The Journal of Eukaryotic Microbiology

ISOP Published by the International Society of Protistologists

Journal of Eukaryotic Microbiology ISSN 1066-5234

ORIGINAL ARTICLE

Protist Diversity and Seasonal Dynamics in Skagerrak Plankton Communities as Revealed by Metabarcoding and Microscopy

Sandra Gran-Stadniczeñko[a], Elianne Egge[a], Vladyslava Hostyeva[b], Ramiro Logares[c], Wenche Eikrem[a,b] & Bente Edvardsen[a]

a Department of Biosciences, University of Oslo, P. O. Box 1066 Blindern, 0316, Oslo, Norway
b Norwegian Institute for Water Research, Gaustadalléen 21, 0349, Oslo, Norway
c Department of Marine Biology and Oceanography, Institut de Ciències del Mar (CSIC), 08003, Barcelona, Catalonia, Spain

Keywords
Biovolume; high-throughput sequencing; Oslofjorden; richness.

Correspondence
B. Edvardsen, Department of Biosciences, University of Oslo, P. O. Box 1066 Blindern, 0316 Oslo, Norway
Telephone number: +47-22-85-70-38;
FAX number: +47-22-85-47-26;
e-mail: bente.edvardsen@ibv.uio.no

Received: 30 April 2018; revised 15 October 2018; accepted October 18, 2018.
Early View publication November 29, 2018

doi:10.1111/jeu.12700

ABSTRACT

Protist community composition and seasonal dynamics are of major importance for the production of higher trophic levels, such as zooplankton and fish. Our aim was to reveal how the protist community in the Skagerrak changes through the seasons by combining high-throughput sequencing and microscopy of plankton collected monthly over two years. The V4 region of the 18S rRNA gene was amplified by eukaryote universal primers from the total RNA/cDNA. We found a strong seasonal variation in protist composition and proportional abundances, and a difference between two depths within the euphotic zone. Highest protist richness was found in late summer-early autumn, and lowest in winter. Temperature was the abiotic factor explaining most of the variation in diversity. Dinoflagellates was the most abundant and diverse group followed by ciliates and diatoms. We found about 70 new taxa recorded for the first time in the Skagerrak. The seasonal pattern in relative read abundance of major phytoplankton groups was well in accordance with microscopical biovolumes. This is the first metabarcoding study of the protist plankton community of all taxonomic groups and through seasons in the Skagerrak, which may serve as a baseline for future surveys to reveal effects of climate and environmental changes.

PROTISTS are unicellular and multicellular algae and protozoans with a wide range of ecological functions (Massana 2015). Microalgae play key roles in coastal ecosystems contributing significantly to carbon flux through the microbial loop (Not et al. 2012), and are the main suppliers of photosynthetic products that higher trophic levels of the marine food web depend upon. Protists are morphologically and genetically diverse, and are present in all types of marine habitats (Massana 2015). Phytoplankton communities on continental shelves are dominated in biomass by diatoms, dinoflagellates, and haptophytes (Simon et al. 2009). In temperate seas, community composition and abundance undergo strong seasonal changes as a result of alterations in abiotic factors, such as irradiance, temperature and nutrient levels, and biotic factors, such as grazing, pathogens, and competition.

The Skagerrak, off the coasts of Norway, Sweden and Denmark, undergoes strong seasonal environmental variations due to changes in meteorological and hydrological conditions, and irradiance. The balance of hydrological forces from brackish Baltic currents, saline North Atlantic currents, and land runoff lead to considerable salinity and temperature fluctuations and seasonal water column stratification. The water currents also bring in allochtonous plankton, which further contribute to a species rich phytoplankton community in this area (Andersen et al. 2001). In addition, variation in nutrient availability and grazing pressure cause inter-annual variations in the protist species composition (Braarud et al. 1953) with different environmental preferences. The Outer Oslofjorden monitoring location in the Skagerrak is considered to represent the Southern Norwegian

coastal waters (Braarud and Bursa 1939; Dragsund et al. 2006).

Studies on protist taxonomic composition in the Skagerrak area have been carried out for over a century with a focus on diversity and dynamics based on light-, electron-, and epifluorescence microscopy and flow cytometry (Backe-Hansen and Throndsen 2002; Braarud et al. 1953; Bratbak et al. 2011; Dittami et al. 2013; Hasle and Smayda 1960; Hjort and Gran 1900; Kuylenstierna and Karlson 1994). These studies have revealed the dynamics and distribution of organisms belonging to different trophic (auto-, mixo- and heterotrophs) and taxonomic groups such as dinoflagellates, diatoms, haptophytes, cryptophytes, prasinophytes, dictyochophytes, and euglenoids. The overall seasonal pattern that has emerged can be described as follows: Low protist abundances are found in the Outer Oslofjorden during winter due to constant mixing of water masses combined with low solar irradiance (Dittami et al. 2013). An increase in irradiance and heat, together with brackish water inputs from the Baltic Current and river run-offs lead to water stratification in early spring, in February–March. Stratification leads to improved light conditions in the upper mixed layer which triggers the first spring bloom dominated by diatoms (mainly *Skeletonema*, *Thalassiosira*, *Chaetoceros, Pseudo-nitzschia* spp.), where nutrients are supplied from bottom waters (Paasche and Østergren 1980). A second bloom dominated by diatoms may occur in May–June with river run-offs as nutrient source. Strong summer stratification in July–August limits the transport of nutrients from deep waters to the upper water column, resulting in relatively low phytoplankton biomass and a dominance of mixotrophic and heterotrophic flagellates, including dinoflagellates and haptophytes. A third, smaller bloom may occur in August-September, when decreased stratification and wind mixing bring up nutrients to the upper, photic zone. Finally, heavy storms and a decrease in irradiance and temperature occur in late autumn forcing a decline in the general protist community (Braarud and Bursa 1939). The aforementioned microscopy studies were, however, limited to small water volumes (up to 50 ml) and identification to species level of cells larger than ca. 20 μm. Thus, little is yet known about seasonal dynamics of smaller, fragile, or less abundant protists.

New molecular techniques have proven to be an indispensable tool to examine the marine microbial diversity (Medlin and Kooistra 2010) to overcome the limitations of traditional methods, for example, microscopy. They have revealed the existence of an immense variety of novel protists (Epstein and López-García 2008) without the need for isolation or culturing (Medlin and Kooistra 2010). The small subunit (SSU) 18S rRNA gene is the most widely used marker to detect and classify known species present in marine eukaryotic microbial communities and to assess the phylogenetic affiliations of unknown sequences (see López-García et al. 2001). Recently, studies targeting the haptophytes in the Outer Oslofjorden with high-throughput sequencing have elucidated a vast diversity in a greater detail than has previously been obtained by microscopy (Egge et al. 2015a,b; Gran-Stadniczeñko et al. 2017).

Here, we investigate how the protist plankton community in the Skagerrak changes through the seasons by combining high-throughput sequencing (HTS) of the V4 region of the 18S rRNA gene and microscopy analyses of samples taken monthly over two years. We addressed the following questions: (i) How do HTS-inferred community composition and relative abundance change with season and depth? (ii) Which are the main abiotic drivers for these changes? (iii) What is the proportion of heterotrophic and autotrophic protists through the seasons? (iv) Which species dominate in the HTS dataset and what are their seasonal distributions? (v) Does HTS reveal taxa not previously recorded in the area, or taxa novel to science? (vi) How do HTS results compare to microscopy observations? Here, we reveal novel diversity not previously recorded in the Skagerrak, and how major protist components occur through the year. This study also contributes to a better understanding of protist plankton community structure and dynamics.

MATERIALS AND METHODS

Sampling

The sampling was performed as previously described in Egge et al. (2015a,b). Twenty-one coastal sampling campaigns were performed at the OF2 monitoring station (59.17 N, 10.69 E) located in the Outer Oslofjorden, Northern Skagerrak on board R/V Trygve Braarud. Samplings were conducted monthly for 2 yr, between September 2009 and June 2011 (with exception of February 2010 when samples and measurements were collected by the Ferrybox ships of opportunity, due to ice coverage) within the HAPTODIV project. Samples from September 2009 and June 2010 were also parts of the EU project BioMarKs (www.biomarks.org).

A conductivity-temperature-depth sensor (CTD, Falmouth Scientific Inc., Cataumet, MA) attached to a Niskin bottle rosette was used to obtain physicochemical water column profiles (temperature, conductivity/salinity, depth and fluorescence) from 1 to 100 m depth. Niskin bottles were used to collect water samples for nutrients (N, P, Si and Tot-P) and Chlorophyll-a (Chl-a) analysis at eight different depths (1, 2, 4, 8, 12, 16, 20, and 40 m). Water samples for nutrient analysis were frozen and stored in 20-ml scintillation vials until analysed in an autoanalyzer (Bran Luebbe Autoanalyzer 3). For Chl-a analyses, 100–500 ml water from each depth were filtered onto glass-fibre filters (Whatman GF/F, 25 mm, c. 0.8 mm mesh size), transferred to 2-ml cryotubes and frozen in liquid N_2 at −196 °C. Filters were incubated in 10 ml 90% acetone for 30–60 min and Chl-a was fluorometrically quantified with a Turner Designs fluorometer TD-700 (Turner Designs, Sunnyvale, CA) as described by Strickland and Parsons (1972).

Protist communities were collected by filtration onboard ship. At each sampling occasion, 20 liters of sea water was collected with 5 liters Niskin bottles at two different

depths: subsurface (1 m) and the depth for the bottom of the deep chlorophyll maximum (DC) when present, which was determined by visual inspection of the fluorescence on the CTD plots and tables. When no chlorophyll peak was observed, the depth for DC samples was 8 m. To remove large plankton, a prefiltration step was performed through a 45 μm nylon mesh into hydrochloric acid-washed plastic carboys. Protist cells were then collected by fractionated filtration with a peristaltic pump (Masterflex 07523-80; Cole-Parmer, Vernon Hills, IL) at a rate of 0.5–1 l/min, through 142 mm diameter polycarbonate filters (Millipore, Billerica, MA) with pore sizes of 3 and 0.8 μm, in a line giving the size fractions 45–3 μm (nano) and 3–0.8 μm (pico) plankton. To minimise RNA degradation, filtration was conducted for maximum 40 min. Finally, filters were cut in four and each piece was transferred into a 5-ml cryotube, which was frozen in liquid N_2 onboard ship, and stored at −80 °C. During the BioMarKs sampling in September 2009 and June 2010, prefiltration was performed at 20 μm giving a nano size fraction of 3–20 μm.

High-throughput sequencing

Total RNA was extracted and amplified as described in Egge et al. (2015a) using RNA NucleoSpin II (Macherey-Nagel, Düren, Germany). From each sample, ½ of the filter was extracted (representing a 10 liters water sample). Sixty microliters of RNA eluate was obtained and concentration was checked with a NanoDrop spectrophotometer (Wilmington, DE). Standard PCR with universal eukaryote partial 18S rRNA gene primers 1F and 300R (see Edvardsen et al. 2003) was performed to check for residual DNA in the RNA eluates. DNase (TURBO DNA-freeTM kit, Ambion, Austin, TX) treatment was performed with the samples where PCR products were observed by gel electrophoresis, as described in the manufacturer's protocol. To reverse-transcribe the RNA to cDNA, the High-Fidelity first Strand cDNA Synthesis Kit (Agilent, Santa Clara, CA) with random primers was used according to the manufacturer's protocol. In the synthesis reaction, approx. 100 ng of RNA per sample was used. Samples from the BioMarKs project (September 2009 and June 2010) were prepared as specified in Logares et al. (2012). PCR amplification of cDNA was done using the eukaryote specific primers by Stoeck et al. (2010) TAReuk454FWD1 (5′-CCAGCA(G/C)C(C/T)GCGGTAATTCC-3′) and TAReukREV3 (5′-ACTTTCGTTCTTGAT(C/T)(A/G)A-3′), adapted for 454-pyrosequencing. The forward primer contained a sample specific tag (MIDs). PCR was conducted in four separate reactions per sample on an Eppendorf thermocycler (Mastercycler, ep gradient S, Eppendorf). The PCR mixtures (25 μl) contained 5 μl 5× Phusion GC buffer, 0.5 μl of dNTP at a concentration of 10 μM, 0.75 μl of DMSO, 1 μl of each primer at a concentration of 10 μM, 0.25 μl of polymerase (Phusion, Finnzymes, Vantaa, Finland), 1 μl of template cDNA (10–60 ng/μl) and 15.5 μl sterilised PCR water. The PCR-programme included an initial denaturation at 98 °C for 30 s, followed by 30 cycles (denaturation at 98 °C for 10 s, annealing at 55 °C for 30 s, extension at 72 °C for 30 s) and a final extension at 72 °C for 10 min. Pooled PCR reactions were then purified with AMPure beads (BeckmanCoulter, Brea, CA), quantified with NanoDrop and pooled to obtain equal concentrations for sequencing. The samples were prepared for sequencing with Lib-L chemistry and sequenced unidirectionally from the forward primer on ½ of a 454 life sciences GS-FLX Titanum sequencing plate (454 Life Sciences, Branford, CT) at the Norwegian Sequencing Centre at the Department of Biosciences, University of Oslo (http://www.sequencing.uio.no). Raw SFF sequence files were deposited to GenBank under the project number PRJNA497792.

Bioinformatic pipeline

AmpliconNoise v.1.6.0 (Quince et al. 2011) was used to denoise the 454 reads, which were truncated at 400 bp. Reads with > 8 bp homopolymers and/or presenting mismatches in barcode or primers were removed. Perseus (incorporated in AmpliconNoise) was used to identify and remove putative chimeras. Some chimeras were also found by manual inspection by BLASTn against the NCBI nucleotide database and excluded. Clustering and taxonomic assignation of reads were performed with the "pick_open_reference_otus.py" command implemented in QIIME v.1.9.1 (Caporaso et al. 2010). UCLUST v. 1.2.22 (Edgar 2010) was used to cluster the reads into OTUs with 98% sequence identity. An initial taxonomical assignment was performed against the Protist Ribosomal Reference Database (PR2 v.1.0.0, including only sequences from cultures and longer than 800 bp; Guillou et al. 2013, https://github.com/vaulot/pr2database) at > 90% similarity, using the parameter "pick_open_reference_otus.py". Subsequent taxonomic assignments were done with all OTUs that did not initially match any phylum, using BLASTn within the software Geneious (v10.2.2) against the PR2 and then the NCBI databases. By manual BLAST, some of them were found to be chimeras and thus removed. All OTUs assigned to metazoans were removed from the data set. OTUs with less than 10 reads were excluded from the dataset, to remove possible spurious diversity. Scripts for the bioinformatics pipeline in Qiime and statistical analyses in R are found in File S1.

Phylogenetic analyses

The 16 most abundant OTUs (> 1% of total reads) were taxonomically placed by the EUKREF RAxML-EPA (Evolutionary Placement Algorithm) pipeline (del Campo, pers. commun., Berger et al. 2011; Stamatakis 2014) for a more reliable taxonomic assignation. Reference sequences of Gymnodiniales (Dinophyta), Geminigeraceae (Cryptophyta), Mamiellaceae (Chlorophyta), Mediophyceae (Bacillariophyceae), Chrysochromulinaceae and Noelaerhabdaceae (Haptophyta) were selected from the PR2 database, and Stephanoencidae (Choanoflagellata) from NCBI, and then aligned by MAFFT- E-INS-I v.7.300 (Katoh et al. 2009). Phylogenetic analyses using RAxML v.8.0.26 (Stamatakis

2014) based on reference sequences were performed implementing GTRGAMMA model with 100 bootstrap runs. OTUs were aligned to reference sequences by MAFFT -addfragments and added to the reference RAxML best tree with raxmlHPC-PTHREADS-SSE3 using GTRCATI. The analyses were conducted on the Abel cluster at the University of Oslo. Scripts for the phylogenetic analyses are found in File S1.

Identification of novel taxa

To assess new records for the area we compared the taxonomic assignments of the OTUs (to $\geq$ 90% similarity in QIIME) in this study with species lists in the Norwegian Species Information Centre (Artsnavnebasen at Artsdatabanken 2018, http://www2.artsdatabanken.no/artsnavn/Contentpages/Sok.aspx) and the Nordic Microalgae and Aquatic Protozoa Checklist in Sweden and Norway (http://nordicmicroalgae.org). In addition, we checked all novel taxa by manual Blast against NCBI to verify the taxonomic assignation.

Statistical analyses

All statistical analyses and figures were performed in R software v.3.4.1 (R Development Core Team 2017). Treemap plots representing the complete protist community composition at the OF2 station during the 2 years were created based on read abundance and OTU richness, with the treemap package (Tennekes 2017). The Vegan package (Oksanen et al. 2017) was used in all diversity analyses. To compare the communities in the different samples, the dataset was normalised to equal sample sizes by rarefying (i.e. subsampling) using the "rrarefy" function, each of the 82 samples to the lowest number of reads found in a single sample (998 reads). As some OTUs occur in both nano- and pico-size fraction samples, the data from the two size fractions within a sample were pooled after subsampling to give 41 samples. Richness (number of OTUs per sample), proportional abundances and the Shannon diversity index H′ (Shannon 1948) determined by the "diversity" function in R, were used to investigate the seasonal variation in the community structure at the two studied depths. Nonparametric generalised additive model (GAM) was used to fit monthly linear diversity time trends with the "gam" function from the "mgcv" package. To test if the two studied depths were significantly different with respect to richness and diversity, Welch Two Sample *t*-test was applied. Bray–Curtis distances (Bray and Curtis 1957) were generated and used to produce a dissimilarity matrix based on OTU presence-absence data. Non-Metric Multidimensional Scaling (NMDS) analyses based on the dissimilarity matrix were performed to explore community patterns applying the monoMDS function. ANOSIM (Analysis of similarity) were used to test differences in composition between seasons. To analyse the correlations between environmental factors and community changes, canonical correspondence analysis (CCA), Mantel test and PERMANOVA (Permutational multivariate analyses of variance) were conducted. Similarity percentage analysis (SIMPER) was performed with the "simper" function to identify the OTUs that drove most of the differences in seasonal assemblages. To compare relative read abundance obtained by HTS with relative biovolume measured by microscopy of specific taxonomic groups (Bacillariophyceae, Chrysophyceae, Dictyochophyceae, Dinophyceae and Euglenophyceae), Welch Two Sample *t*-test was applied.

Microscopy

Total water samples (100 ml) were dispensed into flasks directly from the Niskin bottles and preserved in Lugol's solution (1% final concentration, Throndsen et al. 2007). The phytoplankton cell concentrations were determined in a 10-ml sub-sample that was allowed to settle overnight and subsequently counted in an inverted microscope (Nikon Diaphot 300; Nikon Corporation, Tokyo, Japan) in accordance with the method of Utermöhl (1958). Qualitative inspections were also made on vertical (0–20 m depth) and horizontal phytoplankton net samples (mesh size 10 μm) preserved with Lugol's solution (1% final conc.). Phytoplankton taxa were identified to the lowest level possible with light microscopy (LM) according to Throndsen et al. (2007). Biovolumes were estimated from cell counts using the HELCOM 2006 protocol (Olenina et al. 2006).

RESULTS

The outer Oslofjorden is a dynamic locality with respect to hydrographical conditions and protist composition and abundance. Here, we examined the community structure of the eukaryotic pico- and nano-plankton (passing a nylon sieve with 45 μm mesh size) at a monitoring station (OF2) during 2 years (2009–2011) with monthly samplings, and at two depths. This is the first paper on seasonal dynamics of the total planktonic protist community in the Oslofjorden using metabarcoding and microscopy.

Seasonal variations of environmental factors

The physicochemical parameters temperature, salinity, density, and chlorophyll fluorescence at the OF2 station in the upper water column (0–40 m) showed seasonal variations as shown in Table 1, Fig. S1 and File S2, and previously described by Egge et al. (2015b) and Dittami et al. (2013). The chlorophyll *a* concentration was highest in the upper 4 m of the water column at all times and usually higher in 1 and 2 m than at 4 m. To compare whether there was a difference in the small protist community composition and structure within the well-lit eutrophic zone, a sample at 1 m depth (subsurface, SS) and bottom of the chlorophyll fluorescence peak, here called "deep chlorophyll" and shortened DC were sampled and analysed. The depth for the DC samples ranged between 5 and 22 m depth. The hydrographical conditions in the upper 40 m, including the upper mixed layer and the

Table 1. Physical and chemical conditions at the OF2 station in the sampling depths subsurface (1 m) and bottom of the deep chlorophyll maximum (DC) in the sampling period September 2009 to June 2011

		Temperature (°C)		Salinity (PSU)		Density (σt)		Chl-a (μg/l)		N [μM]		Si [μM]		P [μM]		Tot-P [μM]	
Sample	Depth (m) DC	1 m	DC	1 m	DC	1 m	DC	1 m	DC	1 m	DC	1 m	DC	1 m	DC	1 m	DC
Sep. 09	20	15.47	16.05	25.25	28.68	18.38	20.97	2.76	0.9	0.66	0.94	2.08	2.83	0.06	0.03	0.69	0.66
Oct. 09	5	11.83	13.1	31.41	32.79	23.86	24.71	0.6	0.3	4.44	4.17	5.47	3.93	0.11	0.1	0.81	0.82
Nov. 09	14	6.94	7.5	19.11	21.92	13.03	15.57	0.13	2.5	5	1.53	3.91	4.15	0.48	0.42	0.98	1.39
Dec. 09	8	5.6	6.37	22.93	26.62	18.08	20.94	0.62	0.48	8.92	4.43	14.71	5.2	0.45	0.45	1.13	0.96
Jan. 10	12	−1.25	1.27	24.2	29.46	19.42	23.63	7.71	2.95	0.64	4.43	1.05	2.82	0.45	0.26	0.94	0.93
Feb. 10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Mar. 10	12	5.25	5.78	32.75	33.11	25.87	26.14	2.49	1.95	5.43	6.28	5.21	4.59	0.62	0.7	0.96	0.88
Apr. 10	6	5.76	4.1	24.83	27.72	19.56	22.03	2.33	1.33	2.93	1.09	6.62	0.57	0.19	0.16	0.23	0.21
May. 10	10	7.65	6.36	26.3	29.05	20.5	22.86	1.63	0.98	1.13	0.75	2.72	1.48	0.06	0.13	0.39	0.44
Jun. 10	10	14.93	11.8	22.16	29.71	16.12	22.57	0.38	0.35	0.14	0.16	3.03	1.41	0.11	0.07	0.36	0.32
Aug. 10	20	18.44	14.12	23.31	31.74	16.25	23.73	1.85	0.19	0.14	1.68	1.17	2.42	0.06	0.16	0.26	0.39
Sep. 10	20	15.76	15.76	25.12	25.13	18.22	18.31	1.44	1.38	0.14	0.14	0.72	0.66	0.06	0.08	0.29	0.32
Oct. 10	8	10.35	11.14	25.1	26.52	19.2	20.19	1.58	0.65	2.75	2.19	2.8	2.99	0.15	0.13	0.22	0.24
Nov. 10	8	8.09	8.96	29.87	30.52	23.24	23.66	0.43	0.1	5.48	2.59	4.86	2.75	0.28	0.23	0.31	0.26
Dec. 10	8	0.31	3.52	28.37	30.96	22.74	24.66	0.59	0.57	3.64	2.57	5.18	3.83	0.81	0.48	9.52	4.04
Jan. 11	8	0.69	1.79	29.12	29.78	23.34	23.84	0.89	0.78	3.71	2.28	6.55	4.55	0.39	0.29	0.74	0.61
Feb. 11	15	0.56	2.9	29.01	31.84	23.22	25.45	5.42	1.32	0.36	5.43	0.32	4.46	0.26	0.45	0.17	0.71
Mar. 11	22	−0.26	2.12	24.97	31.06	20.01	24.93	0.8	0.19	0.43	4.93	1.17	3.08	0.48	0.74	0.48	0.81
Apr. 11	8	5.12	3.61	16.1	20.95	10.82	14.42	0.36	0.7	6.04	3.48	5.74	0.67	0.49	0.16	3.92	1.36
May. 11	10	12.32	11.86	17.84	20.58	13.25	15.51	0.49	0.46	1.19	0.87	6.17	5.23	0.25	0.12	0.47	0.26
Jun. 11	10	16.06	10.85	19.31	28.04	13.71	21.43	1.9	0.37	1.71	1.87	13.15	2.59	0.17	0.11	0.3	0.14

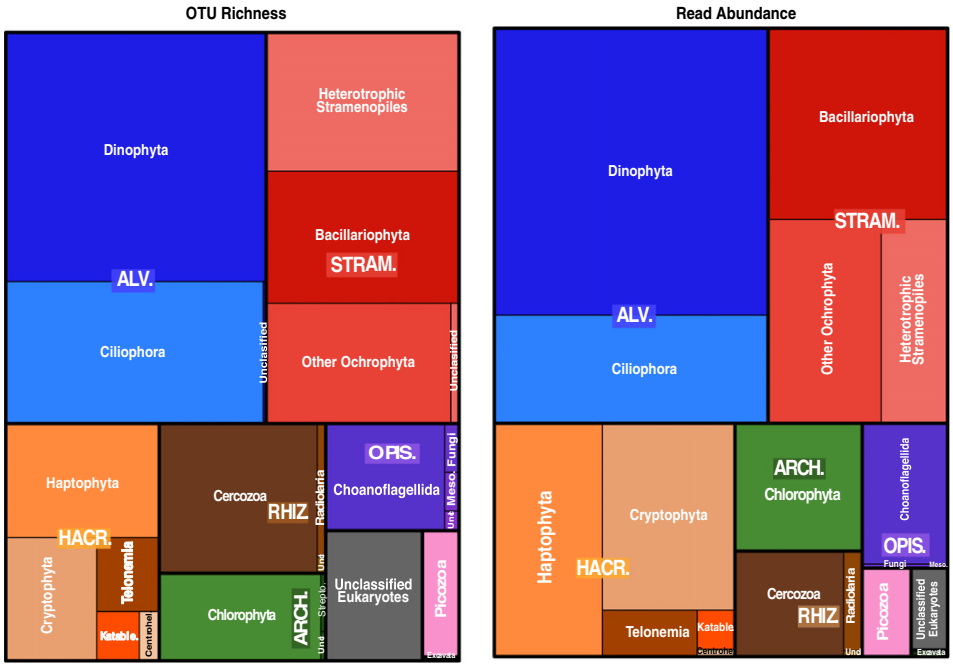

Figure 1 Tree maps displaying the taxonomic composition of the complete Outer Oslofjorden OF2 station protist HTS dataset at supergroup and phylum levels: OTU richness (left) and proportional read abundance (right) of supergroups Alveolata (ALV.), Stramenopila (STRAM.), Hacrobia (HACR.), Rhizaria (RHIZ.), Archeplastida (ARCH.), Opistokonta (OPIS.), Excavata and unclassified groups.

pycnocline, presented strong fluctuations. The seawater temperature increased during spring and summer up to 18.4 °C and decreased during autumn and winter to a minimum of −1.2 °C. An opposite pattern was observed for salinity and density. Highest values were found in winter or early spring with salinities up to 32.8, whereas lowest salinities were registered during late-spring and summer with minimum 16.1. Temporal patterns were also found in the Chl-*a* concentrations with highest values (up to 7.7 µg/l) during the main spring-bloom in late January 2010 and in February 2011. Concentrations of inorganic nutrient peaked during winter.

Taxonomic composition and relative abundance

After initial filtration of reads in the QIIME pipeline, including denoising by AmpliconNoise, we obtained 670,886 reads with average fragment length 375 bp, ranging between 184 and 400 bp. A second filtering step (removal of chimeras by Perseus, metazoan OTUs and OTUs with less than 10 reads in the whole dataset), resulted in 613,031 reads assigned to 2,032 OTUs (File S3). The taxonomic classification and absolute number of reads per OTU in each sample are presented in Table S1. Of the total OTUs, 1,791 were rare, with < 0.05% of the reads per OTU, while 44 OTUs were typically abundant, with > 0.5% of the reads per OTU, comprising 13.8% and 55.6% of the total reads, respectively. Most (95%) of the OTUs could be taxonomically assigned to one of 18 major micro-eukaryotic taxonomic taxa (superphylum to subphylum; Fig. 1). The remaining 5% were assigned as unclassified eukaryotes.

The infrakingdom Alveolata dominated the communities both in richness (36% of OTUs) and abundance (41% of reads). All alveolate OTUs except three were classified to a phylum: dinoflagellates or ciliates. Dinoflagellates were the most abundant phylum within alveolates, accounting for 67% of the total alveolate reads. Gymnodiniales was the most abundant order within dinoflagellates and the only order found in all samples (Fig. S2). Abundant taxa within the order Gymnodiniales were *Karenia* spp., *Karlodinium* sp., *Lepidodinium* sp., *Gyrodinium helveticum* and *Akashiwo sanguinea*. The second most abundant dinoflagellate group was Syndiniales, divided into the clades MALV I–V (marine alveolates without a cultured representative). MALV clades I–III were more abundant (> 0.5% of reads) than MALV IV and V (< 0.1%). Besides that, a few reads were assigned to Dinophysiales, Gonyaulacales, Noctilucales, Peridiniales, Prorocentrales, Suessiales, and Thoracosphaerales. Ciliates were both diverse (12.9% of total OTUs) and abundant, representing 33% of the alveolate reads (Fig. 1). Spirotrichea was the most represented class within ciliates (Fig. S2). The five most abundant ciliate OTUs had best match to the family Strombidiidae (Table S1).

Stramenopiles were the second most OTU-rich (26% of total OTUs) and abundant (20% reads) high rank group (subkingdom) after the alveolates (Fig. 1). Stramenopile OTUs were found in all samples, with highest abundances observed during spring. Two thirds were assigned to phylum Ochrophyta, and one-third was assigned to entirely

heterotrophic stramenopile phyla. Diatoms (Bacillariophyta) were the most diverse and abundant stramenopile group (36% of the total stramenopile reads) and were found during the entire sampling period (Fig. S2). The most important diatoms were the centric *Skeletonema marinoi*, *Thalassiosira nordenskioeldii* and *Chaetoceros neogracilis*. Other abundant diatoms were the centric *Chaetoceros debilis, Ch. calcitrans, Minidiscus trioculatus, Eucampia zoodiacus, Brockmanniella brockmannii, Ditylum brightwellii, Porosira pseudodenticulata, Leptocylindrus minimus, L. aporus, Proboscia alata,* and members of the pennate genus *Pseudo-nitzschia* (Table S1). Other ochrophyte groups present in all samples were Dictyochophyceae and Chrysophyceae. The class Dictyochophycea (silicoflagellates) was mainly represented by *Dictyocha speculum* (OTU 6). The next most abundant dictyochophyte was the picoflagellate *Florenciella parvula* present in 80% of the sampling dates, followed by the potentially ichthyotoxic species *Pseudochattonella verruculosa*, which was present in 50% of the sampling dates.

Within the heterotrophic stramenopiles, the most abundant groups were MAST-1, -3 and -7, which consist of marine stramenopiles without a cultured representative. Operational taxonomic units assigned to the heterotrophic stramenopile groups Bicoecea, Labyrinthulea, Oomyceta, Pirsonia, MAST-4, -6, -8, -9, -10, -12 and the phototrophic MOCH (marine ochrophyte without cultured representative) were also present in our dataset (Fig. S2).

The subkingdom "Hacrobia" was also considerably rich and abundant, with haptophytes, cryptophytes and telonemians contributing to 91% of the total Hacrobia reads (Fig. 1). Prymnesiales was the most abundant, frequently detected and diverse haptophyte order. Within this order, OTUs with best match to *Emiliania huxleyi* and *Chrysochromulina simplex* were the most abundant, followed by *Chrysochromulina acantha*. Other abundant haptophytes were assigned to *Prymnesium faveolatum, Imantonia rotunda*, and the bloom-forming species *Phaeocystis pouchetii* (Table S1). Two cryptophytes, *Teleaulax amphioxeia* and *Teleaulax gracilis,* were among the most abundant protists. Of the 29 telonemian OTUs found, only one had match with *Telonema antarcticum*, the rest belonged to unclassified *Telonemia* Group 1 and 2 (Table S1). Members of the heterotrophic phyla Katablepharida and Centroheliozoa were found in low proportions (Fig. S2).

Archaeplastida was primarily represented by chlorophytes (~7.6% total reads). Within Chlorophyta the most abundant and diverse group was Mamiellophyceae represented by 4.6% of all reads. *Micromonas commoda,* a picoflagellate belonging to this class, was among the most abundant OTUs. OTUs assigned to *Micromonas* spp. represented 6.7% of the reads in the pico size fraction. Other major components of the Chlorophyta community belonged to Pycnococcaceae (*Pycnococcus provasolii*), Trebouxiophyceae (*Amphikrikos nanus*), Pyramimonadales (*Pyramimonas* spp., *Pterosperma cristatum*), Nephroselmidophyceae (Fig. S2; *Nephroselmis pyriformis*), which are all pico- or small nanoplankton.

A total of 101 Opisthokonta OTUs were detected. Most of them were identified as choanoflagellates, mainly represented by the order Acanthoecida, which was present in all our samples (Fig. S2). The most abundant opisthokont OTU was placed close to *Calliacantha* spp. in our RAxML phylogeny. An additional BLAST against NCBI showed it to be nearly identical to the sequence KU587842 of *Calliacantha natans,* differing in only two bases, one being in a homopolymer. Five Fungi and four Mesomycetozoa OTUs were also found in low abundances (< 0.2% of total reads).

Rhizarians, mainly cercozoans, were diverse (178 OTUs) and detected in relative high abundances (6% of total reads; Fig. 1). Picozoa was found to be rather diverse (33 OTUs), and quite abundant (~2% reads). They are known as heterotrophic picoplankton with only one described species (*Picomonas judraskeda*). We found an unknown picozoa (OTU 21), differing from the only described species in seven positions, to be among the 25 most abundant. It was however identical to the sequence JN934893 of an uncultured picozoa isolated from Maine, USA. Finally, Excavata was represented by one abundant OTU, with 100% match to *Eutreptiella gymnastica* (accession number KF559331).

Most abundant OTUs

The 16 most abundant OTUs with > 1% of total reads per OTU, were more accurately taxonomically placed by RAxML-EPA. Separate trees for each taxon are presented in Fig. S3. Five of these OTUs were assigned to the dinoflagellate order Gymnodiniales, placed close to *Karenia papillonaceae* (OTU 1), *Karenia/Karlodinium* sp. (OTU 3), *Lepidodinium chlorophorum/L. viride*. (OTU 7), *G. helveticum* (OTU 8) and *A. sanguinea* (OTU 14) respectively. OTU 1 had identical sequence to *K. papillonaceae* whereas OTU 3 differed in one base pair from these, and in one other base pair from *Karlodinium micrum*. Two cryptophyte OTUs were identical to reference sequences of *T. amphioxeia* (OTU 2) and *T. gracilis* (OTU 13). The four most represented stramenopile OTUs had identical sequences to the diatoms *S. marinoi*, *T. nordenskioeldii* and *Ch. neogracilis* (OTUs 4, 9, and 16, respectively) and the dictyochophyte *D. speculum* (OTU 6). The fifth most abundant (OTU 5) was phylogenetically placed closest to *M. commoda*, differing in three base pairs. Three haptophyte OTUs were among the 16 most abundant. OTUs 10 and 15 were identical to *E. huxleyi* and *C. acantha* reference sequences respectively, whereas OTU 12 differed in two base pairs to that of *C. simplex*. The most abundant opisthokont OTU (OTU 11) was phylogenetically placed closest to *C. natans/C. longicaudata,* differing in one base pair from *C. natans*.

Seasonal variation in taxonomic groups at two depths as revealed by HTS

Succession of the 18 major taxonomic groups (from superphylum to subphylum) at the two studied depths through

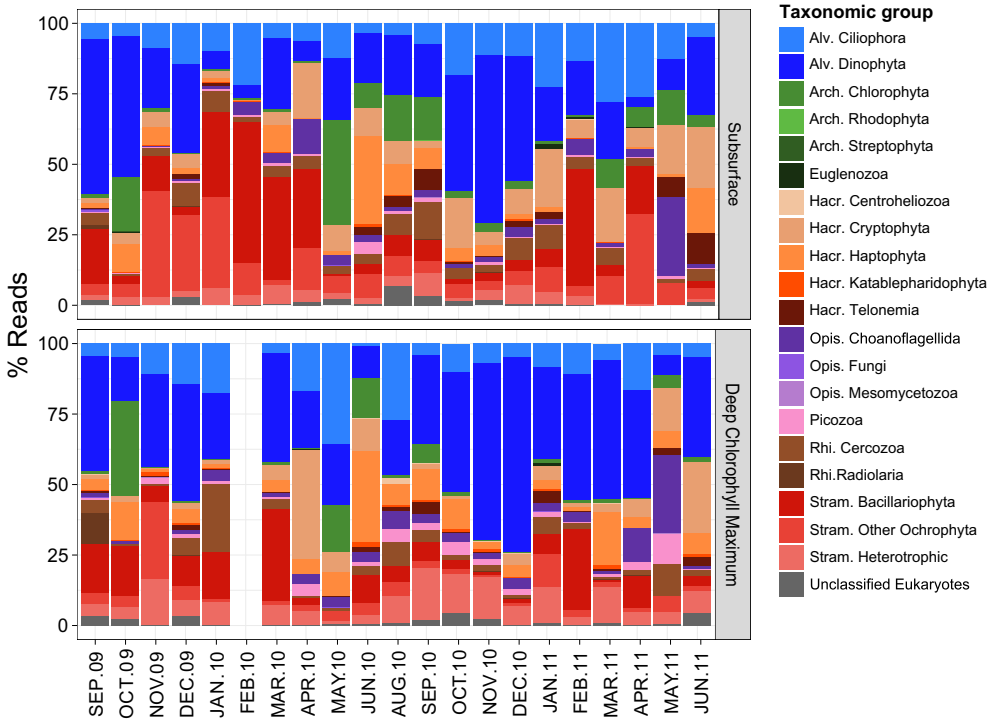

Figure 2 Succession of proportions of reads of the 18 major taxonomic groups across the 21 temporal samples at OF2 station.

the sampling period is shown as proportion of reads in Fig. 2 and OTUs in Fig. S4. All groups were found at both depths, but relative read abundance varied through the year and between the two depths (Fig. 2). There was no clear seasonal trend in proportion of OTU richness of the different groups, nor was there a consistent difference between the depths (Fig. S4). At both depths, alveolates were as a rule the most abundant group during autumn and early winter. Dinoflagellates were usually the most important alveolate group through the year, except for five samples where ciliates showed highest relative abundance (Fig. 2). The proportion of dinoflagellate reads was higher at the DC than at the SS from September 2010 to April 2011. This pattern was not clearly observed the year before. At lower taxonomic levels, reads representing *Lepidodinium* sp. and *A. sanguinea* were more abundant at the SS than at DC. Contrarily, *G. helveticum* and *K. papillonaceae* were more dominant at DC (Fig. 3), which partly explains the relatively high dinoflagellate abundances in March 10 and September 10 to April 11, respectively. The heterotrophic MALV clades I–III were present during the entire study period with few exceptions. MALV IV and V, however, appeared only in a few samples during summer-autumn and in very low abundances (< 0.3%). There were no clear differences in MALV distributions between the two depths.

Stramenopiles varied through the year and had highest proportional abundance during the winter–spring 2010. Their dominance was less pronounced during winter–spring 2011 (Fig. 2). This seasonal trend was more marked at the SS than at the DC. Diatoms and other ochrophytes showed higher proportions at the SS than DC. The phototrophs *D. speculum*, *T. nordenskioeldii* and *Ch. neogracilis* were amongst the most dominant taxa during the winter-spring 2010, and *S. marinoi* during the spring bloom in 2011 (Fig. 3), indicating the importance of these species during the spring blooms. Heterotrophic stramenopiles (mainly MAST OTUs; Fig. S2) and Picozoa showed higher proportions at the DC compared to SS. Members of Picozoa were present on all sampling occasions and depths.

Chlorophytes, haptophytes, and cryptophytes showed highest relative abundance during spring and summer, and similar at both depths (Fig. 2). The heterotrophic groups katablepharids, telonemians, choanoflagellates, and cercozoans were also present during all seasons but in low relative abundances (max. ~5% of reads per group in each sample). Exceptions were the cercozoans and choanoflagellates that showed a peak in January 2010 and May 2011, respectively. The choanoflagellate peak was due to the high proportion of *C. natans/C. longicaudata* reads (OTU 11) found in May 2011 at both depths (Fig. 3).

Seasonal dynamics of functional groups

We classified the OTUs into three trophic groups based on their taxonomic assignation: heterotrophs (choanoflagellates, picozoa, heterotrophic non-ochrophyte stramenopiles, ciliates, telonemia, radiolarians, katablepharids, cercozoans, fungi, centroheliozoans, mesomycetozoans, and members of class Syndiniophyceae within Dinophyta), autotrophs (cryptophytes, ochrophytes, haptophytes, rhodophytes, euglenoids,

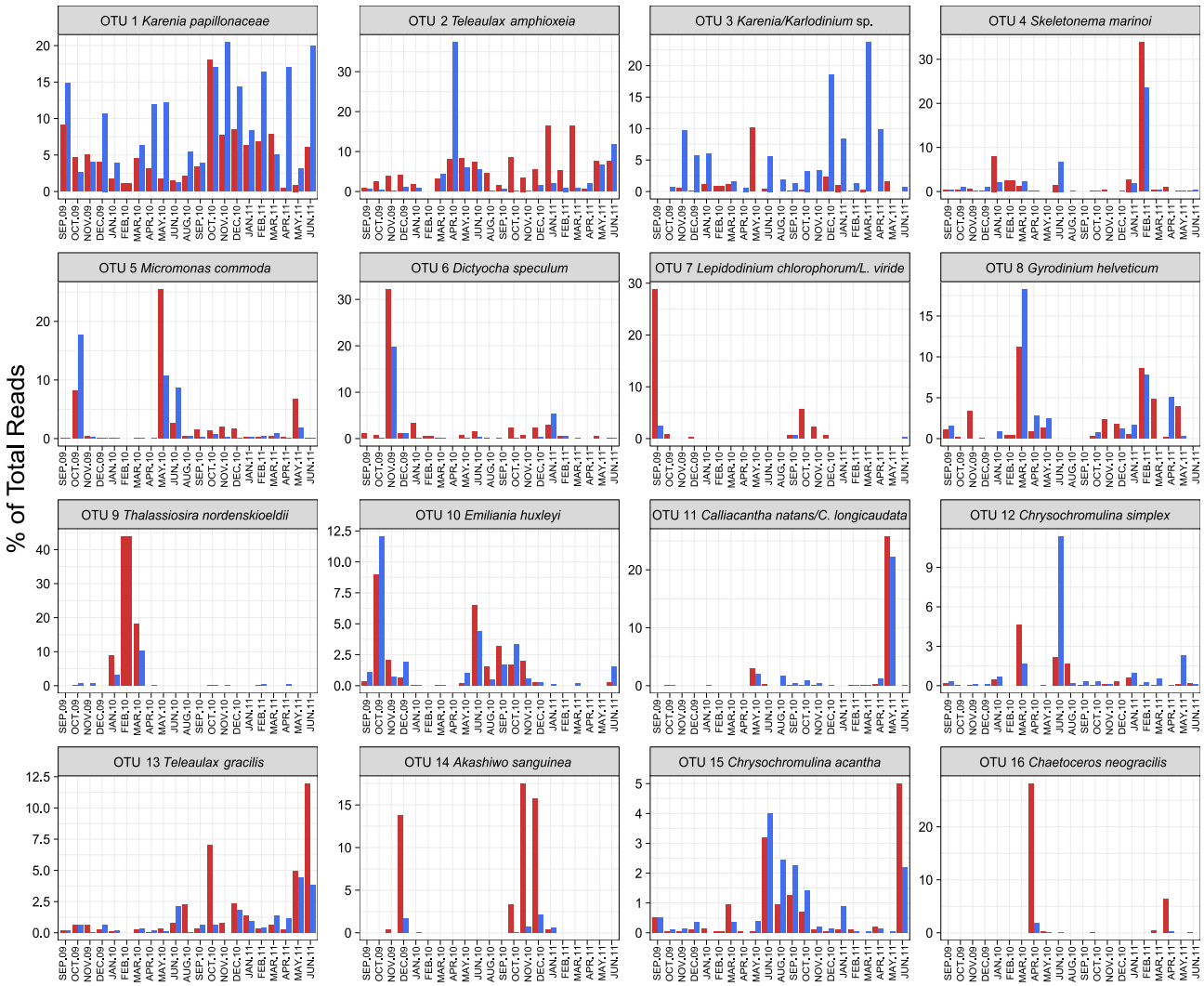

Figure 3 Temporal dynamics of the 16 most abundant OTUs (> 1% of total reads per OTU) at subsurface (SS, red) and bottom of deep chlorophyll maximum (DC, blue).

chlorophytes and streptophytes), and dinoflagellates consisting of autotrophic, mixotrophic and heterotrophic taxa, except the heterotrophic class Syndiniophyceae (Fig. 4). Within some of the autotrophic groups, some members have, however, lost their photosynthetic ability or may be mixotrophic.

Heterotrophs generally contributed more to the richness (per cent of total OTUs) than autotrophs through the sampling period at both depths, except in June 2010 and 2011 (SS), and October 2009 and June 2010 (DC), when the autotrophic community was more diverse (Fig. S5). The ratio of autotrophic- to heterotrophic OTU richness was rather similar through the entire study period. The dinoflagellate contribution of the OTUs varied between 6% and 18% and showed no clear pattern over the year and was similar at the two depths.

When comparing proportional abundances among trophic modes, a seasonal pattern was observed (Fig. 4). Reads corresponding to autotrophic groups dominated at the SS during the winter to summer period (January–August 2010, ~50% of reads) and dropped considerably during autumn (September–December 2011, ~36% of reads). In autumn, dinoflagellates reached their highest proportional abundance, especially the phototrophic *A. sanguinea* (Fig. 3). At the DC, fluctuations in read proportions were observed for all trophic groups. At DC autotrophs dominated in the autumn 2009 and spring and summer 2010, coinciding with high proportions of reads observed for *M. commoda*, *D. speculum*, *E. huxleyi*, and *Chrysochromulina* spp. (Fig. 3, 4). Heterotrophs were proportionally more abundant during winter 2009, autumn 2010 and spring 2011, and dinoflagellates took over during the autumn 2010 and winter 2011. Welch two-sample *t*-test showed significant differences in autotrophic proportional read abundances between the two depths ($P = 0.007$). However, no differences were found for proportional read abundances of heterotrophs and dinoflagellates between the two depths ($P = 0.07$ and $P = 0.12$, respectively). Autotrophic relative abundance was also higher in the subsurface than in

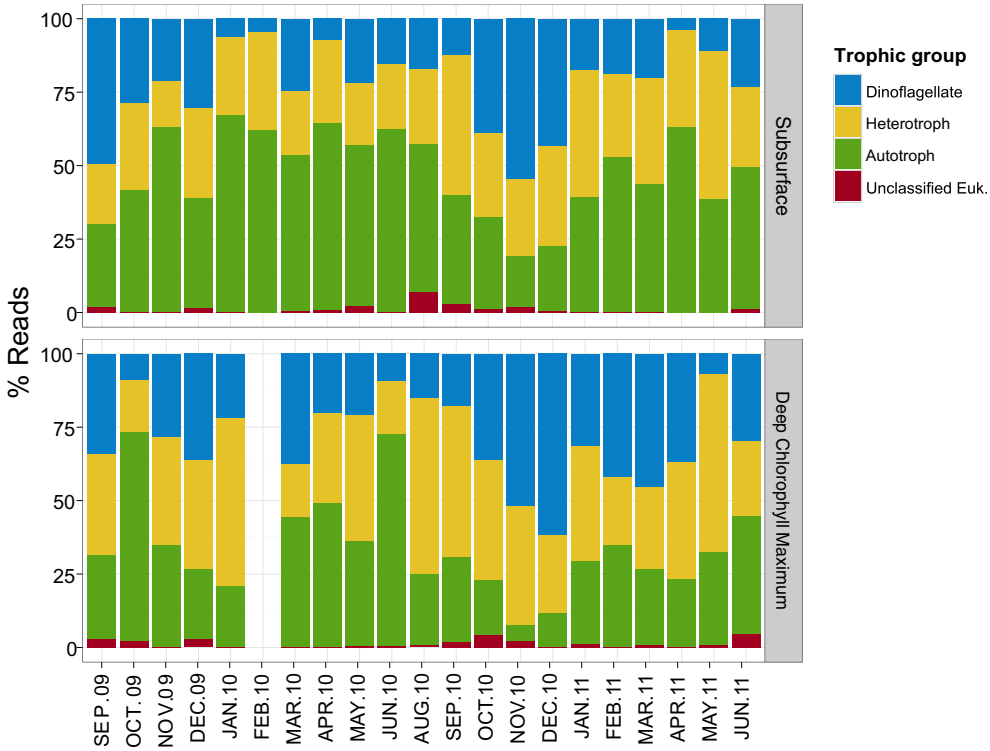

Figure 4 Succession of proportions of reads of the different trophic groups across the 21 temporal samples at OF2 station.

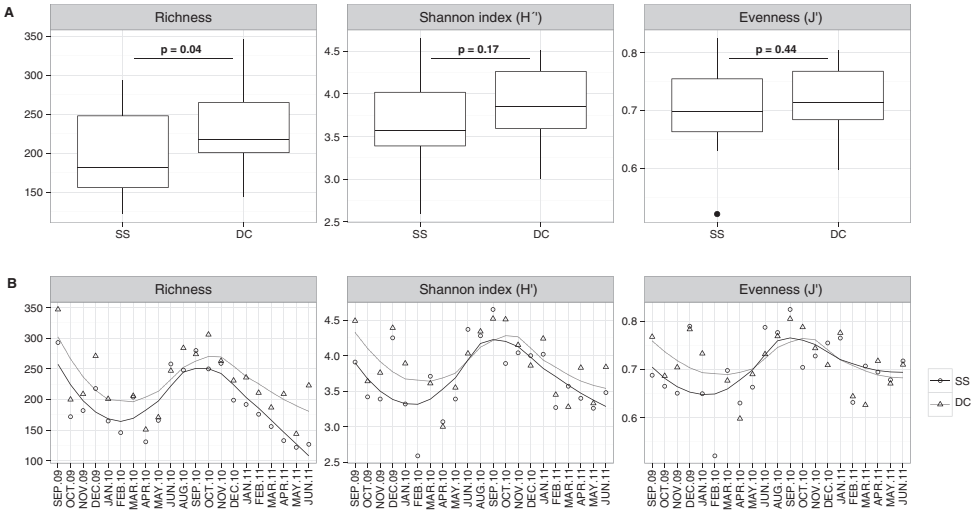

Figure 5 OTU richness, Shannon (H′) index, and Evenness (J′) index at both depths for the study site. (**A**) Boxplots. (**B**) Generalised additive model (GAM) smoothing curves fitted to the diversity indices temporal trends.

the DC, with peaks coinciding with the peaks of Chl-a observed in the isopleths (Fig. S1b).

Community structure in relation to environmental factors

Protist communities at OF2 showed a seasonal pattern. The richness median at the SS (182 OTUs, range 122–293) was significantly different than at the DC (217 OTUs, range 144–347; *t*-test $P = 0.04$; Fig. 5A and Table S2). In contrast, the SS and DC presented similar mean values of Shannon diversity index (3.57 and 3.85 respectively [*t*-test $P = 0.17$]) as well as Pielou's evenness index (0.70 and 0.71 respectively [*t*-test $P = 0.44$]). Richness and the Shannon index strongly differed between samples. They displayed a similar seasonal pattern at both depths

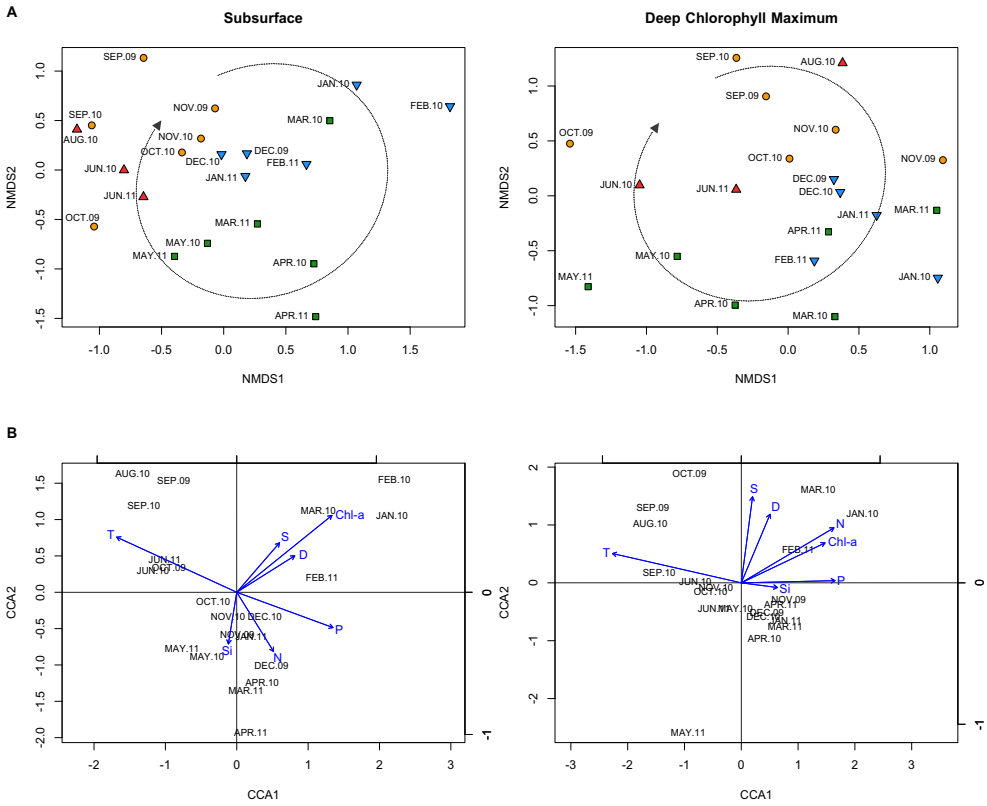

Figure 6 Ordination plots for both studied depth. (**A**) The diagram shows seasonal changes in the protist community composition. Seasons are indicated by different colours: spring (green), summer (red), autumn (yellow) and winter (blue). (**B**) Canonical Correspondence Analyses (CCA) plots showing correlations between seasonal communities and environmental factors.

reaching highest values during the summer-autumn seasons (Fig. 5B). Evenness also showed a seasonal pattern but the range was not large (0.52–0.82 for SS, 0.6–0.81 for DC).

Seasonality at both depths was also inferred by the ordination analyses based on Bray–Curtis dissimilarities (Fig. 6A), where protist communities presented four distinct seasonal clusters placed in a circular pattern. Summer and winter communities were more different at SS than at DC (ANOSIM: $R = 0.4558$, $P = 0.001$ and $R = 0.1976$, $P = 0.017$ for SS and DC respectively). The CCA (Fig. 6B) analyses were run to detect possible correlations between the environmental factors and the variations in protist communities. Temperature was found to be the most significant factor at both depths. In agreement with CCA, PERMANOVA results indicated that 20% of the seasonal variation in the SS protist community could be explained by the temperature ($P = 0.001$) and phytoplankton biomass (Chl-*a*; $P = 0.002$). In addition, salinity accounted for 8% of the variability, but this effect was not significant ($P = 0.08$; Table S3). In contrast, only temperature was a significant factor ($P = 0.001$) at the DC community explaining 10% of its variability.

The Venn diagram shows the percentage OTUs that are unique or shared between seasons (Fig. S6). The proportions were similar at both depths. Only ~7.5% of the total OTUs recorded during the 2-yr sampling were shared among the four seasons. Those consisted mostly of dinoflagellates, ochrophytes, heterotrophic stramenopiles, and ciliates OTUs.

SIMPER analyses results showed that eight OTUs (representing *Ch. neogracilis*, *S. marinoi*, *T. nordenskioeldii*, *D. speculum*, Pelagophyceae sp., *M. commoda*, *T. amphioxeia*, *A. sanguinea*, *K. papillonaceae*, and *Lepidodinium* sp.) contributed the most to the separation between seasons in the SS protist community (> 3% contribution per OTU; Table 2). Members of the infrakingdom Stramenopila contributed most to the community composition variation by season. At the DC, seven OTUs from five different phyla were the main responsible for the seasonal differences in community composition (*S. marinoi*, *M. commoda*, Colpodea sp., *T. amphioxeia*, *G. helveticum*, *Karenia* spp.), with the alveolates being the most important.

Novel records for Scandinavian waters

We detected 69 potentially new species and 40 potentially new genera of protists for Scandinavian coastal waters (see Table S4a) as compared to existing species lists for Norwegian and Scandinavian waters. For diatoms, 10 new records were found based on HTS, not yet observed by

Table 2. Contribution of variance of top OTUs between seasons by SIMPER analysis

Depth	Season	OTU ID	Av. Diss	SD	Contr. %	Cum contr. %	Taxonomy
Subsurface	Spring–Summer	OTU 16	0.03	0.05	3.44	3.44	Bacillariophyta; Bacillariophyceae; *Chaetoceros neogracilis*
		OTU 5	0.03	0.04	3.29	6.72	Chlorophyta; Mamiellales; *Micromonas commoda*
	Summer–Autumn	OTU 7	0.03	0.05	3.991	3.991	Dinophyta; Gymnodiniales; *Lepidodinium chlorophorum/L. viride*
		OTU 6	0.03	0.06	3.755	7.746	Dictyochophyceae; Dictyochales; *Dictyocha speculum*
		OTU 1	0.03	0.02	3.332	11.079	Dinophyta; Gymnodiniales; *Karenia papillonaceae*
	Autumn–Winter	OTU 9	0.04	0.08	5.38	5.38	Bacillariophyta; Thalassiosirales; *Thalassiosira nordenskioeldii*
		OTU 4	0.04	0.06	4.74	10.11	Bacillariophyta; Thalassiosirales; *Skeletonema marinoi*
		OTU 14	0.03	0.04	3.91	14.03	Dinophyta; Gymnodiniales; *Akashiwo sanguinea*
		OTU 7	0.03	0.05	3.84	17.87	Dinophyta; Gymnodiniales; *Lepidodinium chlorophorum/L. viride*
		OTU 6	0.03	0.06	3.74	21.61	Dictyochophyceae; Dictyochales; *Dictyocha speculum*
		OTU 17	0.03	0.05	3.12	24.72	Pelagophyceae; Pelagophycea sp.
	Winter–Spring	OTU 9	0.05	0.08	6.24	6.24	Bacillariophyta; Thalassiosirales; *Thalassiosira nordenskioeldii*
		OTU 4	0.04	0.06	4.66	10.90	Bacillariophyta; Thalassiosirales; *Skeletonema marinoi*
		OTU 2	0.03	0.02	3.61	14.52	Cryptophyta; Pyrenomonadales; *Teleaulax amphioxeia*
		OTU 16	0.03	0.05	3.56	18.08	Bacillariophyta; Bacillariophyceae; *Chaetoceros neogracilis*
		OTU 5	0.03	0.05	3.32	21.40	Chlorophyta; Mamiellales; *Micromonas commoda*
		OTU 17	0.03	0.05	3.17	24.56	Pelagophyceae; *Pelagophycea* sp.
		OTU 14	0.02	0.04	3.01	27.58	Dinophyta; Gymnodiniales; *Akashiwo sanguinea*
Deep chlorophyll maximum	Spring–Summer	OTU 2	0.05	0.05	5.66	5.66	Cryptophyta; Pyrenomonadales; *Teleaulax amphioxeia*
		OTU 1	0.04	0.03	4.87	10.53	Dinophyta; Gymnodiniales; *Karenia papillonaceae*
		OTU 3	0.03	0.04	3.87	14.40	Dinophyta; Gymnodiniales; *Karenia/Karlodinium* sp.
		OTU 64	0.03	0.04	3.65	18.06	Ciliophora; Colpodea sp.
	Summer–Autumn	OTU 1	0.04	0.03	5.54	5.54	Dinophyta; Gymnodiniales; *Karenia papillonaceae*
		OTU 64	0.03	0.04	3.77	9.31	Ciliophora; Colpodea sp.
		OTU 2	0.03	0.02	3.57	12.88	Cryptophyta; Pyrenomonadales; *Teleaulax amphioxeia*
		OTU 5	0.02	0.03	3.12	16.00	Chlorophyta; Mamiellales; *Micromonas commoda*
	Autumn–Winter	OTU 1	0.04	0.02	4.63	4.63	Dinophyta; Gymnodiniales; *Karenia papillonaceae*
		OTU 3	0.03	0.03	4.15	8.78	Dinophyta; Gymnodiniales; *Karenia/Karlodinium* sp.
		OTU 4	0.03	0.05	3.60	12.38	Bacillariophyta; Thalassiosirales; *Skeletonema marinoi*
	Winter–Spring	OTU 3	0.04	0.03	5.54	5.54	Dinophyta; Gymnodiniales; *Karenia/Karlodinium* sp.
		OTU 2	0.04	0.06	5.47	11.01	Cryptophyta; Pyrenomonadales; *Teleaulax amphioxeia*
		OTU 1	0.03	0.02	3.63	14.64	Dinophyta; Gymnodiniales; *Karenia papillonaceae*
		OTU 4	0.03	0.04	3.61	18.25	Bacillariophyta; Thalassiosirales; *Skeletonema marinoi*
		OTU 8	0.02	0.03	3.16	21.42	Dinophyta; Gymnodiniales; *Gyrodinium helveticum*

Av. Diss = average dissimilarity between seasons; SD = standard deviation; Contr. % = percentage contribution of variance OTU; Cum. Contr. % = cumulative contribution of OTU in per cent. The taxonomy is based on EPA phylogenetic placement (see text). Method for taxonomic classification was EPA phylogeny for the 16 most abundant OTUs (OUT 1 – OTU 16), for the rest UCLUST against the PR2 database was used.

microscopy, for example, the genera *Eunotogramma* and *Tenuicylindrus*. Of the dinoflagellate species, eight new records were registered based on HTS; for example, the genera *Luciella* and *Adenoides*. Within the alveolates there were also 23 new ciliate records (Table S4b). Of other phytoplankton groups can be mentioned three new records of Chlorophyta, one new each of Cryptophyta, Pelagophyceae (genus *Ankylochrysis*) and Raphidophyceae (genus *Haramonas*), and Bolidophyceae (*Bolidomonas pacifica*). No new records of chrysophyceae, dictyochophyce, or euglenophytes were identified.

Community structure revealed by metabarcoding versus light microscopy

With metabarcoding, we targeted the protist-plankton community in the size range ca. 0.8–45 μm. The light microscopy (LM) cell counts were done on the total water sample, including all size groups. Cells smaller than c. 15–20 μm could, however, not be taxonomically identified to species under the light microscope. The comparison between methods was therefore done at the class and not species level. The main phytoplankton groups identified and counted by microscopy were the diatoms, with 51 recorded taxa or categories, and dinoflagellates (Dinophyta), with 59. In addition, two dictyochophyte, two euglenophyte and one chrysophyte taxa were observed and counted. Microscopy cell counts were transformed into biovolume to allow the comparison to relative read abundance (Table S5 and Fig. S7).

Comparisons of HTS reads and light microscopy biovolumes were performed on five major phytoplankton groups possible to identify by both methods. Ratios of the taxonomic groups (Fig. 7A) showed that Chrysophyceae and

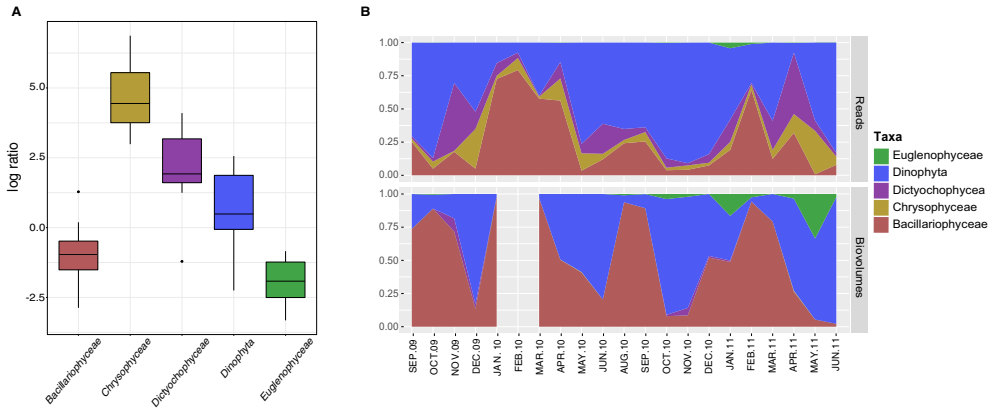

Figure 7 Comparison of metabarcoding reads versus microscopy data (biovolumes) for five phytoplankton groups. (**A**) Boxplots representing the log-ratio between reads and biovolumes. (**B**) Temporal variation in read and biovolume proportions.

Dictyochophyceae were overrepresented in HTS compared to LM, whereas Bacillariophyceae and Euglenophyceae were underrepresented. Welch *t*-test, however, showed that significant differences between the methods were only found for Bacillariophyceae ($P = 0.007$), Dictyochophyceae ($P = 0.002$), and Chrysophyceae ($P < 0.001$; Table S6). Dinoflagellates were not significantly different between methods ($P = 0.222$). The proportions of the five phytoplankton classes through the study period showed both similarities and differences between the two methods (Fig. 7B). Bacillariophyceae was the most abundant group during the winter and early spring in both years using both LM and HTS. There was also a diatom peak during summer (August–September) 2010 observed with both methods, but more pronounced by microscopy, where the diatoms were assessed to constitute 92% of the biovolume compared to 25% the reads. Microscopical counts showed that the dominating diatoms in September–November 2009 and August–September 2010 belonged to *Chaetoceros* species forming large chains that were probably removed to some extent by prefiltration prior to HTS.

Dinoflagellates showed a similar pattern, with peaks with both methods during spring to early summer (May–June) and late autumn (October–November) 2010, and May-June 2011. Dinophyta was the dominating group (> 60% of reads) in most HTS samples. In the microscopy counts this group dominated during late-autumn and late-spring, following the diatom blooms. Chrysophyceae was detected with HTS in all samples at low proportions, except for December 2009 and May 2011. With LM, Chrysophyceae, represented by the colony-forming *Dinobryon* sp., was barely detected in six samples. Dictyochophyceae reads were recovered in all HTS samples and was the group with highest proportions in November 2009 and April 2011. They were assigned to the genera *Florenciella*, *Pseudochattonella* and *Apedinella,* as determined by classification against PR2. In addition, 14 OTUs classified to unknown Dictyochophyceae were more than 90% similar to *Dictyocha* spp. as revealed by BLAST against NCBI. With microscopy only one dictyochophyte genus was detected: *Dictyocha,* found in nine samples at low biovolumes. Euglenophyceae was only detected in a few samples by both methods but seemed to be better detected by LM (January 2011 and May 2011).

DISCUSSION

This is the first long-term study of the protist community of the Oslofjorden and the Skagerrak by metabarcoding. A total of 2,026 OTUs from different trophic groups were revealed. This amount of OTUs is almost three times the number of taxa previously recorded for the Norwegian coastal waters through morphological observation (about 700 phytoplankton species according to Throndsen et al. (2007)). Our OTUs were defined at a 98% similarity level, as this has been found to be suitable for species-level distinctions of most protist groups (Caron et al. 2009). We found, however, many OTUs matching reference sequences from the same species, thus, the number of OTUs probably represents an inflated estimate of the true species diversity. Such a result indicates that a lower similarity level is needed to estimate the true diversity for some microeukaryotic groups. However, in some taxonomic groups, such as diatoms and haptophytes, different species may have identical V4 18S rRNA gene sequence, and a higher clustering level than 98% is needed to separate to species level (Egge et al. 2013). Furthermore, some microeukaryotic taxa are difficult to cultivate and/or identify through microscopy, and therefore, no molecular references are available. This may explain the number of unclassified taxa (4% of OTUs) obtained in our study.

Community composition

Alveolates, Stramenopiles and "Hacrobia" were the dominating supergroups in this study. Dinoflagellates were the most abundant phylum. They are, after diatoms, considered the most important primary producers in the ocean, reaching their highest abundances in estuaries and coastal

marine waters (Not et al. 2012). The high dinoflagellate contribution at the outer Oslofjorden is thus in accordance with Not and co-authors. Members of the genera *Karenia, Karlodinium* and *Akashiwo,* three of the most abundant genera in our waters, may form blooms associated to mortalities of fish or birds in marine coastal waters (Jones et al. 2017; Tangen 1977). *Karenia papillonaceae* was the most abundant OTU but has not previously been recorded in the Skagerrak or Norwegian waters by microscopy. This species has a second stage of small cells in culture (Carmelo Tomas, pers. commun.) that cannot be identified to species under the light microscope. It may thus have been misidentified or overlooked in past microscopical surveys. The uncultured marine alveolates group named MALV was first described by López-García et al. (2001) in 18S rRNA gene marine molecular surveys by environmental clone libraries. Members of MALV I-V have been phylogenetically placed in the dinoflagellate order Syndiniales and renamed Syndiniales group I–V (Guillou et al. 2008). Syndiniales group I and II are all assumed to be parasitic, and within group II we find the genus *Amoebophrya*. Notably, members of the MALV clades I–V were detected in all occasions at OF2, and mainly clades MALV-II and III. MALV-II has been described as a predominant group in marine metabarcoding surveys (Koid et al. 2012; Massana et al. 2011) and as a potential parasite of the class Dinophyceae (Park et al. 2004), which is similar to our results.

The most abundant diatoms in the HTS dataset, *S. marinoi*, *T. nordenskioeldii*, and *Ch. neogracilis,* are described as common diatom species in temperate coastal waters (Throndsen et al. 2007). The TARA Oceans survey included samples across the global ocean euphotic zone south of the 44°N latitude. In that survey, *Thalassiosira* and *Chaetoceros* were also two of the most diverse and cosmopolitan diatom genera, whereas *Skeletonema* was underrepresented compared to microscopy (Malviya et al. 2016). All the abundant diatoms found by HTS in this study are well-known species from Norwegian coastal waters (Throndsen et al. 2007; see Table S1). *Leptocylindrus aporus* was previously named *L. danicus* var. *aporus* but renamed by Nanjappa et al. (2013) and was found in our HTS-dataset.

Dictyochophyceae was the second most represented stramenopile class, represented mainly by *D. speculum*. This is a cold-water species with cosmopolitan distribution (Chang et al. 2003; Glezer 1970; Rigual-Hernández et al. 2010). It can be a major component in coastal and estuarine waters and has previously been linked to fish mortalities (Henriksen et al. 1993). *Dictyocha speculum* is a common species in the Oslofjorden (Throndsen et al. 2007). Another dictyochophyte recorded here, the picoflagellate *F. parvula,* was first described from the English Channel in 2004 (Eikrem et al. 2004). The ichthyotoxic dictyochophyte *P. verruculosa,* found in this study, was recorded for the first time in northern Europe (Germany) in 2000 (Riisberg and Edvardsen 2008). The cold-water species *Pseudochattonella farcimen*, that has formed ichthyotoxic blooms in the Skagerrak since 1998 (Edvardsen et al. 2007) was not recorded in our dataset. These two *Pseudochattonella* species differ in only one position within the V4 18S rRNA gene region (Riisberg and Edvardsen 2008) and may have been clustered together as *P. verruculosa*.

Marine Stramenopiles (MAST) include a large number of predominantly heterotrophic groups and are well represented both in the plankton and in sediments, playing a key role in marine ecosystems (Logares et al. 2012). MAST-1, -3, -4, and -7 have previously been found in open ocean and coastal systems. Although MAST-1 and 3 were the most abundant in our dataset, MAST-4 was present in small abundances. MAST-4 is a dominant group in most oceans but is absent in waters < 4 °C (Massana et al. 2006). Although the Oslofjorden waters are below such temperatures during half of the year, we detected MAST-4-related OTUs in our dataset. This may be due to the presence of the cyanobacteria *Synechoccocus,* which seems to be a prey for MAST-4 (Lin et al. 2012). The MAST-1, -6, -9, and -12, recorded in this study, are important clades in both planktonic and sediment samples (Massana et al. 2014).

In a previous study using metabarcoding with haptophyte-specific primers and with the same samples (Egge et al. 2015a,b), Prymnesiales was the most abundant, frequent, and diverse haptophyte order. As also found by Egge et al. (2015b), the most abundant haptophyte OTUs had best match to *E. huxleyi* and *C. simplex*. Members of Prymnesiales are abundant in the Skagerrak coastal waters and usually have densities over one million cells per litre during summer (Lekve et al. 2006) corresponding to our findings. *Prymnesium faveolatum* was among the most abundant haptophytes in this study, a species not previously recorded by microscopy in the Skagerrak (Artsdatabanken 2018).

The two cryptophytes *T. amphioxeia* and *T. gracilis* were among the 16 most abundant OTUs. *Teleaulax amphioxeia* is well know from brackish waters in Europe (Throndsen et al. 2007), whereas *T. gracilis* was first described in 2012 from the Atlantic coast of Spain (Laza-Martínez et al. 2012) and has not been recorded by microscopy in Norwegian waters (Artsdatabanken 2018).

Within Chlorophyta the species *M. commoda,* a picoflagellate belonging to class Mamiellophyceae, was among the 16 most abundant OTUs. This newly described species was recently separated from *Micromonas pusilla* (Simon et al. 2017). This is the first time *M. commoda* has been recorded from the Skagerrak. *Micromonas pusilla* has been shown to dominate the eukaryotic picoplankton in North Atlantic coastal and Arctic waters (Not et al. 2004, 2005). Our findings, however, suggest that it is *M. commoda,* and not *M. pusilla* that dominates in Oslofjorden (Fig. S3). OTUs assigned to *Micromonas* spp. represented almost 7% of the read abundance in the pico-size fraction in our study, and thus was less dominant compared to the findings by Not et al. (2004), where they used fluorescence in situ hybridisation (FISH) for quantification. The remaining major components of the Chlorophyta community, all pico- or small nanoplankton, have previously been recorded from the North Atlantic, except

for *A. nanus,* which is usually found in freshwater (John et al. 2002).

Within Opisthokonta we found an abundant choanoflagellate OTU differing in only two bases (one being in a homopolymer) from *C. natans,* recently genetically characterized by (Nitsche et al. 2017). This difference could be explained as a sequencing error. This species has not previously been recorded from the Oslofjorden, but was found to be a dominant choanoflagellate species in the winter–spring community in the southern Kattegat (Thomsen et al. 2016). In addition, it was the second-most globally abundant choanoflagellate in the Tara Oceans data, exhibiting highest relative abundances at cold-water stations (Nitsche et al. 2017).

Temporal variation

The richness index and Shannon diversity index (H) showed seasonal fluctuations at both depths. They reached maximum values during late summer-early autumn (June–September) and were generally higher at the deep chlorophyll than at the subsurface. Evenness varied slightly through the seasons, with generally high values (> 0.6), which allows the detection also of rare taxa (Caporaso et al. 2012). The high diversity in late summer-autumn has been proposed to be due to the influence of the North Atlantic current that brings in allochtonous plankton taxa (Andersen et al. 2001).

Marked seasonal variations in the protist community were observed at both depths, with a distinct separation between summer and winter. This seasonality revealed by HTS coincides with microscopy cell counts in this study, as well as previous microscopy-based studies in the Oslofjorden (Hasle and Smayda 1960). Taxonomic groups with marked increase in richness in June-September are Haptophyta, Chlorophyta and Cercozoa (Fig. S4). This agrees with Hasle and Smayda (1960) showing coccolithophore haptophytes to be present mainly during June–November and with Egge et al. (2015b) showing highest haptophyte diversity in this period by metabarcoding using haptophyte specific primers. Pico- and nanoplanktonic chlorophytes requires electron microscopy or molecular methods for identification, and there are no previous seasonal surveys on richness or number of taxa of this group from the Skagerrak. Diatoms were found in this study to have highest richness during autumn-winter. Lange et al. (1992) similarly found the highest diversity of diatoms during autumn–winter, which was explained by the period of major advection of foreign species introduced into the Skagerrak by the Jutland and Dooley currents from the North Atlantic. Highest diversity of dinoflagellates was in the autumn (September–December, Fig. S4 and Table S5), which has also previously been found by microscopy (Throndsen et al. 2007).

Large changes in proportional abundance of the major taxonomic groups were observed between samplings. This could be explained by the long (monthly) sampling intervals (Countway et al. 2005). The seasonal dynamics here are consistent with previous observations such that diatoms dominate during the spring bloom, whereas dinoflagellates have their highest proportion in autumn–winter (Hasle and Smayda 1960). Haptophyte proportional abundance peaked in June (Fig. 2). Members of *Chrysochromulina* were most abundant during the summer (Fig. 3), which corresponds with previous findings (Lekve et al. 2006) suggesting that they are favoured by low nutrient concentrations and high freshwater influence (Edvardsen and Paasche 1998).

Profound differences in community composition by season were also indicated at both depths by ordination analyses (NMDS), where four clear clusters were observed (Fig. 6B). According to SIMPER analyses 10 OTUs generated most of the differences between seasons, corresponding to the most abundant OTUs. CCA ordination and PERMANOVA analyses showed that temperature and salinity influenced the community structure. Temperature and salinity displayed negative correlations with nutrient concentrations which indicate that terrestrial and riverine inputs bring nutrients to the Outer Oslofjorden. As proposed by Simon et al. (2015), the detection of few correlations may result from biotic factors (e.g. predation, mutualism, parasitism and virus) not being included in this study.

Trophic status through the season and by depth

Ratios between percentage of autotrophic and heterotrophic OTUs was similar through the two years (Fig. S5). Heterotrophs were more diverse than autotrophs through most of the sampling period except in June 2010 and 2011. Similar results were found in the TARA Oceans expedition where heterotrophic groups contributed more to the richness than autotrophic (de Vargas et al. 2015). The relative abundance of trophic groups showed a clear seasonal pattern, especially at the subsurface. The proportion of reads assigned to autotrophic groups was highest during winter to spring, the period with highest chl-*a* concentrations (> 2 μg/l), and lowest surface water temperatures (−1 to + 5 °C), and lowest in the late autumn to early winter when water temperature was gradually decreasing from 12 to 0 °C. The opposite pattern was found for dinoflagellates. This pattern is similar to that found by Piredda et al. (2017) studying protist plankton communities in Gulf of Naples, Italy.

Novel records for Scandinavian waters

We recorded 69 potentially new species and 40 potentially new genera for the Skagerrak area that are not registered in the Norwegian Biodiversity Information Centre (Artsdatabanken 2018) nor the Nordic Microalgae and Aquatic Protozoa (NOD) database (Karlson et al. 2015). The number of pelagic and benthic protist species recorded in Norwegian marine waters based on microscopy are estimated to ca. 1,200, according to the Norwegian Species Information Centre (Artsdatabanken 2018, Antall arter i norsk natur 2016). About 1,020 of these species belong to a phylum with microalgal representatives. Throndsen et al.

(2007) estimated that more than 700 phytoplankton species may be present in Norwegian coastal waters. The approximately 2,000 planktonic protist OTUs reported here in the Skagerrak, passing a 45 μm sieve and collected on 0.8 μm pore size filters, and after a strict read filtering, are considerably more than the number of protists observed in the microscopy through all times, but still of the same magnitude. Furthermore, some closely related species have identical V4 18S rRNA gene region (e.g. among haptophytes Egge et al. 2015a, and dinoflagellates, Edvardsen et al., unpubl. data), and one OTU may thus represent more than one species. In our study, however, several OTUs represent the same species, which reduces the number of taxa. This study focuses on the smallest protists taxa that are difficult to identify in the light microscope. Many species of the important groups, dinoflagellates, diatoms and ciliates may be larger than 45 μm, and are not included in metabarcoding data. Since previous studies are mainly based on morphological techniques, some parasitic groups (e.g. Syndiniales) have been overlooked and recorded for the first time with metabarcoding.

High-throughput sequencing versus microscopy comparison

A few previous studies have compared metabarcoding to microscopy quantitative surveys, focusing on Arctic lakes (Majaneva et al. 2012) or targeting a specific group (Bachy et al. 2013; Young et al. 2014). Our study is one of the first comparing microscopy cell counts and HTS data of several protist classes from marine waters during a long-term time series. Our results showed some clear differences between the two approaches. Bacillariophyceae and Euglenophyceae proportional abundances were underrepresented by HTS of the nano-picoplankton, and the latter was almost overlooked by this approach. The only euglenoid genus found with both approaches was *Eutreptiella*, which ranges between 12 and 115 μm in cell size (Throndsen et al. 2007). The underestimation by metabarcoding can be explained by the < 45 μm prefiltration of the water samples for RNA extraction. It can also partly be explained by the V4 18S rRNA gene PCR primers used by Stoeck et al. 2010; that seem to be poor in amplifying members of Euglenophyta compared to amplification using chloroplast gene targeting primers (Amaral-Zettler et al. 2011). Bacillariophyceae was found abundant by both approaches, but significant differences through the year were found between the methods. Such discrepancies were also found for diatoms in freshwater studies (Xiao et al. 2014). Seven large, chain-forming *Chaetoceros* species were only detected by microscopy in our study. Prefiltration can explain this underestimation by metabarcoding also in this case.

In contrast, the classes Chrysophyceae and Dictyochophycea were favoured in the metabarcoding compared to microscopy. Of Chrysophyceae, only the genus *Dinobryon*, which forms large colonies, was observed by microscopy. *Dinobryon* sequences were included in our reference sequence database, but it was not detected by HTS. Many other OTUs were however assigned to Chrysophyceae.

The class Dictyophyceae was also detected by both methods but favoured by HTS in both relative read abundance and richness. Only *D. speculum* was observed by microscopy, while by metabarcoding, *Florenciella*, *Pseudochattonella, Apedinella*, and several unclassified dictyocophytes were also detected.

Dinoflagellates have a wide size distribution. In LM, we included all size groups that could be identified under a light microscope (larger than c. 15 μm), whereas in HTS we analysed the 45–0.8 μm size fraction. Compared to taxa with similar cell size, dinoflagellates have large genomes (Hackett et al. 2004) and putatively high rRNA gene copy number (Prokopowich et al. 2003) thus, an overrepresentation in metabarcoding surveys based on DNA may be expected. However, in this study, rRNA was isolated from the plankton, converted to cDNA by RT-PCR and then cDNA was amplified in the PCR, which is expected to reduce the bias for organisms with large genome size (Not et al. 2009). Indeed, the log-ratio of proportion of reads to cell count-based biovolume of Dinophyceae was 0.99, and the dinoflagellates were less overrepresented in the HTS dataset compared to microscopy, than Cryptophyceae and Dictyocophyceae.

Another important aspect in this comparison is that in LM only a small volume was analysed (10 ml) allowing very few species to be observed compared to HTS, where 20 liters were filtrated and RNA from ca. 10 liters were used in the further processing (RT-PCR to cDNA).

CONCLUDING REMARKS

A high diversity of protists was revealed by metabarcoding compared to previous surveys by microscopy through a decade. About 70 protist taxa was recorded for the first time. Metabarcoding can reveal a detailed protist composition and allows large sample sizes. The protist community composition and relative abundance in the Oslofjorden show large variation though the year. There was a difference in protist community structure between the two sampled depths, with higher proportional abundance of autotrophs found in the subsurface than at the deep chlorophyll maximum. The seasonal pattern in relative read abundance of major phytoplankton groups was well in accordance with microscopy biovolumes of the same groups. However, when comparing proportion of reads with biovolumes for major phytoplankton groups, some are overrepresented and other underrepresented in HTS versus microscopy. As neither method shows the full picture, they should be used complementary to each other. More reference sequences, connecting a genotype to a morphotype, are needed to enable a more precise taxonomic identification and reducing the number of OTUs with best match to an "uncultured marine eukaryote". This may also improve the assessment of the actual taxon richness instead of OTU richness. This study may serve as a baseline for future studies and

monitoring to reveal effects of environmental and climate change.

ACKNOWLEDGMENTS

Financial support was given by the University of Oslo, MN Faculty to SGS and BE, by the Research Council of Norway through grant 190307 HAPTODIV to BE, EDE, and WE, by grant 225956/E10 MIKROPOLAR to EDE and BE, and by the EU project BioMarKs (2008-6530, ERA-net Biodiversa, EU) to BE and RL. We thank Rita Amundsen, captain Sindre Holm and the crew onboard R/V Trygve Braarud for assistance during sampling, Sissel Brubak and Berit Kaasa for analysing Chlorophyll a and nutrient samples, Colomban de Vargas for leading the BioMArks project, Anette Engesmo for calculating phytoplankton biovolumes from cell numbers and Tom Andersen for statistical advice. High-throughput sequencing was performed at the Norwegian Sequencing Centre and bioinformatics analyses on the University of Oslo Abel cluster.

LITERATURE CITED

Amaral-Zettler, L. A., Zettler, E. R., Theroux, S. M., Palacios, C., Aguilera, A. & Amils, R. 2011. Microbial community structure across the tree of life in the extreme Río Tinto. *ISME J.*, 5:42–50.

Andersen, I., Berge, J. A., Andersen, J. H., Berntsson, I., Danielssen, D., Foverskov, S., Fyrberg, L., Gjøsæter, J., Granmo, A., Green, N., Hansen, O. S., Hylland, K., Håkansson, B., Johannessen, T., Karlson, B., Knutsen, J. A., Knutzen, J., Magnusson, J., Molvaer, J., Pedersen, B., Sjöberg, B., Szaron, J., Torstensen, E., Tveite, S. & Ærtebjerg, G. 2001. The Skagerrak – environmental state and monitoring prospects. Göteborg. 118 p.

Artsdatabanken. 2018. The Norwegian Biodiversity Information Centre, Norway. https://www.artsdatabanken.no/.

Bachy, C., Dolan, J. R., López-García, P., Deschamps, P. & Moreira, D. 2013. Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *ISME J.*, 7:244–255.

Backe-Hansen, P. & Throndsen, J. 2002. Pico- and nanoplankton from the inner Oslofjord, eastern Norway, including description of two new species of *Luffisphaera* (incerta sedis). *Sarsia North Atl. Mar. Sci.*, 87:55–64.

Berger, S. A., Krompass, D. & Stamatakis, A. 2011. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.*, 60:291–302.

Braarud, T. & Bursa, A. 1939. The phytoplankton of the Oslo Fjord 1933–1934. *Hvalrådets Skr.*, 19:6–63.

Braarud, T., Gaarder, K. R. & Grøntved, J. 1953. The phytoplankton of the North Sea and adjacent waters in May 1948. Conseil permanent international pour l'exploration de la mer, Copenhague. 93 pp.

Bratbak, G., Jacquet, S., Larsen, A., Pettersson, L. H., Sazhin, A. F. & Thyrhaug, R. 2011. The plankton community in Norwegian coastal waters-abundance, composition, spatial distribution and diel variation. *Cont. Shelf Res.*, 31:1500–1514.

Bray, R. & Curtis, T. 1957. An Ordination of the upland forest communities of Southern Wisconsin. *Ecol. Monogr.*, 27:325–349.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Gonzalez Peña, A., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*, 7:335–336.

Caporaso, J. G., Paszkiewicz, K., Field, D., Knight, R. & Gilbert, J. A. 2012. The Western English Channel contains a persistent microbial seed bank. *ISME J.*, 6:1089–1093.

Caron, D. A., Countway, P. D., Savai, P., Gast, R. J., Schnetzer, A., Moorthi, S. D., Dennett, M. R., Moran, D. M. & Jones, A. C. 2009. Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Appl. Environ. Microbiol.*, 75:5797–5808.

Chang, F. H., Zeldis, J., Gall, M. & Hall, J. 2003. Seasonal and spatial variation of phytoplankton assemblages, biomass and cell size from spring to summer across the north-eastern New Zealand continental shelf. *J. Plankton Res.*, 25:737–758.

Countway, P. D., Gast, R. J., Savai, P. & Caron, D. A. 2005. Protistan diversity estimates based on 18S rDNA from seawater incubations in the Western North Atlantic. *J. Eukaryot. Microbiol.*, 52:95–106.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Acinas, S. G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M. E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P. & Karsenti, E. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348:1261605-1-11.

Dittami, S. M., Hostyeva, V., Egge, E. S., Kegel, J. U., Eikrem, W. & Edvardsen, B. 2013. Seasonal dynamics of harmful algae in outer Oslofjorden monitored by microarray, qPCR, and microscopy. *Environ. Sci. Pollut. Res.*, 20:6719–6732.

Dragsund, E., Aspholm, O., Tangen, K., Bakke, S. M., Heier, L. & Jensen, T. 2006. Overvåking av eutrofitilstanden i Ytre Oslofjord. Femårsrapport 2001–2005, nr. 2006-0831. Oslo. 127 p.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26:2460–2461.

Edvardsen, B., Eikrem, W., Shalchian-Tabrizi, K., Riisberg, I., Johnsen, G., Naustvoll, L. & Throndsen, J. 2007. *Verrucophora farcimen* gen. et sp. nov. (Dictyochophyceae, Heterokonta) - a bloom-forming ichthyotoxic flagellate from the Skagerrak, Norway. *J. Phycol.*, 43:1054–1070.

Edvardsen, B. & Paasche, E. 1998. Bloom dynamics and physiology of *Prymnesium* and *Chrysochromulina*. *NATO ASI Ser. G. Ecol. Sci.*, 41:193–208.

Edvardsen, B., Shalchian-Tabrizi, K., Jakobsen, K. S., Medlin, L. K., Dahl, E., Brubak, S. & Paasche, E. 2003. Genetic variability and molecular phylogeny of *Dinophysis* species (Dinophyceae) from norwegian waters inferred from single cell analyses of rDNA. *J. Phycol.*, 39:395–408.

Egge, E., Bittner, L., Andersen, T., Audic, S., de Vargas, C. & Edvardsen, B. 2013. 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. *PLoS ONE*, 8:e74371.

Egge, E. S., Eikrem, W. & Edvardsen, B. 2015a. Deep-branching novel lineages and high diversity of haptophytes in the Skagerrak (Norway) uncovered by 454 pyrosequencing. *J. Eukaryot. Microbiol.*, 62:121–140.

Egge, E. S., Johannessen, T. V., Andersen, T., Eikrem, W., Bittner, L., Larsen, A., Sandaa, R. A. & Edvardsen, B. 2015b. Seasonal diversity and dynamics of haptophytes in the Skagerrak, Norway, explored by high-throughput sequencing. *Mol. Ecol.*, 24:3026–3042.

Eikrem, W., Romari, K., Latasa, M., Gall, F., Le Throndsen, J. & Vaulot, D. 2004. *Florenciella parvula* gen. et sp. nov. (Dictyochophyceae, Heterokontophyta), a small flagellate isolated from the English Channel. *Phycologia*, 43:658–668.

Epstein, S. & López-García, P. 2008. "Missing" protists: a molecular prospective. *Biodivers. Conserv.*, 17:261–276.

Glezer, Z. I. 1970. Silicoflagellatophyceae. *In:* Gollerbakh, M. M. (ed.), Cryptogamic plants of the U.S.S.R, Vol. 7. Israel Program for Scientific Translation, Jerusalem. p. 1–363.

Gran-Stadniczeñko, S., Šupraha, L., Egge, E. D. & Edvardsen, B. 2017. Haptophyte diversity and vertical distribution explored by 18S and 28S ribosomal RNA gene metabarcoding and scanning electron microscopy. *J. Eukaryot. Microbiol.*, 64:514–532.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H., Lara, E., Le Bescot, N., Logares, R., Mahé, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A. L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P. & Christen, R. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.*, 41:D597–D604.

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., Scanlan, D. J. & Worden, A. Z. 2008. Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ. Microbiol.*, 10:3349–3365.

Hackett, J. D., Anderson, D. M., Erdner, D. L. & Bhattacharya, D. 2004. Dinoflagellates: a remarkable evolutionary experiment. *Am. J. Bot.*, 91:12.

Hasle, G. R. & Smayda, T. J. 1960. The annual phytoplankton cycle at Drøbak, Oslofjord. *Nytt Mag. Bot.*, 8:53–75.

Henriksen, P., Knipschildt, F., Moestrup, Ø. & Thomsen, H. A. 1993. Autecology, life history and toxicology of the silicoflagellate *Dictyocha speculum* (Silicoflagellata, Dictyochophyceae). *Phycologia*, 32:29–39.

Hjort, J. & Gran, H. H. 1900. Hydrographic-biological investigations of the Skagerrak and the Christiania Fiord. *Report on Norwegian Fishery- and Marine-Investigations*, Vol. I No 2. 49 pp.

John, D. M., Whitton, B. A. & Brook, A. J. (ed.) 2002. The freshwater algal flora of the British Isles: an identification guide to freshwater and terrestrial algae, Vol. 1. Cambridge University Press, Cambridge. 702 pp.

Jones, T., Parrish, J. K., Punt, A. E., Trainer, V. L., Kudela, R., Lang, J., Brancato, M. S., Odell, A. & Hickey, B. 2017. Mass mortality of marine birds in the Northeast Pacific caused by *Akashiwo sanguinea*. *Mar. Ecol. Prog. Ser.*, 579:111–127.

Karlson, B., Andreasson, A., Johansen, M., Karlberg, M., Loo, A. & Skjevik, A.-T. 2015. Nordic Microalgae. http://www.nordicmicroalgae.org

Katoh, K., Asimenos, G. & Toh, H. 2009. Multiple alignment of DNA sequences with MAFFT. *In:* Posada, D. (ed.), Bioinformatics for DNA Sequence Analysis. Humana Press, Totowa, NJ. p. 39–64.

Koid, A., Nelson, W. C., Mraz, A. & Heidelberg, K. B. 2012. Comparative analysis of eukaryotic marine microbial assemblages from 18S rRNA gene and gene transcript clone libraries by using different methods of extraction. *Appl. Environ. Microbiol.*, 78:3958–3965.

Kuylenstierna, M. & Karlson, B. 1994. Seasonality and composition of pico- and nanoplanktonic cyanobacteria and protists in the Skagerrak. *Bot. Mar.*, 37:17–33.

Lange, C. B., Hasle, G. R. & Syvertsen, E. E. 1992. Seasonal cycle of diatoms in the Skagerrak, North Atlantic, with emphasis on the period 1980–1990. *Sarsia*, 77:173–187.

Laza-Martínez, A., Arluzea, J., Miguel, I. & Orive, E. 2012. Morphological and molecular characterization of *Teleaulax gracilis* sp. nov. and *T. minuta* sp. nov. (Cryptophyceae). *Phycologia*, 51:649–661.

Lekve, K., Bagøien, E., Dahl, E., Edvardsen, B., Skogen, M. D. & Stenseth, N. C. 2006. Environmental forcing as a main determinant of bloom dynamics of the *Chrysochromulina* algae. *Proc. Biol. Sci.*, 273:3047–3055.

Lin, Y. C., Campbell, T., Chung, C. C., Gong, G. C., Chiang, K. P. & Worden, A. Z. 2012. Distribution patterns and phylogeny of marine stramenopiles in the North Pacific Ocean. *Appl. Environ. Microbiol.*, 78:3387–3399.

Logares, R., Audic, S., Santini, S., Pernice, M. C., de Vargas, C. & Massana, R. 2012. Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *ISME J.*, 6:1823–1833.

López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C. & Moreira, D. 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, 409:603–607.

Majaneva, M., Rintala, J. M., Piisilä, M., Fewer, D. P. & Blomster, J. 2012. Comparison of wintertime eukaryotic community from sea ice and open water in the Baltic Sea, based on sequencing of the 18S rRNA gene. *Polar Biol.*, 35:875–889.

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., Wincker, P., Iudicone, D., de Vargas, C., Bittner, L., Zingone, A. & Bowler, C. 2016. Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl Acad. Sci.*, 113:E1516–E1525.

Massana, R. 2015. Protistan diversity in environmental molecular surveys. *In:* Ohtsuka, S., Suzaki, T., Horiguchi, T., Suzuki, N. & Not, F. (ed.), Marine Protists: Diversity and Dynamics. Springer, Tokyo. p. 3–21.

Massana, R., del Campo, J., Sieracki, M. E., Audic, S. & Logares, R. 2014. Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. *ISME J.*, 8:854–866.

Massana, R., Pernice, M., Bunge, J. A. & del Campo, J. 2011. Sequence diversity and novelty of natural assemblages of picoeukaryotes from the Indian Ocean. *ISME J.*, 5:184–195.

Massana, R., Terrado, R., Forn, I., Lovejoy, C. & Pedros-Alio, C. 2006. Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ. Microbiol.*, 8:1515–1522.

Medlin, L. K. & Kooistra, W. H. C. F. 2010. Methods to estimate the diversity in the marine photosynthetic protist community with illustrations from case studies: a review. *Diversity*, 2:973–1014.

Nanjappa, D., Kooistra, W. H. C. F. & Zingone, A. 2013. A reappraisal of the genus *Leptocylindrus* (Bacillariophyta), with the addition of three species and the erection of *Tenuicylindrus* gen. nov. *J. Phycol.*, 49:917–936.

Nitsche, F., Thomsen, H. A. & Richter, D. J. 2017. Bridging the gap between morphological species and molecular barcodes - Exemplified by loricate choanoflagellates. *Eur. J. Protistol.*, 57:26–37.

Not, F., del Campo, J., Balagué, V., de Vargas, C. & Massana, R. 2009. New insights into the diversity of marine picoeukaryotes. *PLoS ONE*, 4:e7143.
Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D. & Simon, N. 2004. A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Appl. Environ. Microbiol.*, 70:4064–4072.
Not, F., Massana, R., Latasa, M., Marie, D., Colson, C., Eikrem, W., Pedrós-Alió, C., Vaulot, D. & Simon, N. 2005. Late summer community composition and abundance of photosynthetic picoeukaryotes in Norwegian and Barents Seas. *Limnol. Oceanogr.*, 50:1677–1686.
Not, F., Siano, R., Kooistra, W. H. C. F., Simon, N., Vaulot, D. & Probert, I. 2012. Diversity and ecology of eukaryotic marine phytoplankton. *In:* Piganeau, G. (ed.), Advances in Botanical Research, Vol. 64. Academic Press, Waltham, MA. p. 1–53.
Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H. & Wagner, H. 2017. vegan: Community Ecology Package. R package version 2.4-3.
Olenina, I., Hajdu, S., Edler, L., Wasmund, N., Busch, S., Göbel, J., Gromisz, S., Huseby, S., Huttunen, M., Jaanus, A., Kokkonen, P., Ledaine, I. & Niemkiewicz, E. 2006. Biovolumes and size-classes of phytoplankton in the Baltic Sea. *HELCOM Balt. Sea Environ. Proc.*, 106.
Paasche, E. & Østergren, I. 1980. The annual cycle of plankton diatom growth and silica production in the inner Oslofjord. *Limnol. Oceanogr.*, 25:481–494.
Park, M. G., Yih, W. & Coats, D. W. 2004. Parasites and phytoplankton, with special emphasis on dinoflagellate infections. *J. Eukaryot. Microbiol.*, 51:145–155.
Piredda, R., Tomasino, M. P., D'Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., Kooistra, W. H. C. F., Sarno, D. & Zingone, A. 2017. Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS Microbiol. Ecol.*, 93:1–14.
Prokopowich, C. D., Gregory, T. R. & Crease, T. J. 2003. The correlation between rDNA copy number and genome size in eukaryotes. *Genome*, 46:48–50.
Quince, C., Lanzen, A., Davenport, R. J. & Turnbaugh, P. J. 2011. Removing noise from pyrosequenced amplicons. *BMC Bioinformatics*, 12:38.
R Development Core Team. 2017. R: A language and environment for statistical computing v.3.4.1. *R Found. Stat. Comput.*
Rigual-Hernández, A. S., Bárcena, M. A., Sierro, F. J., Flores, J. A., Hernández-Almeida, I., Sanchez-Vidal, A., Palanques, A. & Heussner, S. 2010. Seasonal to interannual variability and geographic distribution of the silicoflagellate fluxes in the Western Mediterranean. *Mar. Micropaleontol.*, 77:46–57.
Riisberg, I. & Edvardsen, B. 2008. Genetic variation in bloom-forming ichthyotoxic *Pseudochattonella* species (Dictyochophyceae, Heterokonta) using nuclear, mitochondrial and plastid DNA sequence data. *Eur. J. Phycol.*, 43:413–422.
Shannon, C. E. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423.
Simon, N., Cras, A. L., Foulon, E. & Lemée, R. 2009. Diversity and evolution of marine phytoplankton. *C. R. Biol.*, 332:159–170.
Simon, N., Foulon, E., Grulois, D., Six, C., Desdevises, Y., Latimier, M., Le Gall, F., Tragin, M., Houdan, A., Derelle, E., Jouenne, F., Marie, D., Le Panse, S., Vaulot, D. & Marin, B. 2017. Revision of the genus *Micromonas* Manton et Parke (Chlorophyta, Mamiellophyceae), of the type species *M. pusilla.* (Butcher) Manton & Parke and of the Species *M. commoda* van Baren, Bachy and Worden and description of two new species based on the genetic and phenotypic characterization of cultured isolates. *Protist*, 168:612–635.
Simon, M., López-García, P., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P. & Jardillier, L. 2015. Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. *ISME J.*, 9:1941–1953.
Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30:1312–1313.
Stoeck, T., Bass, D., Nebel, M., Christen, R. & Meredith, D. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.*, 19:21–31.
Strickland, J. D. H. & Parsons, T. R. 1972. A Practical Handbook of Seawater Analysis. Fisheries Research Board of Canada, Ottawa. 328 pp.
Tangen, K. 1977. Blooms of *Gyrodinium aureolum* (Dinophygeae) in North European waters, accompanied by mortality in marine organisms. *Sarsia*, 63:123–133.
Tennekes, M. 2017. treemap: Treemap Visualization. R package version 2.4-2.
Thomsen, H. A., Nitsche, F. & Richter, D. J. 2016. Seasonal occurrence of loricate choanoflagellates in Danish inner waters. *Protist*, 167:622–638.
Throndsen, J., Hasle, G. R. & Tangen, K. 2007. Phytoplankton of Norwegian coastal waters. Almater Forlag AS, Oslo.
Utermöhl, H. 1958. Zur Vervollkommung der quantitativen Phytoplankton-Methodik. *Int. Ver. Theor. Angew. Limnol. Verh.*, 9:1–38.
Xiao, X., Sogge, H., Lagesen, K., Tooming-Klunderud, A., Jakobsen, K. S. & Rohrlack, T. 2014. Use of high throughput sequencing and light microscopy show contrasting results in a study of phytoplankton occurrence in a freshwater environment. *PLoS ONE*, 9:e106510.
Young, J. R., Liu, H., Probert, I., Aris-Brosou, S. & de Vargas, C. 2014. Morphospecies versus phylospecies concepts for evaluating phytoplankton diversity: the case of the Coccolithophores. *Cryptogam. Algol.*, 35:353–377.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Temporal variation in physicochemical parameters measured at the study site.
Figure S2. Heatmaps representing the temporal variations of all studied protist subgroups at the Outer Oslofjorden OF2 station.
Figure S3. Maximum likelihood RAxML-EPA (Evolutionary Placement Algorithm) trees of the 16 most abundant OTUs at the OF2 station.
Figure S4. Succession of proportions of OTUs of the 18 major taxonomic groups across the 21 temporal samples at station OF2.
Figure S5. Succession of proportions of OTUs of the different trophic groups across the 21 temporal samples at station OF2.
Figure S6. Venn diagram showing the unique and shared operational taxonomic units (OTUs) at OF2 station among the four different seasons during the study period.

Figure S7. Tree maps displaying the taxonomic composition of the complete Outer Oslofjorden OF2 station protist microscopy dataset at class levels.
Table S1. Protist V4 18S rRNA gene OTUs recorded at station OF2 during the sampling period September 2009 to June 2011 with number of reads in each sample and taxonomic placement. Available from FigShare at the site: https://figshare.com/s/31d2ea0e5a303fd8813e.
Table S2. Richness, Shannon, and Evenness diversity results with Welch two-tailed test to check correspondence between two studied depths (subsurface and bottom of chlorophyll maximum).
Table S3. PERMANOVA results with permutation of environmental factors for two studied depths (subsurface and bottom of chlorophyll maximum).
Table S4. New taxa recorded for the Skagerrak. Available from FigShare at the site: https://figshare.com/s/31d2ea0e5a303fd8813e.
Table S5. Microscopical biovolumes and cell counts at station OF2. Biovolumes (in μm^3) were estimated from cell counts using the HELCOM 2006 protocol (Olenina et al. 2006). Available from FigShare at the site: https://figshare.com/s/31d2ea0e5a303fd8813e.
Table S6. Welch two sample *t*-test results, to test correspondence between light microscopy cell counts and metabarcoding proportional read abundance.
File S1. Scripts for the bioinformatics pipeline in Qiime, phylogenetic analyses and statistical analyses in R. Available from FigShare at the site: https://figshare.com/s/31d2ea0e5a303fd8813e.
File S2. Environmental data; salinity, temperature, density, fluorescence, nutrients by depth. Available from FigShare at the site: https://figshare.com/s/31d2ea0e5a303fd8813e.
File S3. OTU fasta file with DNA sequences of all OTUs with more than 10 reads. Available from FigShare at the site: https://figshare.com/s/31d2ea0e5a303fd8813e.