Arctic Science

OPEN ACCESS | Research Article

The community of marine alveolate parasites in the Atlantic inflow to the Arctic Ocean is structured by season, depth, and water mass

Elianne Egge[a,b], **Daniel Vaulot**[b,c], **Aud Larsen**[d], **and Bente Edvardsen**[b]

[a]Norwegian Institute for Water Research (NIVA), Økernveien 94, NO-0579 Oslo, Norway; [b]University of Oslo, Department of Biosciences, Section for Aquatic Biology and Toxicology, P.O. Box 1066 Blindern, NO-0316 Oslo, Norway; [c]UMR7144, CNRS, Sorbonne Université, Station Biologique de Roscoff. Place Georges Teissier, FR29680 Roscoff, France; [d]NORCE Norwegian Research Centre AS, P.O. Box 7810, NO-5020, Bergen, Norway

Corresponding author: **Bente Edvardsen** (email: bente.edvardsen@ibv.uio.no)

Abstract

The marine alveolates (MALVs) are a highly diverse group of parasitic dinoflagellates, which may regulate populations of a wide range of hosts, including other dinoflagellates, copepods, and fish eggs. Knowledge on their distribution and ecological role is still limited, as they are difficult to study with morphological methods. In this work, we describe the taxonomic composition and seasonal and depth distribution of MALVs in the Arctic Ocean west and north of Svalbard, based on 18S V4 rRNA metabarcoding data from five cruises. We recovered amplicon sequence variants (ASVs) representing all major groups previously described from environmental sequencing studies (Dino-Groups I–V), with Dino-Groups I and II being the most diverse. The community was structured by season, depth, and water mass. In the epipelagic zone, the taxonomic composition varied strongly by season; however, there was also a difference between Arctic and Atlantic water masses in winter. The spring and summer epipelagic communities were characterized by a few dominating ASVs present in low proportions during winter and in mesopelagic summer samples, suggesting that they proliferate under certain conditions, e.g., when specific hosts are abundant. Mesopelagic samples were more similar across months, and may harbor parasites of deep-dwelling organisms, little affected by season.

Key words: Svalbard, Syndiniales, parasites, biodiversity, metabarcoding

Introduction

Parasitism is one of the most successful life strategies on earth, as evidenced by its independent evolution in most branches in the tree of life (Poulin and Morand 2000). In the marine environment, parasitism by unicellular eukaryotes (protists) may regulate population sizes of both primary producers and consumers, and thus impact the food web structure and the carbon cycle (Skovgaard 2014; Berdjeb et al. 2018). Consequently, new knowledge on the diversity and environmental distribution of marine protistan parasites ultimately has the potential to inform and improve models of trophic transfer and biogeochemical cycling (Anderson et al. 2024).

The enigmatic marine alveolates (commonly referred to as MALVs or 'Syndiniales') is a polyphyletic assemblage of parasitic dinoflagellates, which are considered to be the dominating group of parasites in marine microbial food webs (Bjorbækmo et al. 2019; Holt et al. 2023; Anderson et al. 2024). In environmental sequencing studies, MALVs are generally reported to have high abundance and richness in marine pelagic environments (e.g., Guillou et al. 2008; Massana et al. 2011; Koid et al. 2012; Vargas et al. 2015), including polar waters (López-García et al. 2001; Lovejoy et al. 2006; Cleary and Durbin 2016; Clarke et al. 2019). The existence of high genetic diversity within this group is mostly inferred from environmental sequencing, where a high number of unique rRNA SSU gene sequences are phylogenetically placed next to parasitic dinoflagellate species (Groisillier et al. 2006; Guillou et al. 2008; Holt et al. 2023). Such parasitic dinoflagellates were first described in the 1920s (Chatton 1920) and later in the 1960s (e.g., Cachon 1964). Among the described genera are *Amoebophrya* and *Euduboscquella*, which infect other dinoflagellates or ciliates (Coats 1999; Coats and Park 2002; Farhat et al. 2021; Yoo et al. 2024), *Hobagella* that infects ciliates (Yoo et al. 2024), *Syndinium*, known to cause mortality in several copepod species (Kimmerer and McKinnon 1990; Skovgaard et al. 2005), *Hematodinium*, which includes at least three species infecting a wide range of crustaceans, impacting fisheries and aquaculture (e.g., Davies et al. 2022), and *Ichthyodinium* that infects fish eggs.

Altogether, AlgaeBase lists 43 morphologically described species of MALVs (Guiry and Guiry 2024), but not all of

Table 1. Distribution of amplicon sequence variants between the Dino-Groups, in both size fractions.

Order	Both	Only pico	Only nano-micro	nASVs pico	nASVs nano-micro	nASVs total	PR^2 ref. sequences
Dino-Group-I	221	33	156	254 (14.6%)	377 (21.5%)	410 (19.1%)	230 (24.7%)
Dino-Group-II	1034	357	198	1391 (79.8%)	1232 (70.4%)	1589 (73.9%)	632 (68%)
Dino-Group-III	50	7	31	57 (3.3%)	81 (4.6%)	88 (4.1%)	28 (3%)
Dino-Group-IV	23	3	17	26 (1.5%)	40 (2.3%)	43 (2%)	13 (1.4%)
Dino-Group-V	13	1	1	14 (0.8%)	14 (0.8%)	15 (0.7%)	27 (2.9%)
Syndiniales_X	2	0	4	2 (0.1%)	6 (0.3%)	6 (0.3%)	NA
Sum	1343	401	407	1744	1750	2151	903

Note: "Both" = Number of amplicon sequence variants (ASVs) found in both size fractions. "Only pico" = Number of ASVs only found in the pico fraction (0.4–3 μm). "Only nano-micro" = Number of ASVs only found in the nano-micro fraction (3–200 μm). The taxonomic distribution of the PR^2 reference sequences is given for comparison. nASVs = number of ASVs. Syndiniales_X are unclassified marine alveolates sequences falling outside the known Dino-Groups.

them have a known reference sequence. Due to low visibility within the host and the small size of the free-living dinospores, culture-independent methods, including molecular tools, are often required to study the MALVs. Such techniques have identified host-specific infections in high-biomass blooms of dinoflagellates (Chambouvet et al. 2008). In addition to parasitism, symbiotic relationships between this group and other protists have been proposed (Worden et al. 2015). MALV sequences have also been obtained from single-cell isolates of diatoms (Sassenhagen et al. 2020), Cercozoa, and radiolarians (Bråte et al. 2012); however, it is not known whether these associations are parasitic or symbiotic.

The life-cycle of the described parasitic MALV species is typically an alternation between a biflagellated infectious spore stage and an intracellular stage, during which the infecting spores divide and grow to a multinucleated structure which can fill the entire host cell (Cachon 1964). For *Amoebophrya* the intracellular stage is called the trophont—the trophont will elongate to form the vermiform which may be larger than 50 μm, and which eventually ruptures the host membrane. Within a few hours, the vermiform disintegrates into the free-living dinospores (Chambouvet et al. 2008) which are usually 1–12 μm. Species of *Syndinium* have a similar cycle, but the vermiform stage has not been reported, instead the spores escape as free-swimming zoospores upon release from the host (Skovgaard et al. 2005). For some *Syndinium* species the spores may be up to 20 μm in diameter. Infection by MALVs almost always kills the host, except for some larger crustaceans (Davies et al. 2022).

The MALV group was long thought to be a monophyletic sister group of dinokaryotes (i.e., dinoflagellates with permanently condensed chromosomes, Orr et al. (2012)), and is often collectively referred to as "Syndiniales" (Holt et al. 2023). Based on the rRNA SSU gene, the MALV group is split into two main groups, known in the literature as MALV I and II or Dino-Group (DG) I and II (Groisillier et al. 2006; Guillou et al. 2008). Guillou et al. (2008) also established the smaller clades DG-III–V. Since then, the MALV has been shown to be polyphyletic, with DG-II and DG-IV forming a sister clade to the dinokaryotes and DG-I placed basal to these clades in the dinoflagellate phylogeny (Strassert et al. 2018). Recently, phylogenomics based on single-cell transcriptomics has shown that DG-I and DG-II/IV evolved independently, from two distinct free-living ancestors (Holt et al. 2023). The described genera in DG-I are *Ichthyodinium*, *Eudubosquella*, and *Hobagella* (Holt et al. 2023; Yoo et al. 2024). *Amoebophrya* and *Hematodinium*, which are both characterised by the lack of any trace of a plastid, are placed in DG-II/IV, along with *Syndinium* (*Amoebophrya* in DG-II and *Hematodinium* and *Syndinium* both in DG-IV). Holt et al. (2023) suggest retaining the name "Syndiniales" for the group DG-II/IV, and using the order "Ichthyodinida" for DG-I, echoing an earlier suggestion by Cavalier-Smith (2018). No described species have been assigned to DG-III and DG-V, and it is not yet clear whether these groups also belong to "Ichthyodinida". Based on SSU rDNA phylogenies, DG-I and II/IV are further divided into 8 and 57 subclades, respectively, but the relationship between these clades and the described genera is sometimes unclear; e.g., *Amoebophrya* sequences are placed in several subclades (Guillou et al. 2008). Recent whole genome-sequencing of *Amoebophrya* strains suggests that this genus contains cryptic species (Farhat et al. 2021). In the taxonomy of the Protist Ribosomal Reference database (PR^2, Guillou et al. 2013; Vaulot 2022), the different DGs are assigned at the "order" level, and the subclades are assigned at the "family" level. In the literature these subclades are numbered and referred to as e.g., DG I Clade 1, and we will hereafter refer to them as "clades". DG-II/IV is generally considered the most diverse since it has the highest number of clades, and also the highest fraction of environmental sequences assigned to it (Table 1; Groisillier et al. 2006; Guillou et al. 2008).

Communities of microbial parasites in general, and MALV in particular, in a given layer of the water column may be shaped by biological factors such as access to hosts, and physical factors such as mixing or stratification of the water column, and sinking when attached to hosts or particles (Anderson et al. 2024). Early meta-studies of environmental sequencing datasets suggested that MALV clade composition differs between the photic and aphotic zone (Guillou et al. 2008). A recent metabarcoding study of MALV composition in the water column in the Sargasso sea found depth-structuring which was repeated over several years, and in this temperate region the MALV community was more strongly structured by depth than season (Anderson et al. 2024). However, the seasonal variation of the pelagic MALV community composition is not well studied (Anderson and Harvey 2020),

especially in polar regions and at mesopelagic depths, where sampling during the winter is challenging due to darkness and rough weather. In the Arctic, seasonality in primary production, and thus in protist assemblages, is strong in the epipelagic waters due to the extreme variation in light regime (e.g., Marquardt et al. 2016). Furthermore, during the spring bloom, zooplankton such as copepods migrate to the surface to feed on the phytoplankton bloom (Hobbs et al. 2020). Thus, there is high seasonal variability in the assemblage of potential MALV hosts in this environment. The mesopelagic zone in the Arctic has been much less studied, and little knowledge is available on parasite–host interactions in the deep ocean. Potential hosts at these depths include migratory copepods, and deep-dwelling heterotrophic members of Cercozoa and Radiolaria, which are known to harbour a wide diversity of MALVs (Bråte et al. 2012). The mesopelagic zone is of particular interest regarding carbon sequestration, as it is an important processing zone for sedimenting euphotic production (Terrado et al. 2009). Microbially mediated processes in these deep waters are considered to be crucial for organic matter remineralization, and thus impact the oceanic carbon pump (Nagata et al. 2010). To understand the impact of parasitism by MALV on the marine food web and carbon cycling in the Arctic, knowledge of the dynamics and distribution with time and depth is necessary.

The present study was conducted in the northern Svalbard region of the Arctic Ocean which was sampled in 2014 during five cruises representing the full seasonal cycle, at three to four depths from the surface down to 1000 m. Seawater samples were size-fractionated into two to four size fractions between 0.4–200 μm (including pico-, nano-, and microplankton). The taxonomic composition of the protist community was determined from metabarcodes (ASVs) of the 18S rRNA gene V4 region. The full dataset is presented in Egge et al. (2021). In the present paper, we focus on the MALV community to shed light on the processes that drive its diversity and distribution at epi- and mesopelagic depths in the Arctic Ocean. Our analyses reveal different seasonal distributions in the epipelagic and mesopelagic zones, and different seasonal changes in the pico- and nano-micro fractions. During winter and early spring, we also observed differences in community composition in the epipelagic zone between stations influenced by water flowing south from the Arctic basin, and stations dominated by deep convection of Atlantic water. In spring (May), the photic zone was dominated by a few clades and ASVs that were almost absent in winter and below the photic zone, whereas the aphotic zone community in spring and summer resembled the winter community. We found evidence for a unique MALV community at 1000 m depth, indicated by a high number of ASVs not recovered from other samples and low seasonal variation. In terms of clade composition, our data are similar to studies from lower latitudes, with high abundance of DG-I-Clade 1 and DG-I-Clade 5 in the photic zone under sunlit conditions, and DG-II-Clade 6 and 7 at mesopelagic depths, suggesting a wide distribution and high degree of adaptation within these clades. Finally, we identified ASVs which are abundant in the surface samples during the spring bloom and which could potentially represent parasites of key species in the Arctic pelagic food web.

Materials and methods

A detailed description of sampling, collection of environmental data, molecular lab work and bioinformatic processing can be found in the data paper from the MicroPolar project (Egge et al. 2021) (see also Paulsen et al. 2016; Wilson et al. 2017; Sandaa et al. 2018). The methods are briefly repeated here. Data from Egge et al. (2021), including sampling dates, station coordinates and sampled depths, can be found at https://doi.org/10.17882/79823. Interactive figures of environmental factors and the taxonomic composition of the metabarcoding data can be accessed via an online Shiny app (https://micropolar-protists.metapr2.org/micropolar-protists/).

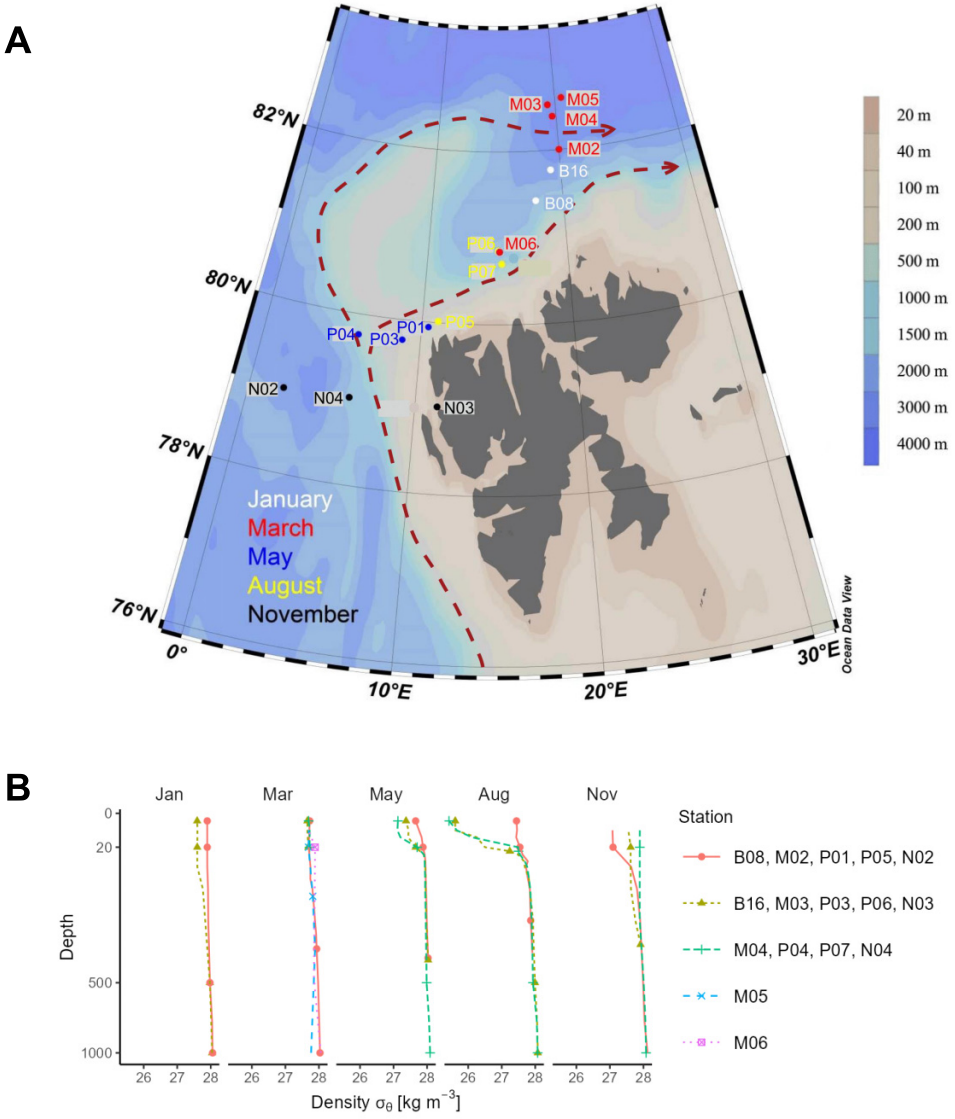
Sampling

Water samples were collected during five cruises west and north of the Svalbard archipelago in the Atlantic and Arctic Oceans in 2014 as part of the MicroPolar project (Paulsen et al. 2016; Wilson et al. 2017; Sandaa et al. 2018). During January (6 January to 15 January), March (5 March to 10 March), and August (7 August 18 August) cruises, samples were collected from the southern branch of the Western Svalbard Current, which transports water into the Arctic Ocean. During May (15 May to 2 June), August (7 August to 18 August), and November (3 November to 10 November) transects across the core of the Atlantic water inflow were made between 79°N and 79.4°N (Fig. 1A). The sampling area and locations were largely determined by the sea ice cover. During each cruise, 3–6 stations were sampled along a transect at four depths: in the epipelagic zone at 1 m and at a depth between 15 and 25 m (for the spring and summer samples, this is where the deep chlorophyll maximum, DCM, was found), and in the mesopelagic zone at two depths, in general 500 and 1000 m, or 10 m above the bottom at shallower stations.

Sampling preparation for DNA extraction

Fifty liters of seawater were sampled from each station and depth. During the January and March cruises, the samples were pre-filtered through a 180 μm mesh, and size fractionated into the 0.4–3 μm (picoplankton) and 3–180 μm (nano- and micro-plankton) fractions by peristaltic pumping (Masterflex 07523-80, Cole Parmer, IL, USA), through serially connected 3 and 0.4 μm polycarbonate filters (142 mm diameter, Millipore), mounted in stainless-steel tripods (Millipore, Billerica, MA, USA). The filters were placed in cryovials with AP1 DNA preservation buffer (DNeasy Plant mini kit, Qiagen, Hilden, Germany), flash frozen in liquid N_2 and kept at −80 °C until DNA extraction. During the May, August, and November cruises, the water was sequentially poured through 200, 50, and 10 μm nylon mesh, the material on each nylon mesh was collected with sterile filtered seawater into a 50 mL Falcon tube, and collected by filtration on a polycarbonate filter (10 μm pore size, 47 mm diameter, Millipore, USA). The plankton smaller than 10 μm passing through the

Fig. 1. Map of the sampling area and water column density profiles at each sampling. (A) Map, sampling stations are indicated with dots, colored according to sampling month, and labelled with names. The arrows show the trajectory of the West Spitsbergen Current, which splits into two branches near the Yermak Plateau. Bottom depth is indicated by the color scale. The map is adapted from Egge et al. (2021). (B) Density profiles of the water column at each station. January: B08 and B16, March: M02, M03, M04, M05 and M06, May: P01, P03 and P04, August: P05, P06, P07, November: N02, N03, N04.



nylon mesh system was fractionated into the 3–10 μm (small nanoplankton) and 0.4–3 μm size fractions by serial filtration through 142 mm diameter polycarbonate filters. The filters were preserved as described above. On May, August, and November, a higher number of size fractions were included to ensure that the full volume of water could be filtered without clogging.

DNA extraction, PCR, and sequencing

DNA was extracted with the DNeasy Plant mini kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol, except that frozen samples were incubated at 95 °C for 15 min, and then shaken in a bead-beater 2 × 45–60 s to rupture cell coverings. The V4 region of the 18S rRNA gene was amplified with primers TAReuk454FWD1 (5′-CCAGCASCYGCGGTAATTCC-3′) and V4 18S Next.Rev (5′-ACTTTCGTTCTTGATYRATGA-3′) (Piredda et al. 2017). The amplification primers were tagged with internal barcodes to allow sample multiplexing. PCR reactions were performed with KAPA HiFi HotStart ReadyMix 2x (KAPA Biosystems, Wilmington, MA, USA), in triplicate for each sample, and pooled prior to purification and quantification. The PCR products were purified with AMPure XP beads (Beckman Coulter, Brea, USA), quantified with NanoDrop and subsequently pooled in equal concentrations. The pools were sent to library preparation at the Norwegian Sequencing Centre (NSC, Oslo, Norway) and GATC GmbH (Konstanz, Germany) with the KAPA library amplification kit (KAPA Biosystems, Wilmington MA, USA). Due to delivery problems with the Illumina MiSeq chemistry in the spring of 2015, the

sequencing was done with a modified HiSeq protocol on two HiSeq runs at GATC GmbH. After initial analysis of the HiSeq data, samples with low number of reads were re-sequenced with Illumina MiSeq v.3 (300 bp, paired-end) at the NSC.

Sequence processing

Reads were processed into ASVs and chimeric ASVs were removed using the R library dada2, v1.16. (Callahan et al. 2016), as described in detail in Egge et al. (2021). The HiSeq and MiSeq runs were processed with dada2 separately. The ASVs were taxonomically assigned with function *assignTaxonomy*, the dada2 implementation of the naive Bayesian classifier method (Wang et al. 2007), against the PR^2 database (Guillou et al. 2013; Vaulot 2022), version 4.12.0 (https://github.com/pr2database/pr2database/releases/tag/v4.12.0). ASVs with less than 90% bootstrap value at class level and/or which comprised less than 10 reads in total were removed. In PR^2, sequences which are assigned to a taxonomic group at a given level, but which cannot be assigned to lower taxonomic levels are denoted with "_X", e.g., DG-I_X, and we will follow this notation. In the current version of the PR^2, "Syndiniales" is still used as a synonym for the entire MALV group, and thus reference sequences in PR^2 which could not be assigned to any of the Dino Groups are assigned to "Syndiniales_X". We therefore use this group name in the results.

Statistical analyses

All statistical analyses were performed in R (version 3.3.1, http://r-project.org), with the packages phyloseq, v. 1.38.0 (McMurdie and Holmes 2013), microbiome, v. 1.12.0 (Lahti and Shetty 2019) vegan 2.5-7 (Oksanen et al. 2021), and mvabund 4.1.12 (Wang et al. 2022), unless otherwise stated. Figures were created with the R-packages ggplot2 (Wickham 2016) and ragg (Pedersen and Shemanarev 2022).

Preparation of MALV ASV tables

For the picoplankton size fraction (0.4–3 μm), MALV ASVs were extracted from the 0.4–3 μm samples in the original ASV table (i.e., not rarefied, entries given as read counts). To be able to compare the May, August, and November samples between 3 and 200 μm with the January and March 3–180 μm samples, we merged these into the size fraction 3–200 μm, in the following referred to as "nano-micro". To ensure proportionate contribution of reads from each size fraction, the total protist ASV table was subsampled to a number of reads corresponding to the sample with fewest reads, as follows: 88 000 reads from 3–180 μm; 40 000 reads from each of 3–10 and 10–50 μm, and 8000 reads from 50 to 200 μm. Subsampling to equal read number was performed 100 times, and the average read number per ASV was used, rounded to 0 decimals. Subsampling was done with the function *rrarefy* from the vegan package. The low number of protist reads in the 50–200 μm fraction was due to a high proportion of Metazoan reads in this fraction. MALV ASVs were then extracted from these rarefied samples, and the fractions 3–10, 10–50, and 50–200 μm were merged into the fraction 3–200 μm by taking the sum of the read number for each ASV. Since metabarcoding data are inherently compositional, we use the term "abundance" to refer to the relative abundance of a taxon, in order not to overcharge the text.

Alpha diversity

For the alpha diversity analyses of the MALV community (richness, evenness, and Shannon diversity), all samples in the pico- and nano-micro size fractions were subsampled to 5000 MALV reads, with the *rrarefy*-function, as described above. ASV richness and Shannon diversity were calculated with the functions *specnumber* and *diversity*, respectively. As recommended by Borcard et al. (2018), Shannon diversity is given as the Hill number N_1 (=*exp(H)*, where H is that Shannon entropy), expressed as an ASV number equivalent. N_1 can then be interpreted as the number of ASVs needed to obtain the observed Shannon entropy, if all ASVs had equal proportion. Evenness is defined as N_1/N_0, where N_0 is the number of observed ASVs.

Beta diversity

For the beta diversity analyses (i.e., comparisons of community composition), we used the un-rarefied MALV ASV tables, following recommendations by McMurdie and Holmes (2014). Amplicon sequencing data are inherently compositional, i.e., the proportions of the different ASVs are not independent of each other (Gloor et al. 2017). This may lead to spurious correlations between ASVs, and incorrect estimates of similarity between samples (Jackson 1997). According to the recommendations of Gloor et al. (2017), the MALV read counts were therefore transformed to centred log-ratio values (CLR), with the function *transform* from the microbiome package. Similar to the original proportion, the CLR value indicates how dominating an ASV is in a particular sample. In addition, CLR values for a given ASV can be compared between samples (Gloor et al. 2017). Similarity in CLR-transformed ASVcomposition between samples was estimated by the Aitchison distance, calculated with the function *vegdist* in package vegan. Principal component analysis was then performed with the function *prcomp*. The samples were clustered according to Aitchison distance with the function *hclust* with complete linkage clustering. Clusters of interest were then delineated after visual inspection of the dendrogram, to identify clusters that corresponded to certain environmental conditions, e.g., combinations of season and depth. Clade composition was visualised as bar charts and the CLR of dominating ASVs were visualised as heatmaps. Plots of shared ASVs between sample clusters were created with the package UpSetR v. 1.4.0 (Conway et al. 2017). To assess which ASVs were significantly differently distributed between these clusters, we used the function *anova.manylm* from the mvabund package. This function fits multivariate linear models to the CLR-transformed abundance table and tests whether the ASVs have significantly different CLR-values between groups of samples (in this case clusters). Homogeneity of variances was checked by plotting residuals against the fitted values. Due to some random variation inherent to the algorithm, *anova.manylm* was performed 10 times, and ASVs that had p-value < 0.05

in 5 or more of the trials were considered to have significant differential abundance between the clusters.

Community composition in relation to environmental factors

The PCA and clustering results indicated differing seasonal variation between the communities in the epi- and mesopelagic zones, and between the two size fractions. We analysed the relationship between community composition and environmental factors in two ways: with variation partitioning using the function *varpart* in vegan, and by testing the correlation between environmental variables and the first two principal component axis scores (PC1 and PC2). As microbial communities in the epipelagic zone at higher latitudes are known to be strongly seasonal, we included 'season', expressed as $\sin(2*pi*d/365) + \cos(2*pi*d/365)$, where d is the julian day of the year, as a covariate (Grover and Chrzanowski 2006). For each size fraction, variation partitioning was performed both on all samples, and on the epi- and mesopelagic samples separately. The significance of the terms was tested with Analysis of variance (anova) of the corresponding redundancy analysis (rda) object. The correlations between individual ASVs and environmental factors were tested by calculating the Spearman's correlation between CLR values and the environmental factors.

Correlation between MALV and host group communities

Correlations between the community composition of MALVs and the putative host groups Dinophyceae + Ciliophora and Radiolaria were assessed by performing mantel tests of the distance matrices. Dinophycaee and Ciliphora were considered together as they are both members of the division Alveloata. As most of these taxa are >3 μm, distance matrices of these groups were based on the 3–200 μm size fraction. The distances were calculated the same way as for the MALV (described above).

Phylogenetic analyses

To assess whether closely related ASVs had similar distributions by season and depth, we performed phylogenetic analyses of the 200 ASVs with highest relative abundance in the total dataset. The ASVs were aligned with MAFFT v. 7, with the G-INS-1 option (mafft.cbrc.jp; Katoh et al. 2019) and trimmed with trimAl (Capella-Gutiérrez et al. 2009). The phylogenetic tree was created with RAxML (Stamatakis 2014) with the GTR-CAT model, based on 340 sites and 100 times bootstrap resampling. The tree was imported into R using functions from the treeio package (Wang et al. 2020), and visualised with functions from the package phyloseq. To simplify visualisation, we used clades DG-I-Clade-1 and DG-II-Clade-6 as examples.

Biogeographic distribution of ASVs and clades

The biogeographic distribution of particular ASVs and clades of interest was investigated using the metaPR^2 online shiny application (Vaulot et al. 2022; https://shiny.metapr2.org, version 1.0). For each ASV, the sequence was used as query in a BLAST-like search, and the biogeographic distribution was taken as the combined distribution of the metaPR^2 ASVs with 100% identity to the queried ASV. Similarly, clade distribution was assessed by using the taxon search.

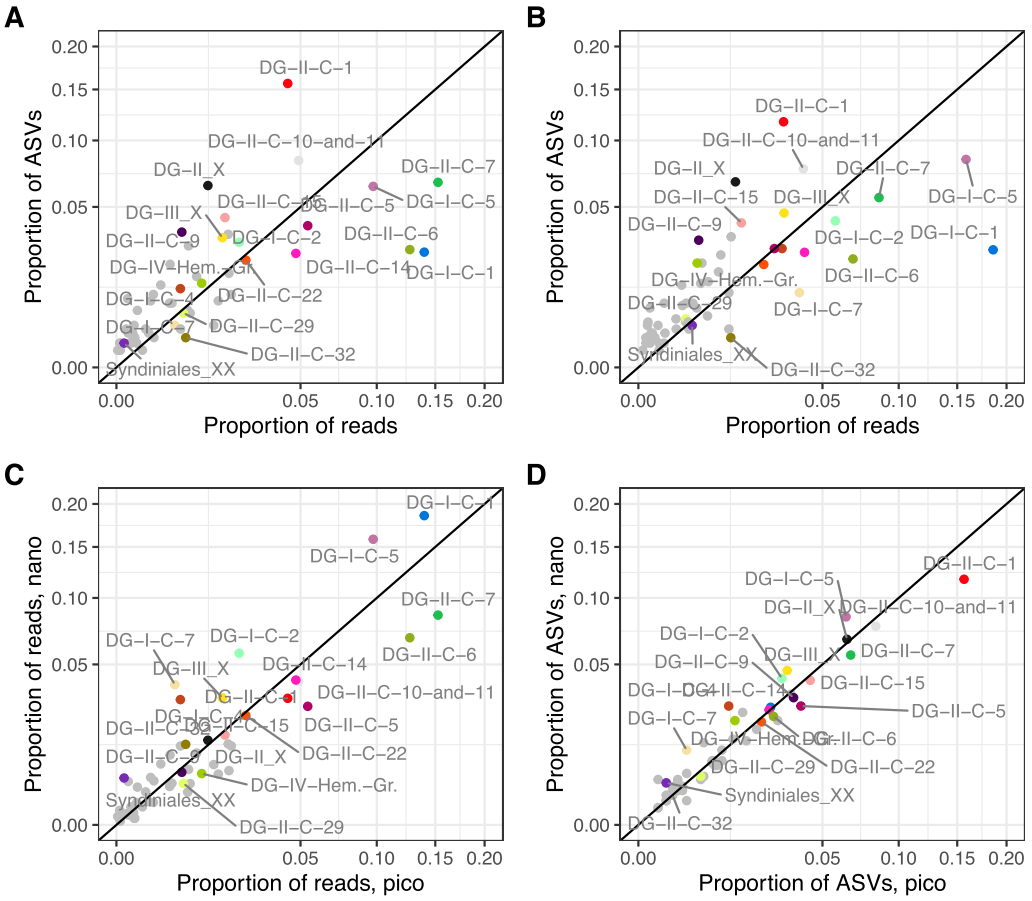
Results and discussion

High taxonomic diversity of MALV in the Arctic

MALV was the most diverse protistan group in the metabarcoding dataset, constituting 33% of the unique amplicon sequence variants (ASVs). It was the most abundant group in the picoplankton size fraction (0.4–3 μm), where it constituted between 6% and 99% of the reads in each sample (Egge et al. 2021, Fig. S1). Considering both size fractions together, we recovered in total 2,151 MALV ASVs. DGs I and II were the most diverse, with 410 and 1589 ASVs, respectively, corresponding to 19% and 74% of the MALV ASVs (Table 1). The order-level clades DGs DG-III, DG-IV, DG-V, and "Syndiniales_X" (PR^2 sequences assigned to MALVs which cannot be placed in any DinoGroup) had 88, 43, 15, and 6 ASVs, respectively (corresponding to 4%, 2%, 0.1%, 0.7%, and 0.3%). This is similar to the global taxonomic distribution of the PR^2 MALV reference sequences (Table 1), but with a higher percentage of DG-II ASVs, and lower percentage of DG-I and DG-V. The number of ASVs was similar in the two size fractions, with 1744 in the pico fraction (0.4–3 μm) and 1750 in the nano-micro size fraction (3–200 μm) (Table 1). The pico fraction had higher richness of DG-II and lower of DG-I compared to the nano-micro fraction (80% and 15% of ASVs in this fraction vs. 70% and 22%, respectively). This was mostly due to higher richness of DG-II-Clade 1 in the pico fraction. Higher richness of DG-II in the pico fraction compared to the larger fractions is also found in the global metaPR^2 dataset (Vaulot et al. 2022).

At lower taxonomic levels, within DG-I, we detected sequences from all the eight subclades delineated in Groisillier et al. (2006) and Guillou et al. (2008), and in addition 7 ASVs which could not be placed in a clade within DG-I. From DG-II we detected sequences assigned to 40 out of the 57 subclades included in PR^2 (Guillou et al. 2013), and 139 ASVs which could not be placed in a clade. In both fractions, DG-II-Clade-1 had the highest number of ASVs, with in total 305 ASVs (14.2% of the ASVs) (273 and 205 in the pico- and nano-micro fractions, respectively, corresponding to 16% and 12%). DGI-Clade 5, DG-II-Clade 10 and 11, DG-II_X, and DG-II-Clade 7 had between 120 and 160 ASVs each, corresponding to 5.5%–7% of the ASVs (Fig. 2; Table S1). The correlation between proportion of reads and proportion of ASVs was c. 0.5 in both size fractions, meaning that for some clades there was a difference in the proportion of ASVs versus proportion of reads. For example, DG-II-Clade 1 was the most ASV-rich in both size fractions, but comprised less than 4% of the reads. Conversely, DG-I Clade 1 and 5 had lower richness, but high proportional abundance. Possible reasons for this discrepancy may be that clades with a low proportion of reads but high ASV richness consist of species with small cell size and/or low 18S copy number, or that these clades contain several species which occur in low cell numbers. Due to the lack of

Fig. 2. Marine alveolates clade representation in reads versus amplicon sequence variants (ASVs), and in the pico- versus the nano-micro fraction. (A and B) Proportion of reads versus proportion of ASVs in the pico- and nano-micro fractions, respectively. (C and D) comparison between the pico- and nano-micro fractions in proportion of reads and proportion of ASVs, respectively.



knowledge on MALVs in general (few identified species with whole genomes sequenced and little knowledge on variation in copy number of the ribosomal operon) it is currently not possible to disentangle these factors.

The MALV community in the Arctic is shaped by season, depth, and water mass

The taxonomic composition of the samples was related to a combination of season, depth zone, and water mass in both size fractions, as shown by principal component ordination of taxonomic similarity between samples (Figs. 3A and 3B). Along the first principal component there was a separation between samples taken in the epipelagic zone in spring and summer, when there was sufficient light for primary production, and samples taken during winter and/or in the mesopelagic zone where there was little light. Consistent with this observation, the PC1 axis scores of the epipelagic samples in both size fractions had a significant correlation with Chl *a* concentration (Table 2). Chl *a* expained 7% and 11% of the total variation in the pico- and nano-micro epipelagic MALV communities, respectively (Table 2). Along the second axis, there was a separation within the winter and/or mesopelagic samples, where the samples taken at 1000 m were separated from the rest, in both size fractions. Furthermore, the samples from the epipelagic zone in January and March clustered according to the origin of the water mass in the surface, i.e., whether it was Atlantic or Arctic. In both size fractions, the PC2 axis scores of the epipelagic samples were correlated with temperature (Table 2). This correlation was mainly driven by the distinct ASV assemblage in the January and March samples of Arctic origin. Temperature explained 2.2% and 7.2% of the total variation in the epipelagic zone in the pico- and nano-micro fractions. Salinity was correlated to the PC2 axis in the mesopelagic samples in both size fractions; however, this factor explained little of the total variation. It should be noted that it is not possible to disentangle the effects of each environmental variable on the MALV community without a much longer time series spanning several years. "Season" can be considered a factor encompassing a variety of abiotic and biotic variables which change throughout the year, and was modeled by a sinusoidal function of day-of-year. "Season" accounted for 77% and 79% of the variation along the PC1 axis in the epipelagic samples from the pico- and nano-micro fractions, respectively, and 15% and 26% of the total variation in these samples. The seasonal variation in PC1 and PC2 is shown in Fig. S2. In the mesopelagic samples this factor accounted for less of the total variation than in the epipelagic (8.3% and 4.4% of total variation). The overall correlation of MALV beta diversity to the beta diversity patterns of Dinophyceae + Ciliophora and Radiolaria was 0.31 and 0.21 for the MALV pico fraction, and 0.77 and 0.67 for

Fig. 3. Structuring of marine alveolates community composition. (A and B) Principal component analysis of similarity in amplicon sequence variant composition between samples with % of variation explained by each axis given next to the axis name. (A) Pico (0.4–3 μm) fraction, (B) nano-micro (3–200 μm) fraction. (C and D) Barplots showing taxonomic composition at the clade level (corresponding to "family" level in PR^2). Clades represented by less than 10% of the reads in all samples are grouped into the category "Other". The samples are grouped according to hierarchical clustering.

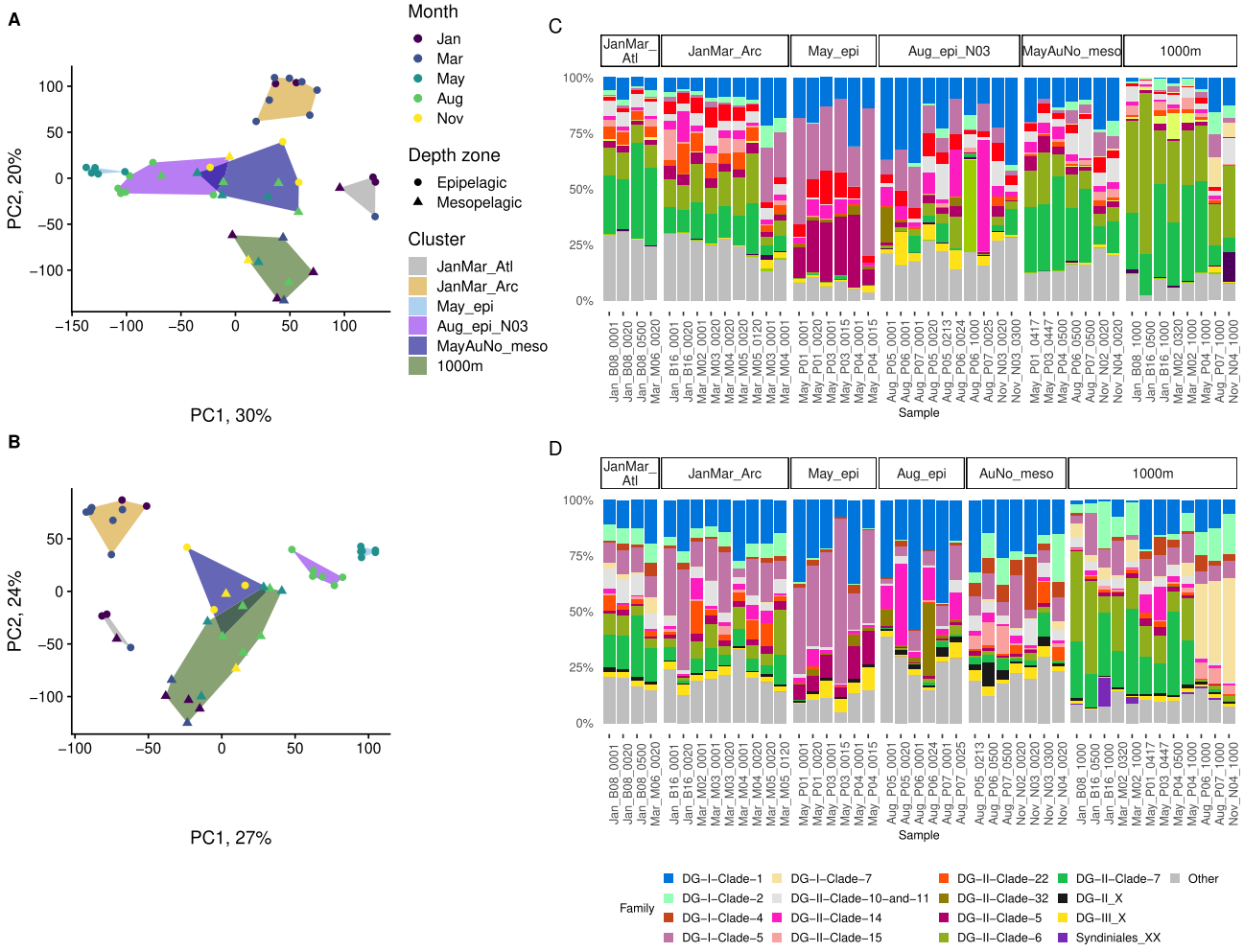


Table 2. Correlations between ordination axes and environmental factors (Gradient) or total variation explained (Varpart).

Gradient	Season (d.o.y)	Chl. *a*	HNF	Large HNF	Temperature	Salinity
Pico, epi, PC1	0.77	0.44	0.3	0.33		
Pico, epi, PC2	0.37		0.18		0.55	
Pico, meso, PC1			0.36		0.59	0.67
Pico, meso, PC2						
Nano-micro, epi, PC1	0.79	0.47	0.33	0.35	0.15	
Nano-micro, epi, PC2		0.38			0.52	
Nano-micro, meso, PC1	0.49	0.32				
Nano-micro, meso, PC2	0.28	0.27	0.4	0.32	0.6	0.86
Varpart						
Pico, epi	15%	7%		6%	2.2	
Pico, meso	8.3%					
Nano-micro, epi	26%	11.1%	7%	6%	7.2	
Nano-micro, meso	4.4%	1.4%			0.2	2.7

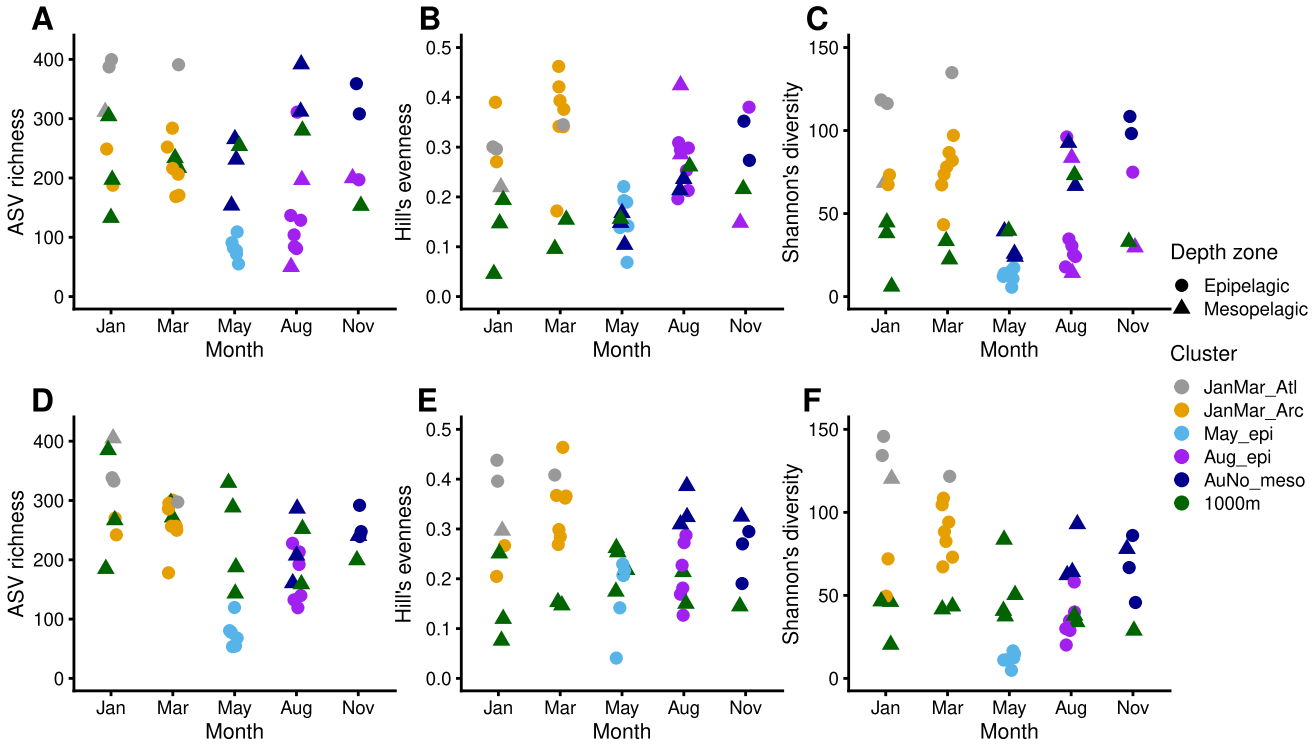
Note: The correlations are given as the square of Pearson's ρ, if $p \leq 0.05$, the variation partitioning values are given as % of total variation explained if $p \leq 0.05$. HNF = Heterotrophic nanoflagellates (between 2–20 μm diameter), Large HNF = HNF with cell size up to 50 μm.

Table 3. Designation of sample clusters.

Cluster name	Pico fraction	Nano-micro fraction
JanMar_Atl	The samples taken at station B08 in January, except 1000 m, and M06 in March. Both stations were influenced by upwelling of Atlantic water.	Same as the pico fraction
JanMar_Arc	The samples taken at station B16 in January, except 500 and 1000 m, and M02-M05 in March, all surface samples influenced by Arctic water.	Same as the pico fraction
May_epi	All samples from the epipelagic zone taken in May	Same as the pico fraction
Aug_epi_N03/Aug_epi	All samples from the epipelagic zone taken in August, and in addition Aug_P05_0213, and both samples from November station N03	All samples from the epipelagic zone taken in August
MayAuNo_meso/AuNo_meso	Samples taken in the mesopelagic zone above 1000 m from May and August, and Nov_N02 and N04, both from 20 m	Samples taken in the mesopelagic zone above 1000 m in August, and Nov_N03 20 and 300 m, and Nov_N04 20 m
1000 m	All samples taken at 1000 m, except Aug_P06_1000 m	All samples taken at 1000 m, and in addition all samples from 400500 m in May, Jan_B16_500 m and Mar_M02_320 m

Note: When 2 names are present in the name column, they correspond to the pico and nano fractions, respectively.

Fig. 4. Alpha diversity. (A, D) amplicon sequence variant (ASV) richness, (B, E) Hill's evenness, (C, F) Shannon's diversity. Top row: pico (0.4–3 μm) fraction, bottom row: nano-micro (3–200 μm) fraction.



the MALV nano-micro fraction (all permutation significance values < 0.01).

Beta diversity patterns of these host groups are shown in Fig. S3.

Complete linkage hierarchical clustering allowed to delimit six main sample clusters in each of the size fractions. The sample composition in these clusters is described in Table 3. In the pico fraction, 19 ASVs had significantly different CLR-values between the clusters, whereas the nano-micro fraction had 114 (Table S2). Most of these ASVs had higher CLR in the "dark" samples (i.e., samples from the mesopelagic zone, or epipelagic zone in winter; Fig. S4).

Evidence for a distinct MALV community at 1000 m

Strikingly, most of the samples taken at 1000 m clustered together and thus had similar ASV composition irrespective of season, in particular in the pico fraction (Fig. 3). This suggests that the community of MALV at great depths is relatively stable over time, and may be parasitic of deepdwelling organisms that are little affected by the strong seasonal variation in the Arctic epipelagic zone. Richness in this cluster was similar to the winter and summer mesopelagic samples above 1000 m (Fig. 4), but it had overall the highest number of unique ASVs (i.e., only detected in this cluster, Fig. S5). A

Table 4. The most abundant clades within each of the sample clusters shown in Fig. 3.

Size fraction	Cluster	Clades with highest mean read abundance
Pico	JanMar_Atl	DG-II-Clade-7 (31.7), DG-II-Clade-6 (10.7), DG-II-Clade-21 (7.6), DG-I-Clade-1 (6.4), DG-II-Clade-10-and-11 (6.3)
	JanMar_Arc	DG-II-Clade-6 (13), DG-II-Clade-7 (12.6), DG-I-Clade-1 (10.4), DG-II-Clade-22 (6.4), DG-I-Clade-5 (6.1)
	May_epi	DG-I-Clade-5 (36.5), DG-II-Clade-5 (21.7), DG-I-Clade-1 (17.6), DG-II-Clade-1 (8.1), DG-II-Clade-14 (4.4)
	Aug_epi_N03	DG-I-Clade-1 (24.3), DG-II-Clade-14 (10.6), DG-I-Clade-5 (10.2), DG-II-Clade-1 (5.3), DG-III_X (5)
	MayAuNo_meso	DG-II-Clade-7 (25.7), DG-I-Clade-1 (15.2), DG-II-Clade-6 (12.1), DG-II-Clade-10-and-11 (8.2), DG-I-Clade-5 (4.3)
	1000 m	DG-II-Clade-6 (34.3), DG-II-Clade-7 (26.6), DG-I-Clade-1 (5), DG-II-Clade-10-and-11 (4.6), DG-I-Clade-2 (3.3)
Nano-micro	JanMar_Atl	DG-II-Clade-7 (18), DG-I-Clade-1 (14.2), DG-I-Clade-5 (10.2), DG-II-Clade-10-and-11 (9), DG-I-Clade-2 (7.4)
	JanMar_Arc	DG-I-Clade-5 (20), DG-I-Clade-1 (17.8), DG-II-Clade-6 (7.7), DG-II-Clade-7 (6.6), DG-I-Clade-2 (5.6)
	May_epi	DG-I-Clade-5 (41.7), DG-I-Clade-1 (23.9), DG-II-Clade-5 (9.6), DG-II-Clade-1 (6.2), DG-III_X (5.5)
	Aug_epi	DG-I-Clade-1 (33.1), DG-II-Clade-14 (13.2), DG-I-Clade-5 (7.8), DG-II-Clade-32 (7.4), DG-II-Clade-1 (5)
	AuNo_meso	DG-I-Clade-1 (21.7), DG-I-Clade-5 (10), DG-I-Clade-4 (8.7), DG-I-Clade-2 (8.5), DG-II-Clade-10-and-11 (5.8)
	1000 m	DG-II-Clade-7 (17.7), DG-II-Clade-6 (15.5), DG-I-Clade-7 (12.4), DG-I-Clade-5 (9.1), DG-I-Clade-2 (9)

high number of unique MALV ASVs at depths 800–1000 m has also been found in oligotrophic, tropical oceans (Anderson et al. 2024). The most abundant clades in this cluster were DG-II Clades 6 and 7, and the most abundant ASVs were also assigned to these clades (Table 4). Clades 6 and 7 are among the most diverse in DG-II in PR^2, and Clade 7 has been found to be the most diverse within DG-II in samples from the aphotic zone (Guillou et al. 2008). In metaPR^2 both DG-II Clade 6 and 7 are detected from tropical to polar latitudes, and are most abundant in mesopelagic samples (Table S3). They are also reported in high proportions in deep samples in Antarctic as well as temperate oceanic regions (Cleary and Durbin 2016; Anderson et al. 2024). So far no representative of the deep-dwelling DG-II-Clade 6 and DG-II-C 7 have been isolated, and their hosts are unknown. Members of DG-II-Clade 7 are speculated to be able to parasitize deep planktonic organisms belonging to Cercozoa and Radiolaria (Guillou et al. 2008; Anderson et al. 2024), which are known to harbour a wide diversity of MALV parasites (Bråte et al. 2012). Principal component analysis of the community of Radiolarians in the nano-micro size fraction revealed a very stable community in the samples taken at 500–1000 m (Fig. S3). Closer inspection of these samples revealed consistent presence of ASVs assigned to the Radiolarian taxon Chaunacanthida at depths below 200 m, constituting c. 80% of the radiolarian reads in these samples (Fig. S6). An association between members of DG-II and Chaunacanthida has previously been found in samples from the Hudson Bay (Jacquemot et al. 2022). Members of Chaunacanthida are heterotrophic, and known to form cysts which are thought to descend to mesopelagic depths to release swarmers (Decelle et al. 2013). The cysts are usually >200 μm and would thus not be recovered in our samples, whereas the swarmers are 2–3 μm and would be detected in the nano-micro fraction, which would explain why Chaunacanthida appears to be most abundant below 200 m in our data. Thus, to determine whether this correlation between Chaunacanthida and DG-II MALV clades indicates a host–parasite relationship, imaging and culturing is needed.

In the nano-micro fraction, the August and November 1000 m samples were distinguished by high abundance of DG-I-Clade 7 (mainly in the 3–10 μm fraction), which shows that variation also occurs at these depths. In metaPR^2, DG-I-Clade 7 is detected from tropical to polar latitudes, generally in low abundance, in the surface as well as at mesopelagic depths. We may thus speculate that this clade infects hosts with vertical migration. Terrado et al. (2009) found seasonal variation in the protist community at 200 m depth, which the authors suggest was due to advected water masses with a distinct community. However, since we only observed the change in community composition at 1000 m in one size fraction and at geographically separated stations, advection is a less likely explanation in this case. Furthermore, ASVs assigned to DG-I-Clade 7 have been found in high relative abundance in amphipods (Savage et al. 2023). The abundant Arctic amphipod *Apherusa glacialis* is known to perform seasonal vertical migration, and may overwinter in the deep within the Atlantic-water inflow near Svalbard (Kunisch et al. 2020; Drivdal et al. 2021), but further investigations are needed to establish whether this particular species is susceptible to MALV infection. Other potential hosts at these depths are Arctic copepod species (e.g., *Calanus glacialis* and *C. hyperboreus*) that are known to descend to depths as far as below 800 m for diapause during winter (Kvile et al. 2019, and references therein). *C. hyperboreus* spawns during winter down to 1000 m (Hirche and Niehoff 1996). The composition of metazoan taxa in our metabarcoding data from the mesopelagic zone displayed considerable variation between the months and stations (Fig. S7). The three samples from August and November taken at 1000 m with high abundance of DG-I-C7 had different taxonomic compositions of metazoa, and were dominated by reads assigned to *Oithona similis*, Maxillopoda sp. and *Calanus* sp., respectively. It should be noted that the metabarcoding data represent size fractions <200 μm, and our sampling and bioinformatic strategy was not optimized to capture the full diversity potential metazoan MALV hosts.

Similar communities in samples taken during the winter and below the photic zone in spring and summer

In November and January, polar night conditions (i.e., no daylight) prevailed. In March, day length was 6–8 h, but the chlorophyll *a* values were <0.1 μg.L^{-1}, indicating that the spring bloom had not yet started (Fig. S8; Randelhoff et al. 2018). The relative abundance of MALV was generally high in both the samples from January, March, and November in the epipelagic zone, and in the samples taken in the mesopelagic zone regardless of season, constituting up to 99% of the protist reads in the pico fraction and up to 77% in the nano-micro (Fig. S1). High relative abundance of MALV reads in Arctic surface waters during winter has previously been reported from Isfjorden, Svalbard (Marquardt et al. 2016). Furthermore, our Arctic data from mesopelagic depths are consistent with previous studies finding high relative abundance of MALV at great depths in various oceanographic regions (Terrado et al. 2009; Pernice et al. 2016; Xu et al. 2017). The samples taken at mesopelagic depths above 1000 m in summer and winter, and in the winter surface had taxonomic compositions that were a mixture between the surface samples from spring and summer (described below) and the 1000 m samples. Similar to the surface samples from spring and summer, these samples had high proportions of DG-I-Clade 1, and in some samples DG-I-Clade 5, in particular in the nano-micro fraction (Fig. 3). Similar to the 1000 m cluster, DG-II was dominating, especially DG-II-Clade 6 and DG-II-Clade 7 (Table 4). DG-II-Clade 7 had the highest number of ASVs with significantly different CLR between clusters in both size fractions, and these ASVs had typically higher CLR-values in the clusters with "dark" samples (Fig. S4).

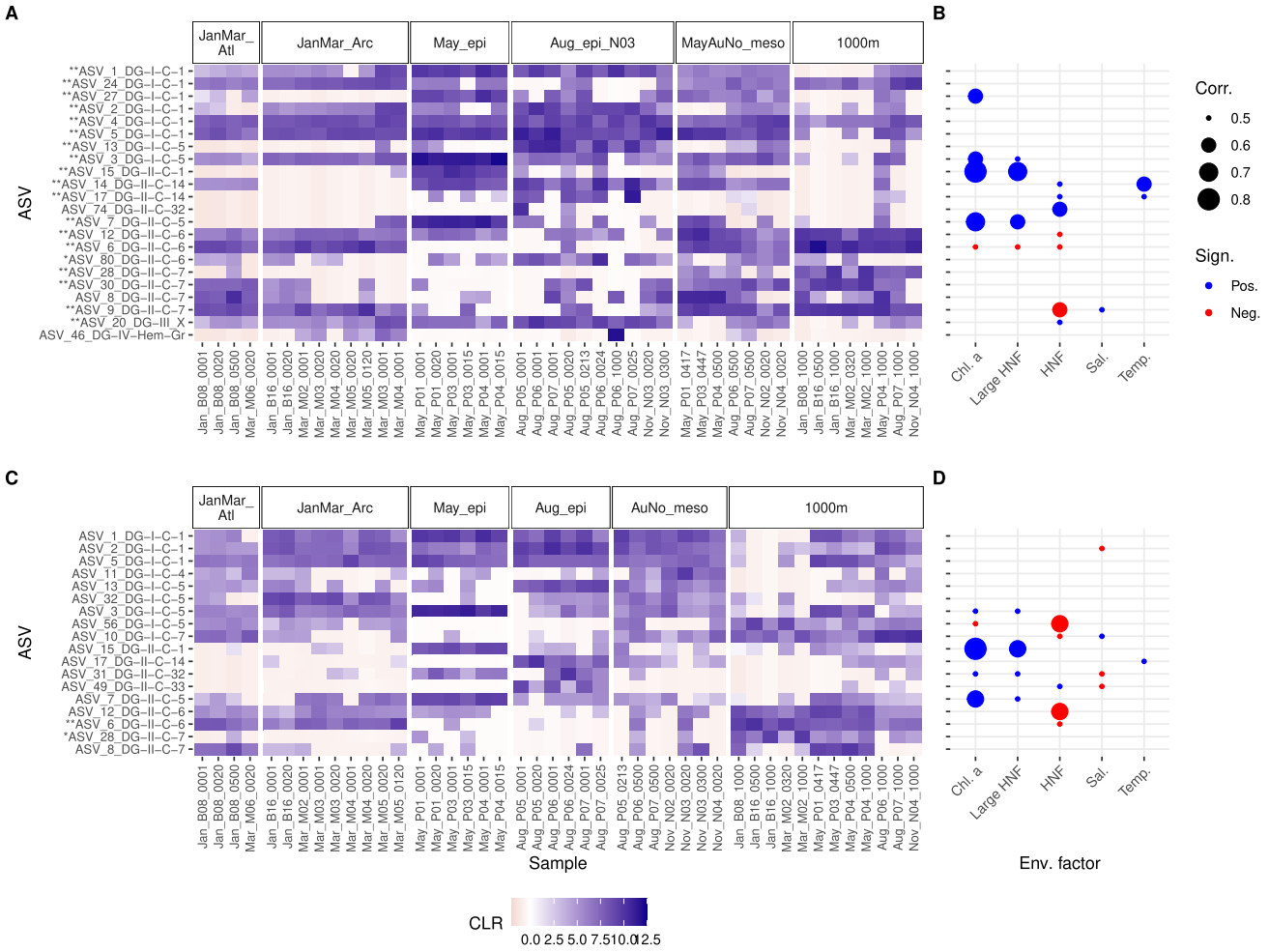
Difference between Arctic and Atlantic water masses in the surface during winter

The MALV communities in the epipelagic and mesopelagic zones during winter were relatively similar (Fig. 3). However, we found geographical differences which could be related to the origin of the water mass at the sampling station. In January and March 2014, there was still open water along the north-west coast of Svalbard. During winter, the water masses on the shelf slope north of Svalbard is dominated by Atlantic water with strong vertical convection, whereas the water masses further off the slope is characterised by a surface layer of fresher Arctic water and some stratification even in the winter (Randelhoff et al. 2018). At the station January B08, which was located closer to the shelf compared to January B16, the MALV communities at 1, 20, and 500 m depth were similar. At B16, which was influenced by Arctic water, as indicated by lower salinity and density in the surface (Fig. 1B; see also Egge et al. 2021), there was a difference in the MALV community between the epi- and mesopelagic zones. At the March station M06, which was located closer to the shelf, the MALV community in the M06_20 m sample (the only sample from this station) was similar to the B08 samples from 1–500 m. The other March stations, which were located further North-East, had a fresher layer down to below c. 150 m, and the samples from this layer were similar to the surface samples from Jan B16. The MALV community in the putative Arctic water masses during winter were characterised by higher relative abundance of DG-II-Clade 1 and DG-II-Clade 22 compared to the winter community of Atlantic origin. Also the overall protist community differed between these water masses, with higher abundance of Picozoa in the samples taken further off the shelf.

Community composition in the epipelagic zone—change between May and August

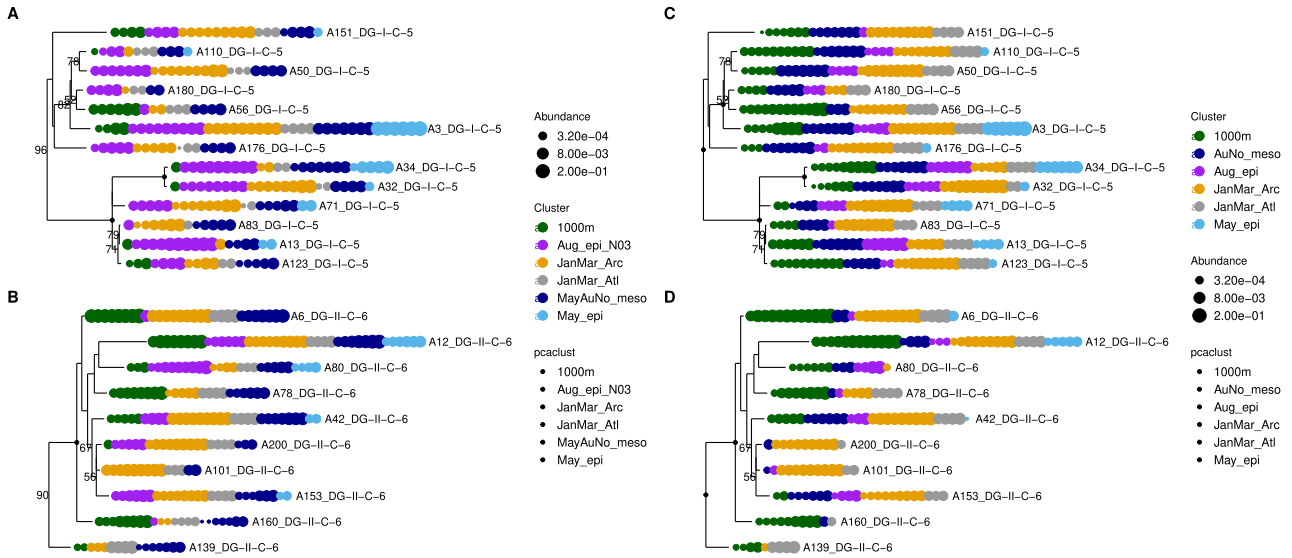
The May and August epipelagic samples were taken under sunlit conditions, but representing different stages in the phytoplankton seasonal succession, as indicated in the change in Chl *a* concentration from May to August. The samples in the May_epi cluster were characterised by relatively low ASV richness and evenness (Fig. 4), and high abundance of DG-I-Clade 5 and DG-II-Clade 5 (Fig. 3). This was due to the very abundant ASVs ASV_3_DG-I-C5 and ASV_7_DG-II-C5, which constituted up to 70% and 26% of the reads, respectively, in these samples (Fig. 5). These ASVs have previously only been detected in surface samples in metaPR^2 (Table S4). In the nano-micro fraction, ASV_15_DG-II-C1 was also abundant in the May samples. The May_epi samples had high levels of Chl *a* with up to 15 μg L^{-1}, and high levels of large heterotrophic nanoflagellates compared to the other samples (Fig. S8; Randelhoff et al. 2018; Sandaa et al. 2018). Thus, the most prominent May ASVs were positively correlated with Chl *a* concentration, and in some cases also the abundance of large HNF (Figs. 5B and 5D). Similar to our results, Clarke et al. (2019) found high abundance of a particular MALV operational taxonomic unit (OTU) associated with high Chl *a* values in the Antarctic ocean. This OTU was affiliated with DG-I-Clade 1, which was also relatively abundant in the May epipelagic samples, but not as dominating as DG-I-Clade 5. It is nevertheless interesting that MALV genotypes from different subclades may dominate during periods of high phytoplankton biomass in different oceanographic regions. DG-I-Clade 5 has been found to be one of the most abundant DG-I clades in the euphotic zone, and in high relative abundance in the surface in oligotrophic tropical oceans (Guillou et al. 2008; Anderson et al. 2024). DG-II-Clade 5 is also mainly found in the same layer, and is detected in the metaPR^2 database from tropical to polar latitudes in low relative abundance, but relative abundance seems to increase with latitude (Table S3). In August, the abundance of DG-I-C5 and DG-II-C5 was reduced, and in particular DG-II-Clade 14 was abundant. This clade has been found both in oceanic surface and deep waters (Groisillier et al. 2006). In metaPR^2 it was detected from tropical to polar latitudes, generally in low abundance (Table S3). Members of this clade have been demonstrated to infect *Heterocapsa triquetra* (Chambouvet et al. 2008), a thecate dinoflagellate which may be abundant in Arctic waters (e.g., Seuthe et al. 2011). In the metabarcoding data, the genus *Heterocapsa* was relatively abundant in the fractions smaller than 10 μm all year (Egge et al. 2021). Furthermore, certain ASVs had higher abundance in August than in May such as ASV_13_DG-I-C5 (pico), and ASV_25_DG-I-C1 and ASV_17_DG-II-C14 (nano-micro). This indicates that the MALV community

Fig. 5. Abundant marine alveolates amplicon sequence variants (ASVs): distribution and correlation to environmental factors. (A and C) Heatmaps of ASVs with CLR (centered log-ratio) above 10 in at least one sample. The intensity of the colour indicates the centred log-ratio (CLR) of the ASV in a given sample. (A) Pico (0.4–3 μm) fraction, (C) nano-micro (3–200 μm) fraction. ∗ indicates whether an ASV has significantly different CLR between clusters. ∗ $p \leq 0.1$, ∗∗ $p \leq 0.05$. (B) and (D) Correlations between ASVs and environmental factors, for ASVs in the pico- and nano-micro fractions, respectively. The size indicates the strength of the correlation, and the colour indicates whether it is positive or negative. Chl. *a* = chlorophyll *a* concentration, HNF = Heterotrophic nanoflagellates, Sal. = Salinity, Temp. = Temperature.



also follows the succession of the phytoplankton community from spring to summer, possibly as a response to a change in the host community. Strong seasonality of MALV taxonomic composition in sunlit waters has previously been observed in a coastal pond (Sehein et al. 2022) and in subtropical coastal waters, where summer samples were distinct from those of colder months (Anderson and Harvey 2020). In the epipelagic May samples there was a high relative read abundance of phytoplankton, dominated by diatoms, *Phaeocystis pouchetii* and the genus *Micromonas* (Egge et al. 2021), which in addition to the high Chl *a* concentrations indicates a spring bloom situation. The "bloom" dynamic of certain MALV ASVs suggests that while they are present in low abundance at other times of the year and at other depths, these ASVs are favored by the conditions in the surface during spring. This pattern may be due to an increased abundance of hosts, either blooming species, of e.g., diatoms or dinoflagellates, or grazers, such as copepods. The diatoms were dominated by *Chaetoceros socialis* and *Detonula confervacea* in May, which were replaced by a more diverse community in the August surface samples (Egge et al. 2021). A somewhat similar pattern was seen for dinoflagellates, with a few taxonomic groups dominating in May (unclassified Dinophyceae, *Gyrodinium*, *Woloszynskia*), and higher taxonomic richness in August (Fig. S9). Another parasitic dinoflagellate genus, *Chytriodinium*, which is not a member of the MALV, but harbours known parasites of copepod eggs, was dominating the dinophyceae reads in the 50–200 μm fraction in May, and had a strongly reduced abundance in August (Egge et al. 2021). Furthermore, the dominating taxon in the metazoan metabarcoding reads shifted from the calanoid copepod genus *Calanus* in May, to the cyclopoid genus *Oithona* in August (Fig. S7). Calanoid and cyclopoid copepod species

Fig. 6. RAxML trees of the most abundant amplicon sequence variants within the marine alveolates clades Dino-Group-I-Clade-5 and Dino-Group-II-Clade-6. (A and B) Pico fraction, (C and D) nano-micro fraction. Each dot at the leaves corresponds to a sample. The size of the dot corresponds to relative abundance in the sample. Samples are colored according to cluster. Bootstrap values > 50 are shown on the nodes.



have been shown to harbour different eukaryotic parasites and symbionts, and in particular, different MALV parasites (Savage et al. 2023).

Difference between epi- and mesopelagic samples in spring and summer

Stratification began to develop in May and was strong in August at the two northernmost stations (P06 and P07; 1 B). Parasites, either in the form of infected hosts or spores, may aggregate on particles and sink out of the photic zone (Anderson et al. 2024), which would then result in similar compositions of the environmental DNA (eDNA) pools in these zones. However, in May and August the communities in the epi- and mesopelagic zone were very different. Depth-structuring of marine protist communities in general, and MALV communities in particular, has also been found at lower latitudes (Ollison et al. 2021; Anderson et al. 2024). Our results thus support the hypothesis that the MALV community in the mesopelagic zone, as revealed by analysis of eDNA, is shaped by the environment and availability of hosts at these depths, rather than passive sinking transport of surface-dwelling taxa (Anderson et al. 2024).

Difference between the pico- and nano-micro fractions

MALV sequences from the smallest size fractions in environmental sequencing studies likely come from the dinospores (Guillou et al. 2008), but could also come from larger infected host cells that are disrupted during filtration. In addition to infecting larger hosts, dinospores may theoretically also be grazed by microzooplankton. Thus, in larger fractions obtained by sequential filtration, the presence of a MALV ASV may represent an infected host, dinospores recently ingested by microzooplankton, freeliving spores caught in the material on the filter, or potentially the vermiform stage. While dinospores are generally short-lived, with a life span of 2–3 days (Coats and Park 2002), some MALV species may have evolved longer-lasting spores, potentially as an adaptation to life in a cold environment, as supported by observations in other parasitic dinoflagellate groups (Coats 1999). If samples in the pico size fraction reflect recent spore releases from hosts in the nano-micro fraction, we would expect that the size fractions had the same taxonomic composition, possibly with different proportional abundance, and that the sample similarity relationships would be similar. This was mostly the case, as we observed similar clustering patterns in the pico- and nano-micro fractions (Fig. 3). However, leakage between size fractions, as described above, cannot be ruled out. We may speculate that the ASVs and clades that are more abundant in the pico fraction, such as DG-II-Clade 6 and 7, represent spores released from hosts larger than 200 μm. Interestingly, the correlations to the host communities Dinophyceae + Ciliophora and Radiolaria were higher for the MALV nano-micro fraction than for the pico fraction (as mentioned above), which may suggest that the MALV metabarcoding data from this fraction to a larger extent represent current infections.

Distribution of phylogenetically similar asvs

The composition of MALV ASVs in each sample cluster usually contained several clades from both DG-I and II/IV (Fig. S4). Previous studies have found that parasites of the same host type may come from several different clades within MALV (Guillou et al. 2008), e.g., both DG-I and DG-II/IV have members parasitizing other dinoflagellates. We also observed different distributions of ASVs within the same clade, e.g.,

within DG-I-C5, ASV_3 was relatively abundant in all sample clusters, whereas the closely related ASV_50, ASV_56, and ASV_180 were not detected in the May_epi samples in either size fraction (Figs. 6A and 6C). Within DG-II-C6, ASV_101, and ASV_200 had very different distributions from the rest of the ASVs, as they were almost only detected in the Jan-Mar_Arc cluster, especially in the nano-micro fraction (Figs. 6B and 6D). This illustrates a complex pattern of environment and host preference both within and between clades of MALV.

Conclusions

Much work remains to be done to elucidate the impact of MALV parasitism in the marine microbial food web, in particular identifying host–parasite relationships and the extent of host specificity. Our study contributes to understanding the environmental preferences of several MALV clades and ASVs. Future work based on this study could include designing fluorescence in situ hybridisation probes based on the most abundant ASVs, and trying to detect them in potential hosts by epifluorescence microscopy. To further clarify the putative interactions between MALV and blooming photoautotrophs, we suggest targeting the ASVs dominating in the spring and summer surface samples. Identifying the hosts of these abundant MALV ASVs in the Arctic Ocean will contribute to quantifying the impact of parasitism on the Arctic food web.

Acknowledgements

We would like to thank all members of the MicroPolar and CarbonBridge consortia, and the crews onboard R/V Lance and R/V Helmer Hanssen for their efforts during the sampling cruises.

Article information

History dates

Received: 1 March 2024
Accepted: 1 December 2024
Accepted manuscript online: 23 January 2025
Version of record online: 26 March 2025

Notes

This paper is part of a collection entitled “Thirty years of Earth System Science in high-Arctic Svalbard: status, trends, and future recommendations”. https://cdnsciencepub.com/topic/as-unis.

Copyright

Data availablity

The raw 18S rRNA V4 reads are available from the European Nucleotide Archive repository under project number PRJEB40133. ASV tables and R scripts from this study are deposited in GitHub at https://github.com/EEgge/micropolar_syndinialespaper. Interactive figures of environmental factors and taxonomic distribution of the metabarcoding data can be accessed via the online Shiny app (https://micropolar-protists.metapr2.org/micropolar-protists).

Author information

Author ORCIDs

Elianne Egge https://orcid.org/0000-0001-5790-6426
Daniel Vaulot https://orcid.org/0000-0002-0717-5685
Aud Larsen https://orcid.org/0000-0001-6927-5537
Bente Edvardsen https://orcid.org/0000-0002-6806-4807

Author contributions

Conceptualization: EE, DV, AL, BE
Data curation: EE, DV
Formal analysis: EE
Funding acquisition: AL, BE
Investigation: EE, AL, BE
Methodology: EE, AL, BE
Project administration: AL, BE
Resources: AL
Supervision: DV, BE
Validation: EE, DV, AL
Visualization: EE
Writing – original draft: EE, DV, AL, BE
Writing – review & editing: EE, DV, AL, BE

Competing interests

The authors declare no competing financial interests.

Funding information

The study was conducted as part of the Research Council of Norway project Micropolar—Processes and Players in Arctic Marine Pelagic Food Webs—Biogeochemistry, Environment, and Climate Change No. 225956/E10. Daniel Vaulot was supported by ANR contract PhytoPol (ANR-15-CE02-0007).

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/as-2024-0016.

References

Anderson, S.R., Blanco-Bercial, L., Carlson, C.A., and Harvey, E.L. 2024. The role of Syndiniales parasites in depth-specific networks and carbon flux in the oligotrophic ocean. ISME Communications, **4**. doi:10.1093/ismeco/ycae014.

Anderson, S.R., and Harvey, E.L. 2020. Temporal variability and ecological interactions of parasitic marine syndiniales in coastal protist communities. mSphere, **5**. doi:10.1128/mSphere.00209-20.

Berdjeb, L., Parada, A., Needham, D.M., and Fuhrman, J.A. 2018. Short-term dynamics and interactions of marine protist communities during the spring-summer transition. The ISME Journal, **12**: 1907–1917. doi:10.1038/s41396-018-0097-x.

Bjorbækmo, M.F.M., Evenstad, A., Røsæg, L.L., Krabberød, A.K., and Logares, R. 2020. The planktonic protist interactome: where do we stand

after a century of research? The ISME Journal, **14**: 544–559. doi:10.1038/s41396-019-0542-5.

Borcard, D., Gillet, F., and Legendre, P. 2018. Numerical Ecology with R. Use R! Springer International Publishing.

Bråte, J., Krabberød, A.K., Dolven, J.K., Ose, R.F., Kristensen, T., Bjørklund, K.R., and ShalchianTabrizi, K. 2012. Radiolaria associated with large diversity of marine alveolates. Protist, **163**: 767–777. doi:10.1016/j.protis.2012.04.004. PMID: 22658831.

Cachon, J. 1964. Contribution a l'etude des peridiniens parasites. Cytologie, cycles evolutifs. Annales des Sciences Naturelles, Paris, **6**: 1–158.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nature Methods, **13**: 581–583. doi:10.1038/nmeth.3869. PMID: 27214047.

Capella-Gutiérrez, S., Silla-Martínez, J., and Gabaldón, T. 2009. A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics, **25**: 1972–1973. doi:10.1093/bioinformatics/btp348. PMID: 19505945.

Cavalier-Smith, T. 2018. Kingdom Chromista and its eight phyla: a new synthesis emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. Protoplasma, **255**: 297–357. doi:10.1007/s00709-017-1147-3. PMID: 28875267.

Chambouvet, A., Morin, P., Marie, D., and Guillou, L. 2008. Control of toxic marine dinoflagellate blooms by serial parasitic killers. Science, **322**: 1254–1257. doi:10.1126/science.1164387. PMID: 19023082.

Chatton, E. 1920. Les Peridiniens parasites: morphologie, reproduction, ethologie. Archives de zoologie expérimentale et générale, **59**: 1–475.

Clarke, L.J., Bestley, S., Bissett, A., and Deagle, B.E. 2019. A globally distributed Syndiniales parasite dominates the Southern Ocean microeukaryote community near the sea-ice edge. The ISME Journal, **13**: 734–737. doi:10.1038/s41396-018-0306-7. PMID: 30367123.

Cleary, A.C., and Durbin, E.G. 2016. Unexpected prevalence of parasite 18S rDNA sequences in winter among Antarctic marine protists. Journal of Plankton Research, **38**: 401. doi:10.1093/plankt/fbw005.

Coats, D.W. 1999. Parasitic life styles of marine dinoflagellates. Journal of Eukaryotic Microbiology, **46**: 402–409. doi:10.1111/j.1550-7408.1999.tb04620.x.

Coats, D.W., and Park, M.G. 2002. Parasitism of photosynthetic dinoflagellates by three strains of *Amoebophrya* (Dinophyta): parasite survival, infectivity, generation time, and host specificity. Journal of Phycology, **38**: 520–528. doi:10.1046/j.1529-8817.2002.01200.x.

Conway, J.R., Lex, A., and Gehlenborg, N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics, **33**: 2938–2940. doi:10.1093/bioinformatics/btx364. PMID: 28645171.

Davies, C.E., Thomas, J.E., Malkin, S.H., Batista, F.M., Rowley, A.F., and Coates, C.J. 2022. Hematodinium sp. infection does not drive collateral disease contraction in a crustacean host. Elife, **11**: e70356. doi:10.7554/eLife.70356. PMID: 35179494.

Decelle, J., Martin, P., Paborstava, K., Pond, D.W., Tarling, G., Mahé, F., et al. 2013. Diversity, ecology and biogeochemistry of cyst-forming acantharia (Radiolaria) in the oceans. PLoS ONE, **8**: e53598. doi:10.1371/journal.pone.0053598. PMID: 23326463.

Drivdal, M., Kunisch, E.H., Bluhm, B.A., Gradinger, R., Falk-Petersen, S., and Berge, J. 2021. Connections to the deep: deep vertical migrations, an important part of the life cycle of Apherusa glacialis, an arctic ice-associated amphipod. Frontiers in Marine Science, **8**: 772766. doi:10.3389/fmars.2021.772766.

Egge, E., Elferink, S., Vaulot, D., John, U., Bratbak, G., Larsen, A., and Edvardsen, B. 2021. An 18S V4 rRNA metabarcoding dataset of protist diversity in the Atlantic inflow to the Arctic Ocean, through the year and down to 1000 m depth. Earth System Science Data, **13**: 4913–4928. doi:10.5194/essd-13-4913-2021.

Farhat, S., Le, P., Kayal, E., Noel, B., Bigeard, E., Corre, E., et al. 2021. Rapid protein evolution, organellar reductions, and invasive intronic elements in the marine aerobic parasite dinoflagellate amoebophrya spp. BMC Biology, **19**: 1–21. doi:10.1186/s12915-020-00927-9. PMID: 33407428.

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., and Egozcue, J.J. 2017. Microbiome datasets are compositional: and this is not optional. Frontiers in Microbiology, **8**: 1–6. doi:10.3389/fmicb. 2017.02224. PMID: 28197127.

Groisillier, A., Massana, R., Valentin, K., Vaulot, D., and Guillou, L. 2006. Genetic diversity and habitats of two enigmatic marine alveolate lineages. Aquatic Microbial Ecology, **42**: 277–291. doi:10.3354/ame042277.

Grover, J.P., and Chrzanowski, T.H. 2006. Seasonal dynamics of phytoplankton in two warm temperate reservoirs: association of taxonomic composition with temperature. Journal of Plankton Research, **28**: 1–17. doi:10.1093/plankt/fbi095.

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R.M., Kirkham, A.R., Massana, R., et al. 2008. Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). Environmental Microbiology, **10**: 3349–3365. doi:10.1111/j.14622920.2008.01731.x. PMID: 18771501.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Research, **41**: D597–D604. doi:10.1093/nar/gks1160. PMID: 23193267.

Guiry, M.D., and Guiry, G.M. 2024. AlgaeBase. World-wide electronic publication, University of Galway. Available from https://www.algaebase.org [accessed 10 October 2024].

Hirche, H.-J., and Niehoff, B. 1996. Reproduction of the Arctic copepod *Calanus hyperboreus* in the Greenland Sea-field and laboratory observations. Polar Biology, **16**: 209–219. doi:10.1007/BF02329209.

Hobbs, L., Banas, N.S., Cottier, F.R., Berge, J., and Daase, M. 2020. Eat or sleep: availability of winter prey explains mid-winter and spring activity in an arctic *Calanus* population. Frontiers in Marine Science, **7**. doi:10.3389/fmars.2020.541564.

Holt, C.C., Hehenberger, E., Tikhonenkov, D.V., Jacko-Reynolds, V.K., Okamoto, N., Cooney, E.C., et al. 2023. Multiple parallel origins of parasitic Marine Alveolates. Nature Communications, **14**: 7049. doi:10.1038/s41467-023-42807-0. PMID: 37923716.

Jackson, D.A. 1997. Compositional data in community ecology: the paradigm or peril of proportions? Ecology, **78**: 929–940. doi:10.2307/2266070.

Jacquemot, L., Vigneron, A., Tremblay, J.-É., and Lovejoy, C. 2022. Contrasting sea ice conditions shape microbial food webs in Hudson Bay (Canadian Arctic). ISME Communications, **2**: 104. doi:10.1038/s43705-022-00192-7. PMID: 37938285.

Katoh, K., Rozewicki, J., and Yamada, K.D. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics, **20**: 1160–1166. doi:10.1093/bib/bbx108. PMID: 28968734.

Kimmerer, W., and McKinnon, A. 1990. High mortality in a copepod population caused by a parasitic dinoflagellate. Marine Biology, **107**: 449–452. doi:10.1007/BF01313428.

Koid, A., Nelson, W.C., Mraz, A., and Heidelberg, K.B. 2012. Comparative analysis of eukaryotic marine microbial assemblages from 18S rRNA gene or gene transcript clone libraries using different methods of extraction. Applied and Environmental Microbiology, **78**: 3958–3965. doi:10.1128/AEM.06941-11. PMID: 22447590.

Kunisch, E.H., Bluhm, B.A., Daase, M., Gradinger, R., Hop, H., Melnikov, I.A., et al. 2020. Pelagic occurrences of the ice amphipod Apherusa glacialis throughout the Arctic. Journal of Plankton Research, **42**: 73–86. doi:10.1093/plankt/fbz072. PMID: 32025067.

Kvile, K.Ø., Ashjian, C., and Ji, R. 2019. Pan-Arctic depth distribution of Diapausing *Calanus* copepods. The Biological Bulletin, **237**: 76–89. doi:10.1086/704694. PMID: 31714854.

Lahti, L., and Shetty, S. 2019. microbiome. R package version 1.28.0.

López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., and Moreira, D. 2001. Unexpected diversity of small eukaryotes in deep-sea antarctic plankton. Nature, **409**: 603–607. doi:10.1038/35054537.

Lovejoy, C., Massana, R., and Pedrós-Alió, C. 2006. Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. Applied and Environmental Microbiology, **72**: 3085–3095. doi:10.1128/AEM.72.5.3085-3095.2006.

Marquardt, M., Vader, A., Stübner, E.I., Reigstad, M., and Gabrielsen, T.M. 2016. Strong seasonality of marine microbial eukaryotes in a high-Arctic fjord (Isfjorden, in West Spitsbergen, Norway). Applied and Environmental Microbiology, **82**: 1868–1880. doi:10.1128/AEM.03208-15.

Massana, R., Pernice, M., Bunge, J.a., and del Campo, J. 2011. Sequence diversity and novelty of natural assemblages of picoeukaryotes from

the Indian Ocean. The ISME Journal, **5**: 184–195. doi:10.1038/ismej.2010.104.
McMurdie, P.J., and Holmes, S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE, **8**: e61217.
McMurdie, P.J., and Holmes, C. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Computational Biology, **10**: e1003531.
Nagata, T., Tamburini, C., Arístegui, J., Baltar, F., Bochdansky, A.B., Fonda-Umani, S., et al. 2010. Emerging concepts on microbial processes in the bathypelagic ocean–ecology, biogeochemistry, and genomics. Deep Sea Research Part II: Topical Studies in Oceanography, **57**: 1519–1536. doi:10.1016/j.dsr2.2010.02.019.
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. 2021. vegan: community ecology package. R package version 2.5-7.
Ollison, G.A., Hu, S.K., Mesrop, L.Y., DeLong, E.F., and Caron, D.A. 2021. Come rain or shine: depth not season shapes the active protistan community at station ALOHA in the North Pacific Subtropical Gyre. Deep Sea Research Part I: Oceanographic Research Papers, **170**: 103494. doi:10.1016/j.dsr.2021.103494.
Orr, R.J., Murray, S.A., Stüken, A., Rhodes, L., and Jakobsen, K.S. 2012. When naked became armored: an eight-gene phylogeny reveals monophyletic origin of theca in dinoflagellates. PLoS ONE, **7**: e50004. doi:10.1371/journal.pone.0050004.
Paulsen, M.L., Doré, H., Garczarek, L., Seuthe, L., Müller, O., Sandaa, R.—A., et al. 2016. Synechococcus in the Atlantic Gateway to the Arctic Ocean. Frontiers in Marine Science, **3**. doi:10.3389/fmars.2016.00191.
Pedersen, T.L., and Shemanarev, M. 2022. ragg: graphic devices based on AGG. R package version 1.2.2.
Pernice, M.C., Giner, C.R., Logares, R., Perera-Bel, J., Acinas, S.G., Duarte, C.M., et al. 2016. Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. The ISME Journal, **10**: 945–958. doi:10.1038/ismej.2015.170.
Piredda, R., Tomasino, M.P., D'Erchia, A.M., Manzari, C., Pesole, G., Montresor, M., et al. 2017. Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. FEMS Microbiology Ecology, **93**: fiw200. doi:10.1093/femsec/fiw200.
Poulin, R., and Morand, S. 2000. The diversity of parasites. The Quarterly Review of Biology, **75**: 277–293. doi:10.1086/393500.
Randelhoff, A., Reigstad, M., Chierici, M., Sundfjord, A., Ivanov, V., Cape, M., et al. 2018. Seasonality of the physical and biogeochemical hydrography in the inflow to the Arctic Ocean through Fram Strait. Frontiers in Marine Science, **5**: 224. doi:10.3389/fmars.2018.00224.
Sandaa, R.—A., Storesund, J.E., Olesin, E., Paulsen, M.L., Larsen, A., Bratbak, G., and Ray, J.L. 2018. Seasonality drives microbial community structure, shaping both eukaryotic and prokaryotic host—viral relationships in an Arctic marine ecosystem. Viruses, **10**: 715, doi:10.3390/v10120715.
Sassenhagen, I., Irion, S., Jardillier, L., Moreira, D., and Christaki, U. 2020. Protist interactions and community structure during early autumn in the Kerguelen region (Southern Ocean). Protist, **171**: 125709. doi:10.1016/j.protis.2019.125709.
Savage, R.—L., Maud, J.L., Kellogg, C.T.E., Hunt, B.P.V., and Tai, V. 2023. Symbiont diversity in the eukaryotic microbiomes of marine crustacean zooplankton. Journal of Plankton Research, **45**: 338–359. doi:10.1093/plankt/fbad003.
Sehein, T.R., Gast, R.J., Pachiadaki, M., Guillou, L., and Edgcomb, V.P. 2022. Parasitic infections by Group II Syndiniales target selected dinoflagellate host populations within diverse protist assemblages in a model coastal pond. Environmental Microbiology, **24**: 1818–1834. doi:10.1111/14622920.16013.
Seuthe, L., Rokkan Iversen, K., and Narcy, F. 2011. Microbial processes in a high-latitude fjord (Kongsfjorden, Svalbard): II. Ciliates and dinoflagellates. Polar Biology, **34**: 751–766. doi:10.1007/s00300-010-0930-9.
Skovgaard, A. 2014. Dirty tricks in the plankton: diversity and role of marine parasitic protists. Acta Protozoologica, **53**. doi:10.4467/16890027AP.14.006.1443.
Skovgaard, A., Massana, R., Balague, V., and Saiz, E. 2005. Phylogenetic position of the copepodinfesting parasite Syndinium turbo (Dinoflagellata, Syndinea). Protist, **156**: 413–423. doi:10.1016/j.protis.2005.08.002.
Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, **30**: 1312–1313. doi:10.1093/bioinformatics/btu033.
Strassert, J.F.H., Karnkowska, A., Hehenberger, E., del Campo, J., Kolisko, M., Okamoto, N., et al. 2018. Single cell genomics of uncultured marine alveolates shows paraphyly of basal dinoflagellates. The ISME Journal, **12**: 304–308. doi:10.1038/ismej.2017.167.
Terrado, R., Vincent, W.F., and Lovejoy, C. 2009. Mesopelagic protists: diversity and succession in a coastal Arctic ecosystem. Aquatic Microbial Ecology, **56**: 25–39. doi:10.3354/ame01327.
Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. 2015. Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science, **348**: 1261605. doi:10.1126/science.1261605.
Vaulot, D. 2022. pr2database. R package version 4.14.0.
Vaulot, D., Sim, C.W.H., Ong, D., Teo, B., Biwer, C., Jamy, M., and Lopes dos Santos, A. 2022. metaPR2: a database of eukaryotic 18S rRNA metabarcodes with an emphasis on protists. Molecular Ecology Resources, **22**: 3188–3201, doi:10.1111/1755-0998.13674.
Wang, L.—G., Lam, T.T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., et al. 2020. treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. Molecular Biology and Evolution, **37**: 599–603. doi:10.1093/molbev/msz240.
Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology, **73**: 5261–5267. doi:10.1128/AEM.00062-07.
Wang, Y., Naumann, U., Eddelbuettel, D., Wilshire, J., and Warton, D. 2022. mvabund: statistical methods for analysing multivariate abundance data. R package version 4.2.1.
Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York.
Wilson, B., Müller, O., Nordmann, E.—L., Seuthe, L., Bratbak, G., and Øvreås, L. 2017. Changes in marine prokaryote composition with season and depth over an arctic polar year. Frontiers in Marine Science, **4**: 1–17. doi:10.3389/fmars.2017.00095.
Worden, A.Z., Follows, M.J., Giovannoni, S.J., Wilken, S., Zimmerman, A.E., and Keeling, P.J. 2015. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. Science, **347**. doi:10.1126/science.1257594.
Xu, D., Jiao, N., Ren, R., and Warren, A. 2017. Distribution and diversity of microbial eukaryotes in bathypelagic waters of the South China Sea. Journal of Eukaryotic Microbiology, **64**: 370–382. doi:10.1111/jeu.12372.
Yoo, J., Kim, S., and Coats, D.W. 2024. Morphology of Hobagella saltata n. gen. and n. sp. (Syndiniophyceae, Miozoa) infecting the marine dinoflagellate *Cucumeridinium coeruleum* (Dinophyceae, Miozoa) and its potential onshore advection. Frontiers in Marine Science, **10**: 1296836. doi:10.3389/fmars.2023.1296836.