Limnol. Oceanogr., 55(6), 2010, 2381–2396

doi:10.4319/lo.2010.55.6.2381

Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean)

Peter D. Countway,[a,*,1] Patrick D. Vigil,[a,b] Astrid Schnetzer,[a] Stefanie D. Moorthi,[a,c] and David A. Caron[a]

[a] Department of Biological Sciences, University of Southern California, Los Angeles, California
[b] Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan
[c] Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Wilhelmshaven, Germany

Abstract

The structure and genetic diversity of marine protistan assemblages were investigated in the upper 500 m of the water column at a Pacific Ocean time-series station off the coast of Southern California. Deoxyribonucleic acid sequence-based microbial eukaryote diversity was examined in January, April, July, and October of 2001 at four depths (5 m, chlorophyll maximum [CM], 150 m, and 500 m). A total of 2956 partial 18S ribosomal ribonucleic acid gene sequences yielded representatives from most of the major eukaryotic lineages. Notable among the taxonomic groups were recently described lineages of stramenopiles, alveolates, and euglenozoa. A large number of polycystine and acantharean sequences were observed at depth. Pairwise sequence analysis was performed to establish operational taxonomic units (OTUs) that were then used to estimate the unsampled protistan diversity by parametric and nonparametric techniques. A total of 2246 protistan sequences grouped into 377 distinct OTUs, with remaining sequences attributed to metazoa. Protistan richness estimates ranged from ~ 600 to 1500 OTUs when all depths and seasons were combined into a single data set. Seasonal and depth-related trends in the observed protistan diversity were apparent from comparisons of univariate and multivariate analyses. Cluster analysis combined with nonmetric multidimensional scaling and analysis of similarity testing identified distinct protistan assemblages at the shallowest depths (5 m and CM) for each season, which were significantly different ($p < 0.03$) from assemblages at the two deepest depths (150 and 500 m) where seasonal changes in the protistan assemblage were not apparent.

Protistan assemblages are composed of morphologically, genetically, and functionally diverse groups of single-celled eukaryotes that collectively represent one of the greatest pools of biological diversity on Earth. Protists are crucial to the structure and function of marine ecosystems where these taxa fulfill a wide variety of ecological roles including phototrophy, phagotrophy, and diverse symbiotic relationships (Sherr et al. 2007). Marine protists conduct the majority of the ocean's inorganic carbon fixation (photosynthesis) and contribute significantly to biogeochemical nutrient cycling and trophic energy transfer (Liu et al. 2009). Despite their importance, the full extent of protistan diversity is still poorly characterized. This biodiversity has become a focal point in recent years as it has become evident that microbial diversity plays fundamental roles in maintaining the functional stability and resilience of ecosystems (Caron and Countway 2009).

Increasing numbers of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) surveys from many ecosystems have begun to improve our understanding of protistan diversity, and indicate that characterizations of protistan assemblages by microscopy have failed to identify large fractions of the biodiversity (Massana et al. 2004*b*; Lovejoy et al. 2006; Stoeck et al. 2006). Sequence databases in the public domain are expanding rapidly as a consequence of the dramatic increase in genetic information obtained from environmental ribosomal (r)RNA gene sequencing (Pruesse et al. 2007). These databases have been extensively used to test hypotheses regarding the phylogenetic relationships among protists, and these relationships are constantly being redefined as information provided by morphological and genetic data sets merge (Burki et al. 2007; Tekle et al. 2009). In addition, environmental cloning, sequencing, and microbial fingerprinting have provided estimates of the numbers of different microbial phylotypes or operational taxonomic units (OTUs) present in particular environments. These approaches can provide estimates of the relative abundance of the OTUs and have been used to examine both species richness and evenness. Molecular methods for examining diversity are not without their potential biases and artifacts (von Wintzingerode et al. 1997; Potvin and Lovejoy 2009) but they provide potentially powerful tools for characterizing microbial diversity and comparing microbial community structure (Curtis and Sloan 2004). The combination of molecular- and morphology-based approaches has even begun to offer images of previously uncharacterized protistan taxa (Massana et al. 2006; Gilg et al. 2010).

Most studies of protistan diversity using genetic methods have thus far reported ''snapshots'' of diversity at particular locations, depths, or times. Few studies have examined both temporal changes (Romari and Vaulot 2004; Countway et al. 2005) and spatial patterns (Gast et al. 2004; Countway et al. 2007) of protistan assemblages, with one recent exception (Stoeck et al. 2009). The goal of

* Corresponding author: pcountway@bigelow.org

[1] Present address: Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine

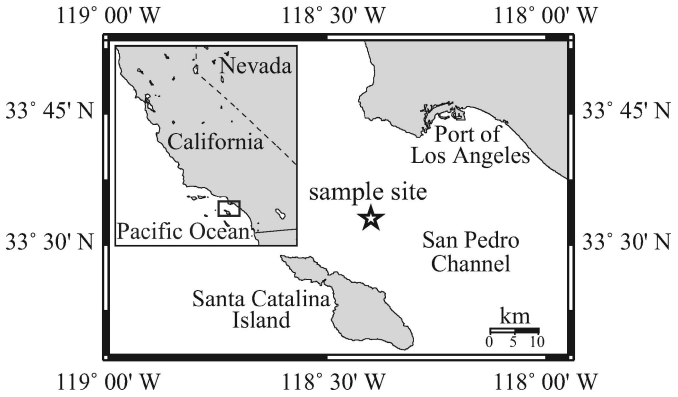

Fig. 1. Station location (33°33′N, 118°24′W), site of sample collection (indicated by the star) for the USC Microbial Observatory and San Pedro Ocean Time-series (SPOT) projects, San Pedro Channel, Pacific Ocean. Map created with online map creation application: http://www.aquarius.ifm-geomar.de/.

our study was to investigate seasonal patterns of marine protistan diversity at several depths through the water column at a coastal time-series station. This was accomplished by DNA-based molecular analyses of environmental 18S rRNA genes using a subset of monthly samples collected throughout a single year at the University of Southern California (USC) Microbial Observatory. A total of 2956 18S rRNA gene clone sequences distributed in 399 OTUs were analyzed, of which 2246 sequences were protistan and comprised 95% of the total OTU count (377), with the remaining 5% of OTUs (21) identified as various metazoa. Overall, protistan diversity estimates were high, with a great breadth of taxonomic representation during all seasons. Seasonal and depth-related patterns of diversity were evident from both cluster analysis and nonmetric multidimensional scaling (MDS). These findings imply that local differences in environmental parameters provided strong local selection of dominant protistan taxa from a large pool of extant but potentially rare taxa.

Methods

Environmental sample collection—Assemblages of planktonic marine protists were collected at the USC Microbial Observatory off the coast of Southern California (33°33′N, 118°24′W) on 16 January, 02 April, 27 July, and 29 October 2001 (Fig. 1). Seawater samples were collected from four depths (5 m, chlorophyll maximum [CM], 150 m, and 500 m) using a conductivity temperature depth (CTD) profiler equipped with 10-liter Niskin bottles (General Oceanics). Samples were prescreened with 200-μm Nitex mesh (Sefar) to reduce the contribution of metazoan plankton to subsequent DNA extracts.

Sample collection and DNA extraction—Triplicate 2-liter samples of the 200-μm prefiltered seawater from each depth were vacuum filtered ($<$ 10 mm Hg) onto 47-mm GF/F filters (Whatman) to collect microbial biomass. Lysis buffer was added to biomass samples before freezing the filters in liquid nitrogen. Extraction of nucleic acids from microbial biomass followed standard cell-disruption techniques as detailed in Countway et al. (2005, 2007).

Environmental polymerase chain reaction (PCR)—18S ribosomal RNA genes were amplified by PCR from DNA extracts with eukaryote primers; Euk-A (5′-AACCTGGTTGATCCTGCCAGT-3′) and Euk-B (5′-GATCCTTCTGCAGGTTCACCTAC-3′). PCR conditions were described previously (Countway et al. 2007). Products from three to four replicate PCRs were pooled for each sample to reduce the effect of potential amplification biases that might be present in individual reactions.

Cloning and DNA sequencing—PCR amplicons were electrophoresed on 1.2% agarose gels (Cambrex) and stained with SYBR Gold (Molecular Probes). PCR amplicons were excised from agarose gels and purified using the UltraClean™ 15 kit (Mo Bio). Cloning of PCR products was accomplished with the TOPO-TA Cloning® kit and TOP10 electrocompetent *Escherichia coli* (Invitrogen) following manufacturer's protocols. Cells and ligation products were mixed and electroporated in 0.10-cm cuvettes using a Gene Pulser Xcell electroporator (BioRad) set to 2500 V, 200 Ω, and 25 μF. Standard protocols were followed for outgrowth, plating, and picking of bacterial clones. Sequencing reactions were performed with purified plasmid DNA, Euk-570F (5′-GTAATTCCAGCTCCAATAGC-3′), and 2 μL of dye terminator cycle sequencing reagent (Beckman). Sequencing was performed on a Beckman CEQ8000, resulting in partial-length reads of 400–700 nucleotides after automated quality-based trimming using a "medium" stringency setting in the Beckman base-calling software. Subsequently, all sequences were manually inspected to check for accuracy of automated base-calling and corrected where possible.

Phylotype assignment and diversity estimation—Putative sequence identities were obtained by basic local alignment search tool (BLAST) analysis (Altschul et al. 1997) of all sequences against local installations of the National Center for Biotechnology Information and Arb-Silva sequence databases using the BLAST-N search algorithm (Benson et al. 2004; Pruesse et al. 2007). Potential sequence chimeras were flagged using results from the ribosomal database project check chimera tool (Cole et al. 2003). It is possible that a small fraction of chimeras were missed because of inherent limitations of chimera detection and our reliance on a single chimera-checking method. Potential chimeras and poor-quality sequences (those with BLAST bit-scores $<$ 200) were automatically removed during upload to a searchable MySQL database. This procedure yielded a total of 2956 sequences that were deposited in GenBank (accession numbers HM856920–HM859875).

Pairwise sequence comparisons were conducted with ClustalW (Thompson et al. 1994) to establish a "global" sequence similarity matrix. Sequences were grouped into OTUs using the microbial eukaryotic species assignment (MESA) program (Caron et al. 2009). MESA assigns OTUs on the basis of our recommended threshold similarity setting of 95% to obtain approximate species-

level OTUs for most taxa. This threshold was determined by our previous analysis of the variability among 18S rRNA gene sequences from morphologically well-defined protists found in GenBank (Caron et al. 2009). Taxonomic affinities for each OTU were assigned by manual inspection of BLAST results for all sequences within an OTU and determined by majority consensus. In most cases taxonomic assignment was restricted to higher taxonomic levels. Finer taxonomic resolution was not assigned unless a majority of sequences in an OTU were $> 97\%$ similar to taxonomically identified reference sequences and had bit-scores from BLAST analysis > 500. OTUs identified as metazoa were removed from further analysis.

Nonparametric estimates of the "unseen" or "unsampled" protistan diversity were calculated from OTU data using the species prediction and diversity estimation program (SPADE; Chao and Shen 2005) and EstimateS (Colwell 2004). Nonparametric estimators included Chao-1 and the abundance-based coverage estimator (ACE-1). In general, nonparametric approaches are considered to be minimal estimates of the total diversity (Hong et al. 2006). Parametric modeling of OTU frequency classes (e.g., numbers of singletons, doubletons, tripletons, etc.) was conducted to provide alternative estimates of the unsampled diversity using a best-fitting model approach that is gaining popularity (Zuendorf et al. 2006; Bunge and Barger 2008). The inverse Simpson's index (D_S^{-1}) was calculated to provide a relatively simple and common diversity statistic that accounts for taxonomic richness (number of OTUs), evenness (relative abundance of sequences within each OTU), and the total number of sequences comprising each sample (clone library). This univariate statistic can range from 1 to the maximum number of OTUs in each sample.

Comparison of protistan assemblages across depths and seasons—Protistan assemblages from different seasons and depths were compared using the Bray–Curtis coefficient of community similarity calculated from square-root-transformed relative OTU abundances. Similarity data were analyzed by cluster analysis and nonmetric MDS using the Plymouth routines in multivariate ecological research (PRIMER v.6) software package (Clarke and Gorley 2006). The similarity profile (SIMPROF) permutation test was conducted in PRIMER v.6 to establish the significance of dendrogram branches resulting from cluster analysis, whereas analysis of similarity (ANOSIM) tests were performed to test for significance among factors such as "depth" or "month" and to aid in the evaluation of relationships among data points in MDS plots.

Results

Physical and chemical setting—The San Pedro Basin is a silled coastal basin (maximum depth $\approx$ 890 m), with restricted circulation in the deep waters. The seasonal temperature range of surface waters during this study was relatively narrow, approximately 14–20°C (Fig. 2; Table 1). Salinity was also stable seasonally (Fig. 2; Table 1). Restricted deep-water movement and terrigenous inputs into the basin result in persistent hypoxia below the surface mixed layer. Dissolved oxygen concentrations below $\sim$ 350 m were less than 1 mL L^{-1} (Fig. 2; Table 1). A distinct subsurface CM was present throughout the year, but the magnitude and depth of this feature varied seasonally. Highest concentrations of chlorophyll (> 6 μg L^{-1}) were observed during April and constituted a broad feature ranging in depth from 5 to 25 m. The least amount of chlorophyll defining a subsurface CM was observed during January (< 2 μg L^{-1}). The depth of the CM was shallowest in April compared with all other sampling months (Fig. 2; Table 1).

Taxonomic distribution of clones—Taxonomic composition of the clone libraries revealed substantial differences in the structure of the protistan assemblages across depths and between seasons (Table 2). A high-resolution taxonomic investigation was not intended in the present study, but information on the taxonomic composition of the assemblages was revealed by BLAST searches against public databases. Cloning and sequencing yielded a total of 2956 high-quality sequences from 16 clone libraries (four seasons $\times$ four depths), which grouped into 399 OTUs on the basis of a 95% sequence similarity threshold. Caron et al. (2009) provides a detailed explanation of the OTU-calling algorithm and a rationale for selecting this threshold value.

Protistan OTUs predominated in our libraries, but metazoan sequences were observed in all samples and were generally a greater proportion at shallower depths (data not shown). Most metazoan sequences probably resulted from small life stages or body fragments passing through the prefiltration step. Metazoan sequences comprised 5% of the OTUs detected (22 OTUs), including several well-populated cnidarian and arthropod OTUs. Removal of metazoan OTUs resulted in 2246 protistan sequences distributed among 377 protistan OTUs (Tables 2, 3). The number of protistan sequences remained relatively high (100–200 per library) after the removal of metazoan sequences, with the exception of 5 m and CM libraries from April, which resulted in just 23 and 37 protistan sequences, respectively (Table 4).

Protistan OTUs were assigned to 21 higher-level phylogenetic groups on the basis of BLAST identifications (Tables 2, 3). Protistan groups detected included Acantharea, Centroheliozoa, Cercozoa, Polycystinea, Ciliophora, Dinophyceae, Ellobiopsidae, group I Alveolata, group II Alveolata, unclassified Alveolata, Cryptophyta, Haptophyceae, Stramenopiles, Chlorophyta, Rhodophyta, Euglenozoa, Choanoflagellida, Fungi, Cryothecomonas, Ichthyosporea, and Telonema. Most of these groups comprised no more than 25% of the sequences in a particular library with the exception of the polycystines in some 150- and 500-m libraries.

The 377 protistan OTUs constituted 1–63 OTUs per higher-level taxonomic group (Table 3). The rhizarian groups Acantharea and Polycystinea were particularly well-represented in deep-water libraries. The most abundant OTU in the study was a spumellarian radiolarian detected exclusively in 150- and 500-m libraries during all

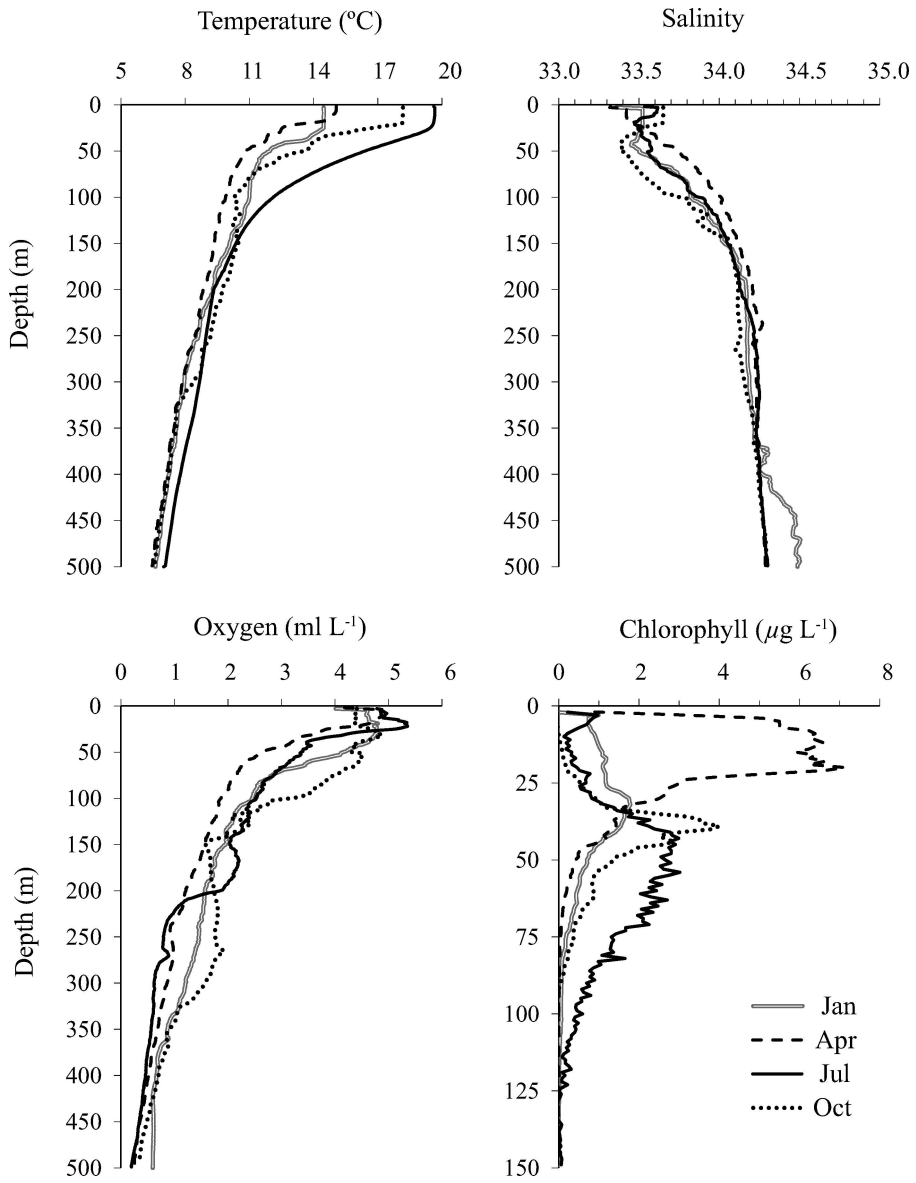

Fig. 2. Continuous CTD profiles of temperature, salinity, dissolved oxygen, and chlorophyll concentration on each of the four sampling dates at the USC Microbial Observatory study site.

seasons. This single OTU was composed of 381 individual sequences or 17% of the total protistan sequence count (Fig. 3). Members of the excavate group Euglenozoa were also detected almost exclusively below the euphotic zone. Most euglenozoan OTUs were most closely related to the diplonemids.

Alveolate lineages represented 57% of the total OTUs. The second most abundant OTU (228 sequences, Fig. 3), detected at all depths and during all seasons, was highly similar ($\geq$ 99%) to sequences from the dinoflagellate *Gyrodinium*. Other dinoflagellate OTUs having close matches to identified taxa were largely restricted to the upper water column (e.g., the ninth most abundant OTU was $\geq$ 99% similar to *Ceratium*) and often showed strong seasonal patterns across months with respect to their relative abundance (Fig. 3). The bloom-forming red tide dinoflagellate *Lingulodinium* was found in July samples at 5 m and the CM. Sequences most similar to several oligotrichous ciliates (*Laboea*, *Parallelostrombidium*, and *Strombidium*) were dominant ciliate OTUs in the clone libraries, with choreotrichs generally displaying lower relative abundances. A large proportion of the unclassified alveolate sequences had best BLAST matches to the parasitic *Amoebophrya*, although similarity values were low (92–95%).

Haptophytes were detected at relatively low abundances during most months and were largely restricted to 5 m and the CM. The picoeukaryote haptophyte *Phaeocystis* frequently grows in seawater enrichment cultures from our study site but it was only represented by a single sequence in clone libraries. *Chrysochromulina* was the most abundant haptophyte, with peak abundance at 5 m in July.

Table 1. Sensor data from continuous CTD profiles representing average values within ± 1 m of the targeted depths (5 m, CM, 150 m, and 500 m). Depths were calculated from CTD pressure measurements, and rounded to the nearest whole meter. nd, chlorophyll values less than the limit of sensor detection, 0.03 μg L^{-1}. nd, not detected.

Variable	5 m	CM	150 m	500 m
January				
Depth (m)	5.0	27.0	150.0	500.0
Temp. (°C)	14.5	14.1	10.1	6.6
Salinity	33.5	33.5	34.0	34.4
Oxygen (mL L^{-1})	4.8	4.8	1.9	0.4
Chlorophyll (μg L^{-1})	0.9	1.8	0.1	0.1
April				
Depth (m)	5.0	14.0	150.0	500.0
Temp. (°C)	15.0	14.5	9.4	6.4
Salinity	33.4	33.4	34.1	34.3
Oxygen (mL L^{-1})	4.9	4.8	1.6	0.2
Chlorophyll (μg L^{-1})	4.5	7.7	0.5	0.5
July				
Depth (m)	5.0	29.0	150.0	500.0
Temp. (°C)	19.5	12.6	9.3	6.7
Salinity	33.6	33.5	34.0	34.3
Oxygen (mL L^{-1})	4.0	4.5	1.9	0.3
Chlorophyll (μg L^{-1})	0.3	2.9	0.1	0.2
October				
Depth (m)	5.0	40.0	150.0	500.0
Temp. (°C)	18.3	14.3	10.5	6.5
Salinity	33.7	33.4	34.1	34.3
Oxygen (mL L^{-1})	4.4	4.5	1.7	0.3
Chlorophyll (μg L^{-1})	nd	2.7	nd	nd

Stramenopiles were not particularly abundant in clone libraries; however, the diversity of stramenopiles was very high, with a total of 47 OTUs detected among 73 sequences. The highest proportions of stramenopiles occurred in samples from 5 m collected during October and January. Diatoms formed the largest group of stramenopiles, with bolidophytes, chrysophytes, pelagophytes, raphidophytes, and silicoflagellates present at much lower abundance. A large proportion of stramenopile clones belonged to marine stramenopile lineages (Massana et al. 2004*b*), most of which are thought to be small heterotrophic flagellates. Clones with $> 99\%$ similarity to the pelagophyte *Aureococcus anophagefferens* appeared in two instances and represented the first documentation of this "brown tide" organism off of Southern California.

Picoeukaryotic chlorophytes comprised several relatively abundant OTUs that were identified with high BLAST similarities ($\geq$ 99%) to the prasinophytes *Bathycoccus*, *Micromonas*, and *Ostreococcus*. These taxa were abundant during January, July, and October in the "shallow" libraries (5 m and CM), and were the numerical dominants among 11 chlorophyte OTUs.

The remaining groups included diverse low-abundance taxa. A small number of fungi OTUs were present, and nearly every fungal sequence defined its own OTU. The greatest proportion of fungal sequences occurred in October at a depth of 5 m. Choanoflagellates were a minor component of clone libraries and were only detected in January and October. The six total choanoflagellate sequences formed five distinct OTUs. Cercozoa were present at a depth of 5 m during July and October; whereas Ichthyosporea (crustacean parasites) were one of the dominant protistan groups at 5 m and the CM during April. Telonemids, a phylogenetically enigmatic group of bacterivorous flagellates (Shalchian-Tabrizi et al. 2007), were most abundant at 5 m during October and January and were represented by 2 OTUs.

Rank abundance of OTUs—A rank abundance curve revealed a few highly populated OTUs and numerous rare taxa (Fig. 3). Seventy percent of the total sequences (1602) belonged to 36 common OTUs (11 or more members per OTU), whereas singletons represented just 9% (208) of the total sequences but 55% of the total OTUs. Relative OTU abundances displayed the proportional contribution from each clone library to a particular OTU as a percentage of the total sequence count (Fig. 3). Notably, most of the abundant OTUs were detected in numerous clone libraries spanning all seasons and depths.

Diversity indices and richness estimators—The inverse of Simpson's index (D_S^{-1}) was calculated for various combinations of the data sets and for each sample individually to evaluate protistan diversity (Table 4). The overall diversity index for all 2246 sequences and 377 OTUs was relatively high, with a D_S^{-1} value of 20.8 (Table 4). Diversity for the entire water column (four depths for each month) was moderate during January (12.4), relatively low during April (8.9), and high during July (23.6) and October (38.1). Values of this index for data sets pooled by depth across all four seasons indicated similar levels of relatively high diversity at 5 m and the CM (28.8 and 23.1, respectively) and substantially lower diversity at 150 and 500 m (6.4 and 10.2, respectively). High values of D_S^{-1} in surface waters (5 m and the CM) were observed during January, followed by very low values during April that gradually increased during July and October. Values of D_S^{-1} for deep assemblages (150 and 500 m) were lowest during January and generally increased in subsequent months.

Total protistan richness was estimated using both nonparametric estimators and parametric estimators. Nonparametric estimates and 95% confidence intervals (CIs) ranged from 682 (587–819) OTUs for Chao-1 to 1105 (859–1476) OTUs for ACE-1, whereas the best parametric estimate and 95% CI was 994 (844–1192) and was described by a three-mixed-exponential model (Table 4). Rarefaction plots of Chao-1 and the number of observed OTUs (S_{OBS}) deviated from a 1 : 1 relationship; however, both continued to increase over the entire range of sample sizes, indicating that additional protistan diversity remained undetected (Fig. 4).

Neither Chao-1 nor ACE-1 estimates of total species richness differed significantly (on the basis of 95% CIs) among libraries pooled by month or by depth (Table 4; values $\sim$ 300–900 OTUs by month; $\sim$ 300–700 OTUs by depth). Few significant differences in species richness were observed among libraries generated for individual depths

Table 2. Distribution of protistan sequences and OTUs (in brackets) across higher-level taxonomic structure for each of the sampled depths (5 m, CM, 150 m, and 500 m) and months (January, April, July, and October) displaying both seasonal and depth-related trends in the protistan assemblage. nd, not detected.

Month and taxon	5 m	CM	150 m	500 m	Total by taxa
January					
Rhizaria					
Acantharea	1 (1)	3 (2)	5 (3)	1 (1)	10 (5)
Centroheliozoa	1 (1)	1 (1)	1 (1)	nd	3 (1)
Cercozoa	3 (3)	1 (1)	nd	nd	4 (4)
Polycystinea	nd	3 (2)	104 (2)	81 (4)	188 (5)
Chromalveolata					
Alveolata; Ciliophora	45 (11)	30 (7)	nd	7 (4)	82 (16)
Alveolata; Dinophyceae	37 (4)	23 (9)	8 (4)	7 (3)	75 (15)
Alveolata; Ellobiopsidae	nd	nd	nd	nd	nd
Alveolata; group I	13 (8)	7 (4)	5 (3)	10 (6)	35 (14)
Alveolata; group II	7 (4)	11 (4)	19 (4)	34 (6)	71 (13)
Alveolata; unclassified	45 (21)	33 (18)	5 (5)	4 (4)	87 (35)
Cryptophyta	5 (2)	6 (4)	nd	nd	11 (5)
Haptophyceae	6 (5)	7 (6)	nd	nd	13 (9)
Stramenopiles	17 (12)	16 (11)	2 (2)	nd	35 (20)
Plantae					
Chlorophyta	35 (8)	32 (5)	nd	nd	67 (8)
Rhodophyta	1 (1)	nd	nd	nd	1 (1)
Excavata					
Euglenozoa	nd	nd	1 (1)	3 (3)	4 (3)
Opisthokonta					
Choanoflagellida	3 (3)	1 (1)	nd	nd	4 (3)
Fungi	nd	1 (1)	nd	1 (1)	2 (2)
Unresolved					
Cryothecomonas	1 (1)	1 (1)	nd	nd	2 (1)
Ichthyosporea	nd	nd	nd	nd	nd
Telonema	6 (2)	2 (2)	nd	nd	8 (2)
April					
Rhizaria					
Acantharea	nd	nd	9 (5)	4 (1)	13 (5)
Centroheliozoa	nd	nd	nd	nd	nd
Cercozoa	1 (1)	1 (1)	nd	nd	2 (1)
Polycystinea	nd	nd	81 (5)	46 (4)	127 (7)
Chromalveolata					
Alveolata; Ciliophora	2 (2)	1 (1)	7 (6)	1 (1)	11 (8)
Alveolata; Dinophyceae	9 (3)	19 (7)	15 (9)	17 (7)	60 (22)
Alveolata; Ellobiopsidae	nd	nd	nd	1 (1)	1 (1)
Alveolata; group I	1 (1)	2 (1)	13 (5)	10 (6)	26 (9)
lveolata; group II	nd	1 (1)	31 (4)	64 (6)	96 (8)
Alveolata; unclassified	1 (1)	2 (2)	11 (8)	4 (4)	18 (12)
Cryptophyta	nd	nd	2 (2)	nd	2 (2)
Haptophyceae	nd	nd	nd	nd	nd
Stramenopiles	1 (1)	2 (2)	2 (2)	nd	5 (5)
Plantae					
Chlorophyta	1 (1)	nd	nd	nd	1 (1)
Rhodophyta	nd	nd	nd	nd	nd
Excavata					
Euglenozoa	nd	nd	4 (4)	7 (4)	11 (8)
Opisthokonta					
Choanoflagellida	nd	nd	nd	nd	nd
Fungi	nd	nd	nd	nd	nd
Unresolved					
Cryothecomonas	nd	1 (1)	nd	nd	1 (1)
Ichthyosporea	7 (1)	8 (1)	nd	nd	15 (1)
Telonema	nd	nd	nd	nd	nd
July					
Rhizaria					
Acantharea	nd	1 (1)	7 (4)	23 (4)	31 (6)
Centroheliozoa	nd	nd	nd	nd	nd

Table 2. Continued.

Month and taxon	5 m	CM	150 m	500 m	Total by taxa
Cercozoa	11 (2)	nd	nd	nd	11 (2)
Polycystinea	1 (1)	nd	22 (4)	39 (5)	61 (8)
Chromalveolata					
Alveolata; Ciliophora	17 (5)	13 (4)	3 (3)	4 (4)	37 (11)
Alveolata; Dinophyceae	35 (10)	74 (9)	19 (6)	40 (4)	168 (22)
Alveolata; Ellobiopsidae	nd	nd	nd	nd	nd
Alveolata; group I	6 (3)	4 (2)	11 (5)	6 (2)	27 (8)
Alveolata; group II	1 (1)	nd	26 (10)	26 (6)	53 (11)
Alveolata; unclassified	28 (8)	12 (7)	9 (5)	4 (4)	53 (17)
Cryptophyta	nd	nd	nd	nd	nd
Haptophyceae	12 (6)	1 (1)	0 (0)	0 (0)	13 (7)
Stramenopiles	2 (2)	5 (5)	3 (3)	1 (1)	11 (10)
Plantae					
Chlorophyta	3 (3)	20 (4)	nd	nd	23 (5)
Rhodophyta	5 (3)	1 (1)	1 (1)	nd	7 (3)
Excavata					
Euglenozoa	nd	nd	5 (4)	9 (5)	14 (8)
Opisthokonta					
Choanoflagellida	nd	nd	nd	nd	nd
Fungi	nd	1 (1)	nd	nd	1 (1)
Unresolved					
Cryothecomonas	nd	nd	nd	nd	nd
Ichthyosporea	nd	nd	1 (1)	nd	1 (1)
Telonema	nd	1 (1)	nd	nd	1 (1)
October					
Rhizaria					
Acantharea	nd	3 (3)	1 (1)	11 (4)	15 (7)
Centroheliozoa	nd	nd	nd	nd	nd
Cercozoa	10 (4)	6 (5)	nd	1 (1)	17 (9)
Polycystinea	nd	11 (2)	47 (3)	25 (7)	83 (10)
Chromalveolata					
Alveolata; Ciliophora	19 (6)	16 (4)	6 (4)	14 (8)	55 (16)
Alveolata; Dinophyceae	28 (9)	47 (9)	15 (9)	17 (6)	107 (24)
Alveolata; Ellobiopsidae	nd	nd	nd	nd	nd
Alveolata; group I	2 (2)	8 (4)	16 (7)	20 (9)	46 (15)
Alveolata; group II	3 (2)	5 (5)	49 (9)	42 (10)	99 (17)
Alveolata; unclassified	37 (11)	29 (11)	18 (9)	10 (10)	94 (29)
Cryptophyta	nd	5 (4)	nd	nd	5 (4)
Haptophyceae	3 (3)	1 (1)	1 (1)	nd	5 (3)
Stramenopiles	10 (7)	6 (6)	2 (2)	4 (4)	22 (18)
Plantae					
Chlorophyta	22 (6)	14 (4)	nd	1 (1)	37 (7)
Rhodophyta	1 (1)	1 (1)	1 (1)	nd	3 (2)
Excavata					
Euglenozoa	1 (1)	1 (1)	16 (8)	23 (9)	41 (14)
Opisthokonta					
Choanoflagellida	nd	1 (1)	nd	1 (1)	2 (2)
Fungi	2 (2)	nd	3 (2)	nd	5 (4)
Unresolved					
Cryothecomonas	nd	nd	nd	nd	nd
Ichthyosporea	1 (1)	nd	nd	1 (1)	2 (2)
Telonema	4 (1)	nd	nd	nd	4 (1)

and seasons when different depths within a single month were compared or when the same depths in different months were compared. For seasonal comparisons at specific depths, the libraries from 5-m samples yielded significantly different Chao-1 estimates for January relative to April and July, and the Chao-1 and ACE-1 estimates for the 150-m libraries from January were significantly different from April, July, and October libraries for the same depth. For depth-to-depth comparisons within a season, Chao-1 and ACE-1 estimates for the 5-m and CM libraries during January were significantly different from the 150-m library.

Trends revealed by parametric richness estimates were in general agreement with those calculated by nonparametric estimates; however, the range of 95% CIs for parametric estimates were in most cases much smaller than those

Table 3. Taxonomic distribution of protistan OTUs and the total number of sequences (*N*) comprising all OTUs within each taxon.

Supergroup	Taxon	N	OTUs
Rhizaria	Acantharea*	69	15
	Centroheliozoa	3	1
	Cercozoa	34	11
	Polycystinea	460	17
Chromalveolata	Alveolata; Ciliophora	185	33
	Alveolata; Dinophyceae	410	62
	Alveolata; Ellobiopsidae	1	1
	Alveolata; group I	134	31
	Alveolata; group II	319	25
	Alveolata; unclassified	252	63
	Cryptophyta	18	8
	Haptophyceae	31	9
	Stramenopiles	73	47
Plantae	Chlorophyta	128	11
	Rhodophyta	11	4
Excavata	Euglenozoa	70	21
Opisthokonta	Choanoflagellida	6	5
	Fungi	8	7
Unresolved	Cryothecomonas	3	1
	Ichthyosporea	18	3
	Telonema	13	2
	Total	2246	377

* Nine probable Sticholonchidae sequences grouped within two acantharean OTUs at 95% sequence similarity. Taxonomic classification of OTUs follows the hierarchy presented in Tekle et al. (2009).

associated with Chao-1 and ACE-1 estimates (Table 4). Parametric estimates were approximately two to four times the number of observed OTUs, with the exception of the 150-m sample from January (less than two times the number) and both pooled July and CM sample data (more than five times the number). Only the estimates from best-fitting parametric models were reported among several that were considered (Bunge and Barger 2008).

Similarity of protistan assemblages—Data transformation minimized the effect of well-populated OTUs from dominating comparisons of assemblage similarity while still allowing OTU evenness to contribute to similarity comparisons. Unweighted pair group method with arithmetic mean cluster analysis of Bray–Curtis coefficients (combined with the SIMPROF significance test, PRIMER v.6) for each pairwise comparison of samples indicated significant ($p < 0.05$) differences in assemblage composition between the shallow (5-m and CM) and deep (150-m and 500-m) protistan assemblages (Fig. 5). Clustering of 5-m and CM assemblages within each month was apparent, as well as clustering between adjacent seasons (months). The assemblages from 150 and 500 m formed separate depth-specific clusters, but the differences among these assemblages were not significant ($p > 0.05$).

Analysis of the Bray–Curtis resemblance matrix by nonmetric MDS also revealed differences between shallow and deep assemblages highlighted in two-dimensional (2-D) space (Fig. 6). Ovals depicting protistan assemblage similarities of at least 15% (on the basis of cluster analysis) were overlaid on the 2-D MDS plot to accentuate groups of similar protistan OTUs. The low stress value (e.g., < 0.10) indicated good correspondence between distances on the plots and the original distances in the Bray–Curtis resemblance matrix (Clarke and Gorley 2006). Agreement between cluster analysis and MDS generally suggests that similarities and differences among assemblages were accurately portrayed by both analyses (Clarke and Gorley 2006). The test-statistic *R*, calculated from Bray–Curtis similarity coefficients using ANOSIM, provided a means of significance testing among groups of data (depths and months) suggested by MDS and cluster analysis. The global ANOSIM test was significant ($p = 0.001$), indicating differences due to depth. Significant differences were detected between all pairwise depth comparisons ($p = 0.03$) except between 5-m and CM assemblages ($p = 0.91$).

Separate analysis of shallow and deep assemblages increased the resolution of distances among closely related samples (Fig. 7A, B). Shallow assemblages from January and October were more similar to one another than they were to other shallow assemblages (Fig. 7A). ANOSIM analysis indicated no significant differences between depths for the subset of shallow data ($p = 0.91$). Differences between months were large enough to obtain a significant global test ($p = 0.01$) for the comparison of shallow samples but low sample numbers resulted in low power for pairwise comparisons, preventing identification of significant differences suggested by the global test. A global ANOSIM test of the 150- and 500-m assemblages (Fig. 7B) indicated no significant differences among deep assemblages when similarities were compared by months ($p = 0.23$) but did reveal a significant difference between depths ($p = 0.03$) that was not apparent in the comparison of all 16 (shallow and deep) assemblages or by cluster analysis.

Discussion

Taxonomic diversity of protistan assemblages—Planktonic organisms are known to be highly patchy with respect to their vertical and horizontal distributions, making it difficult to predict the co-occurrence of different species (Hutchinson 1961). Progress is being made to help explain planktonic distributions from different water masses using a combination of molecular and oceanographic techniques (Hamilton et al. 2008). However, most of the existing samples of microbial diversity in the environment represent only snapshots of the overall diversity at a particular location and are generally considered minimal estimates. To date, studies of the spatiotemporal distributions of protistan diversity have been relegated to distances of 10s–100s of meters in the water column (Stoeck et al. 2003; Behnke et al. 2006) and temporal scales of days (Countway et al. 2005) with the several exceptions spanning a greater range of depths (Countway et al. 2007; Not et al. 2007) or longer time periods (Massana et al. 2004*a*; Romari and Vaulot 2004; Medlin et al. 2006).

It should be noted that the present study relied upon a single PCR primer pair (Euk-A and Euk-B) and internal sequencing primer (570-F) to estimate the total protistan diversity at our study site. Although representatives from

Table 4. Protistan diversity estimates for different combinations of sequences, ranging from estimates for the total database of 2246 sequences and 377 OTUs (95% similarity) to estimates for samples from each month (across all four depths) and depth (across all four months). D_S^{-1} is the inverse Simpson's index. Chao-1 and ACE-1 (95% confidence intervals, CI) are nonparametric richness estimators. Rare OTUs were defined for richness calculations as those with just singletons and doubletons (Chao-1) or those with 10 or fewer members (ACE-1). Best parametric estimates and the models producing them are listed in the far right-hand column.

Sample set	*N*	OTUs	D_S^{-1}	Chao-1 (95% CI)	ACE-1 (95% CI)	Best parametric estimate (95% CI), model
Total	2246	377	20.8	682 (587–819)	1105 (859–1476)	994 (844–1192), three-mixed-exponential
Jan	702	162	12.4	329 (257–455)	532 (374–806)	565 (370–945), two-mixed-exponential
Apr	389	91	8.9	476 (250–1028)	948 (443–2176)	382 (263–582), single-exponential
Jul	513	121	23.6	320 (220–524)	450 (286–777)	717 (328–1840), two-mixed-exponential
Octr	642	184	38.1	472 (350–682)	886 (594–1386)	767 (512–1221), two-mixed-exponential
5 m	513	144	28.8	362 (260–554)	540 (358–876)	544 (349–924), two-mixed-exponential
CM	502	153	23.1	403 (292–603)	689 (447–1131)	844 (445–1787), two-mixed-exponential
150 m	606	120	6.4	305 (216–478)	428 (263–787)	312 (250–402), single-exponential
500 m	625	118	10.2	282 (205–427)	686 (356–1475)	469 (328–704), two-mixed-exponential
Jan 5 m	226	87	22.9	312 (185–603)	455 (248–925)	278 (205–398), single-exponential
Jan CM	178	79	31.3	183 (127–307)	225 (140–428)	222 (160–330), two-mixed-exponential
Jan 150 m	150	25	2.1	39 (29–74)	42 (29–88)	41 (31–66), Poisson
Jan 500 m	148	32	3.5	164 (64–575)	158 (69–457)	79 (55–128), single-exponential
Apr 5 m	23	11	6.0	47 (21–137)	120 (26–791)	Too few data to compute
Apr CM	37	17	6.2	115 (33–602)	501 (100–2835)	Too few data to compute
Apr 150 m	175	50	5.0	266 (116–763)	281 (132–698)	142 (101–217), single-exponential
Apr 500 m	154	34	7.5	347 (90–1782)	387 (126–1394)	111 (69–201), single-exponential
Jul 5 m	121	44	15.5	88 (61–159)	255 (103–800)	145 (95–244), single-exponential
Jul CM	133	36	10.9	91 (53–220)	88 (53–199)	79 (58–120), single-exponential
Jul 150 m	107	46	19.6	148 (82–340)	272 (125–688)	126 (89–195), single-exponential
Jul 500 m	152	35	9.3	90 (52–219)	110 (59–276)	108 (67–199), single-exponential
Oct 5 m	143	56	26.8	152 (92–317)	176 (104–356)	118 (92–161), single-exponential
Oct CM	154	61	14.9	181 (109–363)	325 (163–745)	163 (121–234), single-exponential
Oct 150 m	174	55	14.4	99 (73–164)	154 (90–330)	131 (95–200), single-exponential
Oct 500 m	171	72	33.0	251 (145–506)	461 (236–996)	204 (151–292), single-exponential

most high-level taxonomic groups were detected in our clone libraries, it is entirely possible that diversity was underestimated by our choice of oligonucleotide primers. In fact, recent studies have advocated the use of multiple primer pairs and concluded that single sets of primers may drastically underestimate protistan diversity and bias the apparent contribution of particular taxa (Stoeck et al. 2006; Jeon et al. 2008). The DNA from some protistan taxa (e.g., haptophytes) appears to be resistant to amplification by traditional PCR primers, leading to underestimates of their diversity and relative abundance in 18S rRNA gene libraries (Liu et al. 2009). Although methods for PCR, cloning, and sequencing have varied somewhat among different research groups, several of the environmental DNA sequences detected in the present investigation were very similar to many of the sequences reported in the studies listed above. Collectively, these data are beginning to provide support for the hypothesis that many protistan taxa are globally distributed (Fenchel and Finlay 2004; Finlay and Fenchel 2004). That is not to say that *all* taxa are ubiquitously distributed, but the superior sensitivity of many molecular surveys for detecting rare taxa is revealing the presence of many closely related protists from distant sites. An example from the present study is the detection of the harmful algal bloom-forming brown tide organism, *A. anophagefferens*, which has not been reported previously from Pacific waters.

A large number of novel lineages recently described from other locales were detected, including many group I and group II alveolates (Guillou et al. 2008), stramenopiles (Massana et al. 2002), and an array of diverse phylotypes reported from ''extreme'' habitats (Edgcomb et al. 2002; Stoeck and Epstein 2003). In particular, members of the two novel alveolate groups were recovered at each of the 4 months in our seasonal study. Both novel lineages of alveolates were present in clone libraries from all four depths in January and October. Groups I and II alveolates were generally more common in deeper waters as indicated by their large contributions to all 150-m and 500-m libraries (Table 2). Recent work has suggested that both groups I and II alveolates are part of the Syndiniales order of dinoflagellates, many of which are parasitic (Guillou et al. 2008). The recovery of groups I and II alveolates from depths below the euphotic zone at our site, where oxygen concentrations were substantially lower than surface values, is consistent with observations by Guillou et al. (2008). Additional investigation of both of these alveolate groups is warranted given the large numbers of OTUs detected in the present study, including 31 group I alveolate OTUs and 25 group II alveolate OTUs.

The total number of stramenopile sequences recovered from our clone libraries was relatively small (only 73 of 2246 sequences), but the diversity observed within this modest number of sequences was enormous and was

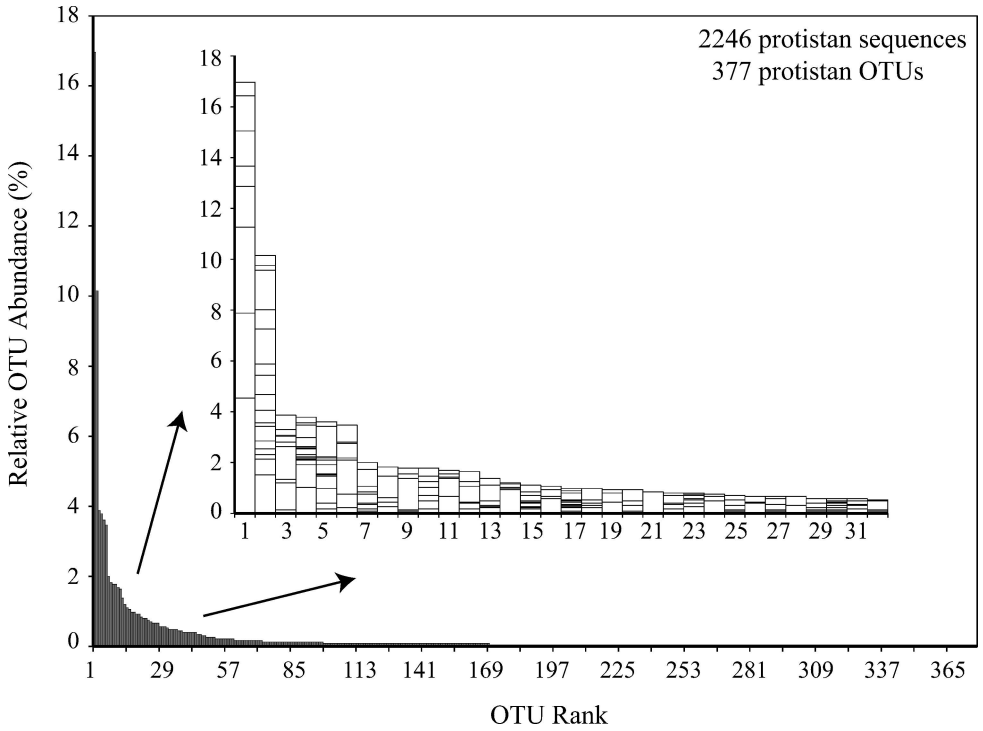

Fig. 3. Relative abundance of 377 protistan operational taxonomic units (OTUs) obtained from a total of 16 individual clone libraries. Inset graphic depicts the first 32 most abundant OTUs (relative abundance summed over all libraries above 0.5%) to display the proportional contributions by sequences from individual clone libraries to the overall OTU abundance (e.g., OTU 2, identified as the dinoflagellate *Gyrodinium*, was present in every season and at every depth depicted by the 16 individual segments within bar).

described by 47 OTUs. Stramenopile OTUs encompassed most of the major morphologically defined taxonomic groups, with more than one-third of the OTUs attributed to diatoms. Previous studies have reported that diatoms typically dominate phytoplankton blooms during the spring and are succeeded by dinoflagellate assemblages in the summer and fall off the coast of Southern California (Moorthi et al. 2006; Schnetzer et al. 2007; Venrick 2009). We identified members of the genera *Minutocellus*, *Cylindrotheca*, *Nitzschia*, *Pseudo-nitzschia*, and *Thalassiosira* on the basis of BLAST similarity scores of at least 97% over the entire length of our partial-length sequences. Although these diatoms are common members of the phytoplankton assemblage in Southern California waters and the San Pedro Channel (D. Caron unpubl.), their relative abundance in the clone libraries did not reflect their typical numerical dominance over dinoflagellates during spring months, possibly suggesting an extraction, amplification, or cloning bias.

A substantial fraction of stramenopile OTUs showed strong affinity to lineages that have been first described during the past decade (Diez et al. 2001; Massana et al. 2002). Novel stramenopiles appeared to be absent from 500-m samples at our site but were present at all other depths during one or more seasons. Romari and Vaulot (2004) detected 26 stramenopile OTUs from a total of 50 stramenopile sequences at an OTU-calling threshold of 98% sequence similarity. Use of this higher similarity threshold in our MESA program increased our stramenopile count to 55 OTUs (up from 47). This modest increase in the number of OTUs for a fairly large increase in sequence similarity (3%) indicates considerable dissimilarity among the 47 OTUs called at 95%.

Polycystines, Acantharea, and Euglenozoa appeared to favor deep-water distributions at our site. Many of these deep-water OTUs were similar to sequences obtained from other deep-sea ecosystems including an anoxic basin in the Caribbean Sea (Stoeck et al. 2003), hydrothermal vents at Guaymas Basin, Gulf of California (Edgcomb et al. 2002), and the deep sea near Antarctica (López-García et al. 2001). Polycystine sequences represented one-fifth of the total number of clones in this study (460) but comprised only 17 OTUs, most of which were $< 95\%$ similar to known taxa. Acantharea comprised a similar number of OTUs (15) but the number of sequences giving rise to these OTUs was much lower (69) than the polycystine count. The presence of these two major heterotrophic groups at depth (particularly 500 m) is not well documented and may be related to the distribution of particular life stages (Gilg et al. 2010). Gilg et al. (2010) detected two novel deep-water acantharean clades (UC1 and UC2) at our study site and for the first time were able to visualize the morphology of an uncultured UC1 acantharean by catalyzed reporter deposition fluorescence in situ hybridization (CARD-FISH). Many morphologically defined species within these groups possess endo-

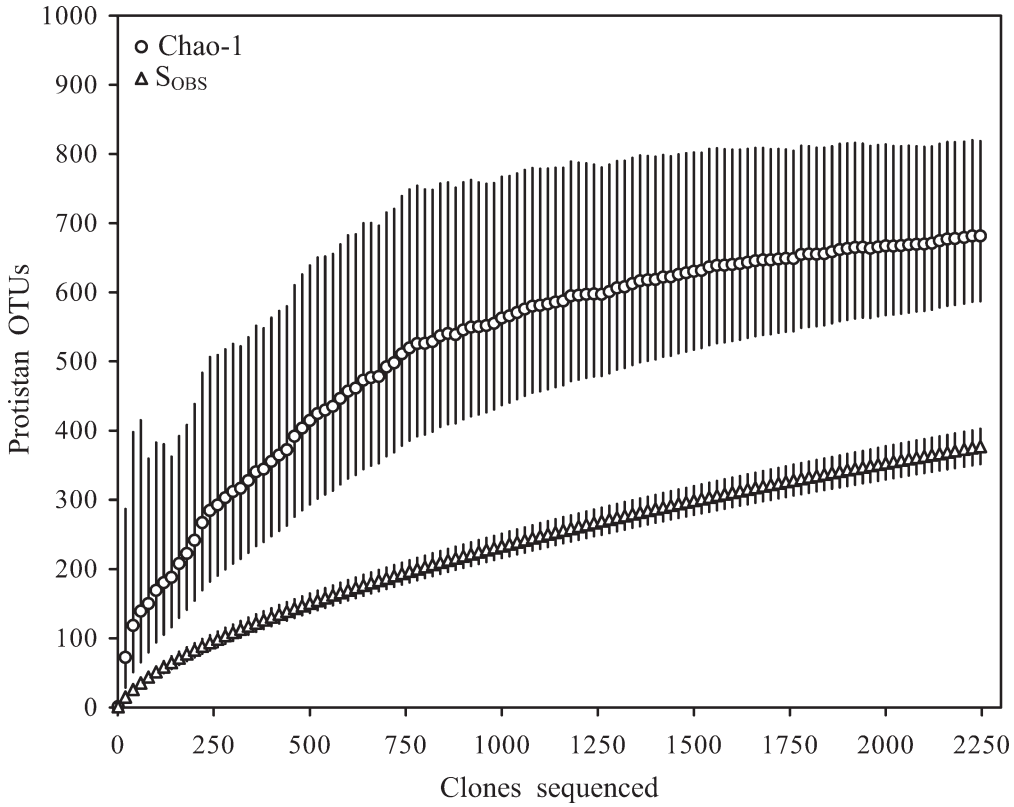

Fig. 4. Rarefaction curves of the Chao-1 diversity estimator (circles) and the number of observed OTUs (triangles) for a given level of sequencing effort using the entire pooled data set of 2246 protistan sequences. Error bars represent the 95% confidence intervals for both sets of data.

symbiotic algae and have been described largely from epipelagic waters (Febvre et al. 2002), but recent studies have also reported large numbers of DNA sequences attributed to these taxa in the deep ocean (Countway et al. 2007; Not et al. 2007).

Euglenozoa were relatively abundant in 150-m and 500-m libraries, with 70 sequences comprising 21 OTUs. This group appears to be more abundant and more diverse in deep-sea ecosystems than previously believed (Lara et al. 2009), with newly described members detected in anoxic ecosystems (Stoeck and Epstein 2003), deep Antarctic waters (López-García et al. 2001), and deep-sea cold seep environments (Buck et al. 2000). Low oxygen tolerance may be one key to explaining the abundance of euglenozoa at 500 m at our study site, where oxygen concentration typically drops to less than 0.50 mL L^{-1}.

In addition to the new lineages described over the past decade, many common protistan lineages were detected in the present study. The prasinophytes *Bathycoccus*, *Micromonas*, and *Ostreococcus* were the most common chlorophytes in our shallow-water clone libraries. This result lends further support to observations by Worden (2006) that these three taxa are dominant members of the picoeukaryote fraction along the coast of Southern California. Evidence suggests that these genera may contribute significantly to primary productivity in this region and at other coastal sites around the world because of their ubiquity and relatively high growth rates (Not et al. 2004; Worden et al. 2004). A quantitative PCR (qPCR)-based study indicated that *Ostreococcus* sp. was highly persistent in the euphotic zone over a 2-yr period at our study site, forming a sizable bloom on at least one sampling date (Countway and Caron 2006). Relative abundance of *Ostreococcus* sequences in our clone libraries mirrored the trends revealed by qPCR in the previous study for all four sampling dates and depths.

Alveolates were a major fraction of most clone libraries, with a high proportion of these lineages attributed to the ciliates (33 OTUs) and dinoflagellates (62 OTUs). Detection of 18 sequences of the red-tide-forming dinoflagellate *Lingulodinium* at the CM in July may provide some of the first evidence of offshore "seed" populations of this genus. It's conceivable that such offshore seed populations may initiate massive blooms if they are advected to the nearshore environment. Unclassified alveolates (likely ciliates or dinoflagellates on the basis of nearest BLAST similarities) were particularly diverse. Many of the unclassified alveolate sequences were most similar to other unclassified alveolate sequences recovered from the English Channel (Romari and Vaulot 2004). Deeper taxonomic surveys of these groups are clearly warranted given their large contributions to our clone libraries.

Observed and estimated protistan diversity—The 2246 protistan sequences reported in this study represent one of the largest protistan sequence data sets for a single

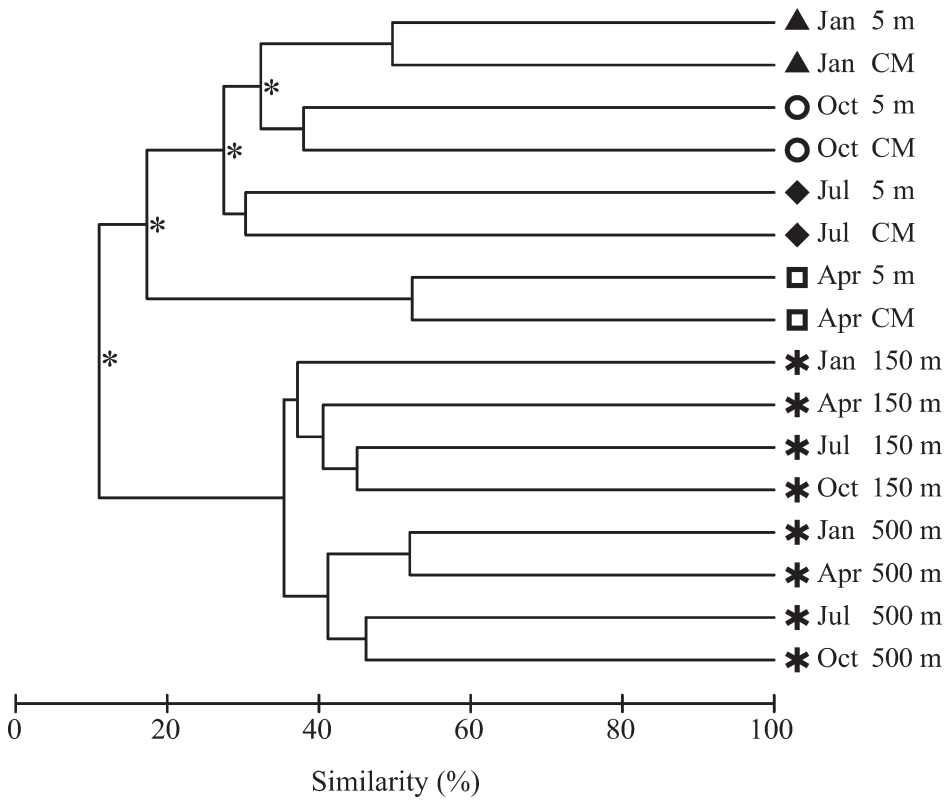

Fig. 5. Cluster diagram of Bray–Curtis similarities calculated from square-root-transformed relative OTU abundances for each clone library. Asterisks at nodes in the dendrogram indicate significant differences between bifurcations ($p < 0.05$). Similar symbols at the end of each branch indicate statistically indistinguishable protistan assemblages.

oceanographic site. Nevertheless, only 36 OTUs were classified as abundant (containing > 10 sequences). Most of the 377 protistan OTUs were present as singletons (208) or doubletons (71), yielding a rank abundance curve possessing a long tail of rare taxa (Fig. 3). Rare taxa dominated the list of OTUs in the rank abundance curves for all individual data sets as well as pooled data sets. This large contribution of the protistan "rare biosphere" (i.e., taxa present in assemblages at very low abundance) has been found to be a common feature of most natural assemblages of microbes, and may represent a unique aspect of microbial community structure and function (Caron and Countway 2009).

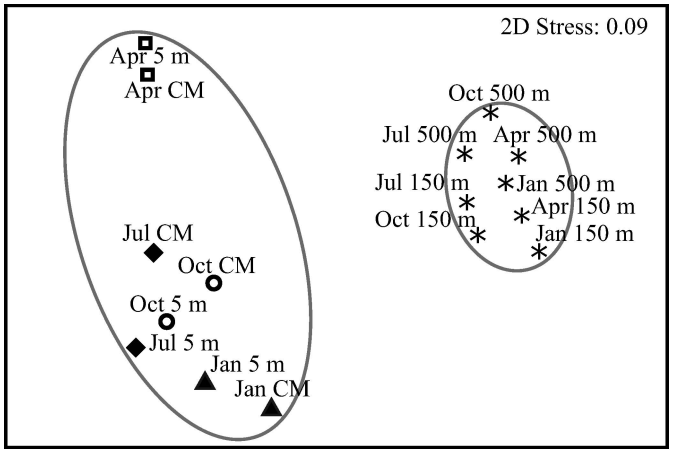

Fig. 6. Nonmetric multidimensional (MDS) scaling plots for the entire data set of 16 clone libraries in two dimensions constructed from a Bray–Curtis similarity matrix of square-root-transformed relative OTU abundances. A similarity value of 15% (solid lines) is depicted overlaying the two-dimensional MDS plot to highlight differences in shallow (5-m and CM) and deep (150-m and 500-m) protistan assemblages. Symbol usage follows that of Fig. 5.

These results imply that, using extant approaches to assess microbial communities, observed diversity will increase with sequencing effort or experimental manipulations because more rare taxa will be detected with increased effort. Indeed, Countway et al. (2005) observed substantially greater protistan diversity in seawater incubations that were repeatedly sampled for 3 d, relative to diversity at the initial time point. Although high-throughput sequencing methods are beginning to provide better sample coverage (Amaral-Zettler et al. 2009; Stoeck et al. 2009), present estimates of the diversity of microbial communities must still rely on the application of a variety of diversity estimators that extrapolate the total diversity from the observed diversity, which has often been collected from relatively small numbers of samples (usually < 10) and limited sample volumes (usually a few liters).

Diversity statistics such as the inverse Simpson index (D_S^{-1}) provide a means of comparing samples analyzed by unequal sampling effort. This index indicated that April

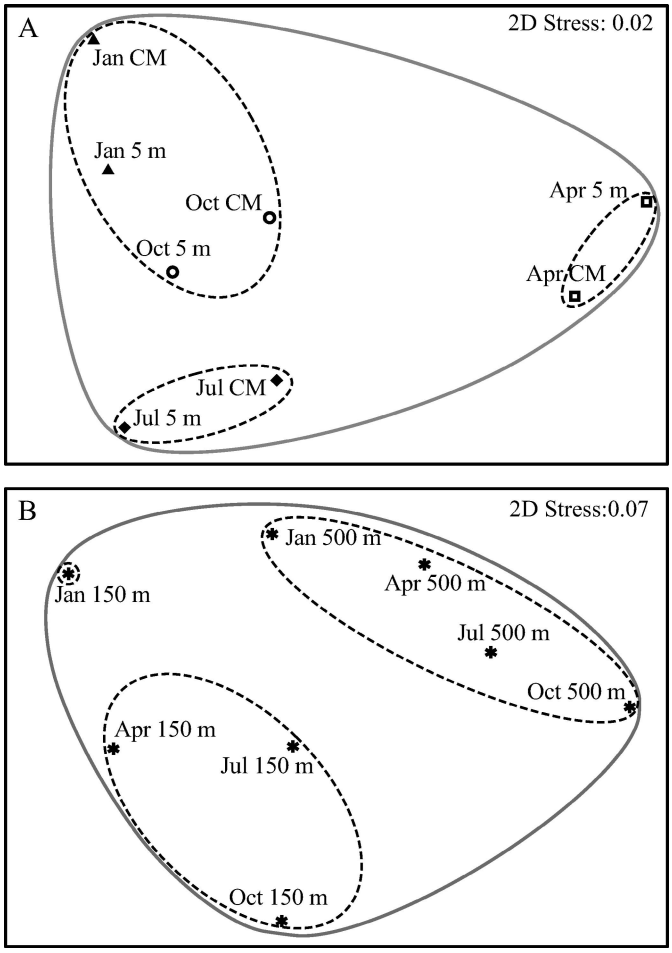

Fig. 7. Nonmetric multidimensional (MDS) scaling plots in two dimensions for the (A) shallow and (B) deep protistan assemblages constructed from a Bray–Curtis similarity matrix of square-root-transformed relative OTU abundances. Assemblage similarity values overlaying the shallow plots are 10% (solid lines) and 30% (dashed lines) and overlaying the deep plots are 10% (solid lines) and 40% (dashed lines). Symbol usage follows that of Fig. 5.

was the least diverse month when OTUs were pooled for each of the four depths within each month (Table 4). One explanation for this trend is the occurrence of a spring phytoplankton bloom at this site (Countway and Caron 2006; Schnetzer et al. 2007) resulting in a protistan assemblage that was strongly dominated by a small number of phytoplankton species and many metazoa. However, April libraries from 150 and 500 m also yielded low values for D_S^{-1} (Table 4), suggesting that the spring bloom may reduce diversity throughout the water column. Overall, deep samples from January were the least diverse, perhaps reflecting the high abundance of polycystine sequences that were distributed among only a few OTUs in these libraries (Table 2).

Values of D_S^{-1} calculated for samples grouped by depth (combining all 4 months) were higher for the 5-m and CM libraries than for the 150- and 500-m libraries (Table 4). These trends are not surprising given that deeper waters experience environmental conditions that are more seasonally stable (Fig. 2). The presence of phototrophic taxa (e.g., stramenopiles and dinoflagellates) in the upper water column also explains a portion of this depth-related diversity gradient. A similar relationship between protistan assemblage diversity and water depth was observed in the oligotrophic North Atlantic across a much larger range of depths and spatial scales (Countway et al. 2007; Not et al. 2007). One exception to the general trend of lower protistan diversity with increasing depth was observed for the sample collected at 500 m during October. This sample had the highest D_S^{-1} value (33.0) for any of the 16 individual clone libraries in our study (Table 4). Many of the major taxonomic groups were well represented at 500 m in October by multiple OTUs, explaining this high value (Tables 2, 4). Higher frequency analysis will be needed to explain periodic increases in diversity.

Estimates of total protistan species richness provided by the nonparametric Chao-1 and ACE-1 estimators were very high for the combined 16 clone libraries in this study (Table 4). These values represent some of the highest values of protistan diversity yet reported for a single oceanographic site. Similar values were reported by Countway et al. (2007) for samples collected in the North Atlantic (694 OTUs by Chao-1 and 773 OTUs by ACE-1) using a 95% similarity threshold for defining OTUs (Caron et al. 2009). These findings imply that the underlying pool of protistan diversity from which samples were drawn was substantially higher than the observed diversity (Fig. 4). Zuendorf et al. (2006) found that ACE-1 provided the best agreement with diversity estimates on the basis of parametric approaches. Parametric methods for assessing the total diversity from a subsample have the advantage of using more of the biological data than nonparametric approaches described above. Protistan diversity in two anoxic fjords was estimated by parametric methods and ranged from 32–143 (Behnke et al. 2006) up to 188 (Zuendorf et al. 2006) at an OTU-calling threshold of 98% sequence similarity. Both of the previous studies observed considerably lower numbers of OTUs compared with diversity estimates predicted by model-based approaches.

Analysis of community similarity—Multivariate approaches including hierarchical cluster analysis and nonmetric MDS have been increasingly used to compare diverse protistan assemblages (Countway et al. 2007; Vigil et al. 2009). Bray–Curtis similarities are typically favored for the comparison of microbial assemblages because joint absences of OTUs do not affect the calculation (Clarke and Gorley 2006). Multivariate data analysis has become a powerful new tool for the comparison of microbial assemblages primarily because it makes use of similarity matrices that reflect differences in the type and relative abundance of each OTU but is also useful for analyzing presence vs. absence data.

Protistan assemblages at the USC Microbial Observatory formed distinct clusters on the basis of the depth of sample collection, the month of collection, and a combination of these factors (Fig. 5). These groupings were further supported by nonmetric MDS plots and statistical testing that indicated significant differences on the basis of

depth (Fig. 6). The stress value for the 2-D MDS plot was low (< 0.1), which indicated unambiguous relationships (i.e., believable distances) among the data points. Assemblages from 5 m and the CM within each month were not significantly different; however, these shallow-water assemblages were significantly different from all deep-water (150 and 500 m) assemblages. Similarities between shallow-water assemblages (5 m and CM) during each month were not surprising since these sampling depths were located above the main thermocline of the water column, were separated in the water column by only 10s of meters, and tended to be dominated by similar guilds of protists. Similar depth-related trends have been reported for phytoplankton assemblages in the central North Pacific, where the communities were composed of many rare microalgal species and relatively few common species that dominated cell counts (Venrick 1990).

Differences between shallow-water protistan assemblages from different seasons were not surprising, and are well supported by studies of planktonic species succession that predict a general progression of the dominant plankton types over time within the euphotic zone (Larsen et al. 2004). Nonmetric MDS plots of the shallow samples alone revealed that assemblages from October and January were more similar to one another than they were to shallow assemblages from April or July (Fig. 7A). In addition, similarity of the shallow assemblages during October and January was reflected by their similar values of D_S^{-1}, which were generally higher than the indices for April and July (Table 4). The large difference in MDS distances between shallow-water assemblages from April and all other months was attributable to the spring phytoplankton bloom that was dominated by a relatively small number of phytoplankton taxa, and to the overall small number of protistan taxa detected in April when copepods dominated clone libraries.

The detection of differences between shallow and deep protistan assemblages was not unexpected given the obvious differences in the trophic structure between the euphotic zone and deep-water communities. Initial MDS plots using the entire data set (16 libraries) yielded a tight coupling of 150- and 500-m assemblages, implying a relatively stable protistan community structure across mid-water depths (Fig. 6). The overall taxonomic composition of the assemblages from these depths revealed interesting similarities among five major taxonomic groups that included group I and group II alveolates, acanthareans, polycystines, and euglenozoa. These five taxonomic groups accounted for more than half of all sequences analyzed from 150- and 500-m libraries and were likely the driving force behind the observed similarities. However, further MDS analysis of the deep protistan assemblages in the absence of shallow samples revealed separation between communities from 150 and 500 m, which was determined to be significantly different by ANOSIM (Fig. 7B) and not revealed by cluster analysis alone. Seasonality was not apparent in the physical and chemical parameters observed at these two depths (Fig. 2), but substantial gradients that existed across the 150–500-m depth interval (in particular, oxygen; Fig. 2) presumably gave rise to the subtle differences observed between assemblages at these two depths. Similarly, Behnke et al. (2006) observed major differences in the structure of protistan assemblages sampled across a narrow depth interval (18 m) spanning sharp environmental gradients. The combined results of the present study and previous ones suggest that many protistan taxa are greatly restricted in their vertical range across the water column, with many deep-sea taxa rarely if ever detected in the euphotic zone (Countway et al. 2007; Not et al. 2007).

In summary, genetic analyses of protistan assemblages in the upper 500 m of the water column at a time-series station in the eastern North Pacific yielded a total of 377 unique protistan OTUs from the analysis of 16 clone libraries containing a total population of 2246 DNA sequences. Nearly 75% of these OTUs were present only as singletons or doubletons, indicating a much higher unsampled diversity from which the observed OTUs were drawn. Conservative estimates of the total unsampled protistan diversity typically ranged from 100 to 300 OTUs for individual clone libraries but increased to approximately 700 to 1000 OTUs for this study site when all libraries were pooled. Multivariate statistical analysis was utilized for comparisons of assemblage diversity across all 16 clone libraries. Significant differences were detected between seasons for shallow-water assemblages (5 m and CM); however, there were no significant differences between these shallow-water assemblages within a given month. Highly significant differences were detected between all shallow libraries and those from 150 and 500 m. Deep-water libraries were not statistically different when libraries from all depths were analyzed; however, a separate analysis of the eight deep-water assemblages indicated subtle differences between these two depths. Our results support the idea that protistan assemblages are highly diverse, and that local environmental conditions and processes select a small number of dominant taxa from a very wide diversity of taxa present at this locale. These selective processes result in compositional differences at different depths and through time.

Acknowledgments

We thank the captain and crew of the RV *Sea Watch* for assistance with the collection of samples. Chemical and physical data were provided by the San Pedro Ocean Time-series (SPOT) project and the Wrigley Institute for Environmental Studies. We are grateful to Tony Michaels for research vessel support provided to the University of Southern California (USC) Microbial Observatory project in conjunction with the long-term oceanographic sampling program at the SPOT. Julie Rose and Isabel Romero provided advice on statistical analyses. Ian Hewson, Josh Steele, and Jed Fuhrman collaborated with us on many aspects of the USC Microbial Observatory project. We thank three anonymous reviewers for providing helpful comments on earlier versions of this manuscript.

This work was supported by National Science Foundation grant MCB-0703159.

References

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402, doi:10.1093/nar/25.17.3389

Amaral-Zettler, L. A., E. A Mccliment, H. W. Ducklow, and S. M. Huse. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE **4:** e6372. doi:10.1371/journal.pone.0006372

Behnke, A., J. Bunge, K. Barger, H. W. Breiner, V. Alla, and T. Stoeck. 2006. Microeukaryote community patterns along an O_2/H_2S gradient in a supersulfidic anoxic fjord (Framvaren, Norway). Appl. Environ. Microbiol. **72:** 3626–3636, doi:10.1128/AEM.72.5.3626-3636.2006

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler. 2004. GenBank: Update. Nucleic Acids Res. **32:** D23–26, doi:10.1093/nar/gkh045

Buck, K. R., J. P. Barry, and A. G. B. Simpson. 2000. Monterey Bay cold seep biota: Euglenozoa with chemoautotrophic bacterial epibionts. Eur. J. Protistol. **36:** 117–126.

Bunge, J., and K. Barger. 2008. Parametric models for estimating the number of classes. Biom. J. **50:** 971–982, doi:10.1002/bimj.200810452

Burki, F., K. Shalchian-Tabrizi, M. Minge, Å. Skjæveland, S. I. Nikolaev, K. S. Jakobsen, and J. Pawlowski. 2007. Phylogenomics reshuffles the eukaryotic supergroups. PLoS ONE **2:** e790, doi:10.1371/journal.pone.0000790

Caron, D. A., and P. D. Countway. 2009. Hypotheses on the role of the protistan rare biosphere in a changing world. Aquat. Microb. Ecol. **57:** 227–238, doi:10.3354/ame01352

———, and others. 2009. Defining DNA-based operational taxonomic units for microbial–eukaryote ecology. Appl. Environ. Microbiol. **75:** 5797–5808, doi:10.1128/AEM.00298-09

Chao, A., and T.-J. Shen. 2005. Program SPADE (Species Prediction And Diversity Estimation). Available online at: http://chao.stat.nthu.edu.tw.

Clarke, K. R., and R. N. Gorley. 2006. Primer v6: User manual/tutorial. PRIMER-E.

Cole, J. R., and others. 2003. The ribosomal database project (RDP-II): Previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. Nucleic Acids Res. **31:** 442–443, doi:10.1093/nar/gkg039

Colwell, R. K. 2004. Estimates: Statistical estimation of species richness and shared species from samples, version 7. Available online at: http://purl.oclc.org/estimates.

Countway, P. D., and D. A. Caron. 2006. Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as revealed by quantitative PCR. Appl. Environ. Microbiol. **72:** 2496–2506, doi:10.1128/AEM.72.4.2496-2506.2006

———, R. J. Gast, M. R. Dennett, P. Savai, J. M. Rose, and D. A. Caron. 2007. Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). Environ. Microbiol. **9:** 1219–1232, doi:10.1111/j.1462-2920.2007.01243.x

———, ———, P. Savai, and D. A. Caron. 2005. Protistan diversity estimates based on 18S rDNA from seawater incubations in the western North Atlantic. J. Eukaryot. Microbiol. **52:** 95–106, doi:10.1111/j.1550-7408.2005.05202006.x

Curtis, T. P., and W. T. Sloan. 2004. Prokaryotic diversity and its limits: Microbial community structure in nature and implications for microbial ecology. Curr. Opin. Microbiol. **7:** 221–226, doi:10.1016/j.mib.2004.04.010

Diez, B., C. Pedrós-Alió, and R. Massana. 2001. Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. Appl. Environ. Microbiol. **67:** 2932–2941, doi:10.1128/AEM.67.7.2932-2941.2001

Edgcomb, V. P., D. T. Kysela, A. Teske, A. D. Gomez, and M. L. Sogin. 2002. Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. Proc. Natl. Acad. Sci. USA **99:** 7658–7662, doi:10.1073/pnas.062186399

Febvre, C., J. Febvre, and A. F. Michaels. 2002. Acantharia, p. 783–803. *In* J. J. Lee, G. F. Leedale, and P. Bradbury [eds.], An illustrated guide to the protozoa, 2nd ed. Allen Press.

Fenchel, T., and B. J. Finlay. 2004. The ubiquity of small species: Patterns of local and global diversity. Bioscience **54:** 777–784, doi:10.1641/0006-3568(2004)054[0777:TUOSSP]2.0.CO;2

Finlay, B. J., and T. Fenchel. 2004. Cosmopolitan metapopulations of free-living microbial eukaryotes. Protist **155:** 237–244, doi:10.1078/143446104774199619

Gast, R. J., M. R. Dennett, and D. A. Caron. 2004. Characterization of protistan assemblages in the Ross Sea, Antarctica, by denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. **70:** 2028–2037, doi:10.1128/AEM.70.4.2028-2037.2004

Gilg, I. C., L. A. Amaral-Zettler, P. D. Countway, S. Moorthi, A. Schnetzer, and D. A. Caron. 2010. Phylogenetic affiliations of mesopelagic Acantharia and Acantharian-like environmental 18S rRNA genes off the Southern California coast. Protist **161:** 197–211, doi:10.1016/j.protis.2009.09.002

Guillou, L., and others. 2008. Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). Environ. Microbiol. **10:** 3349–3365, doi:10.1111/j.1462-2920.2008.01731.x

Hamilton, A. K., C. Lovejoy, P. E. Galand, and R. G. Ingram. 2008. Water masses and biogeography of picoeukaryote assemblages in a cold hydrographically complex system. Limnol. Oceanogr. **53:** 922–935.

Hong, S. H., J. Bunge, S. O. Jeon, and S. S. Epstein. 2006. Predicting microbial species richness. Proc. Natl. Acad. Sci. USA **103:** 117–122, doi:10.1073/pnas.0507245102

Hutchinson, G. E. 1961. The paradox of the plankton. Am. Nat. **95:** 137–145, doi:10.1086/282171

Jeon, S., J. Bunge, C. Leslin, T. Stoeck, S. H. Hong, and S. S. Epstein. 2008. Environmental rRNA inventories miss over half of protistan diversity. BMC Microbiol. **8**: doi:10.1186/1471-2180-8-222

Lara, E., D. Moreira, A. Vereshchaka, and P. Lopez-Garcia. 2009. Pan-oceanic distribution of new highly diverse clades of deep-sea diplonemids. Environ. Microbiol. **11:** 47–55, doi:10.1111/j.1462-2920.2008.01737.x

Larsen, A., and others. 2004. Spring phytoplankton bloom dynamics in Norwegian coastal waters: Microbial community succession and diversity. Limnol. Oceanogr. **49:** 180–190, doi:10.4319/lo.2004.49.1.0180

Liu, H., and others. 2009. Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. Proc. Natl. Acad. Sci. USA **106:** 12803–12808, doi:10.1073/pnas.0905841106

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

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

Massana, R., V. Balague, L. Guillou, and C. Pedrós-Alió. 2004*a*. Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. FEMS Microbiol. Ecol. **50:** 231–243, doi:10.1016/j.femsec.2004.07.001

———, L. Guillou, B. Diez, and C Pedrós-Alió. 2002. Unveiling the organisms behind novel eukaryotic ribosomal DNA sequences from the ocean. Appl. Environ. Microbiol. **68:** 4554–4558, doi:10.1128/AEM.68.9.4554-4558.2002

———, R. Terrado, I. Forn, C. Lovejoy, and C. Pedrós-Alió. 2006. Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. Environ. Microbiol. **8:** 1515–1522, doi:10.1111/j.1462-2920.2006.01042.x

———, and others. 2004*b*. Phylogenetic and ecological analysis of novel marine stramenopiles. Appl. Environ. Microbiol. **70:** 3528–3534, doi:10.1128/AEM.70.6.3528-3534.2004

Medlin, L., K. Metfies, H. Mehl, K. Wiltshire, and K. Valentin. 2006. Picoeukaryotic plankton diversity at the Helgoland time series site as assessed by three molecular methods. Microb. Ecol. **52:** 53–71, doi:10.1007/s00248-005-0062-x

Moorthi, S. D., P. D. Countway, B. A. Stauffer, and D. A. Caron. 2006. Use of quantitative real-time PCR to investigate the dynamics of the red tide dinoflagellate *Lingulodinium polyedrum*. Microb. Ecol. **52:** 136–150, doi:10.1007/s00248-006-9030-3

Not, F., R. Gausling, F. Azam, J. F. Heidelberg, and A. Z. Worden. 2007. Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. Environ. Microbiol. **9:** 1233–1252, doi:10.1111/j.1462-2920.2007.01247.x

———, M. Latasa, D. Marie, T. Cariou, D. Vaulot, and N. Simon. 2004. A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the western English Channel. Appl. Environ. Microbiol. **70:** 4064–4072, doi:10.1128/AEM.70.7.4064-4072.2004

Potvin, M., and C. Lovejoy. 2009. PCR-based diversity estimates of artificial and environmental 18S rRNA gene libraries. J. Eukaryot. Microbiol. **56:** 174–181, doi:10.1111/j.1550-7408.2008.00386.x

Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. G. Ludwig, J. Peplies, and F. O. Glockner. 2007. SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. **35:** 7188–7196, doi:10.1093/nar/gkm864

Romari, K., and D. Vaulot. 2004. Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. Limnol. Oceanogr. **49:** 784–798, doi:10.4319/lo.2004.49.3.0784

Schnetzer, A., and others. 2007. Blooms of *Pseudo-nitzschia* and domoic acid in the San Pedro Channel and Los Angeles Harbor areas of the Southern California Bight, 2003–2004. Harmful Algae **6:** 372–387, doi:10.1016/j.hal.2006.11.004

Shalchian-Tabrizi, K., H. Kauserud, R. Massana, D. Klaveness, and K. S. Jakobsen. 2007. Analysis of environmental 18S ribosomal RNA sequences reveals unknown diversity of the cosmopolitan phylum Telonemia. Protist **158:** 173–180, doi:10.1016/j.protis.2006.10.003

Sherr, B. F., E. B. Sherr, D. A. Caron, D. Vaulot, and A. Z. Worden. 2007. Oceanic protists. Oceanography **20:** 130–134.

Stoeck, T., and S. Epstein. 2003. Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments. Appl. Environ. Microbiol. **69:** 2657–2663, doi:10.1128/AEM.69.5.2657-2663.2003

———, B. Hayward, G. T. Taylor, R. Varela, and S. S. Epstein. 2006. A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples. Protist **157:** 31–43, doi:10.1016/j.protis.2005.10.004

———, G. T. Taylor, and S. S. Epstein. 2003. Novel eukaryotes from the permanently anoxic Cariaco Basin (Caribbean Sea). Appl. Environ. Microbiol. **69:** 5656–5663, doi:10.1128/AEM.69.9.5656-5663.2003

———, and others. 2009. Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. BMC Biol. **7**: doi:10.1186/1741-7007-7-72

Tekle, Y. I., L. W. Parfrey, and L. A. Katz. 2009. Molecular data are transforming hypotheses on the origin and diversification of eukaryotes. Bioscience **59:** 471–481, doi:10.1525/bio.2009.59.6.5

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680, doi:10.1093/nar/22.22.4673

Venrick, E. L. 1990. Phytoplankton in an oligotrophic ocean—species structure and interannual variability. Ecology **71:** 1547–1563, doi:10.2307/1938291

———. 2009. Floral patterns in the California Current: The coastal-offshore boundary zone. J. Mar. Res. **67:** 89–111, doi:10.1357/002224009788597917

Vigil, P., P. D. Countway, J. Rose, D. J. Lonsdale, C. J. Gobler, and D. A. Caron. 2009. Rapid shifts in dominant taxa among microbial eukaryotes in estuarine ecosystems. Aquat. Microb. Ecol. **54:** 83–100, doi:10.3354/ame01252

von Wintzingerode, F., U. B. Gobel, and E. Stackebrandt. 1997. Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. **21:** 213–229, doi:10.1111/j.1574-6976.1997.tb00351.x

Worden, A. Z. 2006. Picoeukaryote diversity in coastal waters of the Pacific Ocean. Aquat. Microb. Ecol. **43:** 165–175, doi:10.3354/ame043165

———, J. K. Nolan, and B. Palenik. 2004. Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. Limnol. Oceanogr. **49:** 168–179, doi:10.4319/lo.2004.49.1.0168

Zuendorf, A., J. Bunge, A. Behnke, K. J.-A. Barger, and T. Stoeck. 2006. Diversity estimates of microeukaryotes below the chemocline of the anoxic Mariager Fjord, Denmark. FEMS Microbiol. Ecol. **58:** 476–491, doi:10.1111/j.1574-6941.2006.00171.x

Associate editor: Michael R. Landry

Received: 19 August 2009
Accepted: 27 May 2010
Amended: 29 July 2010