

---

Faculty of Social Sciences

Faculty Publications

---

Combining multi-marker metabarcoding and digital holography to describe eukaryotic plankton across the Newfoundland Shelf

MacNeil, L., Desai, D. K., Costa, M., & LaRoche, J.

2022

© 2022 Liam MacNeil et al. This is an open access article distributed under the terms of the Creative Commons Attribution License.

<http://creativecommons.org/licenses/by/4.0/>

This article was originally published at:  
<https://doi.org/10.1038/s41598-022-17313-w>

---

Citation for this paper:

MacNeil, L., Desai, D. K., Costa, M., & LaRoche, J. (2022). "Combining multi-marker metabarcoding and digital holography to describe eukaryotic plankton across the Newfoundland Shelf." *Scientific Reports*, 12(13078).  
<https://doi.org/10.1038/s41598-022-17313-w>



OPEN

## Combining multi-marker metabarcoding and digital holography to describe eukaryotic plankton across the Newfoundland Shelf

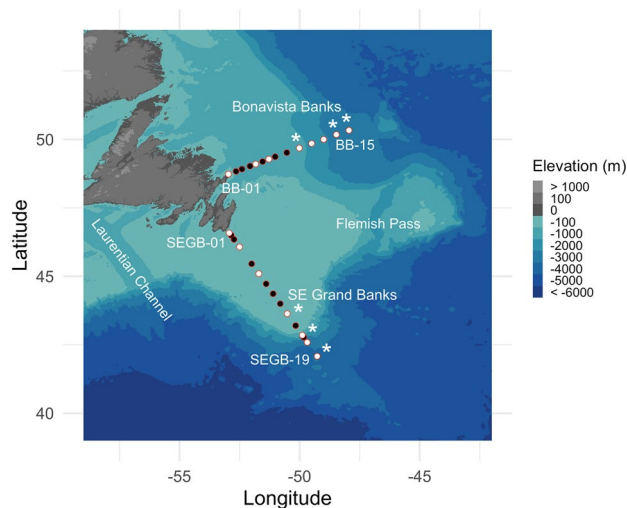
Liam MacNeil<sup>1,4</sup>✉, Dhvani K. Desai<sup>1,2</sup>, Maycira Costa<sup>1,3</sup> & Julie LaRoche<sup>1</sup>✉

The planktonic diversity throughout the oceans is vital to ecosystem functioning and linked to environmental change. Plankton monitoring tools have advanced considerably with high-throughput in-situ digital cameras and genomic sequencing, opening new challenges for high-frequency observations of community composition, structure, and species discovery. Here, we combine multi-marker metabarcoding based on nuclear 18S (V4) and plastidial 16S (V4–V5) rRNA gene amplicons with a digital in-line holographic microscope to provide a synoptic diversity survey of eukaryotic plankton along the Newfoundland Shelf (Canada) during the winter transition phase of the North Atlantic bloom phenomenon. Metabarcoding revealed a rich eukaryotic diversity unidentifiable in the imaging samples, confirming the presence of ecologically important saprophytic protists which were unclassifiable in matching images, and detecting important groups unobserved or taxonomically unresolved during similar sequencing campaigns in the Northwest Atlantic Ocean. In turn, imaging analysis provided quantitative observations of widely prevalent plankton from every trophic level. Despite contrasting plankton compositions portrayed by each sampling method, both capture broad spatial differences between the northern and southern sectors of the Newfoundland Shelf and suggest complementary estimations of important features in eukaryotic assemblages. Future tasks will involve standardizing digital imaging and metabarcoding for wider use and consistent, comparable ocean observations.

Eukaryotic plankton are enormously diverse and drift with surrounding water masses, also contributing to ecosystem function through primary and secondary productivity, mediating carbon sequestration along with other oceanic elements, and containing potentially harmful blooming species<sup>1</sup>. Plankton monitoring tools using conventional microscopy are steadily being replaced by a new generation of optics in digital imaging instruments, equipped with algorithms to extract, store, and classify major plankton groups mostly within nano-mesoplankton (2 µm to 20 mm) size ranges and bearing morphologically distinct features<sup>2,3</sup>. Imaging provides quantitative observations of taxonomic and functional diversity while also capturing species behaviour with unintrusive, in-flow sampling designs<sup>4</sup>. Together, the enhanced sampling capacity and stored digital information in image data have tremendous potential for quantitative plankton ecology, especially for morphologically diverse eukaryotes<sup>5,6</sup>.

Numerous imaging instruments and modes currently exist for plankton (e.g., see Table 1 in<sup>1</sup>). Among these, digital holographic imaging based on wavefront diffraction has grown across various disciplines, configured with or without optical lenses and supported by increasingly efficient numerical reconstruction algorithms applied to off-axis and in-line illumination paths for high-resolution amplitude and phase images<sup>7,8</sup>. The ordinary holographic principle to reconstruct the wavefront diffraction of an object from a coherent light source (i.e., a laser) can also be extended to record dimensions of color from incoherent light, including fluorescence<sup>8</sup>. Instrument designs without objective lenses or other opto-mechanical parts are well purposed to sample a wide range of

<sup>1</sup>Biology Department, Dalhousie University, 1355 Oxford St, Halifax, NS B3H 4J1, Canada. <sup>2</sup>Department of Biology and Pharmacology, Dalhousie University, 5850 College St, Halifax, NS B3H 4R2, Canada. <sup>3</sup>Department of Geography, University of Victoria, STN CSC, PO Box 1700, Victoria, BC V8W2Y2, Canada. <sup>4</sup>Present address: GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany. ✉email: L.macneil@dal.ca; Julie.Laroche@dal.ca



**Figure 1.** The Bonavista Banks (BB) and SE Grand Banks (SEGB) sampling transects. Stations (white and black) contain full CTD profiles, white dots indicate Niskin bottle samples for DNA and imaging at 5 m, and the asterisks indicate additional Niskin samples from 20 and 50 m. Note that the ship did not sample every original station, but original station names are kept here, thus SEGB-19 was the 15<sup>th</sup> station sampled at the SE Grand Banks. Maps were generated using the *marmap* R package (v. 1.0.6)<sup>82</sup>.

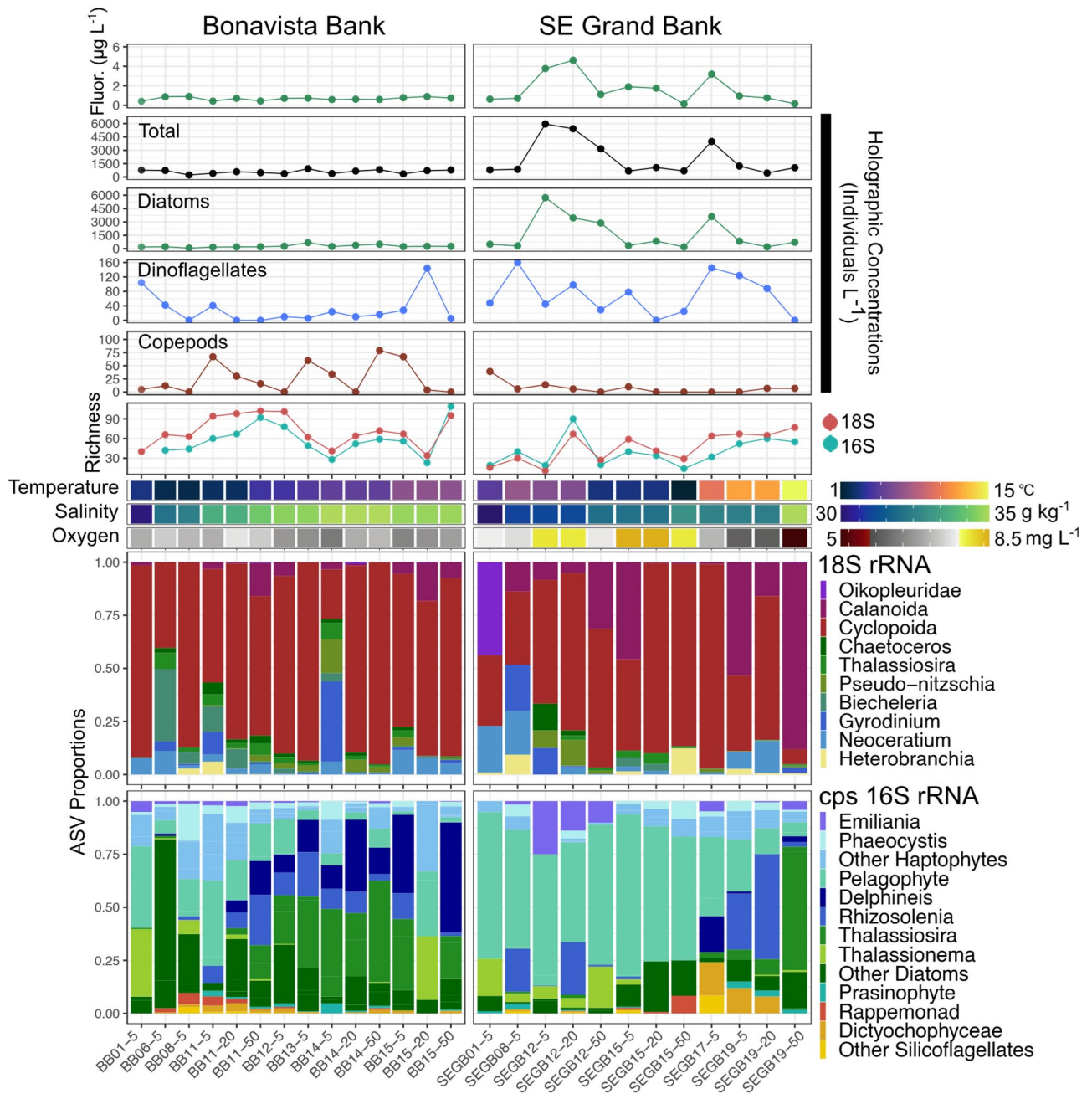
plankton in-situ where sufficient image resolution and throughput are available<sup>9–11</sup>. These requirements lend strongly to digital in-line holographic microscopes, which rely on numerical reconstruction of a wavefront from a coherent light source without an objective lens, effectively increasing the depth of field that can be acceptably focused without sacrificing image resolution<sup>12</sup>. Using a single coherent light source, inflow designs allow in-line holographic microscopes an enhanced depth-of-field to image larger volumes for high-throughput sampling with simple designs for deployment, while simultaneously recording pixel intensity, amplitude, and phase shift at micrometer scales<sup>13,14</sup>. Lensless digital in-line holographic microscopes can produce comparable image resolution to high-end conventional light microscopy with orders of magnitude greater field-of-view<sup>15</sup> and genus-level resolution is often achievable in plankton<sup>16–18</sup>. These traits make digital in-line holography especially applicable for sampling plankton and particulates in their environment<sup>19</sup>.

Any single imaging instrument has limited resolution and records a fraction of the plankton size spectrum, unlike DNA sequencing technologies which have revealed large portions of the oceans hidden, morphologically indistinct, and unculturable diversity<sup>20</sup>. The bulk pool of marine environmental DNA (eDNA) has recovered previously unexpected eukaryotic diversity<sup>21,22</sup>, global patterns of community structure<sup>23</sup>, and specific groups can be targeted using phylogenetic marker genes, known as metabarcoding<sup>24</sup>. Although no universal marker gene currently exists for hyperdiverse eukaryotic lineages, highly conserved markers include a nuclear target located on the small subunit of the 18S rRNA gene<sup>25</sup>, and a bacterial homologue to the cyanobacterial 16S rRNA gene present in the chloroplasts of photosynthetic protists<sup>26,27</sup>. Combining marker gene databases through multi-marker metabarcoding can reveal the often-unmanageable planktonic diversity present amongst the eukaryotes<sup>28–31</sup>. Due to the nature of the sequencing assay, sequence (read) counts do not reflect absolute abundances and must be treated as compositional<sup>32</sup>; reads are further decoupled from absolute abundances due to variable rRNA gene copy numbers, especially in the nuclear 18S rRNA gene, which can vary orders of magnitude in protists<sup>33–35</sup> even among closely related groups<sup>36</sup>. Therefore, despite extensive taxonomically resolved compositions produced by metabarcoding, these methods do not yet recover quantitative assessments or morphological information (e.g., sex, life-stage, behaviour) available in microscopy.

Here we combine digital in-line holographic microscopy (HoloSea S5) and multi-marker metabarcoding (18S V4 and 16S V4-V5) resolved to single nucleotide Amplicon Sequence Variants (ASVs)<sup>37</sup> with conductivity-temperature-depth (CTD) profiles including in-situ chlorophyll-a fluorescence to describe eukaryotic plankton composition > 10  $\mu\text{m}$  in the Newfoundland Shelf (Canada) surface waters (Fig. 1). This approach ensured identical volumes were imaged and sequenced, and together revealed a complex community of eukaryotic plankton supported by quantitative assessments of microplankton (> 20  $\mu\text{m}$ ) and mesoplankton (> 200  $\mu\text{m}$ ) from every trophic level. Plankton compositions displayed notable differences between transects and across the shelf-slope gradient. These results build on recent work characterizing seasonal planktonic diversity in the Northwest Atlantic Ocean, including observations of previously unobserved taxa that contribute significantly to spring blooms off-shelf, by providing taxonomically detailed compositions on this prominent coastal shelf during the 2019 early winter phase of the North Atlantic bloom phenomenon<sup>38</sup>.

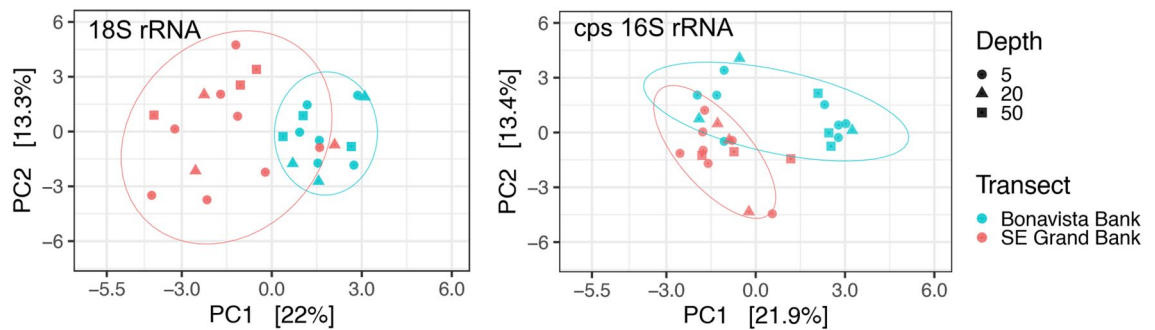
## Results

**Multi-marker sequencing characterizes eukaryotic diversity.** The combined set of CTD-derived physicochemical variables, quantitative imaging, and ASV richness with community composition from 18 and 16S rRNA genes is shown in Fig. 2. Combining 18S and 16S markers revealed rich eukaryotic compositions



**Figure 2.** The full set of observations for both transects including chlorophyll-a fluorescence (Fluor.), imaging concentrations, ASV richness, physicochemical data (temperature, salinity, oxygen) and the multi-marker taxonomic composition. Richness error bars represent standard errors under the breakaway model. The x-axis is oriented by station order from shore to shelf and depth (5–50 m). Comparing transects reveals a contrast between higher plankton richness and lower abundance on Bonavista Banks and vice versa on the SE Grand Banks. Physical gradients spanning on-shelf to off-shelf water masses are also evident in salinity on the Bonavista Banks and temperature on the SE Grand Banks; however the 18S and 16S taxonomic compositions across these gradients are more complex during the early winter sampling period.

on the Newfoundland Shelf. After strict quality filtering (See Methods), 18S sequences (> 332 K) were classified to 237 ASVs and 16S sequences (> 99 K) were classified to 293 ASVs. A full list of unique genera classified from both markers is shown in Fig. S6. Heterotrophs dominated 18S rRNA sequence proportions owing to the copepod order Cyclopoida, which contributed up to 86% on the Bonavista Banks (Fig. 2) but was unclassified to lower taxonomic ranks. Four copepod species were classified in the order Calanoida (*Calocalanus curtus*, *Centropages hamatus*, *Paracalanus parvus*, and *Temora longicornis*) throughout our samples at low proportions (< 1%), although reaching > 80% of total sequence proportions at 50 m off shelf (SEGB-19) at the SE Grand Banks. Animals were also detected belonging to the gastropods (*Heterobranchia*) and Appendicularia (*Oikopleuridae*), constituting a notable fraction of ASVs (> 40%) to the SE Grand Banks coastal station (SEGB-01). A sap-



**Figure 3.** Principal Component Analysis (PCA) of 18S and chloroplast (cps) 16S rRNA markers showing the first (PC1) and second (PC2) most important components (axes) with shapes indicating sample depth (m). The ordination is calculated with an Aitchison's distance<sup>40</sup> a compositionally valid Euclidean distance of the clr-transformed ASVs.

rophyte genus of Labyrinthulomycetes (*Aplanochytrium*) was detected in rare proportions (<1%) of 18S ASVs at off shelf stations on both Bonavista Banks and the SE Grand Banks. Phytoplankton usually composed <30% of 18S ASVs mostly belonging to four dinoflagellate genera: *Triplos* (*Neoceratium*), *Protoperidinium*, *Biechalaria*, and *Pelagodinium*. Diatoms contributed between <1 and 10%, consisting of ASVs classified as *Chaetoceros*, *Guinardia*, *Pseudo-nitzschia*, and *Thalassiosira*.

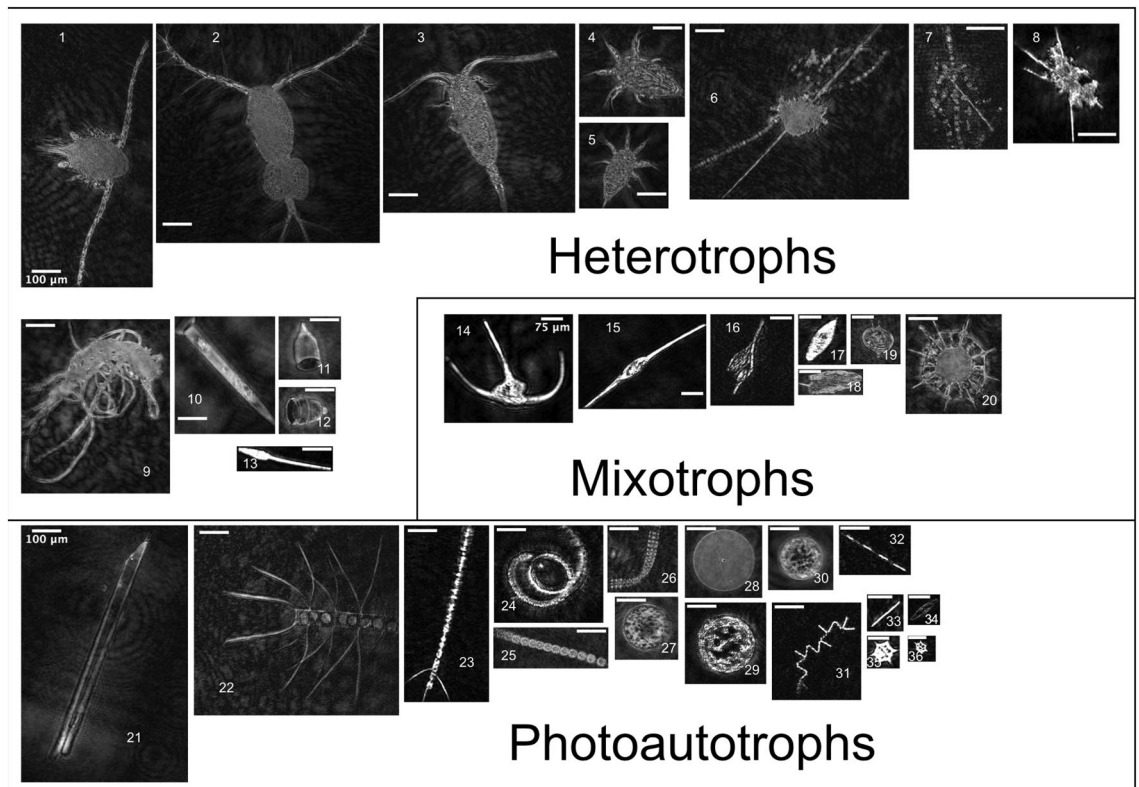
Despite numerous phytoplankton classified in 18S ASVs including dinoflagellates, which were undetected in our chloroplast 16S V4–V5 targets and in other V1–V2 hypervariable regions<sup>39</sup>, the 16S show phytoplankton diversity is drastically underestimated (Fig. 2). Picophytoplankton classified as pelagophytes were widely distributed despite our 10  $\mu\text{m}$  filter and were the main component (>60%) on the SE Grand Banks. The diatoms were more diverse and contributed similar proportions on the Bonavista Banks, mostly belonging to *Thalassiosira*, *Thalassionema*, *Rhizosolenia* and *Delphineis*, among other rarer genera. Haptophytes contributed up to 25% of compositions as *Phaeocystis* and *Emiliania*. Other rare phytoplankton >10  $\mu\text{m}$  were distributed across our samples including raphidophytes and silicoflagellates mostly belonging to the family Dictyochophyceae.

ASV richness from 18S and 16S markers were generally in step, although 18S amplifications were richer overall (Fig. 2). Principal component analysis indicated an apparent clustering of samples within transects regardless of depth or amplicon marker (Fig. 3), indicating samples reflect surface water conditions only and were likely sampled above the mixed layer depth (Fig. S1). However, these results were not statistically significant for 18S ( $p=0.165$ ) or 16S ( $p=0.114$ ) using a permutational analysis of variance. Samples that co-clustered with opposing transects were characterized by relatively lower diversity as seen in two 18S compositions on the Bonavista Banks (BB-01, BB-15 at 20 m) dominated by arthropods (>85%) (Fig. 2), and in several 16S samples where pelagophytes and diatoms contributed overwhelmingly (>90%) to compositions. Distinct 16S compositions on the Bonavista Banks are found at the Northeast Slope on the Bonavista Banks where *Thalassiosira*, *Delphineis*, and *Rhizosolenia* and other diatoms contribute significantly to overall compositions.

**Imaging composition and concentrations.** In total, >55 K holograms were analyzed and >105 K objects were detected. The bulk of objects were small (<50  $\mu\text{m}$ ) non-living particles but >4600 objects were taxonomically identified into 26 groups, ranging from phylum (e.g., Labyrinthulomycetes) to species-level (e.g., *Triplos fusus*), and large (>50  $\mu\text{m}$ ) non-living particulates of marine snow and other aggregates were also counted. A collage is shown in Fig. 4. Imaged phytoplankton were overwhelmingly diatoms, dominated by chain-forming, centric, and rod-shaped cells. Three chained diatom genera (*Chaetoceros*, *Pseudo-nitzschia*, and *Thalassionema*) and two rod-shaped genera (*Proboscia* and *Nitzschia*) were identified, all of which were detected in the paired DNA samples, however most (>80%) chained, centric, and rod-shaped diatoms lacked taxonomically diagnostic features and remained only morphologically identified. Dinoflagellates contained four physically defined genera (*Gyrodinium*, *Procerentrum*, *Protoperidium*, and *Triplos*) but only *Triplos fusus* and *Triplos lineatum* were identified at a species level from holograms and furthermore only *Triplos fusus* was detected as *Triplos fusus f. tenuis* in matching DNA samples; all other *Triplos* (*Neoceratium*) groups were only classified to genus-level in 18S ASVs.

The heterotrophic taxa included larvacean zooplankton (Appendicularia), amoebozoans, copepods, tintinnids, and fungus-like protists. Among these, larvae and amoeba were broadly categorized, and copepods were differentiated by life stage (Fig. 4). The tintinnids contained at least two identifiable genera: *Salpingella* and *Condonella* from the phylum Ciliophora. The fungus-like protists belonged to the Labyrinthulomycetes phyla and were initially grouped with non-living particulates and only identified in images after DNA sequencing revealed their presence in specific samples. The Labyrinthulomycetes are increasingly recognized as ecologically important saprophytes and are described further in the discussion.

**Eukaryotic plankton across the Newfoundland Shelf gradient.** The assemblages of eukaryotic plankton observed here in the early winter transition of the North Atlantic bloom phenomenon<sup>38</sup> displayed notable differences between north and south sections on the Newfoundland Shelf. Temperature and Salinity (T-S) transitioned at both transects from relatively cooler, fresher shelf water towards warmer, saltier water masses off shelf (Fig. 2; Fig. S1). ASV richness was greater on the Bonavista Banks than the SE Grand Banks,



**Figure 4.** A collection of eukaryotic taxa observed on the Newfoundland Shelf. Groups are broadly divided into trophic level with respective scale bars, labelled at image 1, 14, and 21. Heterotrophs included adult copepods (1–3), larval nauplii (4–5), the saprophytic phylum Labyrinthulomycetes (6–8), the Amoebozoa *Platamoeba* (9), tintinnids (10–12), and appendicularians (13). Mixotrophic dinoflagellates (14–19) included several genera of *Triplos* (14–16), with *Triplos fusus* (15) and *Triplos lineatum* (16), *Gyrodinium* (17), *Prorocentrum* (18), and *Protoperidinium* (19). The heterotrophic radiolarian Acantharia (20) commonly bears photosynthetic symbionts, creating mixotrophic nutrition<sup>41</sup>. The photosynthetic autotrophs included diatoms (21–35), and the silicoflagellate genera *Dictyocha* (35–36). The diatom genera included *Proboscia* (21), *Chaetoceros* (22–24), plus taxonomically unresolved chain-forming (25–26) and centric groups (27–30), *Thalassionema* (31), *Pseudo-nitzschia* (32), rod-shaped groups (33), and *Nitzschia* (34). To highlight additional data-driven benefits of holography, partially focused images (e.g., 21 and 23) produced from our methods can be refocused from raw holograms using, for example, an oblique reconstruction to recover obscured features outside the plane perpendicular to the optical axis<sup>42</sup>.

highest towards the Northeast Slope (BB-11, BB-12) and at 50 m off shelf (BB-15) into the increasing gradient of temperature and salinity (Fig. 2, Fig. S1). These samples were characterized by generally low plankton concentrations observed in holographic images indicating smaller nanoplankton contribute to this diversity where microplankton concentrations captured in our images are low (Fig. 2). Conversely, the SE Grand Banks displayed rich phytoplankton samples near the shelf break (SEGB-12) seen in 16S ASVs coincident with high fluorescence ( $> 4 \mu\text{g L}^{-1}$ ) and concentrated diatoms ( $> 4500 \text{ individuals L}^{-1}$ ) composed mostly of chains (Figs. 1, 3). Off-shelf of the SE Grand Banks (SEGB-17) a warm water mass contained a combined elevation of chlorophyll-*a* fluorescence with dinoflagellates ( $> 120 \text{ individuals L}^{-1}$ ) and diatom ( $> 3000 \text{ individuals L}^{-1}$ ) concentrations that steadily declined while richness increased towards the warmest off-shelf station (SEGB-19). Low autotrophic biomass on the Bonavista Banks was evident in both fluorescence and imaged plankton concentrations although no correlation was apparent; however imaged concentrations did co-vary with fluorescence on the SE Grand Banks ( $R^2 = 0.76$ ) (Fig. S7). Elevated plankton concentrations clustered on the shelf break of the SE Grand Banks containing centric ( $> 350 \text{ L}^{-1}$ ), chain-forming ( $> 650 \text{ L}^{-1}$ ) and rod-shaped ( $> 5000 \text{ L}^{-1}$ ) diatoms, cooccurring with relatively high concentrations of the dinoflagellate genera *Triplos* ( $> 80 \text{ L}^{-1}$ ).

## Discussion

The Newfoundland Shelf displayed a physical gradient transitioning across the continental shelf (Fig. 2; Fig. S3), capturing the temperate ocean water mass boundaries delimited by the SE Grand Banks and Bonavista Banks, with subpolar and subtropical water masses prevailing to the north and south, respectively<sup>43</sup>. This separation between shelf and open ocean waters is a ubiquitous feature of the Northwest Atlantic Ocean<sup>44</sup> and can create biologically productive boundaries where (sub)mesoscale dynamics actively modify photosynthetically available radiation, nutrients, and predator encounter frequency<sup>45</sup>. The work described here combines high-throughput

holographic imaging and multi-marker metabarcoding to characterize rich eukaryotic plankton communities (534 unique ASVs) across the Newfoundland Shelf gradient. By complementing imaging and multi-marker metabarcoding, broad quantitative and taxonomic differences were evident between north and south sections on the Newfoundland Shelf: Bonavista Banks contained low biomass, high richness samples characterized by relative higher proportions of diatoms and haptophytes; the SE Grand Banks revealed more complex eukaryotic compositions with notably high diatom concentrations near the continental margin, likely stimulated by upwelling of deep, cold, nutrient rich waters ( $> -50^{\circ}\text{W}$ ; Fig. S1) that frequently characterize this region<sup>46</sup>. Taken together, complementing imaging and multi-marker metabarcoding afforded broad comparisons between plankton abundance, diversity, and taxonomic composition against broadly different environmental profiles.

Seasonal planktonic diversity has recently been characterized in surface waters between  $37$  and  $66^{\circ}\text{N}$  off the Newfoundland Shelf using chloroplast 16S (V1–V2) amplicon sequencing with conventional and imaging flow cytometry<sup>39,47</sup> that provides important context to our findings. The off-shelf plankton compositions from the early winter transition phase (2015) were dominated by cyanobacteria and pico-phytoeukaryotes with large fractions of *Micromonas*, *Bathycoccus*, *Ostreococcus* (Clade I–II) and increasing proportions of pelagophytes northward<sup>47</sup>, for which we observe the opposite trend on the Newfoundland Shelf. The samples here omit most small eukaryotes due to  $> 10\ \mu\text{m}$  filtration, although the pelagophyte picophytoplankton occurring throughout both transects were likely trapped in filter pores clogged by larger cells and in particle-associated forms, instead these results highlight that the numerically rare nano-microeukaryotes are highly diverse, especially diatoms, and patchy across the Newfoundland Shelf. We also recovered multiple dinoflagellate genera through 18S ASVs and imaging samples which were undetected with chloroplast 16S targets<sup>39</sup> and taxonomically unresolvable from pigments<sup>48</sup>. Unravelling mixotrophic feeding strategies in dinoflagellates and other eukaryotic protists (e.g., radiolarians) to estimate feeding rates and biogeochemical fluxes will depend on adequate detection and behavioural observations<sup>49</sup> that are possible with our combined holographic imaging and multi-marker metabarcoding approach. It is also observed that the important chain-forming diatom *Chaetoceros*, which rapidly contributes to phytoplankton compositions upon the initiation of the subpolar spring bloom, is undetected off-shelf in winter-time and is likely transported within water masses or experiences higher success during the winter-to-spring transition<sup>47</sup>; our observations suggest *Chaetoceros* and other springtime blooming plankton could be advected across the Newfoundland shelf-slope boundary, implying a possible pathway within the Labrador current export along the Flemish Cap and SE Grand Banks<sup>50</sup>.

The skewing of planktonic diversity from underlying absolute abundances in compositional sequencing data is due to several technical and biological factors. Principally, there are potentially large differences in sample sources between quantitative imaging and metabarcoding: Images collected here are mostly intact living cells, but both cellular and extracellular DNA exist in the bulk DNA pools of seawater, thus it cannot be determined what fraction of a sequenced sample is living metabolizing cells, dead cells, dormant cysts, or detritus<sup>51,52</sup>. For 18S compositions, arthropods likely dominated due to their multicellularity, although size fractioning through a larger pore size, e.g.,  $> 250\text{--}500\ \mu\text{m}$  as per conventional net tows<sup>53</sup> would largely remove copepods to mitigate their overrepresentation during the amplification stage. Assigning ASVs as unique biological organisms is also limited by incomplete reference databases and biased for organisms of historical importance<sup>54</sup>. Work is still needed to understand biases in each marker gene target but extending to full-length gene sequencing platforms could superiorly resolve taxonomic composition and better account for intragenomic gene copies<sup>55–57</sup>. Altogether, this suggests that until quantitative community level sequencing is achieved<sup>58</sup> it remains a separate, but complementary line of evidence with quantitative imaging towards plankton community composition.

Despite its limitations, metabarcoding benefitted the overall community assessment with a much deeper taxonomic resolution of the eukaryotes than paired images. Only 12 genera were identified across all imaging samples, whereas 78 unique genera were identified in the metabarcoding analysis. This resulted in important microbial eukaryotic groups undetected in our images but classified in ASVs. These include protists belonging to the major phyla Cercozoa in 18S ASVs, and haptophytes in the 16S ASVs from widely abundant coccolithophorid genera *Emiliania* and non-coccolithophorid genera *Phaeocystis*, which forms the densest colonial blooms in the North Atlantic and Southern Oceans<sup>59</sup>. Each group was likely unidentifiable in our holographic images due to their small cell size ( $< 50\ \mu\text{m}$ ) which has previously been difficult for recovering sharp images using the HoloSea<sup>60</sup>, and due to morphologically indistinct features. Metabarcoding also provided potential and validated taxonomic identities of unidentified image objects. Many detected rod-shaped and centric diatoms with distinct features remained taxonomically unresolved, however the corresponding ASVs suggest they likely belong to genera *Guinardia*, *Actinocyclus*, and *Thalassiosira*. The fungus-like Labyrinthulomycetes are the clearest example for improved taxonomy as they were missed during manual identification but classified as the genera *Aplanochytrium* in paired DNA samples; their absence in paired images directed a re-assessment of detected objects which led to 10 identified specimens. The Labyrinthulomycetes identified here belong to the order Labyrinthulida<sup>61</sup>, a dominant marine group also detected in the photic zone during the Tara Oceans expedition<sup>22,61</sup>—although the Tara expedition samples largely omit the Northwest Atlantic Ocean. A new ecological view of the Labyrinthulomycetes is growing with observations showing these protists can exceed prokaryotic biomass and frequently cooccur with fungi to saprophytically break down neutrally buoyant marine snow in the bathypelagic zone, significantly contributing to the metabolism of deep-sea ecosystems<sup>62,63</sup>.

Quantitative imaging by digital in-line holography captured abundant micro and mesoplankton from every trophic level using the pipeline for hologram reconstruction and object detection outlined in<sup>60</sup>. The HoloSea returned highly focused images for quantitative observations of major plankton clades in aquatic ecology (diatoms, dinoflagellates, ciliates, copepods, radiolarians, silicoflagellates, other large protists) but other important groups (coccolithophorids, haptophytes, other small flagellates, fish larvae) are currently omitted due to size, pixel resolution, or morphologically indistinct features. Future studies investigating plankton size structure from the HoloSea or similar holographic microscopes should implement a correction factor to account for the scattering

of coherent light that biases size estimates non-linearly depending on object distance from the point source<sup>64</sup>. The promise for digital in-line holography is shared by digital imaging technologies more broadly, and is typically prohibitive for conventional light microscopy, which is to yield data-rich, high-frequency observations which to date have captured planktonic fish-larvae survivorship under river plume discharge regimes<sup>65</sup>, the evolution of particulate-mediated carbon export in rapidly changing marginal zones<sup>66</sup>, spatial patterns in plankton community compositions through time<sup>67</sup>, at regional scales<sup>4,68</sup> and specific taxonomic lineages at global scales<sup>69</sup>. Novel biotic interactions have recently been observed using in-situ imaging, including pseudopodial feeding in acantharians from the East China Sea likely missed previously due to destructive sampling by conventional plankton nets<sup>4</sup> and the frequent parasitization of the cosmopolitan copepod *Oithona* at the Scripps Pier in the Pacific Ocean<sup>70</sup>. Digital imaging can also capture size-structure patterns as an indicator of functional diversity<sup>71</sup>, which is a deeply evolutionary trait that affects ocean food webs and biogeochemistry<sup>72,73</sup>. The speciality of imagery for morphologically distinct eukaryotes is of particular importance because, unlike bacteria, morphology and behavior are central to eukaryotic diversity and ecology and metabarcoding is blind to these features<sup>74</sup>.

This work presents paired high-throughput imaging, metabarcoding, and physicochemical observations from the productive Newfoundland Shelf to show that the eukaryotic plankton fraction displays subregional variation across physical gradients, including patchy biogeography and biomass along the continental margin during the early winter bloom transition. As a geographic juncture between North Atlantic water masses and ongoing climate changes to the subarctic North Atlantic Gyre<sup>75,76</sup>, the Newfoundland-Labrador shelves deserve a more comprehensive study of planktonic diversity and variability using state-of-the-art molecular and optical tools to improve upon existing seasonal surveys relying on net tows, conventional microscopy, and bulk chlorophyll-a<sup>77</sup> or flow cytometry<sup>78</sup>. The high-throughput, durable, and lensless designs for digital holographic microscopes supported by increasingly automated pipelines will make suitable complements with high-sensitivity molecular and oceanographic sensor platforms and promote further *in-situ* deployments to investigate plankton diversity and structure.

## Methods

**Study area.** The Grand Banks of Newfoundland (Fig. 1) are a group of continental embankments where two major wind-driven Western Boundary Currents—the Arctic-sourced Labrador Current and the North Atlantic Current—converge in the Northwest Atlantic Ocean<sup>79</sup>. The Grand Banks represent the foremost component of the Newfoundland-Labrador Shelves that supports one of the most seasonally productive regions in the Northwest Atlantic Ocean<sup>80</sup>. The data collected here is from two transects taken on the Newfoundland Shelf during the Atlantic Zone Monitoring Program (AZMP) 2019 cruise (*RRS James Cook*; cruise *JC190*; November 18–December 7, 2019). The Bonavista Banks transect contains 14 stations across >400 km extending over the Northeast Slope subject to a majority of Arctic water influence. The Southeast (SE) Grand Banks transect includes 15 stations across >550 km extending onto the Southeast Shoal, a uniquely shallow (<90 m) ancient sandy plateau subject to intensive vertical mixing between the North Atlantic Current and the Labrador Current<sup>81</sup>.

**Oceanographic physicochemical data.** Full water column profiles were obtained by lowering a conductivity-temperature-depth (CTD) rosette equipped with a calibrated fluorometer and a dissolved oxygen sensor (Seabird SBE-9+) from surface waters to 10 m above the seafloor. All CTD files were provided by Fisheries and Oceans Canada and handled in R using the *oce* package (v. 1.4.0)<sup>83</sup>. The CTD files were filtered in a standard fashion: CTD casts were trimmed to remove upcast anomalies, interpolated using the default method at 1 m increments, smoothed with local (Boxcar) averaging and gridded into sections<sup>83</sup>. The chlorophyll-a fluorescence ( $\mu\text{g L}^{-1}$ ) corresponding to the Niskin samples (Table S1) was extracted and averaged within three depth bins (5–6, 20–21, 50–51 m) to approximate conditions at Niskin bottle sample depths.

**Sample collection: paired imaging and filtration.** Seawater was collected using 20 L Niskin bottles attached to the CTD rosette, emptied directly into 10 L cubitainers where a minimum of 2 L sample volume was allocated for paired imaging of micro and mesoplankton and filtration of eDNA. The full dataset includes 27 water samples from 15 stations (See Table S1) at 5 m, seven from the SE Grand Banks and eight from Bonavista Banks; three shelf break stations along each transect were also analyzed at 20 and 50 m. First, seawater was pumped at  $150 \text{ mL min}^{-1}$  using a peristaltic pump (Fisherbrand GP1000) into a digital in-line holographic microscope, the HoloSea S5 (Described below) at 10 frames  $\text{s}^{-1}$  and the seawater was collected in a container for further filtration for DNA extraction. The lower size limit of the object detection algorithm for the HoloSea S5 was set at  $20 \mu\text{m}$  as a conventional lower size threshold to capture microplankton and larger objects<sup>60</sup>. Directly after imaging, samples were pumped through sterile tubing (Masterflex) and filtered onto  $10 \mu\text{m}$  polycarbonate Isopore Membrane filters (Millipore, United States). Each filter was handled with ethanol-sterilized tweezers, promptly stored at  $-80 \text{ }^\circ\text{C}$ .

**Digital in-line holographic imaging.** The HoloSea ( $92 \times 351 \text{ mm}$ , 2.6 kg) is a digital in-line holographic microscope using a solid-state laser (386 nm) as a point source emitted through a  $0.5 \mu\text{m}$  pinhole located 54 mm distance from the monochrome complementary metal-oxide-semiconductor (CMOS) camera with  $7.4 \mu\text{m}$  per pixel in the resultant  $2048 \times 2048$  hologram. Further specifications and theoretical background for the HoloSea and 4-Deep software are described in<sup>60</sup>. To briefly detail the acquisition process of plankton objects, first raw holograms are reconstructed in 4-Deep Octopus software using a Helmholtz-Kirchhoff transformation and subsequently, using the 4-Deep Stingray software, biological Regions of Interest (ROIs) are detected using a global adaptive thresholding and filtered for in-focus objects in each reconstructed plane<sup>60</sup>. These in-focus objects compose our quantitative profiles which were manually identified to the lowest possible taxonomic rank (<sup>84</sup>Table S2).

Concentrations for each sample were calculated by dividing the total number of individuals by the total volume that was imaged, calculated by multiplying the number holograms by the effective volume per hologram (0.063 mL) of the HoloSea<sup>64</sup>.

**DNA extraction.** DNA was extracted from the 10 µm polycarbonate filters using the DNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions with modifications adapted to in-house reagents for the cell lysis procedure detailed in the Supplementary Material. The remaining steps to isolate and elute the DNA contents followed the manufacturer's protocol. The final DNA elution included 100 µL of elution buffer. The DNA concentrations and purity were measured with a NanoDrop 2000 (Thermo Scientific, United States). The final DNA aliquot (27 µL) was stored at – 80 °C until further analysis.

**Library preparation and illumina MiSeq sequencing.** PCR amplification, multiplex library preparation, quantification, and sequencing by Illumina MiSeq were performed at Integrated Microbiome Resources (IMR; Halifax, NS). Raw sequences can be found in under the NCBI Sequence Read Archive BioProject PRJNA803249. Samples were amplified using dual-indexing Illumina fusion primers targeting the V4 508 bp region in the 18S rRNA gene (E572F—CYGCGGTAATTCCAGCTC; E1009R—AYGGTATCTRATCCTT YG)<sup>85</sup> and the V4–V5 371 bp region of chloroplast 16S rRNA gene (515FB—GTGYCAGCMGCCGCGGTAA; 926R—CCGYCAATTYMTTTRAGTTT)<sup>86,87</sup>. Data processing followed the Microbiome Amplicon Sequencing Workflow<sup>88</sup> within the QIIME2 platform (v. 2020.8)<sup>89</sup>. Preprocessing sequence reads for quality control and generating stitched reads follow<sup>88</sup> and are detailed in the Supplementary Material. To derive single nucleotide resolution between high-quality reads and determine ASVs<sup>37</sup>, Deblur<sup>90</sup> was used to subtract erroneous reads based on an upper-bound error profile from the sequence-to-sequence Hamming distances of neighbouring reads. During this step, we also positively filtered for 18S sequences using the PR2 (v. 4.14) 18S rRNA database<sup>91</sup>. The sequences were trimmed at 350 bp for 18S, and 360 bp for 16S sequences to remove low-quality reads<sup>92</sup>.

**Bioinformatic analysis.** The ASVs were taxonomically classified using the multinomial naïve-Bayes QIIME2 q2-feature-classifier plugin<sup>93</sup>. The 18S sequence classifier was pre-trained in-house on the SILVA (v. 138.1) database<sup>94</sup>, and the chloroplast 16S sequences were classified in a two-step process: First sequences were run against SILVA (v. 138.1) database to extract chloroplast sequences to be classified from the PhytoREF database<sup>27</sup>. The classified ASVs were then filtered in two ways: (1) Rare ASVs were removed if they were less than 0.1% of mean sample depth in accordance with Illumina's estimation of bleed-through errors during sequencing, (2) all ASVs only classified to the kingdom level were excluded. The remaining ASVs were exported from QIIME artifacts into R (v. 3.6.2)<sup>95</sup> as a Phyloseq object (v. 1.3.6)<sup>96</sup> for further analysis.

Richness (alpha diversity) was estimated using ASVs as a surrogate in the R package breakaway (v. 4.7.3)<sup>97</sup>. The breakaway model attempts to adjust richness estimates for unobserved taxa and debias samples from differing sequencing depths<sup>98</sup>. To assess differences in community composition between samples (beta diversity), all ASVs were transformed using a centered log-ratio (*clr*) of a composition  $x = (x_1, \dots, x_i, \dots, x_D)$  defined as:

$$clr(x) = \left[ \ln \frac{x_1}{g(x)}, \dots, \ln \frac{x_i}{g(x)}, \dots, \ln \frac{x_D}{g(x)} \right] \quad (1)$$

where each ASV ( $x$ ) is divided by the geometric mean of all ASVs in a sample given by  $g(x) = \sqrt[D]{x_1 \dots x_D}$ <sup>99</sup>. To prevent undefined log-ratios, a pseudo-count of one was added to each zero<sup>100</sup>. To summarize compositional differences between transects a principal component analysis of the Euclidean (Aitchison) distances<sup>40</sup> was performed on log-transformed ASV counts<sup>37</sup> and tested for statistical significance using a permutational analysis of variance<sup>101</sup> in the vegan package (v. 2.5.7)<sup>102</sup>.

## Data availability

18S and 16S sequences are available under the NCBI Sequence Read Archive BioProject PRJNA803249. Additional plankton objects which are representative of major taxa found in this study are available (<https://doi.org/10.6084/m9.figshare.20259756>) and the imaging datasets generated during the current study are available from the corresponding author(s) upon reasonable request.

Received: 26 February 2022; Accepted: 25 July 2022

Published online: 29 July 2022

## References

- Lombard, F. *et al.* Consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.* **6**, 196. <https://doi.org/10.3389/fmars.2019.00196> (2019).
- Sieracki, M. E., *et al.* Optical plankton imaging and analysis systems for ocean observation. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*, 878–885 (2010). <https://doi.org/10.5270/OceanObs09.cwp.81>.
- Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L. & Stemmann, L. Machine learning for the study of plankton and marine snow from images. *Ann. Rev. Mar. Sci.* **14**(1), 277. <https://doi.org/10.1146/annurev-marine-041921-013023> (2022).
- Mars Brisbin, M., Brunner, O. D., Grossmann, M. M. & Mitarai, S. Paired high-throughput, in situ imaging and high-throughput sequencing illuminate acantharian abundance and vertical distribution. *Limnol. Oceanogr.* **65**(12), 2953–2965. <https://doi.org/10.1002/lno.11567> (2020).
- Benfield, M. *et al.* RAPID: Research on automated plankton identification. *Oceanography* **20**(2), 172–187. <https://doi.org/10.5670/oceanog.2007.63> (2007).
- Colin, S. *et al.* Quantitative 3D-imaging for cell biology and ecology of environmental microbial eukaryotes. *Elife* **6**, e26066. <https://doi.org/10.7554/eLife.26066> (2017).

7. Kim, M. K. Principles and techniques of digital holographic microscopy. *J. Photonics Energy*. **1**, 018005. <https://doi.org/10.1117/6.0000006> (2010).
8. Tahara, T., Quan, X., Otani, R., Takaki, Y. & Matoba, O. Digital holography and its multidimensional imaging applications: A review. *Microscopy* **67**(2), 55–67. <https://doi.org/10.1093/jmicro/dfy007> (2018).
9. Jericho, S. K., Garcia-Sucerquia, J. F. W., Jericho, M. H. & Kreuzer, H. J. Submersible digital in-line holographic microscope. *Rev. Sci. Instrum.* **77**(4), 043706. <https://doi.org/10.1063/1.2193827> (2006).
10. Bochdanský, A. B., Jericho, M. H. & Herndl, G. J. Development and deployment of a point-source digital inline holographic microscope for the study of plankton and particles to a depth of 6000 m. *Limnol. Oceanogr. Methods* **11**, 28–40 (2013).
11. Yourassowsky, C. & Dubois, F. High throughput holographic imaging-in-flow for the analysis of a wide plankton size range. *Opt. Express* **22**(6), 6661. <https://doi.org/10.1364/OE.22.006661> (2014).
12. Jericho, M. H. & Kreuzer, H. J. Point source digital in-line holographic microscopy. In *Coherent Light Microscopy* (eds Ferraro, P. et al.) 3–30 (Springer, 2011).
13. Kanka, M., Riesenberger, R. & Kreuzer, H. J. Reconstruction of high-resolution holographic microscopic images. *Opt. Lett.* **34**(8), 1162. <https://doi.org/10.1364/OL.34.001162> (2009).
14. Jericho, M. H., Kreuzer, H. J., Kanka, M. & Riesenberger, R. Quantitative phase and refractive index measurements with point-source digital in-line holographic microscopy. *Appl. Opt.* **51**(10), 1503. <https://doi.org/10.1364/AO.51.001503> (2012).
15. Wu, Y. & Ozcan, A. Lensless digital holographic microscopy and its applications in biomedicine and environmental monitoring. *Methods* **136**, 4–16 (2018).
16. Sun, H. et al. digital holography for studies of marine plankton. *Philos. Trans. R. Soc. A*. **366**, 1789–1806 (2008).
17. Bianco, V. et al. microplastic identification via holographic imaging and machine learning. *Adv. Intell. Syst.* **2**(2), 1900153. <https://doi.org/10.1002/aisy.201900153> (2020).
18. Guo, B. et al. Automated plankton classification from holographic imagery with deep convolutional neural networks. *Limnol. Oceanogr.* **19**(1), 21–36. <https://doi.org/10.1002/lom3.10402> (2021).
19. Nayak, A. R., Malkiel, E., McFarland, M. N., Twardowski, M. S. & Sullivan, J. M. A Review of holography in the aquatic sciences: In situ characterization of particles, plankton, and small scale biophysical interactions. *Front. Mar. Sci.* **7**, 572147. <https://doi.org/10.3389/fmars.2020.572147> (2021).
20. Di Bella, J. M., Bao, Y., Gloor, G. B., Burton, J. P. & Reid, G. High throughput sequencing methods and analysis for microbiome research. *J. Microbiol. Methods* **95**(3), 401–414. <https://doi.org/10.1016/j.mimet.2013.08.011> (2013).
21. Stoeck, T. et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**, 21–31. <https://doi.org/10.1111/j.1365-294X.2009.04480.x> (2010).
22. de Vargas, C. et al. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**(6237), 1261605–1261605. <https://doi.org/10.1126/science.1261605> (2015).
23. Lima-Mendez, G. et al. Determinants of community structure in the global plankton interactome. *Science* **348**(6237), 1262073–1262073. <https://doi.org/10.1126/science.1262073> (2015).
24. Santoferrara, L. et al. Perspectives from ten years of protist studies by high-throughput metabarcoding. *J. Eukaryot. Microbiol.* **67**(5), 612–622. <https://doi.org/10.1111/jeu.12813> (2020).
25. Eickbush, T. H. & Eickbush, D. G. Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* **175**(2), 477–485. <https://doi.org/10.1534/genetics.107.071399> (2007).
26. Kirkham, A. R. et al. Basin-scale distribution patterns of photosynthetic picoeukaryotes along an Atlantic Meridional Transect: Marine photosynthetic picoeukaryote community structure. *Environ. Microbiol.* **13**(4), 975–990. <https://doi.org/10.1111/j.1462-2920.2010.02403.x> (2011).
27. Decelle, J. et al. PhytoREF: A reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. *Mol. Ecol. Resour.* **15**(6), 1435–1445. <https://doi.org/10.1111/1755-0998.12401> (2015).
28. Leray, M. & Knowlton, N. Censusing marine eukaryotic diversity in the twenty-first century. *Phil. Trans. R. Soc. B*. **371**(1702), 20150331. <https://doi.org/10.1098/rstb.2015.0331> (2016).
29. Cowart, D. A. et al. Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS ONE* **10**(2), e0117562. <https://doi.org/10.1371/journal.pone.0117562> (2015).
30. Stefanni, S. et al. Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Sci. Rep.* **8**(1), 12085. <https://doi.org/10.1038/s41598-018-30157-7> (2018).
31. Pappalardo, P. et al. The role of taxonomic expertise in interpretation of metabarcoding studies. *ICES J. Mar. Sci.* <https://doi.org/10.1093/icesjms/fsab082> (2021).
32. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **8**, 2224. <https://doi.org/10.3389/fmicb.2017.02224> (2017).
33. Zhu, F., Massana, R., Not, F., Marie, D. & Vaulot, D. Mapping of picoeukaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* **52**(1), 79–92. <https://doi.org/10.1016/j.femsec.2004.10.006> (2005).
34. Sargent, E. C. et al. Evidence for polyploidy in the globally important diazotroph *Trichodesmium*. *FEMS Microbiol. Lett.* **363**(21), 244. <https://doi.org/10.1093/femsle/fnw244> (2016).
35. Gong, W. & Marchetti, A. Estimation of 18S gene copy number in marine eukaryotic plankton using a next-generation sequencing approach. *Front. Mar. Sci.* **6**, 219. <https://doi.org/10.3389/fmars.2019.00219> (2019).
36. Biard, T. et al. Biogeography and diversity of collodaria (radiolaria) in the global ocean. *ISME J.* **11**, 1331–1344 (2017).
37. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**(12), 2639–2643. <https://doi.org/10.1038/ismej.2017.119> (2017).
38. Behrenfeld, M. J. et al. The North Atlantic aerosol and marine ecosystem study (NAAMES): Science motive and mission overview. *Front. Mar. Sci.* **6**, 122. <https://doi.org/10.3389/fmars.2019.00122> (2019).
39. Bolaños, L. M. et al. Seasonality of the microbial community composition in the North Atlantic. *Front. Mar. Sci.* **8**, 624164. <https://doi.org/10.3389/fmars.2021.624164> (2021).
40. Aitchison, J. The statistical analysis of compositional data. *J. R. Stat. Soc. B* **44**(2), 139–160. <https://doi.org/10.1111/j.2517-6161.1982.tb01195.x> (1982).
41. Decelle, J. & Not, F. Acantharia. *ELS*, 1–10 (2015). <https://doi.org/10.1002/9780470015902.a0002102.pub2>.
42. Yu, L., An, Y. & Cai, L. Numerical reconstruction of digital holograms with variable viewing angles. *Opt. Express* **10**(22), 1250. <https://doi.org/10.1364/OE.10.001250> (2002).
43. Della Penna, A. & Gaube, P. Overview of (sub)mesoscale Ocean dynamics for the NAAMES field program. *Front. Mar. Sci.* **6**, 384. <https://doi.org/10.3389/fmars.2019.00384> (2019).
44. Sverdrup, H. U. *Oceanography for Meteorologists* (Prentice Hall, 1942).
45. Mahadevan, A. The impact of submesoscale physics on primary productivity of plankton. *Annu. Rev. Mar. Sci.* **8**(1), 161–184. <https://doi.org/10.1146/annurev-marine-010814-015912> (2016).
46. Fratantoni, P. S. & Pickart, R. S. The Western North Atlantic shelfbreak current system in summer. *J. Phys. Oceanogr.* **37**(10), 2509–2533. <https://doi.org/10.1175/JPO3123.1> (2007).
47. Bolaños, L. M. et al. Small phytoplankton dominate western North Atlantic biomass. *ISME J.* **14**(7), 1663–1674. <https://doi.org/10.1038/s41396-020-0636-0> (2020).

48. Kramer, S. J., Siegel, D. A. & Graff, J. R. Phytoplankton community composition determined from co-variability among phytoplankton pigments from the NAAMES field campaign. *Front. Mar. Sci.* **7**, 215. <https://doi.org/10.3389/fmars.2020.00215> (2020).
49. Faure, E. *et al.* Mixotrophic protists display contrasted biogeographies in the global ocean. *ISME J.* **13**(4), 1072–1083. <https://doi.org/10.1038/s41396-018-0340-5> (2019).
50. Fratantoni, P. S. & McCartney, M. S. Freshwater export from the Labrador current to the North Atlantic Current at the tail of the Grand Banks of Newfoundland. *Deep Sea Res. I.* **57**(2), 258–283. <https://doi.org/10.1016/j.dsr.2009.11.006> (2010).
51. Torti, A., Lever, M. A. & Jørgensen, B. B. Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Mar. Genom.* **24**, 185–196. <https://doi.org/10.1016/j.margen.2015.08.007> (2015).
52. Jian, C., Salonen, A. & Korpela, K. Commentary: How to count our microbes? The effect of different quantitative microbiome profiling approaches. *Front. Cell. Infect. Microbiol.* **11**, 627910. <https://doi.org/10.3389/fcimb.2021.627910> (2021).
53. Djurhuus, A. *et al.* Evaluation of marine zooplankton community structure through environmental DNA metabarcoding: Metabarcoding zooplankton from eDNA. *Limnol. Oceanogr. Methods* **16**(4), 209–221. <https://doi.org/10.1002/lom3.10237> (2018).
54. del Campo, J. *et al.* The others: Our biased perspective of eukaryotic genomes. *Trends Ecol. Evol.* **29**(5), 252–259. <https://doi.org/10.1016/j.tree.2014.03.006> (2014).
55. Karst, S. M. *et al.* Retrieval of a million high-quality, full-length microbial 16S and 18S rRNA gene sequences without primer bias. *Nat. Biotech.* **36**(2), 190–195. <https://doi.org/10.1038/nbt.4045> (2018).
56. Johnson, J. S. *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **10**(1), 5029. <https://doi.org/10.1038/s41467-019-13036-1> (2019).
57. Callahan, B. J. *et al.* High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Res.* **47**(18), e103–e103. <https://doi.org/10.1093/nar/gkz569> (2019).
58. Lin, Y., Gifford, S., Ducklow, H., Schofield, O. & Cassar, N. Towards quantitative microbiome community profiling using internal standards. *Appl. Environ. Microbiol.* **85**(5), 18. <https://doi.org/10.1128/AEM.02634-18> (2019).
59. Vogt, M. *et al.* Global marine plankton functional type biomass distributions: *Phaeocystis* spp. *Earth Syst. Sci. Data* **5**, 405–443. <https://doi.org/10.5194/essd-5-405-2012> (2012).
60. MacNeil, L., Missan, S., Luo, J., Trappenberg, T. & LaRoche, J. Plankton classification with high-throughput submersible holographic microscopy and transfer learning. *BMC Ecol. Evol.* **21**(1), 123. <https://doi.org/10.1186/s12862-021-01839-0> (2021).
61. Pan, J., del Campo, J. & Keeling, P. J. Reference tree and environmental sequence diversity of labyrinthulmycetes. *J. Eukary. Microbiol.* **64**(1), 88–96. <https://doi.org/10.1111/jeu.12342> (2017).
62. Bochdansky, A. B., Clouse, M. A. & Herndl, G. J. Eukaryotic microbes, principally fungi and labyrinthulmycetes, dominate biomass on bathypelagic marine snow. *ISME J.* **11**(2), 362–373. <https://doi.org/10.1038/ismej.2016.113> (2017).
63. Xie, N., Hunt, D. E., Johnson, Z. I., He, Y. & Wang, G. Annual partitioning patterns of Labyrinthulmycetes protists reveal their multifaceted role in marine microbial food webs. *Appl. Environ. Microbiol.* <https://doi.org/10.1128/AEM.01652-20> (2021).
64. Walcutt, N. L. *et al.* Assessment of holographic microscopy for quantifying marine particle size and concentration. *Limnol. Oceanogr. Methods* **3**, 10379. <https://doi.org/10.1002/lom3.10379> (2020).
65. Axler, K. *et al.* Fine-scale larval fish distributions and predator-prey dynamics in a coastal river-dominated ecosystem. *Mar. Ecol. Prog. Ser.* **650**, 37–61. <https://doi.org/10.3354/meps13397> (2020).
66. Trudnowska, E. *et al.* Marine snow morphology illuminates the evolution of phytoplankton blooms and determines their subsequent vertical export. *Nat. Commun.* **12**(1), 2816. <https://doi.org/10.1038/s41467-021-22994-4> (2021).
67. González, P. *et al.* Automatic plankton quantification using deep features. *J. Plankton Res.* **41**(4), 449–463. <https://doi.org/10.1093/plankt/fbz023> (2019).
68. Briseño-Avena, C. *et al.* Three-dimensional cross-shelf zooplankton distributions off the Central Oregon Coast during anomalous oceanographic conditions. *Prog. Oceanogr.* **188**, 102436. <https://doi.org/10.1016/j.pcean.2020.102436> (2020).
69. Biard, T. *et al.* In situ imaging reveals the biomass of giant protists in the global ocean. *Nature* **532**, 504–507 (2016).
70. Orenstein, E. C. *et al.* The scripps plankton camera system: A framework and platform for in situ microscopy. *Limnol. Oceanogr. Methods* **18**(11), 681–695. <https://doi.org/10.1002/lom3.10394> (2020).
71. Fowler, B. L. *et al.* Dynamics and functional diversity of the smallest phytoplankton on the Northeast US Shelf. *PNAS* **117**(22), 12215–12221. <https://doi.org/10.1073/pnas.1918439117> (2020).
72. Tréguer, P. *et al.* Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* **11**(1), 27–37. <https://doi.org/10.1038/s41561-017-0028-x> (2018).
73. Ryabov, A. *et al.* Shape matters: The relationship between cell geometry and diversity in phytoplankton. *Ecol. Lett.* **24**(4), 847–861. <https://doi.org/10.1111/ele.13680> (2021).
74. Keeling, P. J. & del Campo, J. marine protists are not just big bacteria. *Curr. Biol.* **27**(11), R541–R549. <https://doi.org/10.1016/j.cub.2017.03.075> (2017).
75. Sgubin, G., Swingedouw, D., Drijfhout, S., Mary, Y. & Bennabi, A. Abrupt cooling over the North Atlantic in modern climate models. *Nat. Commun.* **8**(1), 14375. <https://doi.org/10.1038/ncomms14375> (2017).
76. Desbruyères, D., Chafik, L. & Maze, G. A shift in the ocean circulation has warmed the subpolar North Atlantic Ocean since 2016. *Commun. Earth Environ.* **2**(1), 48. <https://doi.org/10.1038/s43247-021-00120-y> (2021).
77. Mitchell, M. R. *et al.* Atlantic zone monitoring program protocol. *Can. Tech. Rep. Hydrogr. Ocean Sci.* **223**, 1–23 (2002).
78. Li, W. K. W., Glen Harrison, W. & Head, E. J. H. Coherent assembly of phytoplankton communities in diverse temperate ocean ecosystems. *Proc. R. Soc. B.* **273**(1596), 1953–1960. <https://doi.org/10.1098/rspb.2006.3529> (2006).
79. Richardson, P. L. Florida current, gulf stream, and Labrador current. In *Encyclopedia of Ocean Sciences* (ed. Steele, J. H.) 1054–1064 (Academic Press, 2001). <https://doi.org/10.1006/rwos.2001.0357>.
80. Henson, S. A., Dunne, J. P. & Sarmiento, J. L. Decadal variability in North Atlantic phytoplankton blooms. *J. Geophys. Res.* **114**(C4), C04013. <https://doi.org/10.1029/2008JC005139> (2009).
81. Han, G., Lu, Z., Wang, Z., Helbig, J. & Chen, N. Seasonal variability of the Labrador current and shelf circulation off Newfoundland. *J. Geophys. Res.* **113**, 10. <https://doi.org/10.1029/2007JC004376> (2008).
82. Pante, E. & Simon-Bouhet, B. marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* **8**(9), e73051. <https://doi.org/10.1371/journal.pone.0073051> (2013).
83. Kelley, D. “The Océ Package” In *Oceanographic Analysis with R* 91–101 (Springer, 2018).
84. Tomas, C. R. *Identifying Marine Phytoplankton* (Academic Press Inc, 1997).
85. Comeau, A. M., Li, W. K. W., Tremblay, J.-E., Carmack, E. C. & Lovejoy, C. Arctic ocean microbial community structure before and after the 2007 record sea ice minimum. *PLoS ONE* **6**(11), e27492. <https://doi.org/10.1371/journal.pone.0027492> (2011).
86. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples: Primers for marine microbiome studies. *Environ. Microbiol.* **18**(5), 1403–1414. <https://doi.org/10.1111/1462-2920.13023> (2016).
87. Walters, W. *et al.* Improved bacterial 16S rRNA gene (V4 and V4–5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *MSystems* <https://doi.org/10.1128/mSystems.00009-15> (2016).
88. Comeau, A. M., Douglas, G. M. & Langille, M. G. I. Microbiome helper: A custom and streamlined workflow for microbiome research. *MSystems* **2**(1), e00127–e216. <https://doi.org/10.1128/mSystems.00127-16> (2017).
89. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotech.* **37**(8), 852–857. <https://doi.org/10.1038/s41587-019-0209-9> (2019).

90. Amir, A. *et al.* Deblur rapidly resolves single-nucleotide community sequence patterns. *MSystems* **2**(2), e00191–e216. <https://doi.org/10.1128/mSystems.00191-16> (2017).
91. Guillou, L. *et al.* The protist ribosomal reference database (PR2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**(D1), D597–D604. <https://doi.org/10.1093/nar/gks1160> (2013).
92. Mohsen, A., Park, J., Chen, Y.-A., Kawashima, H. & Mizuguchi, K. Impact of quality trimming on the efficiency of reads joining and diversity analysis of Illumina paired-end reads in the context of QIIME1 and QIIME2 microbiome analysis frameworks. *BMC Bioinform.* **20**(1), 581. <https://doi.org/10.1186/s12859-019-3187-5> (2019).
93. Bokulich, N. A. *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**(1), 90. <https://doi.org/10.1186/s40168-018-0470-z> (2018).
94. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**(D1), D590–D596. <https://doi.org/10.1093/nar/gks1219> (2012).
95. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2021). <https://www.R-project.org/>.
96. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**(4), e61217. <https://doi.org/10.1371/journal.pone.0061217> (2013).
97. Willis, A. & Bunge, J. Estimating diversity via frequency ratios: estimating diversity via ratios. *Biometrics* **71**(4), 1042–1049. <https://doi.org/10.1111/biom.12332> (2015).
98. Willis, A. D. Rarefaction, alpha diversity, and statistics. *Front. Microbiol.* **10**, 2407. <https://doi.org/10.3389/fmicb.2019.02407> (2019).
99. Quinn, T. P. *et al.* A field guide for the compositional analysis of any-omics data. *GigaScience* **8**(9), 107. <https://doi.org/10.1093/gigascience/giz107> (2019).
100. Silverman, J. D., Roche, K., Mukherjee, S. & David, L. A. Naught all zeros in sequence count data are the same. *Comput. Struct. Biotech. J.* **18**, 2789–2798. <https://doi.org/10.1016/j.csbj.2020.09.014> (2020).
101. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **26**, 32–46 (2001).
102. Oksanen, J., *et al.* *vegan: Community Ecology Package*. R package version 2.5-7 (2020). <https://CRAN.R-project.org/package=vegan>.

## Acknowledgements

We would like to thank Fisheries and Oceans Canada (St. John's, NL) and the Bedford Institute of Oceanography for allowing us to join their AZMP mission, aiding in sample collection, and sharing the relevant physicochemical data. We also thank Catherine Bannon for aiding sample collection. LM was partially funded by the Alexander Graham Bell Canada Graduate Scholarship (Master's) through the Natural Sciences and Engineering Research Council (NSERC) of Canada and the research was also supported by a grant to MC from Marine Environmental Observation, Prediction, and Response (MEOPAR) and by an NSERC Discovery Grant to JLR.

## Author contributions

L.M. and J.L.R. conceived the sampling design. L.M. collected samples; L.M. and D.K.D. performed metabarcoding analyses and bioinformatics; L.M. conducted morphological identification and image analyses with help from J.L.R. L.M., J.L.R., M.C. helped structure the study. All authors contributed to writing and reviewing the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-17313-w>.

**Correspondence** and requests for materials should be addressed to L.M. or J.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022, corrected publication 2022