Reference: *Biol. Bull.* **227:** 175–190. (October 2014)

Ciliate Diversity, Community Structure, and Novel Taxa in Lakes of the McMurdo Dry Valleys, Antarctica

YUAN XU[1,*†], TRISTA VICK-MAJORS[2], RACHAEL MORGAN-KISS[3], JOHN C. PRISCU[2], AND LINDA AMARAL-ZETTLER[4,5,*‖]

[1]*Laboratory of Protozoology, Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao 266003, China;* [2]*Montana State University, Department of Land Resources and Environmental Sciences, 334 Leon Johnson Hall, Bozeman, Montana 59717;* [3]*Department of Microbiology, Miami University, Oxford, Ohio 45056;* [4]*The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, Massachusetts 02543; and* [5]*Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, Rhode Island 02912*

Abstract. We report an in-depth survey of next-generation DNA sequencing of ciliate diversity and community structure in two permanently ice-covered McMurdo Dry Valley lakes during the austral summer and autumn (November 2007 and March 2008). We tested hypotheses on the relationship between species richness and environmental conditions including environmental extremes, nutrient status, and day length. On the basis of the unique environment that exists in these high-latitude lakes, we expected that novel taxa would be present. Alpha diversity analyses showed that extreme conditions—that is, high salinity, low oxygen, and extreme changes in day length—did not impact ciliate richness; however, ciliate richness was 30% higher in samples with higher dissolved organic matter. Beta diversity analyses revealed that ciliate communities clustered by dissolved oxygen, depth, and salinity, but not by season (*i.e.,* day length). The permutational analysis of variance test indicated that depth, dissolved oxygen, and salinity had significant influences on the ciliate community for the abundance matrices of resampled data, while lake and season were not significant. This result suggests that the vertical trends in dissolved oxygen concentration and salinity may play a critical role in structuring ciliate communities. A PCR-based strategy capitalizing on divergent eukaryotic V9 hypervariable region ribosomal RNA gene targets unveiled two new genera in these lakes. A novel taxon belonging to an unknown class most closely related to *Cryptocaryon irritans* was also inferred from separate gene phylogenies.

* Authors contributed equally.

‖ To whom correspondence should be addressed. E-mail: amaral@mbl.edu

† Current address: State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200062, China.

Abbreviations: ELB, East Lobe Bonney; GAST, Global Alignment for Sequence Taxonomy; MCM, McMurdo Dry Valleys; OTU, operational taxonomic unit; WLB, West Lobe Bonney.

Introduction

The McMurdo Dry Valleys (MCM), Antarctica, harbor numerous perennially ice-covered lakes in our planet's coldest and driest desert (Priscu *et al.*, 1998). The physicochemical conditions in these lakes can be extreme, with salinities ranging from fresh to as much as 5 times seawater, temperatures below −4 °C, and oxygen ranging from >200% supersaturation in the upper photic zones to suboxic below permanent chemoclines (Spigel and Priscu, 1998; Vick and Priscu, 2012). In addition to environmental extremes, solar radiation fluctuates from 24-h sunlight in the austral summer to complete darkness in the austral winter, imposing additional stresses on the microbiota that dominate these polar habitats. There are 4 mon of full incident sunlight (November to February) and 4 mon of total darkness (May to August) with twilight in between. The remote geographical nature of the lakes, the <3-m-thick perennial ice covers on the lakes, and the circumpolar oceans surrounding Antarctica create a barrier to colonization (Spaulding *et al.*, 2010) that makes these ecosystems ideal locations

to search for endemic novel taxa in a biogeographically isolated environment.

The lakes contain no fish; microorganisms dominate the food web. Among the microbiota that inhabit the MCM lakes are microbial eukaryotes, mainly protists representing a broad range of evolutionary lineages inclusive of phototrophs, mixotrophs, and heterotrophs (Kong *et al.*, 2012; Thurman *et al.*, 2012). Protist communities are important in carbon and nutrient cycling, occupying both the bottom (primary production) and top (predation) of the truncated MCM food web (Bielewicz *et al.,* 2011). Among the heterotrophs, ciliates, often regarded as "flagship" taxa because of their conspicuous size and morphology (Foissner, 2005), have been the object of past morphology-based biodiversity investigations in the MCM lakes. These studies suggested that ciliate populations in the lakes were more diverse than anticipated (Kepner *et al.*, 1999). While the diversity of bacteria and archaea in MCM lakes has been studied using molecular methods (Brambilla *et al.*, 2001; Karr *et al.*, 2006), the biodiversity of protists, especially ciliates, in these lakes has been largely limited to analysis with morphology-based approaches (Laybourn-Parry *et al.*, 1996; Roberts *et al.*, 2000; Song and Wilbert, 2002), except for a single clone library based sequencing study (Bielewicz *et al.*, 2011).

In aquatic ecosystems, ciliates are a main component of the microbial loop (Azam *et al*., 1983) and play an important role in the microbial food web by controlling bacterial prey populations, redistributing organic sources of nutrients and carbon, and linking the microbial loop with higher trophic levels (Finlay and Esteban, 1998). With short life cycles and delicate external membranes, ciliates may respond more quickly than metazoa to environmental changes (Coppellotti and Matarazzo, 2000) and thus represent sensitive indicator species (Xu *et al*., 2011a, b). Many ciliates can tolerate extremes in environmental conditions and have been reported in a range of extreme environments (Coppellotti and Matarazzo, 2000; Jiang *et al.*, 2007). The food web in the MCM lakes is dominated by the microbial loop in which ciliates likely represent the top predators and therefore are important in carbon and nutrient cycling in these aquatic systems. Indeed, in Lake Fryxell (Fig. 1), ciliates occur in high abundance (Laybourn-Parry *et al.*, 1996), providing a good proxy for the study of biodiversity changes in this environment as a whole.

An estimated 83%–89% of the 4500 known free-living ciliate morphospecies remain to be formally described (Foissner *et al.*, 2009; Xu *et al.*, 2011a). Morphological-based identification has shown that more than 20% of the taxa found in the MCM lakes are novel (Petz *et al.*, 2007). Despite a wealth of morphological features, the identification of ciliates at the species level is laborious, often subjective, and requires skilled taxonomic expertise that can miss rare species (Xu *et al.*, 2011a, b). To gain more information on ciliate diversity and community structure

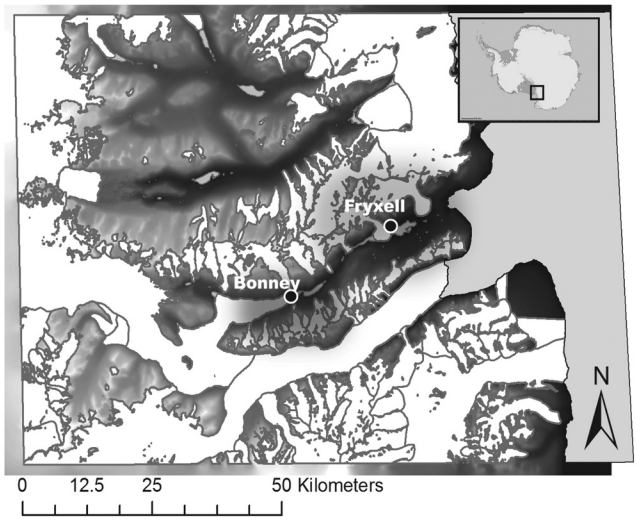

Figure 1. Locations of Lake Fryxell and Lake West Lobe Bonney (WLB) sampled in this study. The white areas represent glaciers; darker shades represent exposed soils. Both lakes are located within the Taylor Valley of the McMurdo Dry Valleys (see inset).

and to determine potentially endemic taxa, we used next-generation amplicon sequencing to examine ciliate diversity and community structure in two MCM lakes, Lake Fryxell and Lake West Lobe Bonney (Fig. 1).

We exploited the unique nature of the MCM lakes, as well as their remote location, to test the following hypotheses regarding ciliate biogeography and species richness: (1) the extreme environmental conditions encountered in the MCM Lakes (*i.e.,* high salinity (>80 PSU), cold, and low oxygen at depth) result in overall low richness of ciliate populations; (2) ciliate richness increases with nutrient concentrations *via* interactions with bacterial communities (*i.e.,* nitrate, dissolved organic carbon, *etc.*), as most ciliates feed on bacteria and organics; (3) according to the day-length hypothesis of Gilbert *et al*. (2010), which postulates that day length might be a primary determinant of the seasonality of marine bacterial communities, richness of ciliate populations is higher in autumn than in summer, a situation that characterizes Antarctic marine bacterioplankton (Ghiglione and Murray, 2012); and (4) given the isolated and extreme physical and chemical nature of the MCM lakes environment, our deep sequencing efforts would identify endemic and/or novel ciliate species.

Materials and Methods

Sample collection and next-generation sequencing

This study was part of the Microbial Inventory Research Across Diverse Aquatic Long Term Ecological Research Sites (MIRADA-LTERS) Project with the broad goal of documenting, describing, and discovering novel diversity at aquatic LTER locations. The MCM-LTER site (MIRADA

Table 1

Contextual data from our study

	Sample ID							
Contextual Data	MCM1	MCM2	MCM3	MCM4	MCM5	MCM6	MCM7	MCM8
Collection Date (UTC)	11/22/07	11/22/07	11/30/07	11/30/07	3/17/08	3/17/08	3/25/08	3/25/08
Depth (Meters)	6	9	13	18	13	18	6	9
Lake (Name)	Lake Fryxell	Lake Fryxell	Lake WLB	Lake WLB	Lake WLB	Lake WLB	Lake Fryxell	Lake Fryxell
Latitude (Decimal Degree)	−77.61031	−77.61031	−77.72006	−77.72006	−77.72006	−77.720072006	−77.61031	−77.61031
Longitude (Decimal Degree)	85.53597	85.53597	84.57732	84.57732	84.57732	84.57732	85.53597	85.53597
Ammonium (μmol l^{-1})	0.3	0.17	0.96	172.9	0.67	157.75	0.08	0.32
Pressure (db)	5.27	8.2	12.23	17.51	12.3	17.42	5.28	8.2
Chlorophyll (μg l^{-1})	3.3	5	4	0.9	5.5	0.8	4	10
Conductivity (mS cm^{-1})	1.36	3.3	12.54	65.13	10.01	64.08	1.83	2.04
Dissolved Inorganic Nitrogen (μmol l^{-1})	0.57	0.34	13.26	189.12	9.94	174.08	0.31	0.58
Dissolved Organic Carbon (μmol l^{-1})	237	472	370	1033	320	758	170	514
Dissolved Organic Nitrogen (μmol l^{-1})	21.39	33.31	6.62	53.84	2.99	56.41	11.58	36.4
Dissolved Oxygen (μmol l^{-1})	802.67	914.54	1591.72	58.55	1562.62	68.72	807.32	764.52
Nitrate (μmol l^{-1})	0.15	0.08	11.21	13.2	9.53	14.68	0.1	0.09
Nitrite (μmol l^{-1})	0.11	0.09	1.09	3.02	0.45	2.32	0.14	0.17
Bacterial Cells ($\times 10^6$ ml^{-1})	1.13	0.84	0.05	0.07	0.09	0.04	1.42	1.25
Particulate Carbon (μmol l^{-1})	30.85	36.45	28.28	20.15	43.16	17.71	44.47	84.54
Particulate Nitrogen (μmol l^{-1})	3.78	4.32	2.01	1.78	1.72	1.57	3.29	5.34
pH (Log H+)	8.02	7.73	7.05	5.87	7.62	5.97	8.18	7.73
Phosphate (μmol l^{-1})	0.13	0.28	0.05	0.27	0.04	0.19	0.06	0.15
Salinity (PSU)	1.26	3.12	13.55	90.02	10.61	88.08	1.82	1.96
Silicate (μmol l^{-1})	200	220	200	220	200	220	200	220
Temperature (°C)	1.42	2.17	1.03	−0.72	1.25	−0.65	1.31	1.83
Water Column Depth (Meters)	18.6	18.6	41.5	41.5	41.5	41.5	18.6	18.6

LTERS, 2008) focuses on the Taylor Valley located in the Transantarctic Mountains (−77.5, 160.0) and lies within the largest ice-free area (~4500 km^2) on the Antarctic continent (Levy, 2012). The perennially ice-covered lakes, ephemeral streams, and extensive areas of exposed soil within the McMurdo Dry Valleys (MCM) are subject to low temperatures and limited precipitation. The MCM is a region where life approaches its environmental limits, and is an end-member in the spectrum of environments included in the LTER Network. We collected two replicate samples of water (between 0.8- and 2-l each) for DNA extraction at two depths in lakes Fryxell and Bonney in November 2007 and March 2008. Samples were collected through holes melted over the central (and deepest) portion of each lake basin—sites used for long-term data collection by the MCM LTER program (Table 1). Lake Fryxell has a maximum depth of about 18.5 m and a surface area of 7.8 km^2. Lake Bonney has two distinct basins referred to as the east (ELB; surface area ~3.3 km^2) and west (WLB; surface area ~1.0 km^2) lobes; our study focused on WLB. The lobes (both ~41.5-m deep) are separated by a sill about 14-m deep that minimizes exchange of waters between the basins, leading to different physical, chemical, and biological characteristics within each lobe. Both lakes Fryxell and WLB have permanent 3–6-m-thick ice covers and are characterized by strong chemoclines and oxyclines at about 9 m in Lake Fryxell and about 15 m in Lake WLB at the time of our study (Vick and Priscu, 2012) (Appendix Table A1). Deep maxima of phytoplankton and bacterioplankton productivity occur in the oxygenated water just above the chemoclines within each lake; water below the chemoclines is suboxic. Samples were collected at 6 m (MCM1 and MCM7) and 9 m (MCM2 and MCM8) in Lake Fryxell and 13 m (MCM3 and MCM5) and 18 m (MCM4 and MCM6) in Lake WLB to capture the microbiology of distinct vertical physicochemical and biogeochemical layers within the trophogenic zone (region where adequate light exists to support phytoplankton photosynthesis) of the two lakes. The November samples were collected when 24 h of sunlight was present and photosynthetically active radiation (PAR) extended to the depth of

the chemocline. No measurable PAR was present in the water column of the lakes during the March sampling. Corresponding environmental data were collected by the MCM LTER program (LTER, 2014). Characteristic physicochemical parameters corresponding to the samples analyzed in this study are summarized in Table 1.

Sampling, 454 next-generation sequencing, and data processing were carried out according to Amaral-Zettler *et al.* (2009) and employed 6% operational taxonomic unit (OTU) cluster widths. Although use of the V9 region to estimate ciliate diversity may cause inflation in richness estimates (Dunthorn *et al.*, 2012), by using a more conservative OTU width of 6% (94% similarity), we recovered numbers of OTUs similar to the numbers of species that had been previously reported in the lakes on the basis of morphological criteria. All sequence data are MIMARKS-compliant (Yilmaz *et al.*, 2011) and have been deposited in the National Center for Biotechnology Information Sequence Read Archives under the accession number SRP028879.

Taxonomic assignment

We employed a combination of reference databases including the Silva-ARB vers. 111 and Protist Ribosomal Reference (PRR) (based on GenBank May 2012 release) databases (Guillou *et al.*, 2013) to assign ciliate taxonomy to our amplicon reads. First, we aligned representative V9 amplicons of all MCM eukaryotic V9 amplicons to a reference database (RefSSU ver. 111 Silva-ARB) of full-length small subunit (*SSU*) *rRNA* genes of known taxonomy. We added our sequences to the reference tree in ARB (Pruesse *et al.*, 2007) using the "quick add taxa to tree" parsimony option and extracted taxa that branched among the ciliate clade. We then classified these ciliate OTUs according to the taxonomy of the most closely related sequences in the ARB tree. In addition to an ARB tree-based taxonomic assignment, we also classified our ciliates using the Global Alignment for Sequence Taxonomy (GAST) method (Huse *et al.*, 2008) with updated reference taxonomies from Silva-ARB (ver. 111) and PRR databases. In most cases, taxonomic assignments based on the position in the ARB tree and GAST taxonomies from Silva-ARB and PRR were consistent with each other; however, about 10% were inconsistent, with some ciliates being classified into another phylum on the basis of ARB GAST and PRR assigned taxonomies. In these cases we made the final taxonomic assignments according to the ARB tree-based taxonomy.

Alpha and beta-diversity calculations

We calculated alpha diversity indices using full data matrices and resampled data matrices, wherein the number of reads per sample was made equal through random resampling (Appendix Table A2). We limited analyses to binary data matrices wherein data were transformed to presence/absence before calculating nonparametric (Chao2) richness estimates using the SPADE program (Chao and Shen, 2003), keeping replicate samples separate because variant rRNA copy numbers (Zhu *et al.*, 2005) can bias abundance-based richness estimation in eukaryotes (Appendix Table A3). For all beta diversity calculations, we pooled data from replicate water samples to create abundance and incidence-based matrices for resampled datasets. All beta diversity analyses were carried out with resampled datasets using Primer-E (Clarke and Gorley, 2006). Resampled datasets involved resampling the whole eukaryotic dataset down to 2811 reads per pooled sample and then removing all ciliate OTUs from the larger resampled matrix.

Permutational analysis of variance (PerMANOVA; Anderson, 2001) was used to examine the importance of 16 environmental factors in partitioning of beta diversity with the *adonis* function calculated in the R library *vegan* (Oksanen *et al.*, 2010). The *dsvdis* function in the R library *labdsv* was used to generate abundance and incidence-based distance matrices (Roberts, 2010).

Multivariate analyses

We employed Hierarchical Clustering and Nonmetric Multidimensional Scaling (NMDS) methods in PRIMER-E ver. 6 (Clarke and Gorley, 2006) to explore ciliate community structure in the two lakes. Similarity profiles analysis (SIMPROF) was used to test for multivariate structure (Clarke and Gorley, 2006). Draftsman Plots in PRIMER-E tested skewness of the environmental data corresponding to our samples. Those data that were considered non-normally distributed were log (X) transformed. Environmental parameters that had less than 90% correlation in Spearman rank resemblance matrices were chosen as available variables. The significance of rank correlations between environmental parameters and ciliate community structure was determined by the global BIOENV match permutation test in PRIMER-E.

Primer design

To validate and characterize MCM ciliate novelty, we designed specific primers targeting the V9 region of the *SSU rRNA* gene with the aim of retrieving longer flanking regions of the *SSU* and *LSU rRNA* genes for down-stream phylogenetic analyses of the novel ciliates (Fig. 2). The criteria we used for designing primers were as follows: we targeted regions that were between 18 and 20 base pairs (bp) in length, contained between 45% and 55% GC content, and had melting temperatures between 50 °C and 60 °C. As a proof of concept, we chose to design primers against an abundant OTU with more than 10% GAST distance based on our reference V9 database. In this case the putative

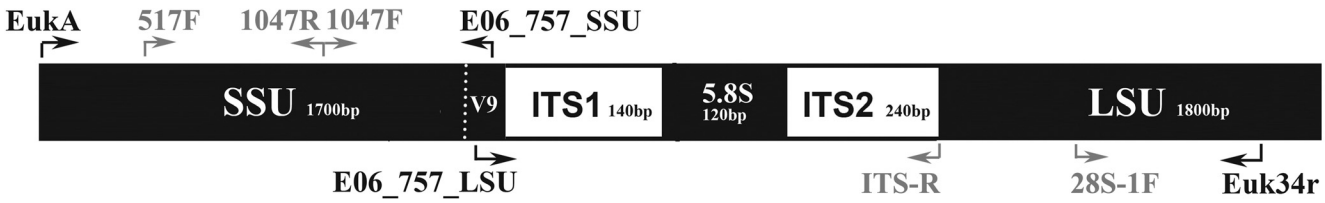

Figure 2. Ciliate-specific primers targeting the V9 region and eukaryotic universal primers used to amplify *SSU* and *LSU rRNA* genes. The approximate length of each region is indicated.

"novel" ciliate came from cluster ID: E06_757 from replicate samples of MCM2 collected at 9 m in Lake Fryxell on 22 November 2007. We used the primer match module in the ARB software package (Ludwig *et al.*, 2004) and the Silva-ARB database ver. 111 (Pruesse *et al.*, 2007) to test the specificity of the primers *in silico*. We also designed primers that would target the 5′ and 3′ ends of the V9 region in order to validate that the V9 region targeted was in fact what we amplified subsequently using a combination of the specific V9 primers in combination with *SSU* and *LSU*-general primers. Our V9-specific primer sequences and locations on the V9 region are shown in Table 2.

PCR amplification and DNA sequencing

Using the V9 region as an anchoring point, we amplified the *ITS1-5.8S-ITS2* and a partial fragment of the *LSU rRNA* gene using the V9-specific forward primer E06_757_LSU (Table 2) and the eukaryotic *LSU rRNA* gene reverse primer Euk34r (5′-GCATCGCCAGTTCTGCTTACC-3′) (Liu *et al.*, 2009). The eukaryotic universal A primer (5′-AAC CTGGTTGATCCTGCCAGT-3′)—targeting the 5′ end of the *SSU rRNA* gene—and V9-specific reverse primer E06_757_SSU (Table 2) amplified the remaining portion of the *SSU rRNA* gene (Medlin *et al.*, 1988). PCR conditions were as follows: initial denaturation at 94 °C for 30 s followed by 30 cycles of denaturation at 94 °C for 15 s, primer annealing at 52 °C for 30 s, and primer extension at 68 °C for 1 min, with a final extension at 68 °C for 5 min. This PCR amplification was performed on replicate samples of MCM2, MCM5, and MCM8; results of next-generation sequencing are shown in Table 3.

We directly sequenced amplicons on an ABI 3730XL (Applied Biosystems, Foster City, CA) capillary sequencer using the BigDye protocol. As the amplicons were around 1600-bp long, we used internal primers 517F (5′-GCCTA AAGCATCCGTAGC-3′), 1047F (5′-GGWGSTGCATG GCYG-3′), and 1047R (5′-CRGCCATGCASCWCC-3′) to sequence the *SSU rRNA* gene, and internal primers ITS-R (5′-CTGATATGCTTAAGTTCAGCGG-3′) (Shang, 2004) and 28S-1F (5′-ACCCGCTGAACTTAAGCAT-3′) (Moreira *et al.*, 2007) to sequence the *LSU rRNA* gene. We performed all sequencing at the Marine Biological Laboratory W. M. Keck Ecological and Evolutionary Genetics Facility following routine protocols. Partial double-stranded sequence reads were obtained to cover the *SSU* and *LSU rRNA* gene targeted regions, and all sequence reads were assembled in Geneious ver. 6.0.5 by Biomatters (Geneious, 2013). All *SSU* and *LSU rRNA* gene sequences of the putative "novel" ciliates were deposited in GenBank under accession numbers KF541342–KF541345.

Phylogenetic analyses

We compared our novel ciliate sequences against existing sequences of *SSU* and *LSU rRNA* gene sequences available in GenBank using BLAST (Altschul *et al.*, 1990) and included sequences of closest relatives for phylogenetic reconstructions. Additional representative ciliate sequences chosen to represent major ciliate clades were also included in our analyses and are identified by GenBank accession numbers after taxon names in Figures 3 and 4. Version 5.19 of the automated SILVA IncremeNtal Aligner (SINA) (Pruesse *et al.*, 2007) guided alignments of partial *LSU* and *SSU rRNA* gene sequences. We then created an automatic filter in ARB with minimal similarity of 50% to mask the

Table 2

Specific primers targeting the novel ciliate SSU rRNA *gene V9 hypervariable region for amplifying* SSU *and* LSU rRNA *genes*

Primer Name	Amplification Region	Primer Sequence (5′–3′)	Length (bp)	Location on V9 Region* (bp)
E06_757_SSU	*SSU rRNA*	AACTTCCCGCAATAGCTAGG	20	50
E06_757_LSU	*ITS* and *LSU rRNA*	TCTGGACTGCGTAGCCTA	18	36

* Location on V9 region (bp) was counted from the 5′ end of the V9.

Table 3

Relative percent recoveries of ciliate operational taxonomic units from MCM samples with average Global Alignment for Sequence Taxonomy (GAST) distance and best taxonomic assignment

			Sample ID*							
Cluster ID	GAST Distance	Taxonomy	MCM1	MCM2	MCM7	MCM8	MCM3	MCM4	MCM5	MCM6
E06_1057	0	*Blepharisma*	8.02%	0%	0.39%	0%	0%	3.18%	0%	0%
U_18433	NA	uncultured litostomatea	0%	0%	0%	0%	0%	1.59%	0%	1.61%
E06_217	0	Mesodiniidae	0%	0%	0.03%	0%	0%	0%	0%	0%
E06_96	0.009	*Didinium*	37.74%	41.38%	32.99%	42.18%	14.88%	20.63%	8.57%	22.58%
E06_630	0	*Haptoria*	0.24%	1.27%	0%	0.56%	2.09%	0%	0%	0%
E06_3357	0.056	*Haptoria*	0%	0%	0%	0%	0%	1.59%	0%	3.23%
E06_2422	0	*Didinium*	0%	0%	0%	0%	0%	1.59%	0%	1.61%
E06_157	0.047	*Balantidion*	0%	0%	0%	0%	0%	0%	0%	1.61%
E06_2139	0.008	*Obertrumia*	0%	0%	0%	0%	0%	0%	0%	1.61%
E06_178	0.017	Scuticociliatia	0%	0%	0%	2.14%	0%	0%	0%	0%
E06_43	0	Scuticociliatia	0%	0%	0%	0%	0%	0%	0%	3.23%
E06_620	0.037	Peritrichia	1.89%	2.90%	2.46%	2.70%	2.35%	1.59%	0%	3.23%
E06_490	0	Peritrichia	0%	0%	0%	0%	0.26%	0%	0%	1.61%
E06_3727	0.074	Peniculia	0%	0%	0%	0%	0%	0%	0%	4.84%
E06_255	0.058	Peniculia	14.62%	3.45%	13.28%	6.13%	3.13%	0%	0%	8.07%
E06_709	0.066	*Apostomatia*	0.24%	2.54%	0.42%	2.96%	0%	12.70%	0%	9.68%
U_956	NA	*Suctoria*	0.24%	0%	0%	0%	0%	0%	0%	0%
E06_250	0.065	*Cryptocaryon*	7.31%	3.45%	2.44%	16.46%	0%	17.46%	2.86%	9.68%
E06_516	0.041	*Cryptocaryon*	5.66%	6.35%	1.75%	2.01%	28.20%	3.18%	8.57%	1.61%
E06_392	0.041	*Cryptocaryon*	1.65%	0.36%	4.48%	1.34%	5.74%	6.35%	5.71%	9.68%
E06_757	0.107	*Cryptocaryon*	0%	24.68%	0%	0.60%	0%	0%	2.86%	0%
E06_605	0	*Cryptocaryon*	0%	0%	0.08%	0%	0%	0%	0%	3.23%
E06_7317	0.081	*Cryptocaryon*	0%	0%	0%	0%	0.26%	0%	0%	0%
U_28539	0.068	*Cryptocaryon*	0%	0%	0%	0.02%	0%	0%	0%	0%
U_26011	0.05	*Cryptocaryon*	0%	0%	0%	0%	0.26%	0%	0%	0%
U_14612	0.058	*Cryptocaryon*	0.24%	0%	0%	0%	0%	0%	0%	0%
E06_6146	0.073	*Cryptocaryon*	0%	0%	0%	0.02%	0%	0%	0%	0%
E06_12	0	uncultured oligotrichia	20.05%	9.44%	34.38%	8.18%	0.78%	7.94%	0%	3.23%
E06_6945	0.048	uncultured oligotrichia	0%	0%	0%	0.02%	0%	0%	0%	0%
E06_6841	0.04	uncultured oligotrichia	0%	0%	0%	0.02%	0%	0%	0%	0%
E06_105	0	*Halteria*	2.12%	4.17%	7.23%	14.67%	26.37%	9.52%	65.71%	4.84%
E06_140	0	*Oxytricha*	0%	0%	0.08%	0%	0.78%	3.18%	0%	1.61%
U_757	NA	Spirotrichea	0%	0%	0%	0%	0%	0%	0%	3.23%
U_169	0	*Euplotes*	0%	0%	0%	0%	14.88%	6.35%	5.71%	0%
E06_180	0	uncultured choreotrichia	0%	0%	0%	0%	0%	3.18%	0%	0%
Total % of ciliates recovered from the full eukaryotic dataset	5.48	7.77	29.70	23.06	3.02	0.49	0.35	0.67		

* Refer to Table 1 for sample identities.

aligned sequences. To create this filter we excluded alignment positions that were predominantly gap characters or ambiguities and assigned equal weight to lowercase and uppercase characters. The datasets used for the primary phylogenetic analyses included 1341 positions and 59 taxa for the *LSU rRNA* gene phylogenies and 1618 positions and 46 taxa for the *SSU rRNA* gene phylogenies. We performed Bayesian (BI) analysis with MrBayes 3.2.1 on XSEDE online on the CIPRES Portal ver. 2.0 (CIPRES, 2014). The chain length for our analysis was 4,000,000 generations with trees sampled every 100 generations and the first 25% discarded as burn-in. Maximum-likelihood (ML) bootstrapping analysis was carried out online with 1000 replicates using RAxML-HPC2 on XSEDE (7.3.2) on the CIPRES

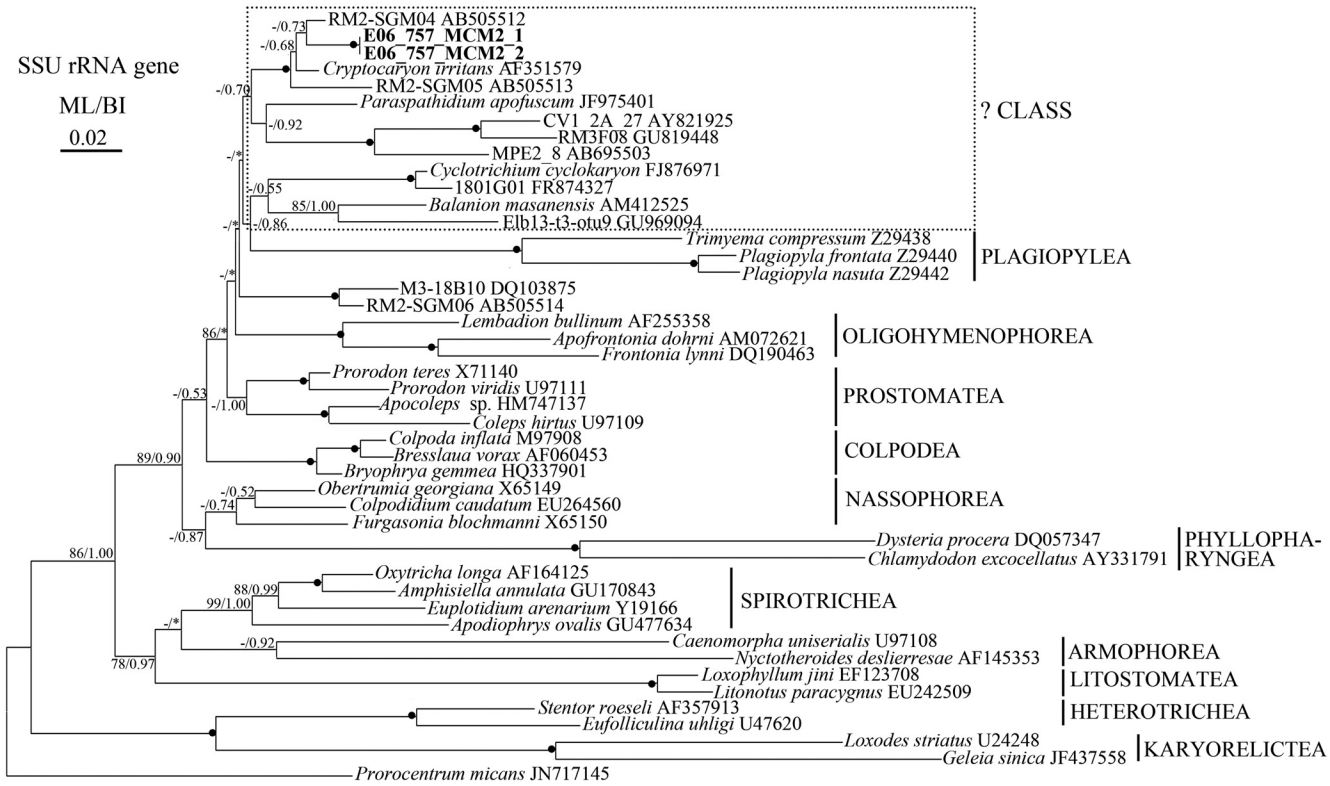

Figure 3. Maximum likelihood (ML) tree inferred from the *SSU rRNA* gene sequences showing the position of the novel species (in bold). Numbers at the nodes represent bootstrap values from ML and posterior probabilities from Bayesian analysis (BI). Black circles indicate full support in both analyses. Clades that differed in topology in the BI tree are indicated with an asterisk (*). Hyphens (-) represent the bootstrap values in the ML tree below 50%. All branches are drawn to scale. The scale bar corresponds to two substitutions per 100 nucleotide positions.

Science Gateway (CIPRES, 2014; Stamatakis *et al.*, 2008). The GTR+I+G model (Zwickl and Holder, 2004) was chosen as the best model by MEGA 5.2.2 (Tamura *et al.*, 2011). *Prorocentrum micans* (Dinophyta) and *Sarcocystis rileyi* (Apicomplexa, Sarcocystidae) were selected as outgroup taxa in the *SSU rRNA* and *LSU rRNA* trees, respectively.

Results and Discussion

Ciliate richness in McMurdo Dry Valleys lakes

The percentage of ciliate sequences of the total eukaryotic community in Lake Fryxell was 19.7% (6.6% in autumn; 25.7% in summer), while in Lake West Lobe Bonney (WLB) it was 1.2% (1.7% in autumn; 0.5% in summer). Roberts *et al*. (2004) also found that ciliates contributed to a relatively higher percentage of the overall eukaryotic community in Lake Fryxell than in Lake WLB. When contrasted with the recovery of ciliate operational taxonomic units (OTUs) in an Arctic study, which varied from 10% to 30% in lacustrine waters (Crump *et al.*, 2012), the concentration of ciliates we measured in Lake WLB is relatively low. We detected 35 different ciliate OTUs collectively in our eight MCM lake samples. Among the 35 OTUs we detected, 14 (40%) were found in both lakes, while eight (23%) and 13 (37%) occurred only in Lake Fryxell and Lake WLB, respectively. The ciliate sequences were phylogenetically diverse and represented six known classes (Heterotrichea, Litostomatea, Nassophorea, Oligohymenophorea, Phyllopharyngea, Spirotrichea) and one unresolved class tentatively assigned to the Prostomatea (see following section *Phylogenetic position of the novel McMurdo Dry Valleys species*) (Fig. 5). According to previous studies (Kepner *et al.*, 1999; Laybourn-Parry and Pearce, 2007; Roberts *et al.*, 2004), 34 ciliate species belonging to eight classes have been reported in Lake Fryxell and Lake WLB. The only class reported previously (Roberts *et al.*, 2004) not found in our study was Colpodea, which could be due to differences in sampling depths together with stratification of different ciliate species with depth (Kepner *et al.*, 1999). The colpodid species was previously reported at 0.3–0.5 m below the surface of the ice, whereas our samples were collected below 6 m (Roberts *et al.*, 2004).

Table 3 summarizes the Global Alignment for Sequence

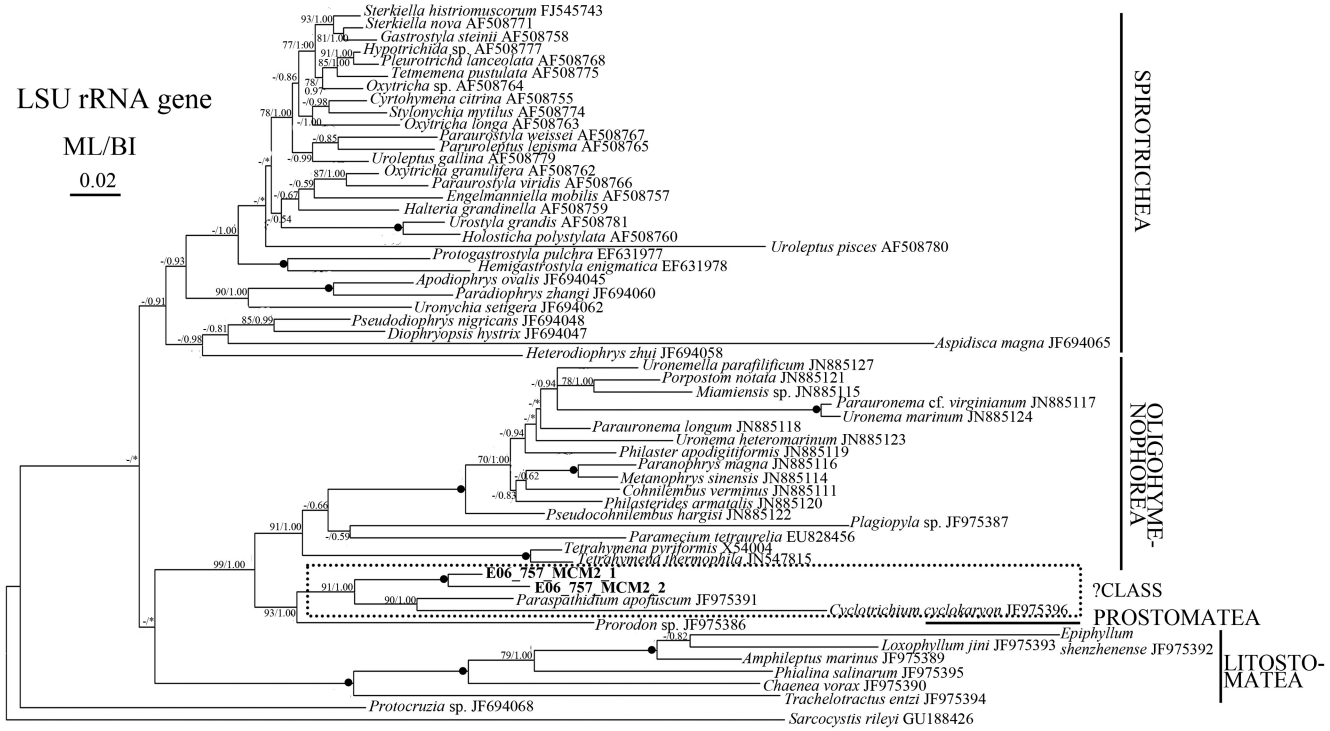

Figure 4. Maximum likelihood (ML) tree inferred from the *LSU rRNA* gene sequences showing the phylogenetic position of the novel species (in bold). Numbers at the nodes represent bootstrap values from ML and posterior probabilities from Bayesian analysis (BI). Black circles indicate full support in both analyses. Clades that differed in topology in the BI tree are indicated with an asterisk (*). Hyphens (-) represent the bootstrap values in the ML tree below 50%. All branches are drawn to scale. The scale bar corresponds to two substitutions per 100 nucleotide positions.

Taxonomy (GAST) values for all of the ciliate OTUs. Twelve OTUs identified as *Blepharisma*, *Mesodinium*, *Didinium*, *Haptoria*, uncultured Scuticociliatia, uncultured Peritrichia, *Cryptocaryon*, *Halteria*, uncultured Oligotrichia, *Oxytricha*, *Euplotes,* and uncultured Choreotrichia had a zero GAST distance, indicating that they were identical to existing sequences in public databases. In those cases, the assigned taxonomy at the genus level was most reliable. Certain species belonging to these groups are planktonic and common in both marine and freshwater habitats. We detected two genera (*Obertrumia* and *Oxytricha*) not previously reported in these lakes.

Alpha diversity of ciliates did not vary systematically across the different lakes nor did it vary significantly by seasons for both full and resampled data sets (Fig. 6). Not surprisingly, the overall diversity was relatively low in both lakes given their ultra-oligotrophic to oligotrophic nutrient status (Laybourn-Parry and Pearce, 2007). As shown in Figure 6, estimated and observed Chao2 richness values for the given lake samples were very similar, indicating that we likely saturated ciliate diversity using our approach. This is to be expected given the low ciliate species richness combined with the application of next-generation sequencing approaches in this study.

Extreme environmental conditions are often associated with overall low species richness (Petz, 2003; Petz *et al.*, 2007). However, our results did not support this assumption. For example, in MCM4 and MCM6 samples (*i.e.,* 18-m Lake WLB) where the salinity was higher than 80 PSU (~2× seawater) and oxygen was low (Table 1), Chao2 estimated and observed richness were both similar to or slightly higher than the less extreme samples (Fig. 6). Furthermore, we found 12 OTUs (E06_1057, E06_96, E06_255, E06_620, E06_709, E06_250, E06_392, E06_516, E06_605, E06_105, E06_12, and E06_140) that occurred in MCM4 and MCM6, as well as in freshwater samples, *i.e.* MCM1, MCM2, MCM3, and MCM4 (Table 3), leading us to conclude that these species may tolerate a range of salinities (Laybourn-Parry and Pearce, 2007). Meanwhile, we also found nine OTUs that occurred only in MCM4 or MCM6, which may suggest there are some salt-tolerant and microaerophilic ciliates that are restricted to these conditions (Bell and Laybourn-Parry, 1999; Yasindi *et al.*, 2002). Our findings indicate that some ciliates either tolerate extreme conditions (high salinity and low oxygen) or are restricted to these conditions. In addition, dissolved organic carbon (DOC) concentration was high in MCM4 and MCM6, which may explain why the richness of ciliates

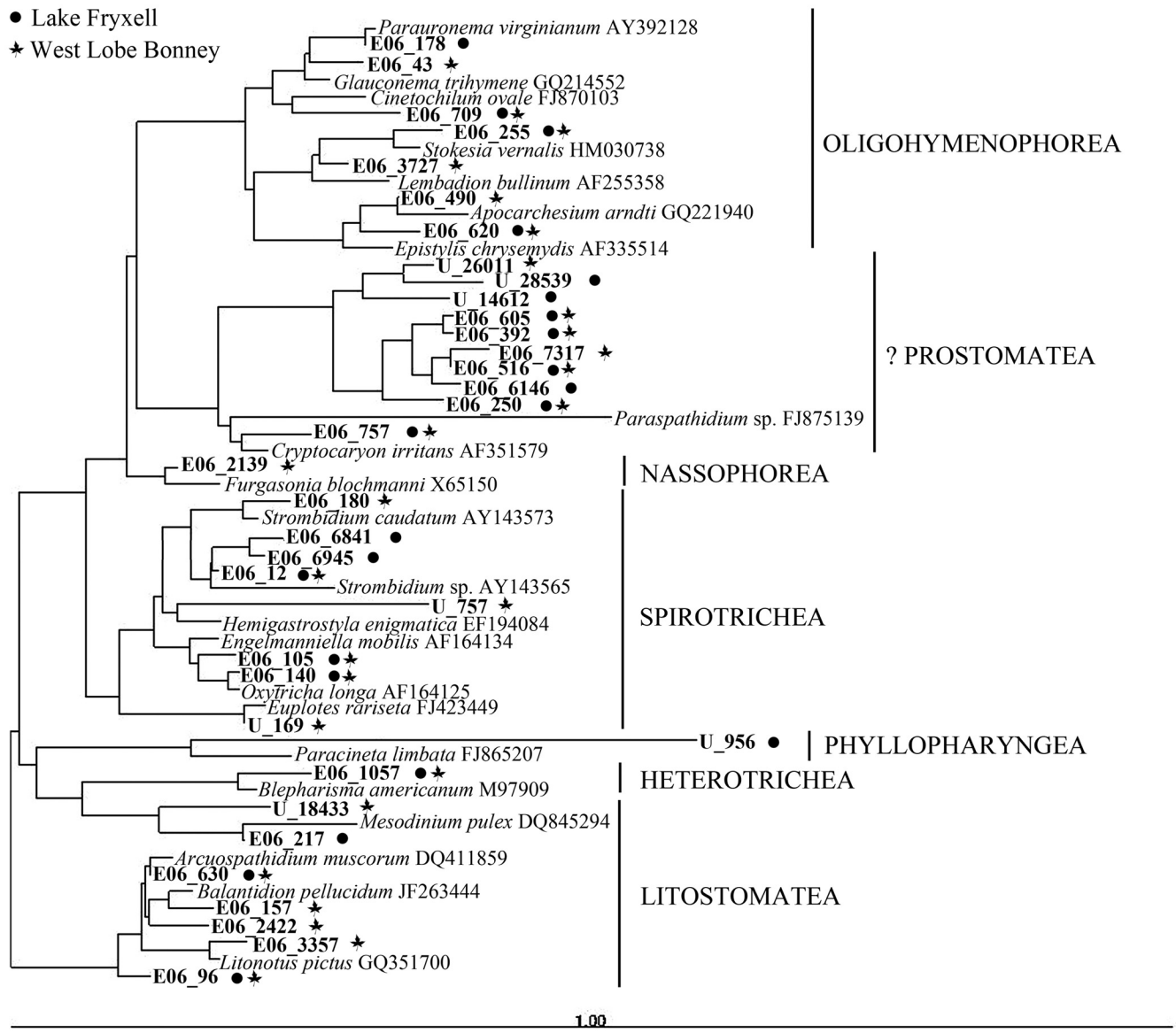

Figure 5. McMurdo Dry Valleys ciliate V9 representative operational taxonomic units found in Lake Fryxell and Lake West Lobe Bonny (in bold) placed within the Silva-ARB ver. 111 *SSU rRNA* gene reference tree. Major ciliate classes are indicated on the tree with capital letters. The question mark (?) points out that the clade including *Paraspathidium* and *Cyclotrichium* used to belong to Prostomatea but now may represent a new class, as suggested by Zhang *et al*. (2012).

in these samples was not as low as in other samples, as they might be able to use some of the dissolved organics to supplement predation.

Given that most ciliates feed on bacteria or organic material, DOC and dissolved organic nitrogen (DON) levels, as well as bacterial concentrations, should influence ciliate richness (Wickham *et al.*, 2004). The MCM lakes are an ideal ecosystem in which to test this hypothesis because nutrient concentrations are typically highly stratified in the water column. To test this hypothesis, we compared ciliate richness between samples with different nutrient concentrations. MCM4 and MCM6 had the highest DOC and DON concentrations (Table 1), which may explain why both estimated and observed richness in those samples were not lower than in other samples, taking confidence bounds into account (Fig. 6). Another possibility is that considering the apparently higher diversity of ciliates and of eukaryotes overall (Vick-Majors *et al.*, 2013) in these same samples (MCM4 and MCM6), the increase in diversity may not be due to trophic activity. MCM4 and MCM6 were sampled at the bottom of the halocline in Lake WLB (Appendix Table A1). Thus, dead eukaryotic cells may also artificially inflate

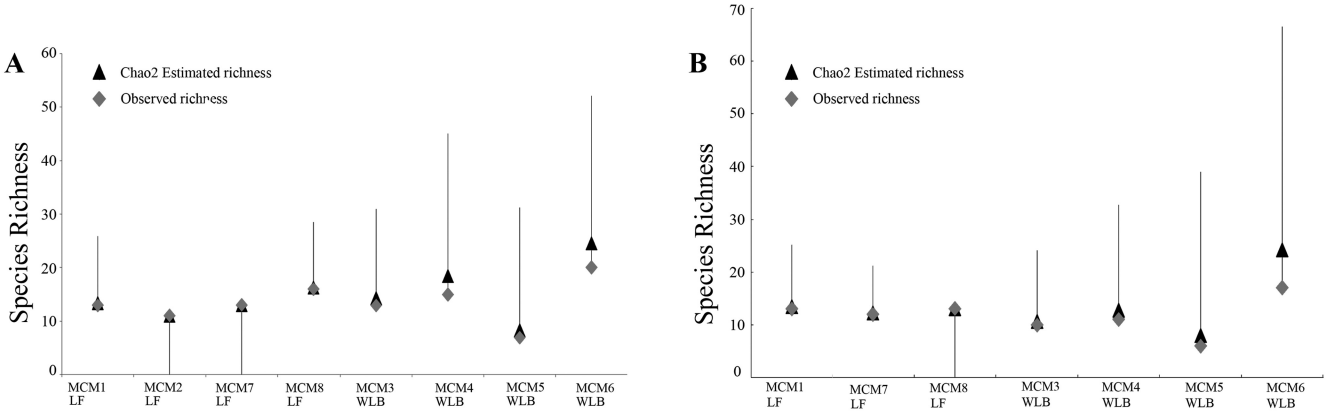

Figure 6. Estimated and observed ciliate species richness for full (A) and resampled (B) datasets with Bonferroni corrected 95% confidence intervals. LF, Lake Fryxell; WLB, West Lobe Bonney. For sample names and metadata refer to Table 1.

richness values by sinking through the fresher upper layers and collecting at around 18 m because they cannot sink through the layer of increased salinity density.

Day length, which varies as a function of season in Antarctica, has been reported to have a strong influence on bacterioplankton and some mixotrophic protist communities (Lizotte *et al.,* 1996; Ghiglione and Murray, 2012; Vick-Majors *et al.*, 2013), as well as on expression of functional genes (Kong *et al*., 2012, 2014), leading us to hypothesize that ciliate richness in samples collected at the same depth in the same lake in autumn should be higher than in summer. However, seasonally separated samples collected in Lake Fryxell had similar richness, while in Lake WLB, richness at 13 m in autumn was 39% lower than in summer but 25% higher at 18 m (Fig. 6). In another study, overall eukaryotic diversity in Lake Fryxell was generally higher in autumn than in summer, while the opposite trend existed in Lake WLB (Vick-Majors *et al.*, 2013). Therefore, the similar richness in Lake Fryxell and inconsistent trend in Lake WLB shows that the richness of ciliates may not be attributed to day length. The richness of ciliates seemed, on the basis of our results, to be more strongly influenced by nutrient concentrations that did not correlate with day length. Similarly, expression of two major photosynthetic genes in Lake Bonney populations of photosynthetic protists during the transition from Antarctic summer to winter (February to April) also exhibited seasonal trends that were thought to be influenced by nutrient availability (Kong *et al*., 2012, 2014).

Community structure

Variation in *rRNA* gene copy numbers among different ciliate groups may confound analyses of ecological diversity studies (Gong *et al.*, 2013). To address this concern, we employed incidence-based (presence/absence) beta diversity measures in our analyses. However, the use of incidence-based approaches and resampled datasets led to a loss of signal in our data (Fig. 7). We tested the statistical significance of the influence of salinity and depth on ciliate communities, which showed patterns similar to those of oxygen status (data not shown). From the Hierarchical Clustering and Nonmetric Multidimensional Scaling (NMDS) for the resampled dataset using Bray-Curtis similarity (Fig. 7A), samples (MCM1, MCM2, MCM7, MCM8) collected from the same depth from Lake Fryxell (median oxygen, low salinity) clustered together regardless of season, followed by samples (MCM3 and MCM5; MCM4 and MCM6) from Lake WLB from different oxygen levels and salinities. From the resampled data of the Jaccard similarity matrix (Fig. 7C), only the samples from 18 m in Lake WLB (MCM4 and MCM6) with high salinity and low oxygen were separated, while other samples clustered together.

The NMDS diagram employing abundance metrics (Fig. 7B) showed that communities in Lake Fryxell shared more than 75% similarity with each other and clustered separately from Lake WLB communities, indicating that the two ciliate communities have few species in common (Laybourn-Parry and Pearce, 2007). Our study supports previous reports that depth, oxygen, and salinity in the lakes all affect the structure of ciliate communities (Kepner *et al.*, 1999) and that ciliate communities are typically stratified in these lakes due to gradients in the stable water column. Ciliate communities failed to cluster by season regardless of the approach or dataset used. The samples collected at the same depth in the same lake, and therefore with the same oxygen status, but in different seasons (summer and autumn) clustered together according to the similarity profiles (SIMPROF) analysis, indicating that seasonal differences (summer *vs.* autumn) did not have a strong influence on ciliate community structure. Even though the SIMPROF

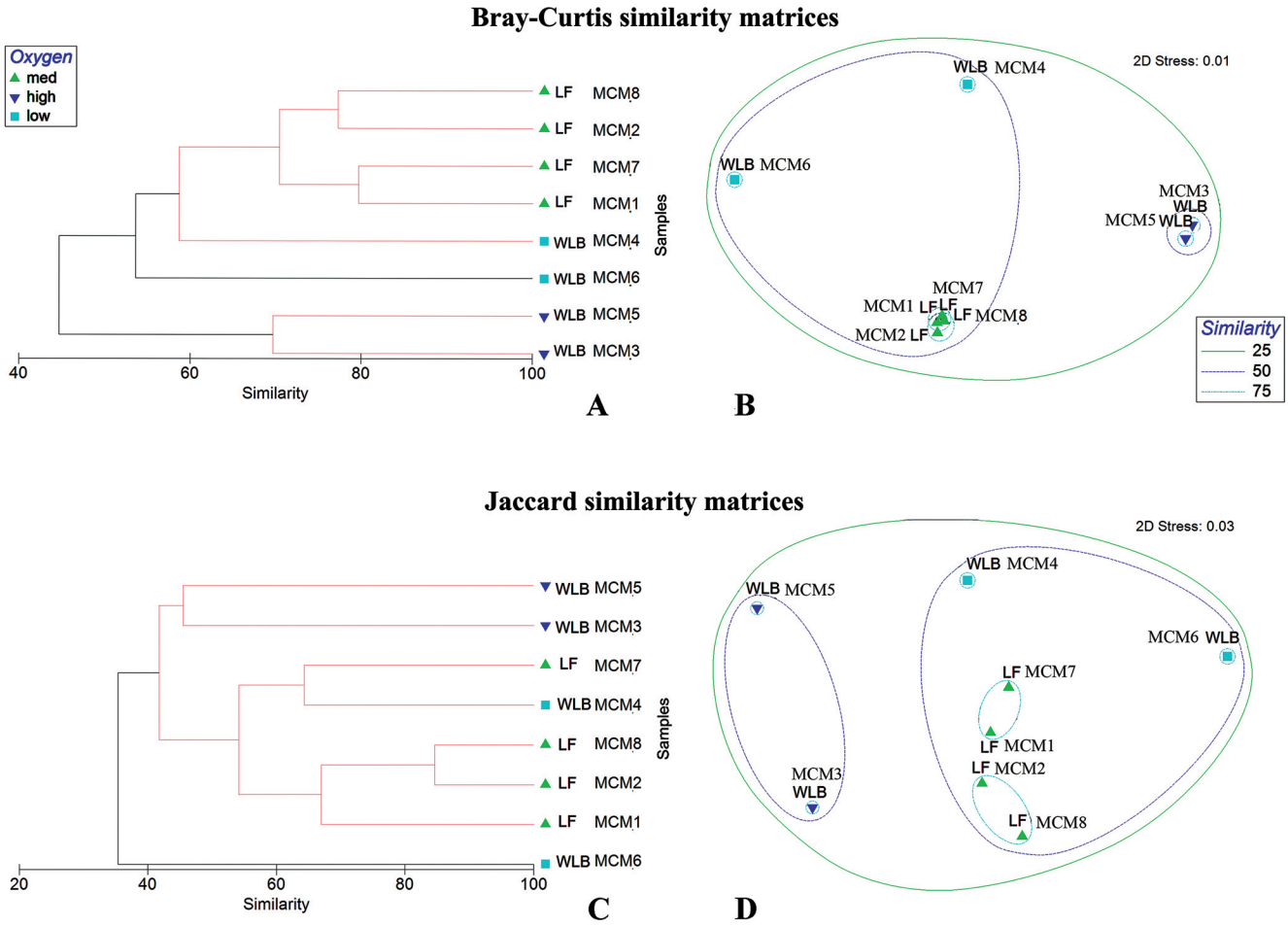

Figure 7. Hierarchical Clustering and Nonmetric Multidimensional Scaling ordination diagrams for Bray-Curtis (A, B) or Jaccard (C, D) similarity matrices. LF, Lake Fryxell; WLB, Lake West Lobe Bonney. The red lines in the cluster diagrams indicate groupings with no statistical evidence for substructure.

model first accounted for the interaction in the ciliate community between lake and depth, season was not significant (in the permutational analysis of variance [PerMANOVA], $P < 0.1$ in resampled data of Bray-Curtis and Jaccard similarity matrices).

In contrast to the trends in ciliate data from our study, the total eukaryotic community analyzed by Vick-Majors *et al.* (2013) from the same lake and depth was different between seasons, mostly due to mixotrophic stramenopiles such as chrysophytes becoming abundant in autumn. Species belonging to the following ciliate groups have been reported in the literature as mixotrophs (Dolan, 1992; Bernard and Rassoulzadegan, 1994): *Strombidium*, *Tontonia*, *Laboea*, *Lohmaniella*, *Paramecium*, and *Mesodinium rubrum*. In our taxonomic assignments, four OTUs (E06_180, E06_6841, E06_6945, E06_12) were identified as *Strombidium* and one (E06_217) as *Mesodinium* (GAST distance = 0). These OTUs did not show seasonal patterns (Table 3). The ability to switch between mixotrophic and heterotrophic life styles may be one reason why seasons do not have strong influence on ciliate populations in these lakes (Vick-Majors *et al.*, 2013).

On the basis of the Draftsman plots and BIOENV analyses, we chose 15 environmental variables (salinity, bacterial cells, chlorophyll, DOC, DON, dissolved oxygen, particulate organic nitrogen, PO_4^{3-}, SiO_2, temperature, total dissolved nitrogen, fluorescence, tritiated leucine incorporation, tritiated thymidine incorporation, and density) for downstream BIOENV analyses. The other seven variables—depth, NH_4^+, atmospheric pressure, conductivity, pH, NO_3^-, and NO_2^-—were excluded because they were correlated with the chosen variables. Although we found dissolved oxygen to be the best explanatory environmental variable for ciliate community structure using the BIOENV global test for resampled data based on both abundance and incidence-based matrices, the P values showed that the result was not significant in either case ($R = 0.706$, $P = 0.09$ for the abundance matrix; $R = 0.652$, $P = 0.13$ for the incidence-based matrix). We think dissolved

oxygen likely drives the variations in ciliate community structure. It is well known that microbial community structure is sensitive to oxygen levels (Laybourn-Parry and Pearce, 2007), and some ciliate groups (including *Plagiopyla, Sonderia*, *Parablepharisma, Metopus, Caenomorpha,* and *Odontostomatida)* have been reported as obligate anaerobes, while others (*e.g., Remanella* and *Loxodes*) are microaerophiles (Finlay and Fenchel, 1989; Fenchel and Finlay, 1991). Therefore, the vertical distribution of ciliates may be tightly coupled to oxygen concentrations at different depths (Berninger and Epstein, 1995). For the resampled data using Jaccard similarity (Fig. 7C, D), dissolved oxygen and particulate organic nitrogen were most strongly correlated. These relationships may be related to ciliate trophic status, that is, mainly feeding on bacteria and microalgae.

For the abundance matrices of resampled data, the PerMANOVA test indicated that depth, dissolved oxygen, and salinity had significant influences on the ciliate community ($P = 0.026$ for depth, $P = 0.015$ for dissolved oxygen, and $P = 0.05$ for salinity), while lake and season were not significant ($P > 0.1$). For the incidence-based matrices of resampled data, depth, dissolved oxygen, salinity, lake, and season were not significant ($P > 0.1$).

Phylogenetic position of the novel McMurdo Dry Valleys species

One of the goals of the Microbial Inventory Research Across Diverse Aquatic Long Term Ecological Research Sites project was to discover novel species by using a next-generation approach based on amplicon sequencing of short hypervariable V9 regions. To accomplish this goal, we selected and pursued further sequencing of the putative novel species (Cluster ID: E06_757) from the MCM2 Lake Fryxell replicate samples (24.68% of all ciliates in MCM2). This novel OTU was also found in the MCM5 Lake WLB and MCM8 Lake Fryxell samples, where it was recovered with lower frequency (2.86% of all ciliates in MCM5, 0.60% of all ciliates in MCM8). Furthermore, by applying a 10% GAST distance cut-off criterion we were able to validate the discovery of a novel taxon.

The E06_757_MCM2_1 and E06_757_MCM2_2 *SSU rRNA* genes are 1666 and 1665 base pairs long, respectively. Phylogenetic trees inferred from *SSU rRNA* gene sequence data using two methods—maximum-likelihood (ML) and Bayesian (BI))—generated similar topologies; therefore, we show only the ML tree in Figure 3. The E06_757_MCM2_1 and E06_757_MCM2_2 sequences branched together in both ML and BI analyses in a clade including E06_757_MCM2_1 and E06_757_MCM2_2, *Cryptocaryon irritans,* and two environmental sequences with 100% bootstrap support and 1.00 BI posterior probabilities. The environmental ciliate sequence (Elb13-t3-otu8) reported in East Lobe Bonney (ELB) (Bielewicz *et al.*, 2011) clustered with *Balanion masanensis* with moderate support (85% in ML, 1.00 in BI) and then formed a sister branch with *Cyclotrichium cyclokaryon* with low bootstrap values (18% in ML, 0.55 in BI). Our tree provides a more accurate placement of this environmental isolate first reported in Bielewicz *et al.* (2011) wherein this environmental sequence clustered with *Cryptocaryon irritans*.

We also sequenced a 1672- and 1671-bp-long section of the *ITS* regions and *LSU rRNA* gene of the E06_757_MCM2_1 and E06_757_MCM2_2 ciliates, respectively. Phylogenetic trees inferred from *LSU rRNA* gene sequence data using ML and BI generated similar topologies, so we present only the ML tree in Figure 4. In both the ML and BI trees, E06_757_MCM2_1 and E06_757_MCM2_2 clustered together with full support and further grouped with *Cyclotrichium cyclokaryon* and *Paraspathidium apofuscum* with strong support (91% in ML, 1.00 in BI), forming a sister branch to the Prostomatea.

Our phylogenetic results indicate that the novel species is related to *Cryptocaryon irritans. Cryptocaryon irritans* is an alveolate parasite of teleost marine fish. Cultured from temperate and tropical seas, it causes disease at temperatures above 15 °C (Wright and Colorni, 2002). Since no fish have been reported in lakes Fryxell or WLB and the genus *Cryptocaryon* includes only one species, we propose that our novel species belongs to a new genus that is free-living and related to *Cryptocaryon*.

According to Takishita *et al.* (2010), the environmental clones RM2-SGM04 and RM2-SGM05, which were collected from the deep-sea sediments (about 1177-m deep) of Sagami Bay, Japan, were the environmental species most closely related to our novel species and could not be assigned to any known ciliate class. Although *Cryptocaryon irritans* belongs to Prostomatea according to the classification of Adl *et al.* (2012), it did not cluster within the Prostomatea clade. Zhang *et al.* (2012) suggested that the clade including *Paraspathidium apofuscum* and *Cyclotrichium cyclokaryon* represents a new class. However, our *SSU rRNA* gene tree shows that this clade is paraphyletic. Therefore we assign this novel species to *incertae sedis* until additional relatives can provide it with a more robust phylogenetic placement.

Acknowledgments

Funding was provided by NSF DEB-0717390 to Linda Amaral-Zettler (MIRADA-LTERS); OPP-1115245, OPP-0838933, OPP-1027284, and OPP-0839075 to John C. Priscu; and OPP-0631659 and OPP-1056396 to Rachael Morgan-Kiss. We would also like to acknowledge the China Scholarship Council (No. [2012] 3013) for fellowship support to Yuan Xu enabling her to study at the Marine Biological Laboratory. The Montana Space Grant Consortium provided additional funding for Trista Vick-Majors. We

thank the McMurdo Microbial Observatory; Sukkyun Han, Chao Tang, Amy Chiuchiolo, and Marie Šabacká for assistance with sample collection; the 2007–2008 McMurdo LTER limnology team for assistance with environmental data collection; and Elizabeth McCliment for assistance with sequencing.

Literature Cited

Adl, S. M., A. G. Simpson, C. E. Lane, J. Lukes, D. Bass, S. S. Bowser, M. W. Brown, F. Burki, M. Dunthorn, V. Hampl, *et al*. 2012. The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* **59:** 429–493.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

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.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **26:** 32–46.

Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyer-Reil, and F. Thingstad. 1983. The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10:** 257–263.

Bell, E. M., and J. Laybourn-Parry. 1999. Annual plankton dynamics in an Antarctic saline lake. *Freshw. Biol.* **41:** 507–519.

Bernard, C., and F. Rassoulzadegan. 1994. Seasonal variations of mixotrophic ciliates in the northwest Mediterranean Sea. *Mar. Ecol. Prog. Ser.* **108:** 295–301.

Berninger, U., and S. Epstein. 1995. Vertical distribution of benthic ciliates in response to the oxygen concentration in an intertidal North Sea sediment. *Aquat. Microb. Ecol.* **9:** 229–236.

Bielewicz, S., E. Bell, W. Kong, I. Friedberg, J. C. Priscu, and R. M. Morgan-Kiss. 2011. Protist diversity in a permanently ice-covered Antarctic lake during the polar night transition. *ISME J.* **5:** 1559–1564.

Brambilla, E., H. Hippe, A. Hagelstein, B. J. Tindall, and E. Stackebrandt. 2001. 16S rDNA diversity of cultured and uncultured prokaryotes of a mat sample from Lake Fryxell, McMurdo Dry Valleys, Antarctica. *Extremophiles* **5:** 23–33.

Chao, A., and T. Shen. 2003. Program SPADE (Species Prediction and Diversity Estimation). Program and user's guide. [Online]. Available: http://chao.stat.nthu.edu.tw/blog/software-download/spade/ [2014, October 2].

CIPRES (Cyberinfrastructure for Phylogenetic Research). 2014. CIPRES Science Gateway. [Online}. Available: http://www.phylo.org [2014, October 2].

Clarke, K., and R. Gorley. 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth, United Kingdom.

Coppellotti, O., and P. Matarazzo. 2000. Ciliate colonization of artificial substrates in the Lagoon of Venice. *J. Mar. Biol. Assoc. UK* **80:** 419–427.

Crump, B. C., L. A. Amaral-Zettler, and G. W. Kling. 2012. Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J.* **6:** 1629–1639.

Dolan, J. 1992. Mixotrophy in ciliates: a review of *Chlorella* symbiosis and chloroplast retention. *Aquat. Microb. Ecol.* **6:** 115–132.

Dunthorn, M., J. Klier, J. Bunge, and T. Stoeck. 2012. Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. *J. Eukaryot. Microbiol.* **59:** 185–187.

Fenchel, T., and B. J. Finlay. 1991. The biology of free-living anaerobic ciliates. *Eur. J. Protistol.* **26:** 201–215.

Finlay, B. J., and G. F. Esteban. 1998. Freshwater protozoa: biodiversity and ecological function. *Biodivers. Conserv.* **7:** 1163–1186.

Finlay, B. J., and T. Fenchel. 1989. Hydrogenosomes in some anaerobic protozoa resemble mitochondria. *FEMS Microbiol. Lett.* **65:** 311–314.

Foissner, W. 2005. Two new "flagship" ciliates (Protozoa, Ciliophora) from Venezuela: *Sleighophrys pustulata* and *Luporinophrys micelae*. *Eur. J. Protistol.* **41:** 99–117.

Foissner, W., A. Chao, and L. Katz. 2009. Diversity and geographic distribution of ciliates (Protista: Ciliophora). Pp. 111–129 in *Protist Diversity and Geographical Distribution*, W. Foissner and D. Hawksworth, eds. Springer, New York.

Geneious. 2014. Bioinformatics software platform. [Online]. Available: http://www.geneious.com/ [2014, October 2].

Ghiglione, J. F., and A. E. Murray. 2012. Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environ. Microbiol.* **14:** 617–629.

Gilbert, J. A., P. Somerfield, B. Temperton, S. Huse, I. Joint, and D. Field. 2010. Day-length is central to maintaining consistent seasonal diversity in marine bacterioplankton. *Nature Precedings* [Online]. Available: http://hdl.handle.net/10101/npre.2010.4406.1 [2014, October 2].

Gong, J., J. Dong, X. Liu, and R. Massana. 2013. Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist* **164:** 369–379.

Guillou, L., D. Bachar, S. Audic, D. Bass, C. Berney, L. Bittner, C. Boutte, G. Burgaud, C. de Vargas, J. Decelle, *et al*. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41:** D597–D604.

Huse, S., L. Dethlefsen, J. Huber, D. Welch, D. Relman, and M. Sogin. 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* **4:** e1000255.

Jiang, J., S. Wu, and Y. Shen. 2007. Effects of seasonal succession and water pollution on the protozoan community structure in an eutrophic lake. *Chemosphere* **66:** 523–532.

Karr, E. A., J. M. Ng, S. M. Belchik, W. M. Sattley, M. T. Madigan, and L. A. Achenbach. 2006. Biodiversity of methanogenic and other archaea in the permanently frozen Lake Fryxell, Antarctica. *Appl. Environ. Microbiol.* **72:** 1663–1666.

Kepner, R. L., Jr., R. A. Wharton, Jr., and D. W. Coats. 1999. Ciliated protozoa of two Antarctic lakes: analysis by quantitative protargol staining and examination of artificial substrates. *Polar Biol.* **21:** 285–294.

Kong, W., D. C. Ream, J. C. Priscu, and R. M. Morgan-Kiss. 2012. Diversity and expression of RubisCO genes in a perennially ice-covered Antarctic lake during the polar night transition. *Appl. Environ. Microbiol.* **78:** 4358–4366.

Kong, W., W. Li, O. Prášil, I. Romancova, and R. M. Morgan-Kiss. 2014. An integrated study of photochemical function and expression of a key photochemical gene (*psbA*) in photosynthetic communities of Lake Bonney (McMurdo Dry Valleys, Antarctica). *FEMS Microbiol. Ecol.* (In press).

Laybourn-Parry, J., and D. A. Pearce. 2007. The biodiversity and ecology of Antarctic lakes: models for evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362:** 2273–2289.

Laybourn-Parry, J., M. R. James, D. M. McKnight, J. Priscu, S. A. Spaulding, and R. Shiel. 1996. The microbial plankton of Lake Fryxell, southern Victoria Land, Antarctica during the summers of 1992 and 1994. *Polar Biol.* **17:** 54–61.

Levy, J. S. 2012. How big are the McMurdo Dry Valleys? Estimating ice-free area using Landsat image data. *Antarct. Sci.* **25:** 119–120.

Liu, H., I. Probert, J. Uitz, H. Claustre, S. Aris-Brosou, M. Frada, F. Not, and C. de Vargas. 2009. Extreme diversity in noncalcifying

haptophytes explains a major pigment paradox in open oceans. *Proc. Natl. Acad. Sci. USA* **106:** 12803–12808.

Lizotte, M. P., T. R. Sharp, and J. C. Priscu. 1996. Phytoplankton dynamics in the stratified water column of Lake Bonney, Antarctica. I. Biomass and productivity during the winter-spring transition. *Polar Biol.* **16:** 155–162.

LTER (long term ecological research). **2014.** McMurdo Dry Valleys. [Online]. Available: http://www.mcmlter.org [2014, October 2].

Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, *et al*. 2004. ARB: a software environment for sequence data. *Nucleic Acids Res.* **32:** 1363–1371.

Medlin, L., H. J. Elwood, S. Stickel, and M. L. Sogin. 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* **71:** 491–499.

MIRADA LTER. 2008. Microbial inventory research across diverse aquatic long term ecological research sites. [Online]. Available: http://www.amarallab.mbl.edu/mirada/mirada.html [2014, October 2].

Moreira, D., S. von der Heyden, D. Bass, P. Lopez-Garcia, E. Chao, and T. Cavalier-Smith. 2007. Global eukaryote phylogeny: combined small- and large-subunit ribosomal DNA trees support monophyly of Rhizaria, Retaria and Excavata. *Mol. Phylogenet. Evol.* **44:** 255–266.

Oksanen, J., F. Blanchet, R. Kindt, P. Legendre, R. O'Hara, G. Simpson, P. Solymos, M. Stevens, and H. Wagner. 2010. vegan: Community Ecology Package. R package version 1.17-1, 2010. [Online]. Available: http://cran.r-project.org/web/packages/vegan/index.html [2014, October 2].

Petz, W. 2003. Ciliate biodiversity in Antarctic and Arctic freshwater habitats—a bipolar comparison. *Eur. J. Protistol.* **39:** 491–494.

Petz, W., A. Valbonesi, U. Schiftner, A. Quesada, and J. Cynan Ellis-Evans. 2007. Ciliate biogeography in Antarctic and Arctic freshwater ecosystems: endemism or global distribution of species? *FEMS Microbiol. Ecol.* **59:** 396–408.

Priscu, J. C., C. H. Fritsen, E. E. Adams, S. J. Giovannoni, H. W. Paerl, C. P. McKay, P. T. Doran, D. A. Gordon, B. D. Lanoil, and J. L. Pinckney. 1998. Perennial Antarctic lake ice: an oasis for life in a polar desert. *Science* **280:** 2095–2098.

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

Roberts, D. 2010. labdsv: Ordination and Multivariate Analysis for Ecology. 2010. *R package version*: 1.4–1. [Online]. Available: http://cran.r-project.org/web/packages/labdsv/index.html [2014, October 2].

Roberts, E. C., J. Laybourn-Parry, D. M. McKnight, and G. Novarino. 2000. Stratification and dynamics of microbial loop communities in Lake Fryxell, Antarctica. *Freshw. Biol.* **44:** 649–661.

Roberts, E. C., J. C. Priscu, C. Wolf, W. B. Lyons, and J. Laybourn-Parry. 2004. The distribution of microplankton in the McMurdo Dry Valley Lakes, Antarctica: response to ecosystem legacy or present-day climatic controls? *Polar Biol.* **27:** 238–249.

Shang, H. 2004. Molecular phylogeny of Scuticociliatia (Ciliophora, Oligohymenophorea). Ph.D. dissertation, Ocean University of China, Qingdao.

Song, W., and N. Wilbert. 2002. Faunistic studies on marine ciliates from the Antarctic benthic area, including descriptions of one epizoic form, 6 new species, and 2 new genera (Protozoa: Ciliophora). *Acta Protozool.* **41:** 23–62.

Spigel, R. H., and J. C. Priscu. 1998. Physical limnology of the McMurdo Dry Valleys Lakes. Pp. 153–187 in *Ecosystem Dynamics in a Polar Desert: the McMurdo Dry Valleys, Antarctica*, J. C. Priscu, ed. American Geophysical Union, Washington, DC. doi: 1029/AR072p0153.

Spaulding, S. A., B. Vijver, D. A. Hodgson, D. M. McKnight, E. Verleyen, and L. Stanish. 2010. Diatoms as indicators of environmental change in Antarctic and subantarctic freshwaters. Pp. 267–283 in *The Diatoms: Applications for the Environmental and Earth Sciences,* J. P. Smol and E. F. Stoermer, eds. Cambridge University Press.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.* **57:** 758–771.

Takishita, K., N. Kakizoe, T. Yoshida, and T. Maruyama. 2010. Molecular evidence that phylogenetically diverged ciliates are active in microbial mats of deep-sea cold-seep sediment. *J. Eukaryot. Microbiol.* **57:** 76–86.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28:** 2731–2739.

Thurman, J., J. Parry, P. J. Hill, J. C. Priscu, T. J. Vick, A. Chiuchiolo, and J. Laybourn-Parry. 2012. Microbial dynamics and flagellate grazing during transition to winter in lakes Hoare and Bonney, Antarctica. *FEMS Microbiol. Ecol.* **82:** 449–458.

Vick, T. J., and J. C. Priscu. 2012. Bacterioplankton productivity in lakes of the Taylor Valley, Antarctica, during the polar night transition. *Aquat. Microb. Ecol.* **68:** 77–90.

Vick-Majors, T. J., J. C. Priscu, and L. A. Amaral-Zettler. 2013. Modular community structure suggests community plasticity during the transition to polar night in ice-covered Antarctic lakes. *ISME J.* **4:** 778–789. doi: 10.1038/ismej.2013.190.

Wickham, S. A., S. Nagel, and H. Hillebrand. 2004. Control of epibenthic ciliate communities by grazers and nutrients. *Aquat. Microbiol. Ecol.* **35:** 153–162.

Wright, A.-D. G., and A. Colorni. 2002. Taxonomic re-assignment of *Cryptocaryon irritans*, a marine fish parasite. *Eur. J. Protistol.* **37:** 375–378.

Xu, H., Y. Jiang, K. S. Al-Rasheid, S. Al-Farraj, and W. Song. 2011a. Application of an indicator based on taxonomic relatedness of ciliated protozoan assemblages for marine environmental assessment. *Environ. Sci. Pollut. Res. Int.* **18:** 1213–1221.

Xu, H., W. Zhang, Y. Jiang, G.-S. Min, and J.-K. Choi. 2011b. An approach to identifying potential surrogates of periphytic ciliate communities for monitoring water quality of coastal waters. *Ecol. Indic.* **11:** 1228–1234.

Yasindi, A., D. Lynn, and W. Taylor. 2002. Ciliated protozoa in Lake Nakuru, a shallow alkaline-saline lake in Kenya: seasonal variation, potential production and role in the food web. *Arch. Hydrobiol.* **154:** 311–325.

Yilmaz, P., R. Kottmann, D. Field, R. Knight, J. R. Cole, L. Amaral-Zettler, J. A. Gilbert, I. Karsch-Mizrachi, A. Johnston, G. Cochrane *et al*. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat. Biotech.* **29:** 415–420.

Zhang, Q., A. Simpson, and W. Song. 2012. Insights into the phylogeny of systematically controversial haptorian ciliates (Ciliophora, Litostomatea) based on multigene analyses. *Proc. R. Soc. B* **279:** 2625–2635.

Zhu, F., R. Massana, F. Not, M. Dominique, and D. Vaulot. 2005. Mapping of picoeukaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* **52:** 79–92.

Zwickl, D. J., and M. T. Holder. 2004. Model parameterization, prior distributions, and the general time-reversible model in Bayesian phylogenetics. *Syst. Biol.* **53:** 877–888.

Appendix

Table A1

Physical and chemical characteristics for MDV lakes used in this study

Abiotic Parameter	West Lobe Bonney	Fryxell
Max depth	41.5 m	20 m
Ice thickness	3–6 m	3–6 m
% Transmission of PAR	7.0 %	1.5 %
Max salinity (PSU)	125	6.2
Temperature (°C)	−5.4 to 3.2 °C	0 to 3.2 °C
Depth of chemocline	15 m	9 m
Molar N:P* -mixo -chemo+mono	164 (37) 670 (94)	2.41 (0.60) 11.0 (0.57)

* For N:P ratios, the integrated mean (standard error) for the mixolimnion (mixo) and the chemocline and monolimnion (chemo+mono) are reported.

Table A2

Full dataset matrix

Cluster ID	Sample ID							
	MCM1	MCM2	MCM3	MCM4	MCM5	MCM6	MCM7	MCM8
E06_1057	34	0	0	2	0	0	15	0
U_18433	0	0	0	1	0	1	0	0
E06_217	0	0	0	0	0	0	1	0
E06_96	160	228	57	13	3	14	1260	1955
E06_630	1	7	8	0	0	0	0	26
E06_3357	0	0	0	1	0	2	0	0
E06_2422	0	0	0	1	0	1	0	0
E06_157	0	0	0	0	0	1	0	0
E06_2139	0	0	0	0	0	1	0	0
E06_178	0	0	0	0	0	0	0	99
E06_43	0	0	0	0	0	2	0	0
E06_620	8	16	9	1	0	2	94	125
E06_490	0	0	1	0	0	1	0	0
E06_3727	0	0	0	0	0	3	0	0
E06_255	62	19	12	0	0	5	507	284
E06_709	1	14	0	8	0	6	16	137
U_956	1	0	0	0	0	0	0	0
E06_250	31	19	0	11	1	6	93	763
E06_516	24	35	108	2	3	1	67	93
E06_392	7	2	22	4	2	6	171	62
E06_757	0	136	0	0	1	0	0	28
E06_605	0	0	0	0	0	2	3	0
E06_7317	0	0	1	0	0	0	0	0
U_28539	0	0	0	0	0	0	0	1
U_26011	0	0	1	0	0	0	0	0
U_14612	1	0	0	0	0	0	0	0
E06_6146	0	0	0	0	0	0	0	1
E06_12	85	52	3	5	0	2	1313	379
E06_6945	0	0	0	0	0	0	0	1
E06_6841	0	0	0	0	0	0	0	1
E06_105	9	23	101	6	23	3	276	680
E06_140	0	0	3	2	0	1	3	0
U_757	0	0	0	0	0	2	0	0
U_169	0	0	57	4	2	0	0	0
E06_180	0	0	0	2	0	0	0	0

Table A3

Ciliate sequences and observed richness per sample from both full and resampled data sets

	Sample ID							
	MCM1	MCM2	MCM3	MCM4	MCM5	MCM6	MCM7	MCM8
Number of ciliate sequences from full data set	424	551	383	63	35	62	3819	4635
Number of ciliate sequences from resampled data set	307	253	160	28	18	36	1612	1318
Observed richness from full data set	13	11	13	15	7	20	13	16
Observed richness from resampled data set	13	—	10	11	6	17	12	13

—Unable to perform analysis.