4 % of the ASVs are assigned to the order Cellvibrionales in the water column, but only ASV74 is identified in the foraminifera as well as the water column. The order Cellvibrionales are gram-positive aerobes that are mesophilic and neutrophilic chemoorganotrophs. However, some members of the Family Halieaceae may additionally be capable of aerobic phototrophic growth using bacteriochlorophyll a, and carotenoids for the harvesting of light. Several strains may also be able to use proteorhodopsin to utilise light as an energy source (Spring et al., 2015).

The final bacterial ASV that contributed to both the significant difference between the water column and the foraminiferal ASV assemblage (Fig. 4) and the core microbiome (Fig. 6) is ASV116, Bradyrhizobium. It constitutes on average 0.09 % of foraminiferal and 0.004 % of water column ASVs, indicating that the relative abundance is an order of magnitude greater in the foraminifera, although actual counts across the foraminifera (1791) versus the water column (86) are extremely low. This genus contains mainly nitrogen fixing species that are part of phylogenetic cluster I of nifH (Chien and Zinder, 1994; Gaby et al., 2014), which encodes the nitrogen fixing enzyme nitrogenase. nifH sequences that cluster with Bradyrhizobium in Cluster I have been isolated from the Central Arctic Ocean water column. In fact, sequences from subcluster IK which includes Bradyrhizobium, made up >50 % of the nifH subcluster sequence abundance in the open waters of the Central Arctic Ocean (Fernández-Méndez et al., 2016), supporting our findings in the polar waters of Baffin Bay.
Three final core ASVs are *Pseudoalteromonas* (ASV26), *Aurantivirga* (ASV27) and *Paraglaciecola* (ASV34) (Fig. 6). None of these ASVs drive the differences between the foraminifera and water column; numbers of ASVs and the relative abundances are substantial and similar between these provenances, and it is therefore highly likely that these species are passively ingested during *N. pachyderma* feeding on POM. As described above, *Pseudoalteromonas* is a hydrocarbon degrading genus, which may explain its presence. *Aurantivirga* is a gram-negative, aerobic, proteorhodopsin-containing, rod-shaped genus of Flavobacteriaceae, described above as feeding on diatom exudates. However, a BLASTn search (NCBI) identifies ASV27 not as *Aurantivirga*, but as 99.6% identical to an alternative Flavobacteriaceae of the genus *Tenacibaculum* including *T. insulae*, *T. haliotis* and *Tenacibaculum* sp. This genus contains many opportunistic fish pathogens, some of which are found to target fish teeth, a high source of calcium shown to promote the bacteria’s growth (Hikida et al., 1979; Frisch et al., 2018). Growth promotion by calcium may be a common feature of the *Tenacibaculum* genus and may be another reason why this ASV is identified with *N. pachyderma* in polar waters may provide a suitable supply of accessible calcium for this genus. Finally, *Paraglaciecola* are a genus of the family Alteromonadaceae. In a BLASTn search this ASV shows 100% identity with *Paraglaciecola psychrophila*, *P. arctica* and several other *Paraglaciecola* sp. sequences. *Paraglaciecola psychrophila* is a gram-negative, psychrophilic, motile rod-shaped bacteria. Identified from the sea ice of the Canadian Basin and the Greenland Sea, it is aerobic and optimum growth is at 12°C. (Zhang et al., 2006). Unable to reduce nitrate, it may be associated with POM as an N-source, and so be ingested by *N. pachyderma* as it feeds on the detritus.

Lastly, of note, although not identified as a core microbiome member, members of the phylum Planctomycetes were often found in greater numbers in their foraminiferal hosts, than in the water column (Fig. A5) and one ASV (101; *Rubripirellula* sp.) contributed to driving the differences between provenances, being more abundant in the foraminifera (Fig. 4). Planctomycetes are widespread in the environment, and are essentially associated with particles including plastics, laminaria seaweeds and POM in the open water column (DeLong et al., 1993; Bondoso et al., 2015; Kallscheuer et al., 2020), again indicating the strong association of *N. pachyderma* with a POM feeding cyst.

### 4.3 Intracellular chloroplast ASVs

On average 55.7% of all 16S rDNA ASVs in the foraminifera belong to chloroplast-containing taxa (Fig 2; Sect. 3.3.1). This contrasts with the 2.97% average proportion found in the water column. Most of the foraminiferal intracellular chloroplast ASVs, are dominated by ASV15, *Chaetoceros gelidus* (BLASTn) from the diatom class Bacillariophyceae. The compositional dominance of ASV15 in the foraminifera reflects the chloroplast ASV composition of the water column (Fig. 5), although found in much higher proportions in the foraminifera (Fig. 3). ASV15 is only missing from specimen Fm176b, one of two specimens thought to potentially have stopped feeding and be in the gametogenic stage of ontogeny. The presence of ASV15, *Chaetoceros gelidus*, in all other specimens indicates its importance to *N. pachyderma* in this location and season. It is characteristic of northern temperate and polar waters (Chamnansinp et al., 2013), and it is a known important biomass fraction in Baffin Bay (Crawford et al., 2018). In fact, eight strains were isolated from Baffin Bay only during bloom development or
bloom peak (Ribeiro et al., 2020) and Chaetoceros’s reputation for bloom forming (Booth et al., 2002) is reflected here in its high abundances compared to other species. The diatom chloroplast 16S ASVs identified in this study (Fig. 5) are also consistent with the diatoms found by Greco et al. (2021), who identified Chaetoceros and Fragilariopsis as major components of the N. pachyderma 18S ASVs from Baffin Bay. Both Chaetoceros (ASV102, ASV15) and Fragilariopsis 16S ASVs (47, 59 and 84) were amongst those ASVs driving the significant difference between the foraminifera and the water column, and both are major constituents of the core microbiome. Intact Fragilariopsis was also identified in the foraminiferal TEM images (Fig. 7a) hinting that, like Ammonia sp. and the miliolid Haurina diversa N. pachyderma may perform intracellular ingestion of the diatom silicate frustules (Jauffrais et al., 2018; Pinko et al., 2023).

4.4 Observation of abundant chloroplasts throughout the cytoplasm of N. pachyderma Type I

To our knowledge this is the first report of large numbers of chloroplasts observed by TEM imaging and recorded via metabarcoding in any planktonic foraminiferal species. The high numbers observed, and the relative abundance of diatom chloroplasts recorded, could indicate a kleptoplastic behaviour in N. pachyderma Type I, a strategy which is well known in several protist lineages such as benthic foraminifera, dinoflagellates, and ciliates. Evidence suggests that photosynthates provided by kleptoplast phototrophy have been important in supporting major evolutionary innovations crucial to the current ecological roles of such protists in the marine environment (Stoecker et al., 2009). Therefore, it is important to assess, and further investigate the role of the observed chloroplasts to understand any contribution they make to N. pachyderma’s evolution and success both now and in the future polar climate.

Foraminifera that eat diatoms (and other algae) such as N. pachyderma would be expected to contain some chloroplasts in their cytoplasm as a byproduct of their grazing. For example, 18S metabarcoding demonstrates that the non-kleptoplastic benthic foraminifer Ammonia sp. (Jauffrais et al., 2016), grazes on diatoms in a comparable way to the kleptoplastic Elphidium sp. and Haynesina germanica, (Chronopoulou et al., 2019), and chloroplasts are indeed observed within the cytoplasm of Ammonia sp. Yet the relative plastid abundance in Ammonia sp., is reported as “rare” compared to “abundant” in Elphidium sp. and Haynesina sp. (Goldstein et al., 2004; Cesbron et al., 2017; Jauffrais et al., 2018) with a high proportion of chloroplasts in Ammonia sp. undergoing degradation (Jauffrais et al. 2018; Lekieffre et al., 2018). In contrast, our TEM images show high numbers of chloroplasts in N. pachyderma, congruent with or greater than the abundance observed in the TEM images of the kleptoplastic foraminifera such as Elphidium sp. and H. germanica (Jauffrais et al., 2018; Fig. 7, Fig. A6).

Kleptoplasty is common amongst benthic foraminifera (Lopez et al 1979; Lee et al 1988; Cedhagen 1991; Tsuchiya et al., 2018; Jauffrais et al., 2018; Pinko et al., 2023). The molecular studies identifying the source of kleptoplasts in benthic foraminifera to date would suggest a diatom source from the family Thalassiosiraceae, but potentially, kleptoplasts from more than one closely related diatom species can be present (Lechliter 2014; Jauffrais et al., 2019a; Tsuchiya et al., 2020; Pinko et al., 2023). More than 20 diatom species have been identified in benthic foraminifera that host intact diatom symbionts, (Lee 1995; Schmidt et al., 2018), with potentially up to three different symbionts within a single foraminifer at the same time (Lee,
2011). In addition, diatom symbiont shuffling appears to be an adaptation to changing environmental conditions such as heat stress (Schmidt et al., 2018). These studies indicate that host-symbiont or host-kleptoplast relationships are not strictly species-specific, supporting our findings of multiple diatom ASVs.

The chloroplasts in *N. pachyderma* are distributed throughout the foraminiferal cytoplasm (Fig. 7d & 7e). In the benthic kleptoplasmic species, chloroplast location is specific to the foraminiferal host species, where some kleptoplasts may be associated with the cellular periphery, while others, as observed here, may be distributed throughout the cell cytosol. Chloroplast placement therefore cannot provide a clear-cut indicator of kleptoplasty (Jauffrais et al., 2018; Pinko et al., 2023).

Ordinarily chloroplasts perform many functions other than photosynthesis. These include amino acid, nucleotide, and fatty acid synthesis as well as N and S assimilation (Bobik and Burch-Smith, 2015). Further, the benthic foraminiferal species *Nonionellina labradorica* retains chloroplasts despite living in sediments below the photic zone. The photosynthetic pathway of their retained chloroplasts is therefore not functional (Jauffrais et al., 2019b) and the reason for chloroplast retention in this species is unknown. However, its importance is reflected by the discovery that the kleptoplast genome in *Nonionella stella*, another benthic species that lives below the photic zone, is transcribed in the host (Gomaa et al., 2021). The role of the retained chloroplasts remains a fascinating question in many species of foraminifera, including *N. pachyderma* and the biological advantages for the host and the impact on their shell geochemistry necessitates further investigation.

The presence of abundant chloroplasts in the cytoplasm of *N. pachyderma* may result from gorging on blooms of *Chaetoceros* spp. (Booth et al., 2002). Yet it is also quite possible that *N. pachyderma* Type I may adopt a hybrid kleptoplasmic lifestyle during the limited summer months, to utilise the carbon fixed by kleptoplasts alongside heterotrophic feeding on diatoms (Mitra et al., 2016), which may also have restricted availability due to plankton patchiness. A wide range of protists exhibit similar mixotrophy, including many planktonic foraminifera which house photosynthesising pelagophyte (Gastrich, 1987; Bird et al., 2018) or dinoflagellate endosymbionts from which they receive fixed carbon (LeKieffre et al., 2018). The presence of abundant chloroplasts in *N. pachyderma* Type I means that a similar mixotrophic life-strategy needs to seriously be considered and investigated. However, a study of six *N. pachyderma* Type VII individuals from the North Pacific using fast repetition rate (FRR) Fluorometry, found no photosynthetic potential in this genotype, or even evidence of non-functional chlorophyll (Takagi et al., 2019). This is surprising given the herbivorous nature of *N. pachyderma* (Spindler and Dieckmann, 1986; Schiebel and Hemleben, 2017; Greco et al., 2021; this study). *N. pachyderma* Type I should be tested using FRR fluorometry to identify whether retained kleptoplasts have photosynthetic potential and also whether it is an obligate or facultative activity.

If other genotypes of *N. pachyderma* have not adopted this habit, it could represent a divergent evolutionary adaptation in *N. pachyderma* Type I to survive and flourish in the extreme Arctic environment. *N. pachyderma* has genetically diversified to inhabit a wide range of extreme environments from the Arctic and Antarctic polar waters to the frontal and upwelling systems of the transitional to tropical zones (e.g. Darling et al., 2008). Type I *N. pachyderma* diverged from its Southern Ocean counterparts during the early Quaternary (Darling et al., 2004), allowing substantial time for distinct adaptations to develop in its North Atlantic and Arctic habitat. Since mixotrophy brings advantages for survival, this adaptation may be a significant factor in its resilience to the extremes of glacial and interglacial cyclicity and its establishment as a true northern polar species.
4.5 Chloroplast storage to facilitate overwintering and reproduction?

Actively photosynthesising kleptoplasts in benthic foraminifera can remain active from a few days to a few months before being digested (Grzymski et al., 2002; Jauffrais et al., 2018 and references therein). Chloroplasts also represent a rich source of amino acids, fatty acids, lipids, vitamin E, pro-vitamin A, lutein, Cu, Fe, Zn and Mn (Gedi et al., 2017). Functioning photosynthetic kleptoplasts and/or the chloroplasts themselves could potentially provide *N. pachyderma* Type I with a substantial additional energy resource in the challenging Arctic environment. Further, if chloroplasts can be retained in the cytoplasm over many months before consumption, they could provide a valuable source of nutrition for the overwintering population.

Ecological processes in the Arctic are largely governed by sea ice and light dynamics. There is a general perception of minimal biological activity in the Arctic marine surface layers during the Arctic winter, due to the low light intensity producing minimal photosynthetic activity. However, studies around Svalbard in January 2012-2015 revealed unexpectedly high biological activity in the Arctic winter, with high respiration rates per unit of biomass in the upper 100 m water column (Berge et al. 2015a, b; Falk-Petersen et al. 2015), and an earlier winter *Calanus* copepod (Arthropoda) presence than previously thought (Espinasse et al., 2022). In Baffin Bay, low but significant phytoplankton growth was also observed during winter under the sea ice at extremely low light levels (Randelhoff et al., 2020). Since *N. pachyderma* Type I are thought to feed on both POM (including Arthropoda) and live diatoms (Greco et al., 2021 and this study), such wintertime POM-producing biological activity combined with stored chloroplasts (whether photosynthesising or not) could provide significant nutritional resources for an overwintering population of foraminifera.

These factors potentially combine to provide *N. pachyderma* with a significant nutritional resource to survive over the winter months, but questions remain about its behaviour in the water column and the form in which it may overwinter. Sediment traps in the Irminger Sea indicate a very low-level population of overwintering *N. pachyderma* and their isotopic signature profiles imply that a dormant noncalciifying population of *N. pachyderma* may remain in the water column during winter (Jonkers et al., 2010). However, it is possible that the *N. pachyderma* population they detected may not fully represent the true winter population size, since sieve sizes of 150 µm would not retain smaller mature/immature *N. pachyderma* specimens. Potentially, *N. pachyderma* could also remain buoyant in the water column as non-reproducing immature cells, slowing down their cellular metabolism as largely quiescent cells during the most challenging winter months. In culture, several specimens of *N. pachyderma* Type I exhibited extended periods of dormancy or inactivity, followed by recovery (Westgard et al, 2023).

4.6 Paleoenvironments, and geochemical signatures

The biological adaptations and interactions of foraminifera have varying influences on the geochemistry of their shell, as photosymbionts are known to influence shell geochemistry (Spero et al., 1991; Bemis et al., 1998; 2002; Anand et al, 2003; Russell et al., 2004), and symbiont-host respiration and potentially respiration of endobiont bacteria may increase the use of metabolic C in their shells (Rink et al., 1998; Wolf-Gladrow et al., 1999; Hönisch et al., 2003; Eggins et al., 2004; Bird et al.,
To fully understand variations in the geochemical signatures of Arctic *N. pachyderma* shells through time in the sediment assemblage, we need to improve our understanding of the ecology and interactions between *N. pachyderma* and the intracellular microorganisms which it hosts. Interactions may exhibit ontogenetic or strong seasonal differences and may be facultative or obligate. Recent geochemical studies have used high resolution single-specimen and even single chamber analyses to investigate both the biological and seasonal influences on shell geochemistry throughout the lifetime of foraminifera (Spindler and Dieckmann, 1986; Takagi et al., 2015, 2016; Lougheed et al., 2018; Pracht et al., 2019; Metcalfe et al., 2019). Single shell analysis of δ¹⁸O isotope values has identified two distinct populations of morphologically identical *N. pachyderma* populations in the north Atlantic during the last deglacial period. Isotope values indicate a temperature difference of about 4°C, potentially due to a bi-modal seasonal population with peak abundances separated temporally in late spring/early summer and late summer (Brummer et al., 2020). Spatial difference in the assemblage water depth, driven by low salinity meltwater (Brummer et al., 2020) may also contribute towards these seasonal differences. Since potential kleptoplasty (this study) could occur seasonally, obligately or facultatively, in *N. pachyderma* Type I, further investigation into the role of retained chloroplasts is required, since photosynthesis is known to influence δ¹⁸O values (Spero & Lea 1993; Bemis et al., 1998).

Improving our understanding of the biology and ecology, including seasonal microbial interactions, of Arctic *N. pachyderma* is required to disentangle the palaeoproxies for this species which is so important in our understanding of the rapidly contracting Arctic biome.
5 Conclusions

6 Appendix A

Figure A1. Bray Curtis dissimilarity NMDS plot of water samples from different depths at three stations. Colours represent depths and shapes represent stations. Depth drives 52% of the variability (Pr = 0.001) compared to Station driving 17% of the variability (Pr = 0.001)
Figure A2. NMDS plot of water column (triangle) and foraminiferal samples (circle) from station 101. Depth is indicated by colour. Note that foraminifera are not defined by depth due to their collection by vertical net tow from 200 m - surface.
Figure A3. Differential abundance testing of ASVs between Provenances at station 101 using DESeq2. The Log2 fold change in ASVs is the log-ratio of the ASV means in the water column and foraminifera. ASVs with positive Log2 fold change are significantly more abundant in the water column assemblage and ASVs with negative values indicate ASVs that are significantly more abundant in the foraminiferal assemblages. The Genus, or the highest level of taxonomic assignment available for each ASV is given on the X-axis.
Figure A4. Bray Curtis dissimilarity NMDS ordination of foraminiferal samples across all stations. There is a degree of clustering of foraminifera by station (*Adonis*: F. model = 3.079, Pr(>F) = 0.001), with 49% of the variation in the foraminifera driven by station.
Figure A5. Box plots showing individual ASV abundance counts. Dots represent single ASV counts within a single sample, grouped according to phyla, and provenance. The median, and upper and lower quartiles are shown. Note the different scales on the y-axis.
**Figure A6.** TEM images of *N. pachyderma* specimens (a) BB1 (b) BB9B, (c) BB11, (d)-(f) BB8, (g)-(h) BB9C and (i)-(j) BB12. Scale bars are 1 µm ((a), (e), (f) and (g)) all others are 2 µm. TEM imaging shows that chloroplasts were observed in all fixed specimens.

**Table A1.** Major ASVs responsible for the significant compositional differences (log2 fold change > +/- 10), between the provenances (water column and the foraminiferal microbiome). Negative values indicate a reduction in the water column mean relative abundance of the ASV compared to the foraminiferal mean relative abundance of the ASV and positive values indicate an increase in ASV in the water column compared to the foraminifera.

<table>
<thead>
<tr>
<th>ASV</th>
<th>Log2fold</th>
<th>padj</th>
<th>Significance in:</th>
<th>Taxonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASV130</td>
<td>-10.64059789</td>
<td>6.57E-29</td>
<td>foraminifera</td>
<td><em>Streptococcus</em> sp.</td>
</tr>
<tr>
<td>ASV42</td>
<td>-10.6621809</td>
<td>4.32E-18</td>
<td>foraminifera</td>
<td><em>Vibrio</em> sp.</td>
</tr>
<tr>
<td>ASV10</td>
<td>-10.84377001</td>
<td>2.79E-38</td>
<td>foraminifera</td>
<td><em>Micrococcus</em> sp.</td>
</tr>
<tr>
<td>ASV107</td>
<td>-23.69049864</td>
<td>1.22E-15</td>
<td>foraminifera</td>
<td><em>Haemophilus</em> sp.</td>
</tr>
<tr>
<td>ASV39</td>
<td>-25.23598127</td>
<td>1.37E-21</td>
<td>foraminifera</td>
<td><em>Haemophilus</em> sp.</td>
</tr>
<tr>
<td>ASV11</td>
<td>-25.70969167</td>
<td>1.22E-41</td>
<td>foraminifera</td>
<td><em>Acinetobacter</em> sp.</td>
</tr>
<tr>
<td>ASV84</td>
<td>-26.68731833</td>
<td>1.14E-25</td>
<td>foraminifera</td>
<td>Bacillariophyceae (diatom class)</td>
</tr>
<tr>
<td>ASV47</td>
<td>-27.43574907</td>
<td>1.77E-20</td>
<td>foraminifera</td>
<td><em>Fragilariopsis</em> sp.</td>
</tr>
<tr>
<td>ASV18</td>
<td>-28.15660108</td>
<td>2.61E-25</td>
<td>foraminifera</td>
<td><em>Cylindrotheca</em> sp.</td>
</tr>
<tr>
<td>ASV6</td>
<td>-28.90488689</td>
<td>2.17E-48</td>
<td>foraminifera</td>
<td><em>Streptococcus</em> sp.</td>
</tr>
<tr>
<td>ASV45</td>
<td>11.46484725</td>
<td>2.82E-48</td>
<td>water column</td>
<td>SAR clade (Marine group B) (Order)</td>
</tr>
<tr>
<td>ASV95</td>
<td>10.69026983</td>
<td>1.26E-29</td>
<td>water column</td>
<td>Pseudohongiella</td>
</tr>
<tr>
<td>ASV71</td>
<td>10.46031032</td>
<td>7.44E-32</td>
<td>water column</td>
<td>Marinimicrobia (Phylum)</td>
</tr>
<tr>
<td>ASV17</td>
<td>10.42038191</td>
<td>4.02E-33</td>
<td>water column</td>
<td>Alphaproteobacteria Clade Ib (Order)</td>
</tr>
<tr>
<td>ASV28</td>
<td>10.34374957</td>
<td>1.44E-38</td>
<td>water column</td>
<td>Candidatus Aquiluna</td>
</tr>
<tr>
<td>ASV20</td>
<td>10.32459242</td>
<td>1.77E-31</td>
<td>water column</td>
<td>Marine Group III (Order)</td>
</tr>
<tr>
<td>ASV122</td>
<td>10.29927908</td>
<td>4.06E-43</td>
<td>water column</td>
<td>Marinimicrobia (Phylum)</td>
</tr>
<tr>
<td>ASV2</td>
<td>10.28800033</td>
<td>3.73E-35</td>
<td>water column</td>
<td>Marine Group II (Order)</td>
</tr>
<tr>
<td>ASV52</td>
<td>10.18517168</td>
<td>3.36E-38</td>
<td>water column</td>
<td>Candidatus Nitrosopumilus</td>
</tr>
<tr>
<td>ASV24</td>
<td>10.09864971</td>
<td>5.62E-19</td>
<td>water column</td>
<td>Candidatus Nitrospopelagicus</td>
</tr>
<tr>
<td>ASV90</td>
<td>10.00382159</td>
<td>8.50E-25</td>
<td>water column</td>
<td>Marine Group II (Order)</td>
</tr>
</tbody>
</table>
Table A2. Taxonomic designations of those ASVs that drive the significant compositional differences between provenances, with a higher relative abundance in the foraminiferal microbiome.

<table>
<thead>
<tr>
<th>Prokaryote Phylum</th>
<th>ASV ID</th>
<th>Taxonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteobacteria</td>
<td>ASV11</td>
<td>Acinetobacter</td>
</tr>
<tr>
<td></td>
<td>ASV107</td>
<td><em>Haemophilus</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV39</td>
<td><em>Haemophilus</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV42</td>
<td><em>Vibrio</em> spp.</td>
</tr>
<tr>
<td></td>
<td>ASV40</td>
<td><em>Vibrio</em> spp.</td>
</tr>
<tr>
<td></td>
<td>ASV116</td>
<td><em>Bradyrhizobium</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV119</td>
<td><em>Colwellia</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV48</td>
<td><em>Colwellia</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV91</td>
<td><em>Alteromonas</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV38</td>
<td><em>Moritella</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV74</td>
<td>Family Halieaceae</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>ASV106</td>
<td><em>Fluvicola</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV3</td>
<td><em>Crocinomix</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV79</td>
<td>Family Cryomorphaceae</td>
</tr>
<tr>
<td></td>
<td>ASV41</td>
<td>Family Cryomorphaceae</td>
</tr>
<tr>
<td></td>
<td>ASV67</td>
<td>Family Flavobacteriaceae</td>
</tr>
<tr>
<td></td>
<td>ASV19</td>
<td>Family Flavobacteriaceae</td>
</tr>
<tr>
<td></td>
<td>ASV29</td>
<td>Family Saprospiraceae</td>
</tr>
<tr>
<td></td>
<td>ASV8</td>
<td>Family Cyclobacteriaceae</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>ASV6</td>
<td><em>Streptococcus</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV130</td>
<td><em>Streptococcus</em> sp.</td>
</tr>
<tr>
<td></td>
<td>ASV110</td>
<td><em>Lactococcus</em> sp.</td>
</tr>
<tr>
<td>Planctomycetes</td>
<td>ASV101</td>
<td><em>Rubripirellula</em> sp.</td>
</tr>
<tr>
<td>Actinobacter</td>
<td>ASV10</td>
<td><em>Micrococcus</em> sp.</td>
</tr>
</tbody>
</table>

7 Data availability

All sequence data has been submitted to NCBI and is freely available. *N. pachyderma* 18S sequences can be found at NCBI, GenBank Accession numbers OR137988-OR138014. The metabarcoding dataset can be found in the NCBI Sequencing Read Archive under BioProject accession PRJNA984332.
8 Author Contribution

CB contributed to the conception and design of the work, gained funding to carry out the molecular work and wrote the manuscript. KD contributed significantly to the writing of the manuscript and conception of the work. AP made a substantial contribution to the conception of the work, the acquisition of funding for sample collection, collected samples in 2017, and contributed to the editing of the manuscript. RT collected samples in 2018 and contributed to the editing of the manuscript.

9 Competing interests

The authors declare that they have no conflict of interest.

10 Acknowledgements

The authors wish to acknowledge the work of Steve Mitchell at the University of Edinburgh TEM facility for substantial contributions to sample processing and TEM analysis.

Sample collection was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2016-05457) awarded to AP. Some of the data presented herein were collected by the Canadian research icebreaker CCGS Amundsen and made available by the Amundsen Science program, which was supported by the Canada Foundation for Innovation and Natural Sciences and Engineering Research Council of Canada. The views expressed in this publication do not necessarily represent the views of Amundsen Science or that of its partners.

Transmission Electron Microscopy costs were covered by the Marine Alliance for Science and Technology for Scotland (MASTS) small grant scheme awarded to CB, and molecular lab work was support by a Carnegie Research Incentive Grant awarded to CB.

11 References


of ecosystem processes during the Arctic polar night, Prog Oceanogr, 139, 258–271, 
https://doi.org/10.1016/j.pocean.2015.08.005, 2015b.


Bobik, K. and Burch-Smith, T. M.: Chloroplast signaling within, between and beyond cells, Front Plant Sci, 6, 781, 


