OXFORD FEMS JOURNALS investing in science

FEMS Microbiology Ecology, 92, 2016, fiw050

doi: 10.1093/femsec/fiw050
Advance Access Publication Date: 2 March 2016
Research Article

RESEARCH ARTICLE

Protistan diversity and activity inferred from RNA and DNA at a coastal ocean site in the eastern North Pacific

Sarah K. Hu[1,*], Victoria Campbell[1,2], Paige Connell[1], Alyssa G. Gellene[1], Zhenfeng Liu[1], Ramon Terrado[1] and David A. Caron[1]

[1]Department of Biological Sciences, University of Southern California, 3616 Trousdale Parkway, Los Angeles, CA 90089-0371, USA and [2]UW Medicine, Division Allergy and Infectious Diseases, 750 Republican St., Seattle, WA 98109, USA

*Corresponding author: Department of Biological Sciences, University of Southern California, 3616 Trousdale Parkway, Los Angeles, CA 90089-0371, USA. Tel: +(213)-821-1800; E-mail: sarah.hu@usc.edu

One sentence summary: Using both DNA and RNA sequencing, this study characterizes the total and active microbial eukaryotic communities with respect to season and location, and documents previously undetected activity in protists.

Editor: Julie Olson

ABSTRACT

Microbial eukaryotes fulfill key ecological positions in marine food webs. Molecular approaches that connect protistan diversity and biogeography to their diverse metabolisms will greatly improve our understanding of marine ecosystem function. The majority of molecular-based studies to date use 18S rRNA gene sequencing to characterize natural microbial assemblages, but this approach does not necessarily discriminate between active and non-active cells. We incorporated RNA sequencing into standard 18S rRNA gene sequence surveys with the purpose of assessing those members of the protistan community contributing to biogeochemical cycling (active organisms), using the ratio of cDNA (reverse transcribed from total RNA) to 18S rRNA gene sequences within major protistan taxonomic groups. Trophically important phytoplankton, such as diatoms and chlorophytes exhibited seasonal trends in relative activity. Additionally, both radiolaria and ciliates displayed previously unreported high relative activities below the euphotic zone. This study sheds new light on the relative metabolic activity of specific protistan groups and how microbial communities respond to changing environmental conditions.

Keywords: microbial ecology; microbial eukaryotes; protistan diversity; 18S rRNA gene sequencing; cDNA sequencing; DNA/RNA comparison

INTRODUCTION

Protists are fundamental for maintaining the functional stability of marine ecosystems (Sherr and Sherr 1994; Caron *et al.* 2012). Microbial eukaryotes serve as important links to higher trophic levels as primary producers and consumers (Sherr and Sherr 2002; Sherr *et al.* 2010). The vast morphological and genetic variability, and consequent physiological diversity among protistan species enables them to fulfill these diverse roles, yet complicates our efforts to study protistan community ecology (Behnke *et al.* 2011; Caron 2013).

Sequencing genes from environmental samples to supplement classical approaches of microscopy and culture for studying the diversity of microorganisms has revolutionized the field of microbial ecology. DNA sequencing has allowed researchers the means to rapidly characterize entire natural assemblages of protists and resolve some taxonomic and phylogenetic relationships among these species (López-García, Rodriguez-Valera and

Received: 2 December 2015; **Accepted:** 25 February 2016
© FEMS 2016. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

Pedros-Alio 2001; Moon-van der Staay, De Wachter and Vaulot 2001; Caron 2013; Massana 2015). High-throughput sequencing of the small subunit 18S rRNA gene is currently a cost-effective method for probing the diversity of these communities. Such studies have revealed immense genetic diversity in virtually all environments sampled to date (Moreira and López-García 2001; Stoeck *et al.* 2010; Caron *et al.* 2012; Armbrust and Palumbi 2015; de Vargas *et al.* 2015; Worden *et al.* 2015).

A number of recent molecular studies have promoted sequencing cDNA (reverse transcribed from extracted total RNA, hereafter referred to as RNA) as a supplement to 18S rRNA gene sequencing (hereafter, referred to as DNA) (Stoeck *et al.* 2007; Not *et al.* 2009; Terrado *et al.* 2011; Charvet, Vincent and Lovejoy 2014; Logares *et al.* 2014; Egge *et al.* 2015; Massana *et al.* 2015). DNA-based sequencing methods provide genetic identifications of Operational Taxonomic Units (OTUs) of the protistan species present in a sample, but they do not necessarily discriminate active organisms from moribund, encysted, metabolically inactive or even non-living genetic material (Stoeck *et al.* 2007; Not *et al.* 2009). DNA is generally more resistant to degradation than RNA (Lorenz and Wackernagel 1987; Karl and Bailiff 1989), and therefore sequence information derived from intact total RNA implies ribosomal activity and the potential for protein synthesis (Poulsen, Ballard and Stahl 1993; Corinaldesi *et al.* 2010; Blazewicz *et al.* 2013; Lejzerowicz, Voltsky and Pawlowski 2013; Egge *et al.* 2015; Massana *et al.* 2015). Consequently, the ratio of RNA to DNA sequences (RNA:DNA) has been used as an index of metabolic activity (Charvet, Vincent and Lovejoy 2014; Logares *et al.* 2014; Massana *et al.* 2015). Results derived from RNA sequencing can provide significant insight into the potential for protein synthesis in the microbial community, as long as the limitations and assumptions of the method are made clear (Blazewicz *et al.* 2013).

Comparisons of RNA and DNA sequence abundances have revealed environmental selection of active species in natural and experimental systems (Poulsen, Ballard and Stahl 1993; Stoeck *et al.* 2007; Not *et al.* 2009; Charvet, Vincent and Lovejoy 2014; Massana *et al.* 2015). In an Arctic freshwater lake study, changes in the RNA-based sequence community correlated with experimental conditions, while DNA-based sequence diversity remained relatively unchanged among treatments (Charvet, Vincent and Lovejoy 2014). Thus, community diversity derived from RNA appears more responsive to environmental conditions than DNA (Terrado *et al.* 2011). RNA–DNA comparisons may be useful for characterizing the dynamics of the protistan rare biosphere (Caron and Countway 2009), where metabolic activity has been detected (Logares *et al.* 2014; Debroas, Hugoni and Domaizon 2015).

Our objectives were to examine differences in community structure based on RNA or DNA sequence results and use RNA:DNA ratios to provide new insight into the dynamics of the functionally active component of the protistan community. Changes in the diversity of the total protistan community (inferred from DNA) and the subset of the community presumed metabolically active (inferred from RNA) were examined at three stations with varying levels of anthropogenic influence in the eastern North Pacific off southern California. Samples were collected seasonally at the San Pedro Ocean Time-series (SPOT) station at four depths (surface, subsurface chlorophyll maximum, 150 and 890 m), and also from surface waters in the Port of Los Angeles (Port of LA), and near Santa Catalina Island (Catalina). RNA and DNA sequences of the V4 hypervariable region were compared in order to infer relative activity of protistan communities temporally (seasonally) and spatially

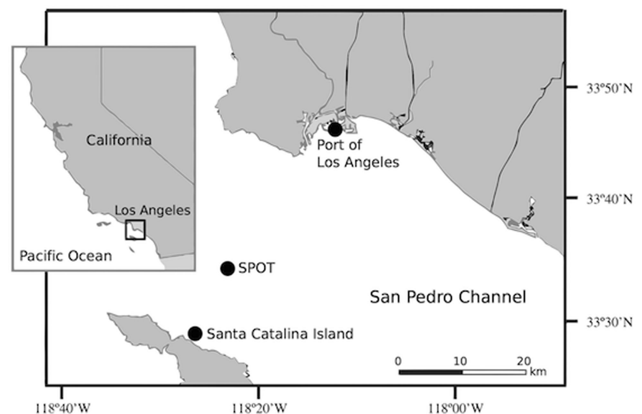

Figure 1. Map of San Pedro Channel, off the coast of southern California. Black circles indicate the three locations sampled; the San Pedro Ocean Time-series (SPOT) station (33°33′ N, 118°24′ W), the Port of Los Angeles (33°42.75′ N, 118°15.55′ W) and offshore from Santa Catalina Island (33°27.17′ N, 118°28.51′ W). Surface (5 m), subsurface chlorophyll maximum, 150 and 890 m were sampled at the SPOT station, while surface water was sampled at the Port of Los Angeles and offshore from Santa Catalina Island. Seawater and environmental parameters were collected at each location quarterly in April, July, October (2013) and January (2014) (Table S1, Supporting Information). Map modified from Schnetzer *et al.* 2011.

(horizontally and vertically). We also address the limitations of evaluating diverse protistan communities using an RNA:DNA ratio approach by separating observations according to individual taxonomic groups. The majority of sequence-based studies to date do not address *in situ* metabolic activity of microbial communities with respect to season or location. This study contributes to our knowledge of the biologically important roles that protists play in marine ecosystems, by evaluating protistan community structure and relative activity.

MATERIALS AND METHODS

Sample collection

Seawater samples were collected in April, July, October and January (2013–14) from the SPOT (33°33′N, 118°24′W), the Port of LA (33°42.75′ N, 118°15.55′ W) and ∼1 km offshore from Catalina (33°27.17′ N, 118°28.51′ W) in the eastern North Pacific (Fig. 1; Table S1, Supporting Information). The SPOT station was sampled from 5 m, the subsurface chlorophyll maximum (SCM), 150 and 890 m using 10 L Niskin bottles mounted on a CTD rosette, during regularly scheduled cruises (https://dornsife.usc.edu/spot/). The depth of the SCM was determined using real-time fluorescence data from the CTD downcast, with seawater collected during the CTD upcast. Surface water from the Port of LA and Catalina was collected within 1 week of the SPOT cruise sampling dates (Table S1, Supporting Information), by submerging acid-cleaned and pre-rinsed 20 L carboys below the seawater surface and capping the containers while underwater to minimize bubble formation and contamination with the surface microlayer.

Seawater from all samples was sequentially pre-filtered through 200 μm and 80 μm Nitex mesh to reduce abundances of multicellular eukaryotes (metazoa). Near-surface and SCM seawater (2 L) and 150 and 890 m seawater (4 L) was filtered onto GF/F filters (nominal pore size 0.7 μm; Whatman, International Ltd, Florham Park, NJ, USA) and immediately flash frozen in liquid nitrogen for later DNA and RNA extraction. Oxygen,

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

temperature, salinity and fluorescence data were obtained from either CTD sensors or a profiling natural fluorometer. Samples for chlorophyll *a* and nutrients were collected at all three surface sites; details can be found in Supporting Information.

Nucleic acid extraction and sequencing

Total DNA and RNA were extracted simultaneously from each sample using the All Prep DNA/RNA Mini kit (Qiagen, Valencia, CA, USA, #80204). Genomic DNA was removed during the RNA extraction with RNase-Free DNase reagents (Qiagen, #79254). Total extracted RNA was checked for residual genomic DNA by performing a polymerase chain reaction (PCR) using DNA specific primers to ensure that no amplified products appeared when run on an agarose gel. RNA was reverse transcribed into cDNA using iScript Reverse Transcription Supermix with random hexamers (Bio-Rad Laboratories, Hercules, CA, USA, #170-8840).

The resulting cDNA and DNA from each sample were PCR amplified using V4 forward (5′-CCAGCA[GC]C[CT]GCGGTA ATTCC-3′) and reverse (5′-ACTTTCGTTCTTGAT[CT][AG]A-3′) primers (Stoeck *et al.* 2010). Duplicate PCR reactions were performed in 50 μL volumes of: 1X Phusion High-Fidelity DNA buffer, 1 unit of Phusion DNA polymerase (New England Biolabs, Ipswich, MA, USA, #M0530S), 200 μM of dNTPs, 0.5 μM of each V4 forward and reverse primer, 3% DMSO, 50 mM of MgCl and 5 ng of either DNA or cDNA template per reaction. The PCR thermal cycler program consisted of a 98°C denaturation step for 30 s, followed by 10 cycles of 10 s at 98°C, 30 s at 53°C and 30 s at 72°C, and then 15 cycles of 10 s at 98°C, 30 s at 48°C and 30 s at 72°C, and a final elongation step at 72°C for 10 min, as described in Rodríguez-Martínez *et al.* (2012). PCR products were purified (Qiagen, #28104) and duplicate samples were pooled. The ~400 bp cDNA and DNA PCR products were quality checked on an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA).

Sequence analysis

A 250 × 250 bp paired-end sequencing was performed on an Illumina MiSeq (V2 chemistry). Nucleotide bases with a Q score lower than 20 for the last 30 bp of each sequence were trimmed. Paired-end sequences were merged using FLASh (Magoc and Salzberg 2011) with a minimum of 10 bp and maximum of 150 bp overlap between each sequence pair. Sequences shorter than 350 bp, longer than 460 bp, or which had an average quality score lower than 25 were discarded using QIIME v1.8 (Caporaso *et al.* 2010). Chimeric sequences were identified and removed, by either *de novo* or reference-based chimera checking (identify_chimeric_seqs.py in QIIME, intersection method). Sequence data have been deposited in the NCBI SRA database under accession number SRP070577.

OTUs were generated using QIIME's subsampled open-reference OTU clustering protocol, which combined closed-reference (using the SILVA v.111 database, Quast *et al.* 2012) and *de novo* (without any reference database) OTU clustering to decrease run time and ensure that the maximum number of sequences were placed into OTUs at 97% sequence similarity (as described in Rideout *et al.* 2014). OTUs with only one (singleton) or two (doubleton) sequences were removed from the dataset for all downstream analyses (Behnke *et al.* 2011).

For each site and month, OTUs with only RNA (RNA-only OTUs) or DNA (DNA-only OTUs) were not included in the analysis. RNA- or DNA-only OTUs may have arisen due to sequencing error, differences in detection level with regard to RNA or DNA copy number, or reverse transcription error [in the case of RNA-only OTUs (Egge *et al.* 2013)]. Defining DNA-only OTUs as dormant was deemed inappropriate, as dormant cells often have detectable levels of RNA (Blazewicz *et al.* 2013). We therefore focused on OTUs with detectable levels of both RNA and DNA.

Protistan community diversity and composition

Whole protistan community composition was evaluated by first compiling all RNA- and DNA-based OTUs into phylum or class level taxonomic designations (SILVA). Each OTU was manually assigned to a major taxonomic group, consisting of dinoflagellates, ciliates, other alveolates (the latter mainly comprised of Syndiniales as defined in Ohtsuka *et al.* 2015), chlorophytes, cryptophytes, haptophytes, MArine STramenopiles (MASTs, as defined in Massana *et al.* 2004), diatoms, other stramenopiles (the latter mainly comprised of *Chrysophyceae* and *Dictyochophyceae*), cercozoa, radiolaria (also known as retaria, Adl *et al.* 2012), metazoa and others.

Bray-Curtis dissimilarity matrices were constructed based on the RNA or DNA sequence results for the three surface stations and four depths at SPOT separately, using the function 'simprof' in the 'clustsig' R package (Clarke, Somerfield and Gorley 2008; Whitaker and Christman 2014). The data were normalized across samples by calculating the relative abundance of each OTU. Then Bray-Curtis distance matrices were visualized by creating dendrograms based on average hierarchical clustering ['hclust' function in R (R Core Team 2014)]. Weighted UniFrac analyses, which incorporated phylogenetic distances among OTUs, were also computed; details can be found in Supporting Information.

Relative activity of protistan groups

RNA:DNA ratios were calculated for each OTU, then average ratios for each major taxonomic group were used as a proxy for relative activity. The relative activity was only compared across sites (spatial) or months (temporal) within an individual taxonomic group because gene copy number per cell can vary widely among different protistan lineages (groups) (Vaulot, Romari and Not 2002; Zhu *et al.* 2005; Godhe *et al.* 2008; Massana and Pedros-Alio 2008). Differences in RNA:DNA ratios by month and site were evaluated using analysis of variance in conjunction with Tukey's honest significant difference test using R (R Core Team 2014) ($P < 0.05$, conf. 0.95).

RESULTS

Environmental parameters at the study sites

The SPOT station is located approximately midway between the Port of LA and Catalina (Fig. 1). Sea surface temperatures at SPOT ranged seasonally from 16°C in April to 20°C in July during the study (Fig. 2A; Table S2, Supporting Information). The depth and intensity of the SCM at SPOT varied with season (Fig. 2B). Chlorophyll *a* at the SCM was highest during the spring and summer (2–4 μg L^{-1}) and lowest during the fall and winter (1.0–1.7 μg L^{-1}, Fig. 2B). The SCM was sampled at 32 m in April, 33 m in July, 34 m in October and 42 m in January. The April SCM sample was obtained from the lower edge of the chlorophyll *a* maximum at 32 m (2 μg L^{-1}) instead of ca. 25 m (Fig. 2B).

SPOT is situated over a deep coastal basin (the San Pedro Basin) with a maximum depth of ca. 900 m. Underwater sills at ca. 740 m to the north and south, and Catalina Island to the west restrict water circulation in the lower half of the water

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

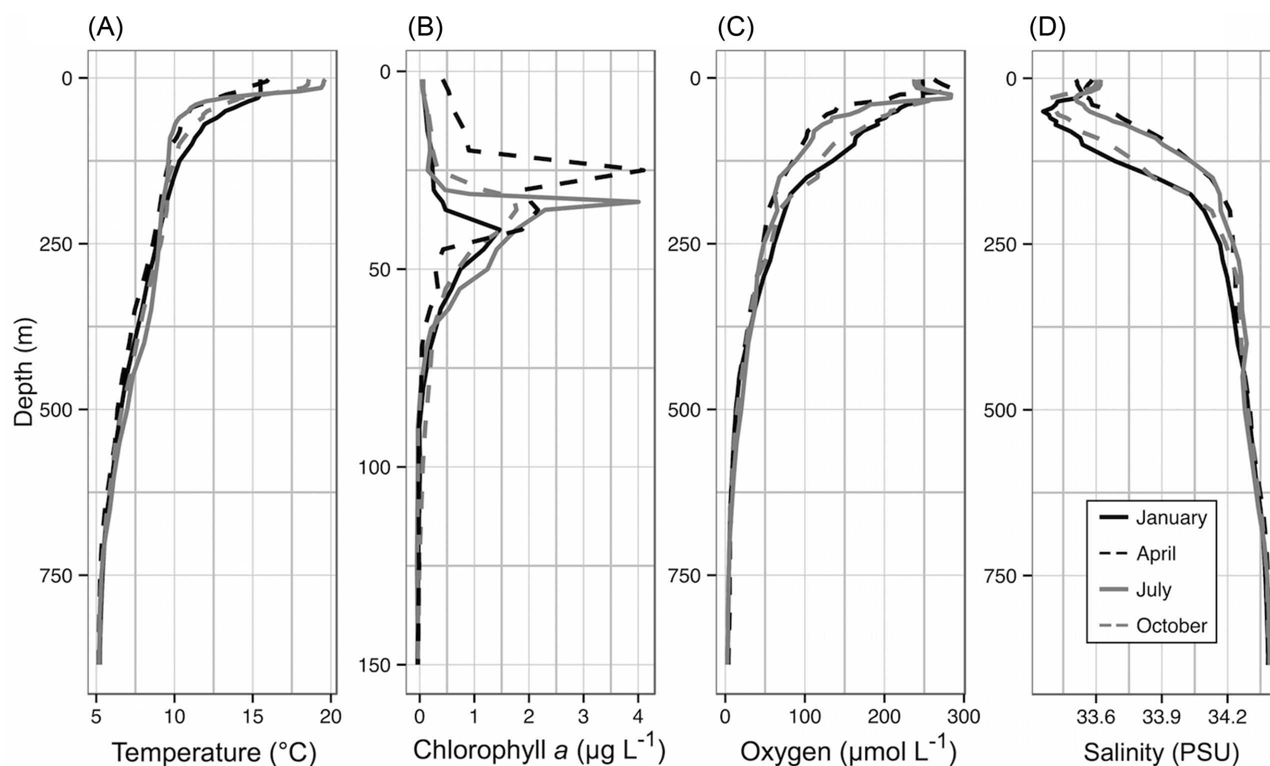

Figure 2. Environmental parameters throughout the water column at the San Pedro Ocean Time-series (SPOT) station during January, April, July and October (Table S1, Supporting Information). Values for (**A**) temperature (°C), (**B**) chlorophyll *a* (μg L^{-1}) (note the change in y-axis in B), (**C**) dissolved oxygen (μmol L^{-1}) and (**D**) salinity (PSU) were obtained from CTD sensor data (Seabird Electronics) during water collection.

column. Temperatures did not fluctuate seasonally below the euphotic zone and remained ca. 9.6°C at 150 m and 5.2°C at 890 m (Fig. 2A). Dissolved oxygen concentrations at 150 m during the study ranged between 90 and 130 μmol L^{-1} (Fig. 2C). There was a persistent oxycline extending from below the euphotic zone to ~350 m, and dissolved oxygen levels were <45 μmol L^{-1}. Below 350 m, the water column remained hypoxic throughout the course of the study (Fig. 2C).

Surface salinity across all three stations and sampling dates remained between 31 and 34 PSU and sea surface temperatures ranged from 14°C to 20°C (Table S2, Supporting Information). The Port of LA had higher concentrations of chlorophyll *a*, phosphate, silicate, nitrite, nitrate and ammonium compared to Catalina and SPOT, indicating the more eutrophic nature of the nearshore station (Table S2, Supporting Information). Chlorophyll *a* concentrations were highest during April at all three surface stations (Table S2, Supporting Information).

Eukaryotic RNA and DNA sequences

High-throughput sequencing initially recovered ca. 8 million RNA and DNA sequences from 24 samples. Approximately 3.1 million of these sequences were discarded following quality checking, paired-end merging and chimera detection (1.6 million chimeric sequences), resulting in 4.88 million high-quality RNA (2.3 million) and DNA (2.5 million) V4 sequences. Sequences were clustered into OTUs at 97% sequence similarity, which generated 30 781 OTUs at approximately species level designations. Following removal of singleton and doubleton OTUs, there were 4.57 million sequences and 18 393 OTUs (See Table S3, Supporting Information for sequence and OTU counts). For all downstream analyses, only OTUs with both RNA and DNA sequences for each sample were used in the analysis, leaving 4.27 million sequences (6.6% decrease in the total number of sequences) and 6931 OTUs (Table S3, Supporting Information). Final RNA and DNA sequence libraries consisted of 2.1 million and 2.2 million sequences, respectively (Table 1). The total number of sequences found in each OTU and taxonomic designation for each OTU can be found in Table S7, Supporting Information. Analyses for community structure and sequence abundances were repeated with all OTUs and showed the same trends as the final culled dataset, see Supporting Information (Table S8 and Fig. S1, Supporting Information).

Protistan community diversity and composition

The same major protistan taxonomic groups (manually designated groups, see Materials and Methods) were detected across all months and sites in both RNA and DNA sequence libraries, although not always in the same relative proportions (Table 1). Alveolates (including dinoflagellates, ciliates and other alveolates) made the largest single contribution to both RNA and DNA sequence libraries (over 40% each, Table 1). After alveolates, chlorophytes, cryptophytes, haptophytes and stramenopiles (the latter includes diatoms, MAST and other stramenopiles) made up a combined 43% of the RNA sequence library, while the same groups contributed only 17% of the DNA sequences (Table 1). Within the rhizaria, less than 2% of the RNA and DNA library were identified as cercozoa (Table 1). Radiolaria comprised less than 2% of the total number of RNA sequences, but contributed 11% of the total number of DNA sequences. Metazoan sequences made up ca. 21% of the DNA sequence library, while fewer than 2% of the RNA sequences were identified as metazoa (Table 1).

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

Table 1. Total number of RNA and DNA sequences, and percentage of the total number of sequences by major taxonomic groups. Values for RNA are shaded.

		RNA		DNA	
		Sequences	Percent	Sequences	Percent
Alveolates	Dinoflagellates	373 698	18.2	825 700	37.2
	Ciliates	501 779	24.4	161 796	7.29
	Other	8999	0.44	8792	0.40
	Chlorophytes	115 167	5.61	81 233	3.66
	Cryptophytes	73 146	3.56	32 586	1.47
	Haptophytes	227 453	11.1	99 109	4.47
Stramenopiles	Diatoms	139 817	6.80	88 447	3.99
	MAST	136 428	6.64	29 227	1.32
	Other	193 555	9.42	38 625	1.74
Rhizaria	Cercozoa	39 954	1.94	30 425	1.37
	Radiolaria	39 145	1.91	244 545	11.0
	Metazoa	26 236	1.28	467 102	21.1
	Other	80 378	3.91	53 632	2.42
	Unassigned	98 919	4.81	57 716	2.60
	Total	2 054 674		2 218 935	

Community dissimilarity based on DNA samples grouped the surface stations approximately by site, Port of LA samples formed a cluster separate from SPOT and Catalina, with the exception of the Port of LA in April and Catalina in January (Fig. 3A). DNA-based diversity in the Port of LA in April was made up of 72% dinoflagellate sequences, while dinoflagellates made up ≤50% of the total number of sequences in other samples (Table S4, Supporting Information). In January, the DNA sample at Catalina revealed comparatively larger abundances of dinoflagellates (44%) and ciliates (20%) relative to other SPOT and Catalina samples (Table S4, Supporting Information).

Bray-Curtis dissimilarity results from the RNA sequence library formed more distinct clusters with respect to both site and season relative to DNA-based estimates (compare Fig. 3A and B). Samples from the Port of LA formed a cluster separate from SPOT and Catalina, with the exception of the Port of LA in April (Fig. 3B). Diatoms generally made up <10% of the total number of RNA sequences in surface samples, but in April, diatoms contributed 45% of the sequences in the Port of LA (Table S4, Supporting Information).

RNA-based communities were not significantly different between SPOT and Catalina within each month sampled ($P < 0.05$; Fig. 3B), although dominant taxonomic groups varied by month (Table S4, Supporting Information). In January, both SPOT and Catalina were dominated by dinoflagellates (SPOT: 21%, Catalina: 15%), ciliates (SPOT: 35%, Catalina: 31%) and haptophytes (SPOT: 18%, Catalina: 21%). The same taxonomic groups were found in April at SPOT and Catalina, but at different relative abundances compared to January: dinoflagellates (SPOT: 11%, Catalina: 16%), ciliates (SPOT: 14%, Catalina: 17%) and haptophytes (SPOT: 17%, Catalina: 22%) (Table S4, Supporting Information). Compared to January and April, dinoflagellates in July and October contributed fewer sequences (between 2%–5%), ciliates were still dominant in July (SPOT: 13%, Catalina: 18%), but less abundant in October (SPOT: 3%, Catalina: 9%), and haptophytes were found to be at lower abundances in July (SPOT: 10%, Catalina: 9%), and high abundances in October (SPOT: 34%, Catalina 26%) (Table S4, Supporting Information). Additionally, the total number of chlorophyte RNA sequences was higher in July (SPOT: 13%, Catalina: 23%) and October (SPOT: 10%, Catalina 23%), relative to January and April (Table S4, Supporting Information).

Analysis of protistan community dissimilarity from four depths at the SPOT station revealed greater agreement between the clustering patterns of the RNA and DNA libraries (Fig. 3C and D) relative to the dissimilarity among the three surface stations (Fig. 3A and B). The clearest trend among these samples was that both RNA- and DNA-based communities from 150 and 890 m grouped apart from shallower depths (5 m and SCM) (Fig. 3C and D). Diversity from shallower depths was mainly comprised of dinoflagellates, ciliates, chlorophytes, haptophytes, stramenopiles and (specific only to DNA) metazoa (Table S4, Supporting Information). Communities at 150 and 890 m were not significantly different from one another ($P < 0.05$; dashed lines in Fig. 3C and D, except for January 890 m DNA sample), and were mainly comprised of dinoflagellates, ciliates, radiolaria and (specific only to DNA) metazoa (Table S4, Supporting Information). The one exception at 890 m where the DNA sample in January did not group with other 890 m samples (Fig. 3C) coincided to substantially higher abundances of chlorophyte (3%), haptophyte (18%) and other stramenopile (11%) sequences, relative to other 890 m depths (Table S4, Supporting Information).

DNA-based communities at 5 m from all months at the SPOT station were not significantly different ($P < 0.05$; Fig. 3C), while season affected the RNA-derived community structure at 5 m (Fig. 3D). RNA samples from April and July were not significantly different from one another ($P < 0.05$; Fig. 3D). There were higher abundances of MAST sequences in April (23%) and July (21%) relative to January (3%) and October (6%) at 5 m (Table S4, Supporting Information). SCM samples from January and October were not significantly different from one another in either DNA- or RNA-derived Bray-Curtis dissimilarity results ($P < 0.05$; Fig. 3C and D). The April sample at SCM did not cluster with other euphotic zone samples (Fig. 3C and D), coinciding with high relative abundances of diatom sequences (RNA: 33%, DNA: 18%, Table S4, Supporting Information).

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

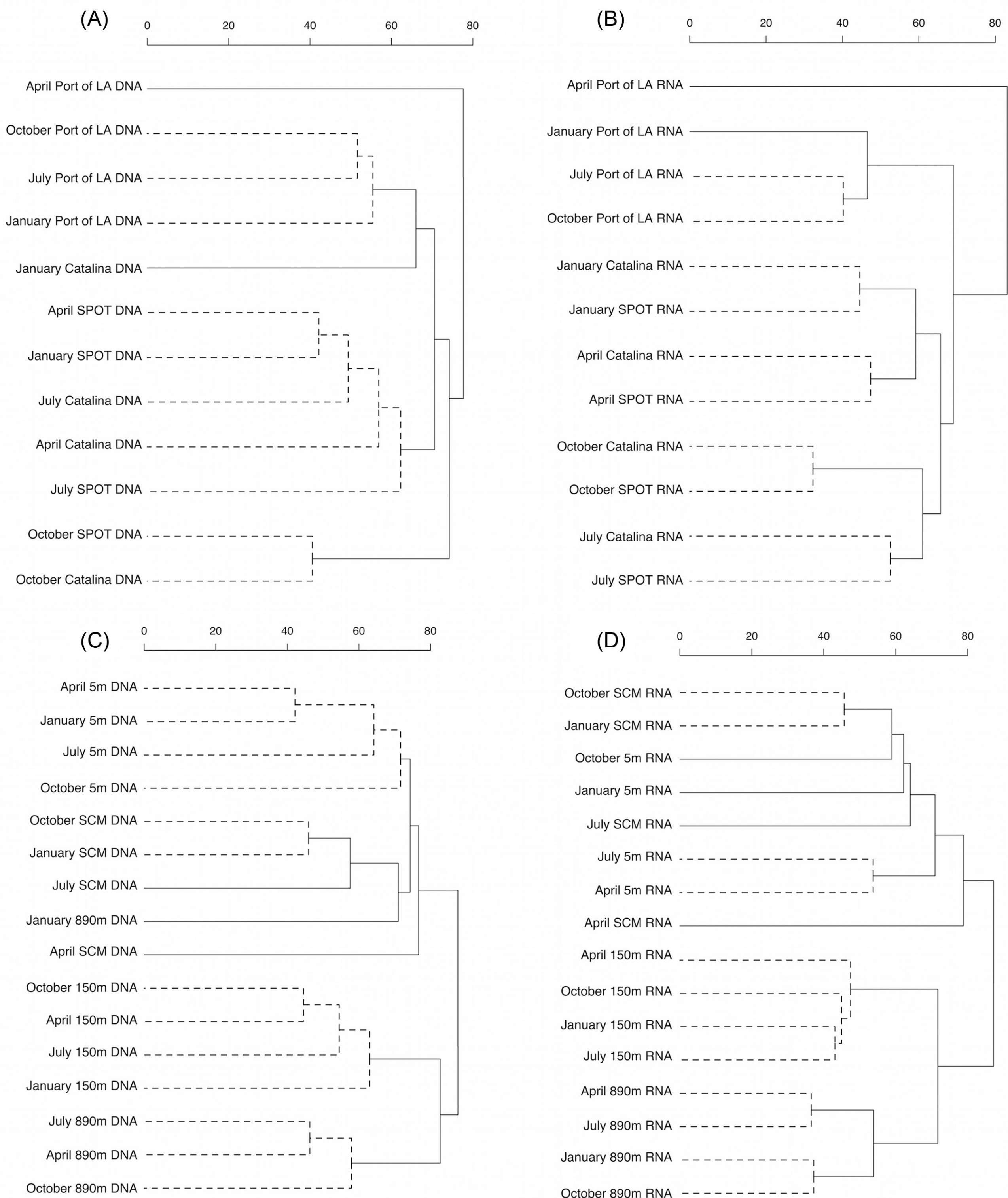

Figure 3. Cluster dendrograms based on average hierarchical clustering. Dendrograms depict level of dissimilarity among (**A**) DNA samples from SPOT, the Port of LA and Catalina (surface stations), (**B**) RNA samples from surface stations, (**C**) DNA samples from four depths at the SPOT station (5 m, subsurface chlorophyll maximum (SCM), 150 and 890 m) and (**D**) RNA samples from the four depths at the SPOT station. Data were normalized by calculating the relative abundance of each OTU, and then Bray-Curtis dissimilarity matrices were constructed for surface stations (A and B) and for each depth sampled at SPOT (C and D), using either DNA (A and C) or RNA (B and D) sequence libraries. The percent dissimilarity among samples is depicted by horizontal axes. Dashed lines represent samples that were not significantly different from one another ($P < 0.05$).

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

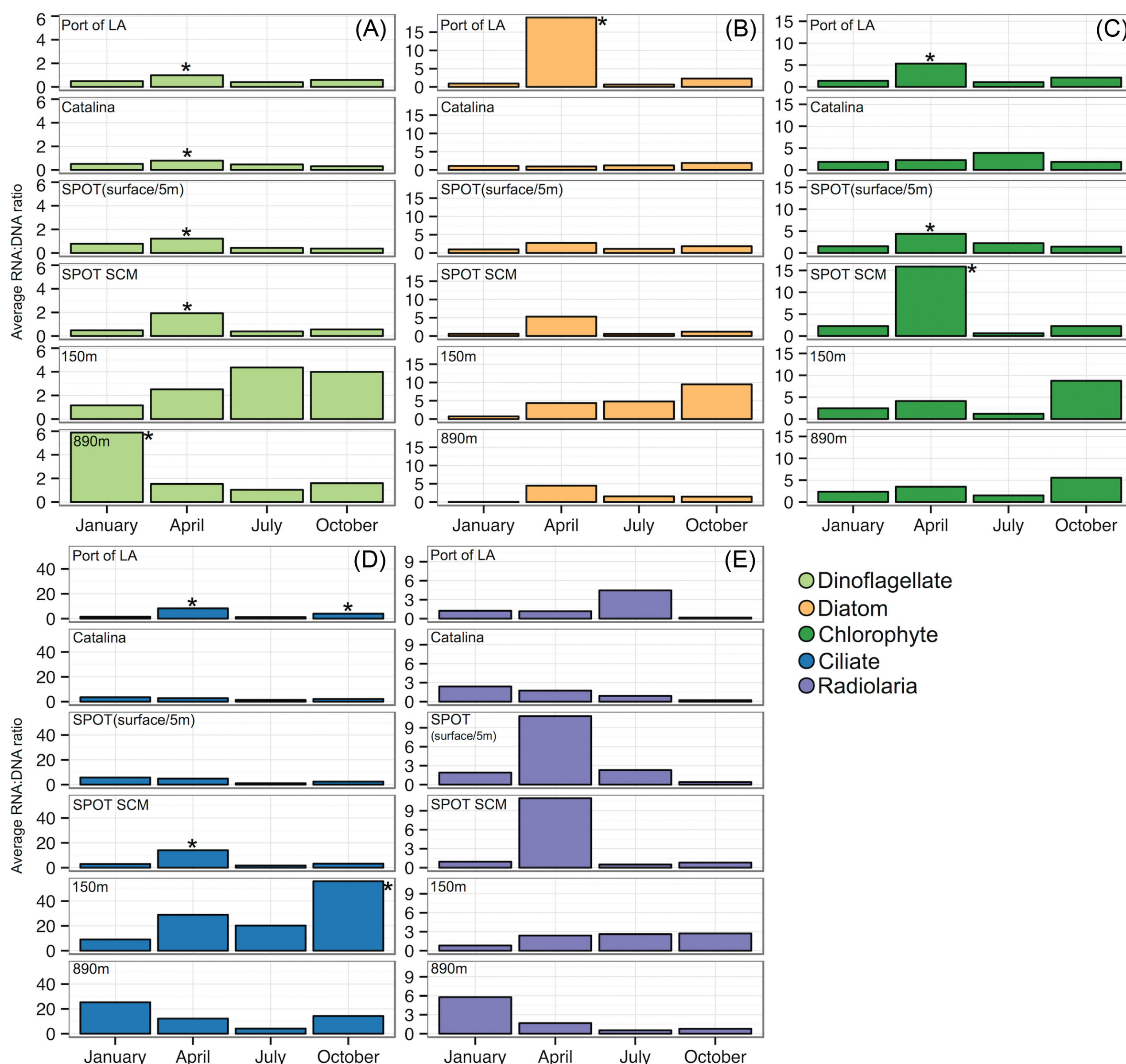

Figure 4. Average RNA:DNA ratios from select taxonomic groups. Ratios represent relative activity at (from top to bottom) the Port of LA, Catalina, SPOT station surface/5 m, SCM, 150 and 890 m, in each month sampled for (**A**) dinoflagellates, (**B**) diatoms, (**C**) chlorophytes, (**D**) ciliates, (**E**) and radiolaria. Average RNA:DNA ratios for all taxonomic groups reported in Table S5, Supporting Information. Asterisks (*) denote significantly higher ($P < 0.05$) average RNA:DNA ratios relative to other months (pairwise comparisons, Table S6, Supporting Information).

Relative activity of protistan groups

Interpretations of relative activity (RNA:DNA ratios) were conducted separately for each major taxonomic group (Table S5, Supporting Information) and graphically visualized for select groups (Fig. 4). Average RNA:DNA ratios for each major group varied widely (Fig. 5). Statistically significant differences in RNA:DNA ratios revealed seasonal and depth-related trends in relative activity ($P < 0.05$; Figs 4, 6 and 7; Tables S5 and S6, Supporting Information).

A common seasonal trend in the euphotic zone (surface and SCM at the SPOT station, Port of LA or Catalina) was that the majority of protistan groups had significantly higher RNA:DNA ratios in April, relative to other months ($P < 0.05$; Fig. 4; Table S5, Supporting Information). Dinoflagellates had significantly higher RNA:DNA ratios in all euphotic zone samples in April ($P < 0.05$; Fig. 4A; Tables S5 and S6, Supporting Information). In April, diatom (Figs 4B and 6) and cercozoa RNA:DNA ratios were significantly higher at the Port of LA, compared to all other months and depths ($P < 0.05$; Tables S5 and S6, Supporting Information). Chlorophytes (Fig. 4C), haptophytes and MASTs had significantly higher relative activity in April at the surface and SCM at the SPOT station and at the Port of LA compared to other months ($P < 0.05$; Tables S5 and S6, Supporting Information). Ciliates had higher levels of relative activity at the SCM (SPOT) and the Port of LA in April ($P < 0.05$; Fig. 4D; Tables S5 and S6, Supporting Information). Higher RNA:DNA ratios at the SCM at SPOT were also observed for cryptophytes ($P < 0.05$; Tables S5 and S6, Supporting Information).

There were fewer seasonal trends in RNA:DNA ratios in deep samples (150 and 890 m) than in the euphotic zone.

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

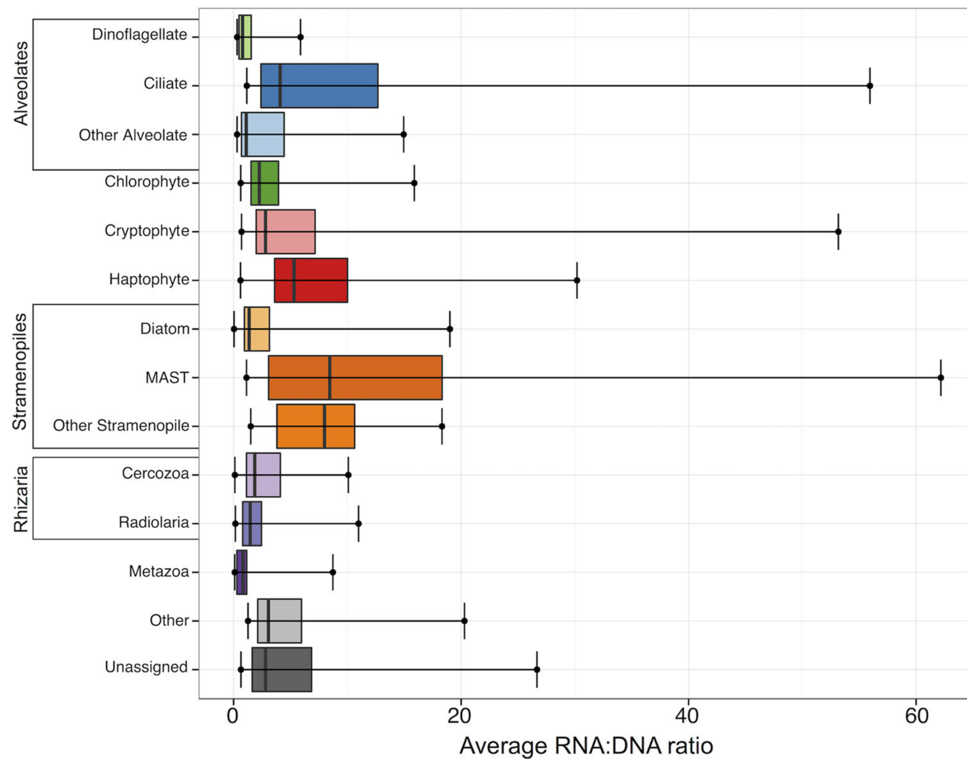

Figure 5. Range of average RNA:DNA ratios for each taxonomic group. Boxplots represent variation among the average RNA:DNA ratios in each sample. Whiskers denote minimum and maximum values. Average RNA:DNA ratios among individual taxonomic groups were highly variable, making direct comparisons between different groups difficult. See Table S5, Supporting Information for all ratios.

Dinoflagellates had significantly higher RNA:DNA ratios at 890 m in January ($P < 0.05$; Fig. 4A; Tables S5 and S6, Supporting Information). Relative activity of diatoms (Fig. 4B) and ciliates (significant $P < 0.05$; Fig. 4D) at 150 m was highest in October relative to other months (Tables S5 and S6, Supporting Information).

Average RNA:DNA ratios revealed vertical trends and a few horizontal spatial trends. Overall, ciliate ratios at the SPOT station were significantly higher at 150 m relative to other depths ($P < 0.05$; Fig. 7A; Tables S5 and S6, Supporting Information). Haptophytes and cercozoa also had higher relative activities at 150 m at the SPOT station, while dinoflagellates, cryptophytes and MASTs had significantly higher ratios at 150 and 890 m relative to the euphotic zone ($P < 0.05$; Tables S5 and S6, Supporting Information). At the SPOT station in April, radiolarian RNA:DNA ratios were significantly higher at the SCM compared to 150 and 890 m, although species richness was greater at deeper depths than at the SCM ($P < 0.05$; Fig. 7B; Tables S5 and S6, Supporting Information).

One major difference between RNA- and DNA-derived community composition in this study was that metazoa were more abundant in DNA samples, while fewer sequences in RNA samples were identified as metazoa (Tables 1, S4 and S7, Supporting Information). High abundances of metazoan sequences in microbial molecular surveys can mask the presence of rare protistan sequences, thus their contribution is often minimized through pre-filtration during seawater collection, or removed during downstream bioinformatic analyses (Countway *et al.* 2010; Schnetzer *et al.* 2011), see Materials and Methods. Differences between the RNA and DNA libraries indicated either the presence of extracellular metazoan DNA (e.g. broken cellular material from the pre-filtration process) or that metazoa have fewer copies of RNA relative to DNA. Regardless, the small numbers of metazoan sequences found in the RNA libraries were considered advantageous for future protistan-focused studies that strive to reduce the presence of metazoan sequences.

DISCUSSION

RNA–DNA comparisons (diversity based on RNA or DNA, RNA:DNA ratios) in this study revealed insights into natural protistan assemblages at different stations and depths in the eastern North Pacific. DNA-based molecular surveys have contributed considerable knowledge of protistan biogeography and diversity (e.g. López-García, Rodriguez-Valera and Pedros-Alio 2001; Caron *et al.* 2012; de Vargas *et al.* 2015; Massana *et al.* 2015), but there is still relatively little information regarding members of the community that are active and how protistan activity is impacted by environmental conditions (Parris *et al.* 2014; Jing *et al.* 2015; Stecher *et al.* 2015).

Observed seasonality across surface stations

A major finding of this study was that RNA-based Bray-Curtis dissimilarity analysis more clearly clustered surface samples according to general environmental and seasonal conditions (chlorophyll and nutrient concentrations, Table S2, Supporting Information) relative to DNA (Fig. 3A and B). Community similarity analysis broadly grouped samples relative to site, and then more finely according to season (Fig. 3B). The ability of the RNA-based Bray-Curtis dissimilarity estimates to differentiate seasons suggests that RNA provided a better snapshot of the active component of the protistan community relative to DNA (Countway *et al.* 2010; Schnetzer *et al.* 2011).

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

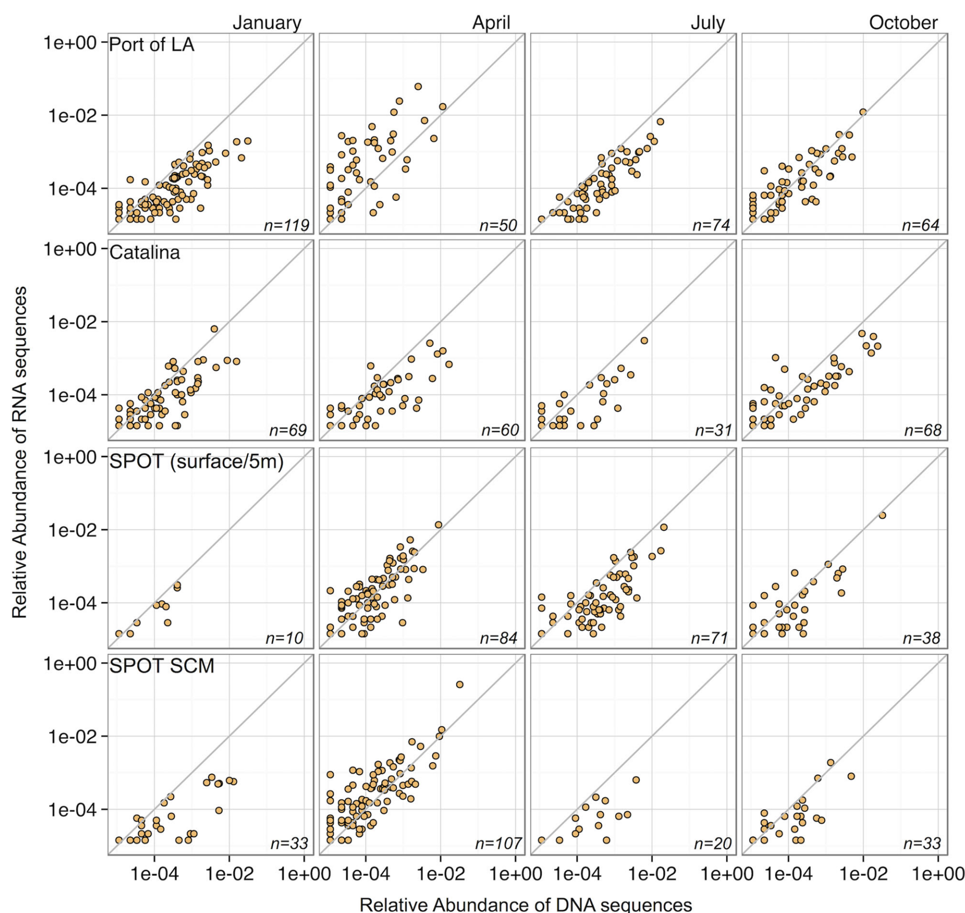

Figure 6. Relative abundance of RNA (y-axis) and DNA (x-axis) sequences for each diatom OTU detected in the euphotic zone. From top to bottom: Port of LA (surface), Catalina (surface), SPOT (surface/5 m) and SPOT subsurface chlorophyll maximum (SCM). Months are depicted from left to right. Each data point depicts an OTU (n = total number of OTUs). RNA:DNA ratios for diatoms were significantly higher at the Port of LA in April, compared to all other months and depths ($P < 0.05$, Tables S5 and S6, Supporting Information).

We acknowledge limitations in the use of RNA:DNA ratios to evaluate microbial activity by making comparisons of relative activity within individual taxonomic groups. Variations in gene copy number (Prokopowich, Gregory and Crease 2003; Godhe *et al*. 2008; Gong *et al*. 2013), cell size (Zhu *et al*. 2005; Godhe *et al*. 2008) and dormant (including cyst-forming) cells make the interpretation of diversity from DNA sequence counts difficult. Similarly, rRNA copy numbers vary as a factor of metabolic state or cell size (Blazewicz *et al*. 2013). Therefore, using a RNA:DNA ratio to infer metabolic activity of a diverse (both metabolically and taxonomically) protistan community is complicated. For example, in the present study, the maximum average dinoflagellate RNA:DNA ratios were ~6, while maximum average RNA:DNA ratios for MASTs were ~62 (Fig. 5; Table S5, Supporting Information). A direct comparison of dinoflagellate and MAST RNA:DNA ratios would imply that dinoflagellates were consistently less active than MASTs. It is prudent to compare RNA:DNA ratios among related species that have comparable cell sizes or metabolisms. For this reason, we used average RNA:DNA ratios within manually designated taxonomic groups as an indication of relative activity (see Materials and Methods).

RNA:DNA ratios provided insights into compositional and seasonal changes in the activity of phytoplankton taxa. Most protistan groups had high RNA:DNA ratios in April. Specifically, relative changes in RNA:DNA ratios of diatoms highlighted an intensification of their activity in April (Fig. 6; Tables S5 and S6, Supporting Information). Springtime increases in diatom absolute and relative abundance have previously been recorded in the Port of LA and the San Pedro Channel (Venrick 2002; Kim *et al*. 2009; Schnetzer *et al*. 2013; Seubert *et al*. 2013). In this study, high relative activity of diatoms in the Port of LA in April co-occurred with the highest chlorophyll *a* measurements observed in this study (Figs 4B and 6; Tables S2 and S5, Supporting Information) and a high biomass of diatoms from microscopy counts (Connell *et al*. in preparation). Diatom OTUs at that time were dominated by a few dominant taxa (Table S7, Supporting Information; Fig. 6).

Chlorophytes in April, comprised mainly of picoeukaryotes *Osterococcus, Micromonas* and *Bathycoccus*, had high relative activities at the SCM at SPOT. This finding was expected as previous studies have noted episodic dominance of minute chlorophytes in the San Pedro Channel (Countway and Caron 2006; Schnetzer *et al*. 2011; Kim *et al*. 2012; Lie *et al*. 2013; Kim *et al*. 2014). Countway and Caron (2006) noted a trend in which high *Ostreococcus* abundance often followed diatom blooms. While both chlorophyte and diatom RNA:DNA ratios were high in the Port of LA and the SCM at the SPOT station in April (Fig. 4B and C; Table S5, Supporting Information), sampling frequency in the current study was not sufficient to resolve temporal coupling of diatom and *Ostreococcus* abundances and activities.

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

Nevertheless, RNA:DNA ratios could provide an additional tool for examining temporal relationships among biologically important phytoplankton groups (Worden 2006).

Seasonal and spatial trends along the vertical profile at SPOT

RNA–DNA comparisons along the vertical profile at SPOT revealed depth-specific trends in community composition and relative activity that were not addressed in DNA-based analyses, emphasizing the value of combining RNA- and DNA-based sequencing efforts. Findings were consistent with previous studies at the SPOT station, where seasonal patterns in community structure were generally restricted to shallower depths (surface and SCM), and shallow and deep (150 and 890 m) protistan communities were distinct (Countway *et al.* 2010; Schnetzer *et al.* 2011; Kim *et al.* 2014) (Fig. 3C and D; Table S4, Supporting Information). However, RNA-based community dissimilarity analysis provided additional insight into how environmental conditions may influence microbial activity. DNA-derived diversity at 890 m in January grouped separate from other deep samples due to a large abundance of chlorophytes, haptophytes and other stramenopiles, which was not seen in the RNA library (Table S4, Supporting Information). This difference appears to indicate that the DNA library was influenced by sinking inactive or non-viable genetic material.

Functional activities of ciliates revealed seasonal trends in the euphotic zone and an uptick in activity at 150 m at the SPOT station (oxycline). RNA:DNA ratios for ciliates found in the euphotic zone were consistently higher in April, relative to other months (Figs 4D and 7A), which may be coupled to seasonal availability of prey, as ciliates are known to be significant grazers of bacteria and other protists (Sherr and Sherr 1994, 2002). Previous studies at SPOT have found evidence of decreases in the relative abundance of ciliate DNA sequences in deep samples (150 and 890 m) relative to shallower samples (surface to SCM) (Schnetzer *et al.* 2011; Kim *et al.* 2014). The same trend was observed in DNA samples from the present study, but relative abundances of ciliate RNA sequences were higher at depth compared to 5 m and SCM (Fig. 7A; Table S4, Supporting Information). Moreover, relative activities of ciliates and species richness was highest at 150 m (Fig. 7A; Tables S4 and S5, Supporting Information), suggesting that ciliates play an important role at the oxycline not previously documented at SPOT. Increased relative abundances and activities have also been described along major transition zones, where a stratified gradient (e.g. sharp oxygen decrease) may promote high prey abundances for ciliate grazers (e.g. Stoeck *et al.* 2007; Lin *et al.* 2008; Stock *et al.* 2009; Edgcomb *et al.* 2011; Wylezich and Jürgens 2011; Anderson, Winter and Jürgens 2012; Edgcomb and Pachiadaki 2014; Parris *et al.* 2014). Relative activities of dinoflagellates (Fig. 4A), haptophytes and MASTs were also higher at 150 m (Table S5, Supporting Information), in concordance with observations by Schnetzer *et al.* (2011). The authors speculated that the oxycline supported diverse microbial physiologies.

Molecular approaches allow us to study the diversity of natural protistan communities using culture-independent methods, and here our RNA and DNA results shed new light on radiolaria—an enigmatic group of which current knowledge is mainly drawn from surface species and which are difficult to rear in the laboratory (Caron and Swanberg 1990; Gilg *et al.* 2010; Decelle *et al.* 2012, 2013; Burki and Keeling 2014; de Vargas *et al.* 2015). Previous DNA-based protistan diversity studies have found that radiolaria make up a substantial component of the community

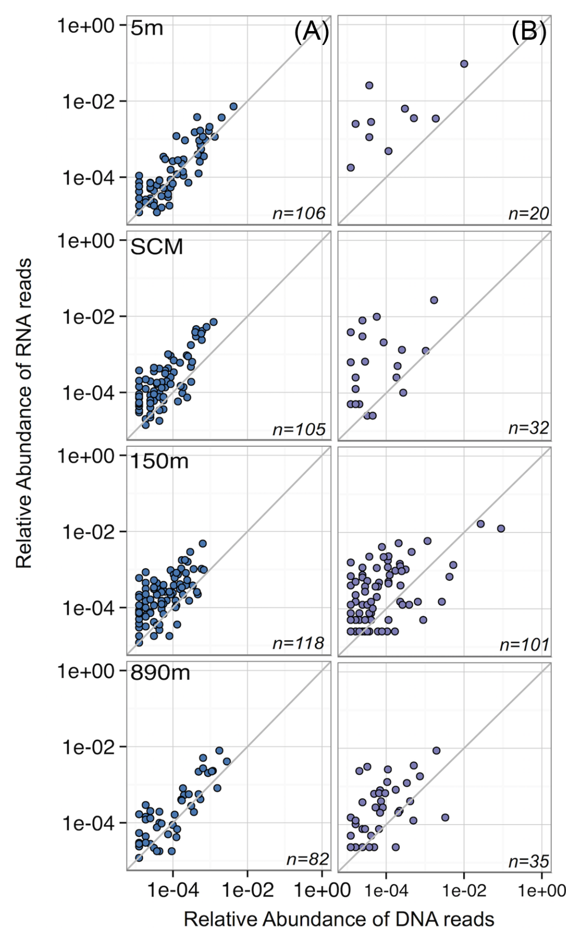

Figure 7. Relative abundance of RNA (*y*-axis) and DNA (*x*-axis) sequences for (A) ciliate and (B) radiolarian OTUs. From top to bottom, 5 m, subsurface chlorophyll maximum (SCM), 150 and 890 m at the SPOT station in April. Each data point depicts **(A)** ciliate or **(B)** radiolarian OTUs (n = total number of OTUs). (A) Ciliate RNA:DNA ratios were significantly higher at the SCM and 150 m at the SPOT station in April ($P < 0.05$; Tables S5 and S6, Supporting Information). Species richness of ciliates was also high at 150 m ($n = 118$). (B) RNA:DNA ratios for radiolaria were significantly higher at the SCM compared to 150 and 890 m at SPOT ($P < 0.05$; Tables S5 and S6, Supporting Information). Radiolarian species richness was highest at 150 m ($n = 101$), which was attributed to a higher number of polycystine OTUs (Table S7, Supporting Information). Generally, radiolarian RNA:DNA ratios were not significantly different among depths (Tables S5 and S6, Supporting Information).

in deep water (e.g. Edgcomb, Teske and Sogin 2002; Countway *et al.* 2007; Not *et al.* 2007). Radiolarian sequences have also been consistently found at the SPOT station below the euphotic zone (Countway *et al.* 2010; Schnetzer *et al.* 2011; Kim *et al.* 2014). However, since the ecology of deep-dwelling radiolaria remains largely undocumented and these studies were DNA-based, it has remained unresolved if radiolarian sequences detected in deep water originated from metabolically active radiolaria or non-viable cellular material sinking from the euphotic zone (Not *et al.* 2007; Gilg *et al.* 2010). Our results support the former, suggesting the presence of a significant assemblage of active radiolaria at 150 and 890 m, indicated by high RNA:DNA ratios (Fig. 7B; Table S5, Supporting Information). Radiolaria species richness was also highest at 150 m (Fig. 7B), which was mostly attributed to a larger abundance of polycystine OTUs compared to other depths (Table S7, Supporting Information).

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

CONCLUSIONS

This study is the first to use differences in RNA:DNA ratios to infer relative activity in individual protistan taxonomic groups on both temporal and spatial scales. We characterized how the protistan community responded to environmental change, with respect to season and spatially distinct locations (i.e. Port of LA versus SPOT and Catalina) by documenting how community composition and potential activity of dominant taxa fluctuated. Frequent sampling for community RNA:DNA ratios is a method to enhance DNA-based approaches for monitoring phytoplankton bloom initiation and demise (e.g. Countway and Caron 2006). The combined RNA and DNA approach provided evidence that ciliates play an important role along the oxycline and demonstrated that RNA:DNA ratios were useful for gathering ecological information from species that are difficult to characterize, such as radiolaria. DNA-based molecular methods characterize the diversity and biogeography of protistan communities, incorporating RNA sequencing enabled us to connect protistan diversity to ecosystem function. Future studies using combined RNA and DNA sequencing to evaluate metabolically active protists must address the limitations of the approach and link relative changes in RNA:DNA ratios to measurements of activities. Characterizing protistan community activity using RNA is one step towards fully understanding ecosystem function and how microbes respond to environmental change.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS

Authors would like to thank the Captain and Crew of the R/V Yellow Fin and the Sundiver Express, Troy Gunderson, Cathy Roney-Garcia, Erin Fichot, Diane Kim and the Wrigley Institute for Environmental Studies for fieldwork support. We would also like to acknowledge Rohan Sachdeva, Benjamin Tully and Jacob Cram for bioinformatics support. The submitted manuscript is original research and has not been previously published or submitted for publication elsewhere.

FUNDING

This work was supported by the National Science Foundation [grant number OCE-1136818].

Conflict of interest. None declared.

REFERENCES

Adl SM, Simpson AGB, Lane CE *et al*. The revised classification of eukaryotes. *J Eukaryot Microbiol* 2012;**59**:429–93.

Anderson R, Winter C, Jürgens K. Protist grazing and viral lysis as prokaryotic mortality factors at Baltic Sea oxic-anoxic interfaces. *Mar Ecol-Prog Ser* 2012;**467**:1–14.

Armbrust EV, Palumbi SR. Uncovering hidden worlds of ocean biodiversity. *Science* 2015;**348**:865–7.

Behnke A, Engel M, Christen R *et al*. Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ Microbiol* 2011;**13**:340–9.

Blazewicz SJ, Barnard RL, Daly RA *et al*. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J* 2013;**7**:2061–8.

Burki F, Keeling PJ. Rhizaria. *Curr Biol* 2014;**24**:R103–7.

Caporaso JG, Kuczynski J, Stombaugh J *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

Caron DA. Towards a molecular taxonomy for protists: benefits, risks, and applications in plankton ecology. *J Eukaryot Microbiol* 2013;**60**:407–13.

Caron DA, Countway PD. Hypotheses on the role of the protistan rare biosphere in a changing world. *Aquat Microb Ecol* 2009;**57**:227–38.

Caron DA, Countway PD, Jones AC *et al*. Marine protistan diversity. *Annu Rev Mater Sci* 2012;**4**:467–93.

Caron DA, Swanberg NR. The ecology of planktonic sarcodines. *Rev Aquat Sci* 1990;**3**:147–80.

Charvet S, Vincent WF, Lovejoy C. Effects of light and prey availability on Arctic freshwater protist communities examined by high-throughput DNA and RNA sequencing. *FEMS Microbiol Ecol* 2014;**88**:550–64.

Clarke KR, Somerfield PJ, Gorley RN. Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. *J Exp Mar Biol Ecol* 2008;**366**:56–69.

Corinaldesi C, Barucca M, Luna GM *et al*. Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments. *Mol Ecol* 2010;**20**:642–54.

Countway PD, Caron DA. Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Appl Environ Microb* 2006;**72**:2496–506.

Countway PD, Gast RJ, Dennett MR *et al*. Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environ Microbiol* 2007;**9**:1219–32.

Countway PD, Vigil PD, Schnetzer A *et al*. Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean). *Limnol Oceanogr* 2010;**55**:2381–96.

de Vargas C, Audic S, Henry N *et al*. Eukaryotic plankton diversity in the sunlit ocean. *Science* 2015;**348**:1261605-1–11.

Debroas D, Hugoni M, Domaizon I. Evidence for an active rare biosphere within freshwater protists community. *Mol Ecol* 2015;**24**:1236–47.

Decelle J, Bittner L, de Vargas C *et al*. An original mode of symbiosis in open ocean plankton. *P Natl Acad Sci USA* 2012;**109**:18000–5.

Decelle J, Martin P, Paborstava K *et al*. Diversity, ecology and biogeochemistry of cyst-forming acantharia (radiolaria) in the oceans. *PloS One* 2013;**8**:e53598.

Edgcomb V, Orsi W, Bunge J *et al*. Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *ISME J* 2011;**5**:1344–56.

Edgcomb VP, Pachiadaki M. Ciliates along oxyclines of permanently stratified marine water columns. *J Eukaryot Microbiol* 2014;**61**:434–45.

Edgcomb VP, Teske A, Sogin ML. Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *P Natl Acad Sci USA* 2002;**99**:7658–62.

Egge ES, Bittner L, Anderson T *et al*. 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. *PloS One* 2013;**8**:e74371.

Egge ES, Johannessen TV, Andersen T *et al*. Seasonal diversity and dynamics of haptophytes in the Skagerrak, Norway, explored by high-throughput sequencing. *Mol Ecol* 2015;**24**:3026–42.

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

Gilg IC, Amaral-Zettler L, Countway PD *et al*. Phylogenetic affiliations of mesopelagic acantharia and acantharian-like environmental 18S rRNA genes off the southern California coast. *Protist* 2010;**161**:197–211.

Godhe A, Asplund ME, Harnstrom K *et al*. Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Appl Environ Microb* 2008;**74**:7174–82.

Gong J, Dong J, Liu X *et al*. Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist* 2013;**164**:369–79.

Jing H, Rocke E, Kong L *et al*. Protist communities in a marine oxygen minimum zone off Costa Rica by 454 pyrosequencing. *Biogeosci Discuss* 2015;**12**:13483–509.

Karl DM, Bailiff MD. The measurement and distribution of dissolved nucleic acids in aquatic environments. *Limnol Oceanogr* 1989;**34**:543–58.

Kim DY, Countway PD, Jones AC *et al*. Monthly to interannual variability of microbial eukaryote assemblages at four depths in the eastern North Pacific. *ISME J* 2014;**8**:515–30.

Kim DY, Countway PD, Yamashita W *et al*. A combined sequence-based and fragment-based characterization of microbial eukaryote assemblages provides taxonomic context for the Terminal Restriction Fragment Length Polymorphism (T-RFLP) method. *J Microbiol Meth* 2012;**91**:527–36.

Kim H-J, Miller AJ, McGowan J *et al*. Coastal phytoplankton blooms in the Southern California Bight. *Prog Oceanogr* 2009;**82**:137–47.

Lejzerowicz F, Voltsky I, Pawlowski J. Identifying active foraminifera in the Sea of Japan using metatranscriptomic approach. *Deep-Sea Res PT II* 2013;**86-87**:214–20.

Lie AAY, Kim DY, Schnetzer A *et al*. Small-scale temporal and spatial variations in protistan community composition at the San Pedro Ocean Time-series station off the coast of southern California. *Aquat Microb Ecol* 2013;**70**:93–110.

Lin X, Scranton MI, Chistoserdov AY *et al*. Spatiotemporal dynamics of bacterial populations in the anoxic Cariaco Basin. *Limnol Oceanogr* 2008;**53**:37–51.

Logares R, Audic S, Bass D *et al*. Patterns of rare and abundant marine microbial eukaryotes. *Curr Biol* 2014;**24**:813–21.

López-García P, Rodriguez-Valera F, Pedros-Alio C. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 2001;**409**:603–7.

Lorenz MG, Wackernagel W. Adsorption of DNA to sand and variable degradation rates of adsorbed DNA. *Appl Environ Microb* 1987;**53**:2948–52.

Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;**27**:2957–63.

Massana R. Protistan diversity in environmental molecular surveys. In:Ohtsuka S, Suzaki T, Horiguchi T *et al*. (eds). *Marine Protists: Diversity and Dynamics*. Tokyo, Japan: Springer, 2015, 3–21.

Massana R, Castresana J, Balague V *et al*. Phylogenetic and ecological analysis of novel marine stramenopiles. *Appl Environ Microb* 2004;**70**:3528–34.

Massana R, Gobet A, Audic S *et al*. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ Microbiol* 2015;**17**:4035–49.

Massana R, Pedros-Alio C. Unveiling new microbial eukaryotes in the surface ocean. *Curr Opin Microbiol* 2008;**11**:213–8.

Moon-van der Staay SY, De Wachter R, Vaulot D. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Lett Nat* 2001;**409**: 607–10.

Moreira D, López-García P. The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends Microbiol* 2001;**10**:31–8.

Not F, del Campo J, Balagué V *et al*. New insights into the diversity of marine picoeukaryotes. *PloS One* 2009;**4**:e7143.

Not F, Gausling R, Azam F *et al*. Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol* 2007;**9**:1233–52.

Ohtsuka S, Suzaki T, Horiguchi T *et al*. *Marine Protists: Diversity and Dynamics*. Japan: Springer, 2015.

Parris DJ, Ganesh S, Edgcomb VP *et al*. Microbial eukaryote diversity in the marine oxygen minimum zone off northern Chile. *Front Microbiol* 2014;**5**:543.

Poulsen LK, Ballard G, Stahl DA. Use of rRNA fluorescence *in situ* hybridization for measuring the activity of single cells in young and established biofilms. *Appl Environ Microb* 1993;**59**:1354–60.

Prokopowich CD, Gregory TR, Crease TJ. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* 2003;**46**:48–50.

Quast C, Pruesse E, Yilmaz P *et al*. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;**41**:D590–6.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2014.

Rideout JR, He Y, Navas-Molina JA *et al*. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* 2014;**2**:e545.

Rodríguez-Martínez R, Rocap G, Logares R *et al*. Low evolutionary diversification in a widespread and abundant uncultured protist (MAST-4). *Mol Biol Evol* 2012;**29**:1393–406.

Schnetzer A, Jones BH, Schaffner RA *et al*. Coastal upwelling linked to toxic *Pseudo-nitzschia australis* blooms in Los Angeles coastal waters, 2005–2007. *J Plankton Res* 2013;**35**:1080–92.

Schnetzer A, Moorthi SD, Countway PD *et al*. Depth matters: microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries. *Deep-Sea Res PT I* 2011;**58**:16–26.

Seubert EL, Gellene AG, Howard MDA *et al*. Seasonal and annual dynamics of harmful algae and algal toxins revealed through weekly monitoring at two coastal ocean sites off southern California, USA. *Environ Sci Pollut R* 2013;**20**:6878–95.

Sherr BF, Sherr EB, Caron DA *et al*. A sea of microbes: oceanic protists. *Oceanography* 2010;**20**:130–4.

Sherr EB, Sherr BF. Bacterivory and herbivory: key roles of phagotrophic protists in pelagic food webs. *Microb Ecol* 1994;**28**:223–35.

Sherr EB, Sherr BF. Significance of predation by protists in aquatic microbial food webs. *A Van Leeuw* 2002;**81**:293–308.

Stecher A, Neuhaus S, Lange B *et al*. rRNA and rDNA based assessment of sea ice protist biodiversity from the central Arctic Ocean. *Eur J Phycol* 2015;**51**:31–46.

Stock A, Jürgens K, Bunge J *et al*. Protistan diversity in suboxic and anoxic waters of the Gotland Deep (Baltic Sea) as revealed by 18S rRNA clone libraries. *Aquat Microb Ecol* 2009;**55**:267–84.

Stoeck T, Bass D, Nebel M *et al*. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* 2010;**19**:21–31.

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016

Stoeck T, Zuendorf A, Breiner H-W *et al*. A molecular approach to identify active microbes in environmental eukaryote clone libraries. *Microb Ecol* 2007;**53**: 328–39.

Terrado R, Medrinal E, Dasilva C *et al*. Protist community composition during spring in an Arctic flaw lead polynya. *Polar Biol* 2011;**34**:1901–14.

Vaulot D, Romari K, Not F. Are autotrophs less diverse than heterotrophs in marine picoplankton? *Trends Microbiol* 2002;**10**:266–7.

Venrick EL. Floral patterns in the California current system off southern California:1990–1996. *J Mar Res* 2002;**60**: 171–89.

Whitaker D, Christman M. Package 'clustsig'. R package version 1.1. 2014. https://cran.hafro.is/web/packages/clustsig/clustsig.pdf.

Worden AZ. Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat Microb Ecol* 2006;**43**:165–75.

Worden AZ, Follows MJ, Giovannoni SJ *et al*. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. *Science* 2015;**347**:735.

Wylezich C, Jürgens K. Protist diversity in suboxic and sulfidic waters of the Black Sea. *Environ Microbiol* 2011;**13**:2939–56.

Zhu F, Massana R, Not F *et al*. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* 2005;**52**:79–92.

Downloaded from http://femsec.oxfordjournals.org/ by guest on April 6, 2016