Journal of Microbiology (2019) Vol. 57, No. 10, pp. 852–864
DOI 10.1007/s12275-019-8668-8
eISSN 1976-3794
pISSN 1225-8873

Structure of bacterial and eukaryote communities reflect *in situ* controls on community assembly in a high-alpine lake[§]

Eli Michael S. Gendron[1,2*], John L. Darcy[3], Katherinia Hell[4], and Steven K. Schmidt[1]

[1]*Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, 80309, USA*
[2]*Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO, 80309, USA*
[3]*Computational Bioscience Program, University of Colorado Anschutz Medical Campus, Aurora, CO, 80045, USA*
[4]*Insitute of Arctic and Alpine Research, University of Colorado, Boulder, CO, 80309, USA*

(Received Dec 6, 2018 / Revised May 15, 2019 / Accepted May 16, 2019)

Recent work suggests that microbial community composition in high-elevation lakes is significantly influenced by microbes entering from upstream terrestrial and aquatic habitats. To test this idea, we conducted 18S and 16S rDNA surveys of microbial communities in a high-alpine lake in the Colorado Rocky Mountains. We compared the microbial community of the lake to water entering the lake and to uphill soils that drain into the lake. Utilizing hydrological and abiotic data, we identified potential factors controlling microbial diversity and community composition. Results show a diverse community entering the lake at the inlet with a strong resemblance to uphill terrestrial and aquatic communities. In contrast, the lake communities (water column and outlet) showed significantly lower diversity and were significantly different from the inlet communities. Assumptions of neutral community assembly poorly predicted community differences between the inlet and lake, whereas "variable selection" and "dispersal limitation" were predicted to dominate. Similarly, the lake communities were correlated with discharge rate, indicating that longer hydraulic residence times limit dispersal, allowing selective pressures within the lake to structure communities. Sulfate and inorganic nitrogen and phosphorus concentrations correlated with community composition, indicating "bottom up" controls on lake community assembly. Furthermore, bacterial community composition was correlated with both zooplankton density and eukaryotic community composition, indicating biotic controls such as "top-down" interactions also contribute to community assembly in the lake. Taken together, these community analyses suggest that deterministic biotic and abiotic selection within the lake coupled with dispersal limitation structures the microbial communities in Green Lake 4.

*For correspondence. E-mail: elge4317@colorado.edu; Tel.: +1-505-948-8029
[§]Supplemental material for this article may be found at http://www.springerlink.com/content/120956.

Keywords: co-occurrence patterns, landscape connectivity, deterministic community assembly, *Hydrurus*

Introduction

Recent research in high-elevation ecosystems such as Green Lakes Valley (GLV) in the Colorado Rockies has shown that these ecosystems are particularly susceptible to disruptions brought on by climate change such as earlier snowmelt, increased drought, and earlier melt out of lake ice (Preston *et al.*, 2016). Understanding how these changes shape aquatic microbial communities is important in order to predict the effects of climate change on the productivity and health of these susceptible ecosystems. Alpine ecosystems are also very susceptible due to the steepness of the terrain resulting in a high degree of connectedness across the landscape. For example, an increase in nitrification in alpine soils has resulted in recent increases in nitrate fluxes to high elevation lakes in the GLV (Ley *et al.*, 2004; Williams *et al.*, 2015). In a similar fashion, high elevation aquatic microbial communities are also likely to be greatly influenced by the degree of connectivity between the aquatic and terrestrial ecosystems (Crump *et al.*, 2012; Ruiz-Gonzalez *et al.*, 2015). This connectivity is also reflected in the Landscape Continuum Model (LCM), which describes the mostly downhill transport of matter and nutrients across landscape units in alpine ecosystems (Seastedt *et al.*, 2004). However, a large unknown within the LCM is the downhill transport of microorganisms and how this transport affects downstream biodiversity, biotic interactions, and ultimately biogeochemical cycles of lakes and streams.

Recent advances in sequencing technologies and high throughput biogeochemical approaches have led to a fairly comprehensive understanding of the spatial distribution and biogeochemical functioning of soil microorganisms in the GLV, especially in the mostly plant-free soils that drain into the headwaters of the basin (Ley *et al.*, 2004; King *et al.*, 2010, 2012; Bueno de Mesquita *et al.*, 2018; Porazinska *et al.*, 2018). There has also been work connecting the microbial communities of the late-season snow pack to the communities found in the soils into which the snowpack melts (Freeman *et al.*, 2009; Naff *et al.*, 2013). Furthermore, there is a wealth of valuable data concerning the hydrology, water chemistry, and phytoplankton communities of the GLV lakes (Flanagan *et al.*, 2009; Miller and McKnight, 2015; Williams *et al.*, 2015). Surprisingly however, there has been no molecular microbiological work examining the microbial communities of the high-elevation lakes in the GLV or work examining the connectivity of these communities to the surrounding ter-

restrial environment.

Recent work and the LCM suggest that the microbial communities of high-elevation and high-latitude lakes can be structured by terrestrial microbes moving downhill in runoff into these lakes (Crump *et al.*, 2012; Ruiz-Gonzalez *et al.*, 2015), in addition to inputs from upstream aquatic microbial communities (Nelson *et al.*, 2009). Like many high-elevation environments, the hydrologic network of GLV has seasonally high flow rates resulting in high connectivity for at least part of the year between upstream and downstream habitats (Williams *et al.*, 1996; Miller and McKnight, 2015). Because of this connectivity and previous work in other cold environments discussed above, we hypothesized that the microbial communities of Green Lake 4 (GL4) would be primarily structured by the flux of microorganisms from uphill habitats.

To determine which factors structure microbial communities in Green Lake 4 GL4, we conducted the first comprehensive 18S and 16S rDNA survey of the microbial communities of the water entering the lake ("inlet"), the lake itself, and upstream terrestrial habitats, examining both temporal and spatial changes in the microbial community. We also measured biotic and abiotic variables to determine which parameters correlate with microbial community patterns. In addition, we examined the degree of co-correlation between bacterial and eukaryotic communities by examining co-occurrence patterns across the aquatic and terrestrial ecosystems of the GLV. Our results show that the microbial community of the lake's inlet shared much of its composition with upstream terrestrial sources, but communities across the rest of the lake were significantly different from the uphill and the inlet communities. There were also strong co-occurrence patterns across trophic groups, as well as a strong correlation in the diversity and composition of the 16S and 18S communities in the lake, indicating that biotic interactions may be important in structuring the communities. Abiotic variables (e.g. hydraulic residence time and inorganic nutrients) also correlated with both 16S and 18S communities, indicating that abiotic factors may also be important in structuring the communities. Whatever the mechanistic underpinnings of the patterns observed, this study clearly shows that the communities present in the lake are determined by selection within the lake as opposed to being structured by the constant flux of organisms from upstream habitats.

Materials and Methods

Site description and field sampling

Green Lakes Valley (GLV) is a glacial valley just east of the Continental Divide located within the Colorado Front Range and is part of the Silver Lake Watershed, which supplies 40% of the water for the City of Boulder, Colorado. Green Lake 4 (GL4) is located in the upper catchment of GLV at 3,561 m above sea level, has a surface area of 5.34 ha, a volume of 215,000 m^3, a maximum depth of 13.1 m, and an average depth of 4.0 m (Gardner *et al.*, 2008). GL4 is an oligotrophic alpine lake that exhibits low primary production and low summer chlorophyll-a contents compared to more eutrophic lakes at lower elevations (e.g. Foley *et al.*, 2012). Thermal stratification beginning in mid-summer at a depth of 8.0 m has been sometimes observed since limnology monitoring started in 2000 and was most accentuated during the 2002 drought (Gardner *et al.*, 2008; Flanagan *et al.*, 2009). However, summer stratification has not been observed in recent years (NWT LTER database http://niwot.colorado.edu/data). GL4 is covered by ice for most of the year, only being ice-free from mid-July to early or mid-October. Water residence times are initially short during melt out, but significantly change over the course of the year, with some of the shortest residence times being as low as one week (Daily discharge: ~27,000 m^3/day) in the early summer, and as high as about 40 (Daily discharge: ~5,300 m^3/day) days in mid-October (Waters, 1999). Our sampling did not take place across this whole range of residence times, but covered residence times in GL4 that ranged from 11.8 (Daily discharge: 18,300 m^3/day)–19.9 days (Daily discharge: 10,804 m^3/day) during the course of this study. Over the course of the study, max air temperatures in the valley ranged from 2–12°C and an average of 3.4°C, with the season's peak air temperature occurring before sampling begun. Average rain storms were seen on Aug. 19th and 26th but no significant spike in discharge was detected. A spike in discharge rates was detected July 29th–31st but rapidly returned to average discharge rates by Aug. 1st. The inlet to GL4 is a stream that drains Green Lake 5 (GL5) and the upper catchment of the GLV, forming a wetland in-between the lakes resulting in extensive hyporheic interactions in the stream reach from GL5 to GL4 (Miller *et al.*, 2006). Water discharge rates were measured at the outlets of GL4 and GL5 by the Niwot LTER throughout the 2014 season (Supplementary data Table S3; NWT LTER database http://niwot.colorado.edu/data). Water input also comes from large snow and ice-fields that cover much of the steep slopes to the north of the lake. Therefore, to characterize the microbial inputs to the lake we sampled the inlet waters and the saturated soils along the edges and under the melting snow field to the north of the lake at sites previously described by Ley *et al.* (2001, 2004) (40.0571556, -105.6229083).

Water samples were collected from GL4 on July 24th, August 7th, and August 21th of 2014. Samples were collected in HDPE Nalgene bottles sterilized with a 90% Ethanol rinse and 15 min exposure in a UV sterilization chamber. Bottles were rinsed with sample water prior to collection. On the first sampling date (07/24/14), one 500 ml sample was taken from each of 5 locations in the lake (inlet at 40.0539306, -105.6221444, outlet at 40.0556278, -105.6173194, and the water column samples of surface, 3 m depth and 9 m depth at 40.0552389, -105.6204889). On the later sampling dates (08/07/14 and 08/21/14), samples were taken in triplicate (250 ml per pseudoreplicate) from each of the same five sampling points in the lake (3 samples per site per date). Samples were placed on ice in a cooler and transported to the University of Colorado, Boulder where they were stored at 4°C and filtered within 60 h of collection. Samples were filtered through 0.2 μm polycarbonate filters using a vacuum filter (Millipore Inc). Between sample filtrations, the filter adapter was rinsed with both 75% ethanol and sterile water, then exposed to high-intensity ultraviolet radiation in a sterilization chamber. Filters were stored at -20°C until DNA extraction. Samples for water chemistry were collected as part of the LTER network's

Fig. 1. Green Lake 4 (GL4) viewed from above. Arrows indicate direction of flow from the inlet (right side) and the outlet (left side). The X in the middle of the lake marks the location from where the water column samples were taken. Talus soil samples were collected northwest of the spot where this photo was taken. Picture taken September 3, 2014 from Niwot Ridge, located atop a south facing ridge made up of unvegetated talus and vegetated alpine tundra soils to the north of the lake.

long-term monitoring of GL4 and were processed by the Kiowa lab (INSTAAR, University of Colorado – Boulder). Samples were analyzed for pH, temperature, conductivity, acid neutralizing capacity, H^+, NH_4^+, Ca^{2+}, Mg^{2+}, Na^+, K^+, Cl^-, NO_3^-, SO_4^{2-}, silica, total dissolved nitrogen, dissolved organic nitrogen and phosphorus, inorganic nitrogen and phosphorus, dissolved organic carbon, zooplankton taxa density and richness, and chlorophyll-a based on protocols established for use by the Niwot LTER (Supplementary data Table S3; Unpublished LTER 2014 data, Miller and McKnight, 2010; McKnight, 2016; Loria, 2019).

To further characterize potential sources of microbial input to the lake, we sampled the saturated soils along the edges and under the melting snow field in the extensive plant-free talus soils to the north of the lake at sites previously described in Ley *et al.* (2001, 2004) (40.0571556, -105.6229083) and Porazinska *et al.* (2018) (Fig. 1). These soil samples (both taken along the edge and from underneath the snow) will be referred to as "talus" in the rest of the manuscript. Talus samples were collected on August 12th and September 3rd of 2014 (18 total talus soil samples spread across the landscape), by scooping a ~100 g sample of soil into plastic bags using a sterile spoon from each of the depths of 0–2 cm, 2–4 cm, and 7–18 cm (maximum depth varied between holes). Community composition across the depths did not significantly vary, and was likely due to the large amount of melt water present in the talus soils at the time which would fill the wells we dug as we gathered samples. We included the talus samples in our analyses even though we did not collect them on the same dates as the water samples in order to show that the soil communities in the water shed are fairly stable across time and are quite different from the lake communities. Furthermore, water from the talus soils drain into GL4 (Liu *et al.*, 2004; Molotch *et al.*, 2008). Samples were placed in a cooler for transport back to the University of Colorado, Boulder and stored at -20°C until DNA was extracted.

DNA extraction and sequencing

DNA extractions from water samples (three replicate extractions per site per date) were performed using the MoBio Inc. PowerWater DNA Isolation Kit. In order to insure enough biomass was present on the filter, each extraction was performed using all of the water from each 250 ml sample we collected. The frozen filters were picked clean of large biomass such as plant debris and zooplankton in order to avoid their DNA overwhelming other sequences during amplification. Bead-beating tubes with filters were warmed at 65°C for 10 min, then bead-beating was performed for 5 min prior to DNA extraction. Soil sample DNA extractions were performed using the MoBio PowerSoil DNA Isolation Kit. ~0.4 g of frozen soil was used for each DNA extraction. Soil was thawed at room-temperature, then bead-beating was performed for 8 min prior to DNA extraction.

Extracted genomic DNA was stored at -20°C. Each extraction was amplified in triplicate targeting 16S and 18S rDNA gene sequences. 16S rDNA sequences were amplified using 515f–806r primers adapted with barcodes and linker sequences (Supplementary data Table S3) (Caporaso *et al.*, 2012), and 18S rDNA sequences were amplified using 1391f-EukBr primers adapted with barcodes and linker sequences (Supplementary data Table S3) (Vestheim and Jarman, 2008; Amaral-Zettler *et al.*, 2009; Caporaso *et al.*, 2012). The primers we used for amplifying 16S rDNA sequences have been shown to poorly detect archaeal taxa (Walters *et al.*, 2016); since the method we chose to implement does not accurately represent the archaeal diversity in our samples we will not be discussing the archaeal taxa that we detected in our samples. Reactions were pooled and concentrations were assayed using Pico Green flurometry on a BioTek Synergy 2 microplate reader. Amplicons were diluted to equimolar concentrations, then sequenced on the Illumina MiSeq platform using paired end 2 × 150 bp chemistry. A 30% phiX spike was added to the run due to limited amplicon variability (Caporaso *et al.*, 2012).

Bioinformatics and statistical analyses

Sequences were de-multiplexed and joined using the QIIME v1.8 bioinformatics package (Caporaso *et al.*, 2010). Joining of the 16S rDNA sequences resulted in 254 bp reads. Joining was not possible for 18S rDNA sequences as the sequenced strands did not overlap, so only single-end 18S reads (157 bp, corresponding to the 1391f primer) were used. 16S and 18S sequences were clustered into operational taxonomic units (OTUs) at 97% similarity utilizing the UCLUST clustering method (Edgar, 2010). Taxonomy was assigned to the 16S OTU representative sequences using the QIIME script parallel_assign_taxonomy_rdp.py which utilizes the Ribosomal Database Project sequences for classification. Taxonomy of the 18S OTU representatives was assigned to OTU representative sequences using QIIME script parallel_assign_taxonomy_blast.py and the ARB SILVA Ref NR 97 database version 119 (www.arb-silva.de). Separate OTU tables were generated for 16S and 18S communities as well as for comparisons among aquatic samples and comparisons between aquatic and terrestrial samples. The 16S OTU tables were filtered of all chloroplast and mitochondrial sequences in order to prevent them from overshadowing community patterns.

The aquatic 16S rDNA OTU table was rarefied to 14,800 sequences per sample, and the aquatic 18S rDNA OTU table was rarefied to 7,000 sequences per sample. Five samples from the aquatic 16S data set that did not have enough sequences for rarefaction were discarded from the analysis; 3 of the 5 samples of the 16S data from July 24th were discarded due to particularly low sequence depths, but a 9 m depth and outlet sample were retained due to their high sequence depth. We decided to keep the 16S data from July 24th in order to have a third time point we could compare against in describing the composition of an alpine growing season community. Furthermore, we were able to retain all of the 18S samples from July 24th and including some 16S data from the same date will provide a more honest comparison of the two communities. The combined aquatic + terrestrial 16S OTU table was rarefied to 5,489 sequences per sample, and 3 samples were discarded due to low sequence counts. The combined 18S OTU table was rarefied to 4,291 sequences per sample.

For all four rarefied OTU tables, representative sequences were aligned using the QIIME script parallel_align_seqs.py and phylogenetic trees were built using FastTree 2 (Price *et al.*, 2010). Rarified OTU tables and phylogenetic trees were used to calculate alpha- and beta-diversity values with QIIME scripts alpha_diversity.py and beta_diversity_through_plots.py. To analyze the dissimilarity in community composition between samples, beta-diversity was calculated using weighted UniFrac. Weighted UniFrac was implemented to help account for both the relative abundance of and evolutionary distances between OTUs in calculating community dissimilarity (Lozupone *et al.*, 2011). Statistical analysis of these values was performed using R v3.2.1 and using the R package vegan v2.3-5. Analysis of variance (ANOVA) was used to test whether alpha-diversity of microbial communities differed by sampling location, and permutational analysis of variance (PERMANOVA) was used to determine significant drivers of community composition (Zapala and Schork, 2006).

In order to determine the proportion of microbial taxa shared between the talus, inlet, and lake communities, all 16S rDNA sequences from the talus, lake, and inlet samples were combined into 3 different meta-samples to represent the overall communities of each environment ("lake", "talus", "inlet") as IDs (Supplementary data Table S5). This table was rarefied to 216,000 sequences per meta-sample, and included all the samples collected except for one inlet sample from July 24th which only contained 16 sequences. The overlap between communities of different environments was calculated as vectors of pair-wise column differences in this table. The number of overlapping sequences was used instead of the number of shared OTUs in order to take relative abundance of the OTUs into account since there would be more sequences from an abundant OTU versus a rare OTU. This process was repeated with 18S rDNA sequences, which were rarefied to 338,000 sequences per meta-sample ("inlet", "lake", "talus"; Supplementary data Table S6). Sequence overlap results were plotted as Venn diagrams using the R package venneuler.

In order to investigate the importance of local environmental factors as correlates of community structure, distance based redundancy analysis (db-RDA) was performed (Legendre and Anderson *et al.*, 1999). Stepwise model selection was performed on the original dbRDA model to produce a final model (Langenheder and Ragnarsson, 2007). Redundancy analysis and stepwise model selection were performed using R v3.2.1 Vegan package v2.3-5. Models were evaluated using Monte Carlo permutation tests (999 permutations under the reduced model). Significance of model terms were evaluated using an analysis of variance (ANOVA) like permutation test for testing the dbRDA model terms.

In order to further investigate the extent of similarity between how the 16S rDNA and 18S rDNA microbial communities were spatially and temporally structured, we analyzed the correlation between 16S and 18S rDNA community dissimilarity. Unifrac distance matrices made from the Green Lake 4 OTU tables (see above) were plotted and tested using a model II linear regression using the package lmodel2 (Legendre, 2014) with 500 permutations in R v3.3.1 (R Core Team, 2016). Furthermore, we also used a model II linear regression to compare the phylogenetic diversity of the 16S and 18S rDNA OTUs. We calculated phylogenetic diversity metric using a standard-effect-size approach to compare diversity between samples (Faith, 1992; Webb *et al.*, 2008; Kembel *et al.*, 2010). Lastly, in order to gain an understanding of how biotic interactions could contribute to the observed spatial structure of microbial communities, we chose to run a co-occurrence analysis on the abundant (> 1% of all observed OTUs) phylotypes observed in our 16S and 18S OTU tables. Since trophic interactions play a major role in determining community structure, we then split our 18S table into an autotrophic and heterotrophic table based on assigned taxonomies. The only abundant phylotypes that passed our 1% cutoff for the 16S table were all heterotrophic. We then ran pairwise Pearson correlations between the autotrophic table and the two heterotrophic tables, followed by FDR (False Discovery Rate) correction of significance levels (Stanish *et al.*, 2013).

In order to determine the relative importance of selection and neutral processes in Green Lake 4, we implemented a neutral model based on Sloan *et al.* (2006)'s adaptation of Hubbell's neutral model (Hubbell, 2001) making use of the approach described in Venkataraman *et al.* (2015). Testing of the neutral model was executed in R v3.2.1 using custom scripts provided by Dr. Venkataraman (Venkataraman *et al.*, 2015). The inlet to Green Lake 4 was defined as the source for the lake body and outlet microbial communities. Utilizing this model, we estimated the distribution of OTUs which were selected for/against in the lake and the distribution of OTUs which were neutrally dispersed into GL4 during the 2014 growing season through random dispersal and ecological drift. Determination of where OTUs fall in this model are dependent on the relative abundance of the OTU in the source site (GL4 inlet) and the frequency at which it is detected in the target site (GL4 lake body and outlet). The goodness of fit of these data to the expected neutral model was assessed using the coefficient of determination (R^2) and 95% confidence intervals were calculated using the R package 'Hmisc'. In this study, the source community was all inlet samples and the target community all lake body and outlet samples together as a 'lake' community.

We used the approach of Stegen *et al.* (2013) to determine the dominant community assembly processes structuring the

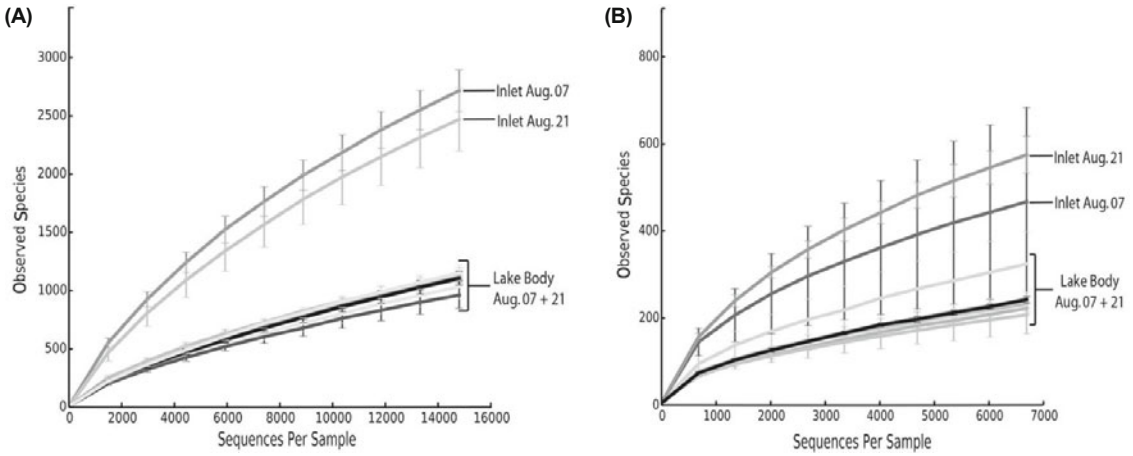

Fig. 2. Alpha-rarefaction plots of GL4 microbial communities. Alpha rarefaction using the observed OTUs (clustered at 97%) metric showed that on average, more OTUs were detected in inlet sites and fewer OTUs were detected in the lake body sites. This pattern is very robust for (A) 16S rDNA data, and for (B) 18S rDNA data from the August 21st samples. August 7th samples are variable (red) but generally the same trend is apparent. Each trend shown is a mean number of OTUs observed for samples from a date and location (labeled at right), and error bars are standard error of the mean.

16S and 18S communities in the lake and inlet. This approach determines the relative contribution of five community assembly processes: i) homogenous selection (abiotic or biotic selection for same OTUs across communities), ii) variable selection (abiotic or biotic selection for different OTUs across communities), iii) homogenizing dispersal (OTUs are easily dispersed between communities), iv) dispersal limitation (OTUs do not disperse easily between communities), and v) undominated (community structure are due to weak selection, weak dispersal, and/or random drift) (Stegen *et al.*, 2015; Whitman *et al.*, 2018). Calculations of β Mean Nearest Taxon Distance (βMNTD) uses the mean phylogenetic dissimilarity between each pair of samples to predict the influence of selection processes (Stegen *et al.*, 2013), and were obtained using the R package picante v1.8 (Kembel *et al.*, 2010). In conjunction, a Raup-Crick metric based on Bray-Curtis dissimilarity (RC_{Bray}) was implemented to predict the influence of dispersal and undominated processes on community assembly (Chase *et al.*, 2011). RC_{Bray} calculations, model construction, and testing were performed using R code used in Stegen *et al.* (2013) (GitHub: Stegen_etal_ ISME_2013/bNTI_Local_Machine.r and Stegen_etal_ISME_2013/Raup_Crick_Abundance.r). The predicted relative influence of each class of assembly processes were calculated 1) across all samples, 2) within only the inlet samples, 3) between inlet and lake samples, and 4) within only the lake samples.

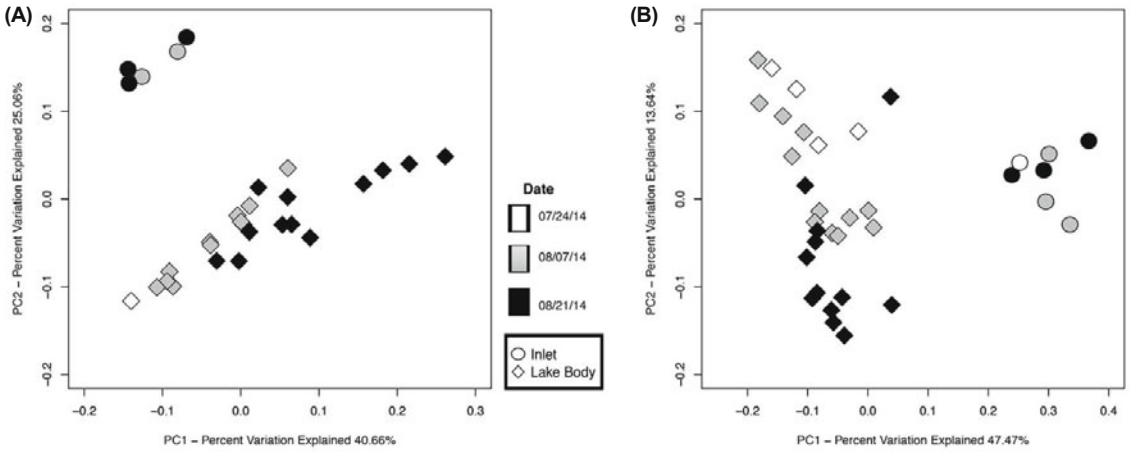

Fig. 3. Principle coordinates analysis of microbial community beta-diversity in GL4 ("Lake Body") and the inlet waters of the lake ("Inlet") across three sampling dates in 2014. (A) 16S rDNA and (B) 18S rDNA community similarity (weighted Unifrac) was calculated using relative abundances of sequences in each taxa and each taxa's presence/absence. For both data sets, inlet (circles) communities clustered separately from lake body (diamonds) communities. This pattern held for both our August 7th sampling date (grey) and our August 21st sampling date (black), indicating that microbial community inlet waters were significantly different than those in the lake body regardless of the sampling date (PERMANVOVA, $P < 0.001$ for 16S and for 18S). Un-replicated samples from July 24th samples are colored in white.

Results

A total of 52,890 16S rDNA OTUs were identified across all samples from the lake (3 dates and 5 sites per date). Alpha-diversity was significantly higher in the inlet compared to the rest of the lake (ANOVA, $P < 0.00005$) (Fig. 2A, Supplementary data Fig. S1A), with a mean Shannon diversity index of 7.9 (± 0.33) compared to 6.3 (± 0.22) for the rest of the lake. Beta-diversity of the 16S communities also showed a striking separation between the community of the inlet compared to the rest of the lake (Fig. 3A). Based on the permutational MANOVA (ADONIS), 16S community composition was significantly different in the inlet compared to the lake ($R^2 = 0.29$, $P < 0.0001$). While we observed a seasonal shift in lake body bacterial communities (Fig. 3A), no such pattern was found in the inlet communities which were significantly different from the lake regardless of seasonality.

A total of 9,982 18S rDNA OTUs (eukaryotic microbes) were identified across all samples from the lake and as with the 16S data, alpha-diversity significantly differed between the inlet and the lake (ANOVA, $P < 5 \times 10^{-10}$) (Fig. 2B, Supplementary data Fig. S1B). The 18S community of the inlet had a mean Shannon diversity index of 5.8 (± 0.11) while the lake had a mean index of 4.2 (± 0.20). Beta-diversity of the 18S communities (Fig. 3B) showed that there was a significant difference in community composition between the inlet and the lake ($R^2 = 0.46$, $P < 0.0001$), despite an observed seasonal shift in the lake's eukaryotic beta-diversity that was not reflected at the inlet (Fig. 3B).

In order to determine which biotic or abiotic variables correlate with beta-diversity patterns, we used Distance-based Redundancy Analysis (dbRDA). The dbRDA results revealed the density of zooplankton taxa, conductivity, daily discharge (Q), inorganic nitrogen (IN), and sulfate (SO_4^{2-}), as significant correlates with beta-diversity for the 16S community (Supplementary data Table S4). 18S beta-diversity was found to significantly correlate with inorganic phosphorus (IP), SO_4^{2-}, Q, IN, and water temperature (Supplementary data Table S4). Zooplankton, conductivity, Q, SO_4^{2-}, and IN accounted for 95.8% of the variance in 16S community beta-diversity (Supplementary data Table S4). While IP, SO_4^{2-}, Q, IN, and water temperature accounted for 77.4% of the variance in 18S community beta-diversity (Supplementary data Table S4). The beta-diversity or composition of the 16S and 18S communities in GL4 were significantly correlated with each other (Fig. 5A), that is, the 16S and the 18S UniFrac community distance matrices showed a strong correlation (Model II linear regression, r = 0.568, $P < 0.001$), suggesting that similar biotic or abiotic variables are structuring both communities and/or that one community may be structuring the other. Comparisons of alpha-diversity also showed a significant relationship between the 16S and 18S communities in GL4 (Fig. 5B). Within sample phylogenetic diversity (PD) showed a significant correlation between the 16S and 18S communities (Model II linear regression, r = 0.371, $P = 0.041$), further highlighting that similar variables are likely responsible for structuring both communities.

To test the idea of connectivity between the terrestrial and aquatic communities, we compared the microbial community of the lake with the community of the soils of an intensively studied talus slope directly uphill from the lake (Ley *et al.*, 2001, 2004; King *et al.*, 2010, 2012). A total of 40,021 16S rDNA OTUs, and 22,758 18S rDNA OTUs were identified from the upslope soil samples (2 dates, 18 samples per date). The alpha-diversity of the soils was similar to the inlet but consistently higher than that of the lake for both 16S and 18S communities. The soil 16S community had a mean Shannon diversity index of 8.4 (± 0.07), and the 18S community had a mean Shannon diversity index of 6.1 (± 0.13). Based on permutational MANOVA (ADONIS) of beta-diversity, both the 16S ($R^2 = 0.77$) and 18S ($R^2 = 0.54$) soil microbial communities were significantly different from both

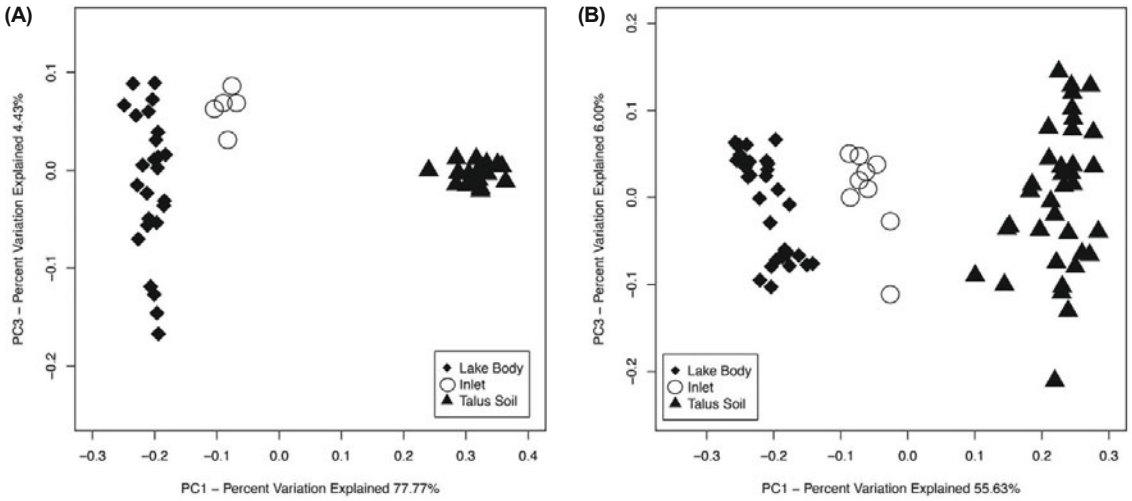

Fig. 4. Principle coordinates analysis of microbial communities in in GL4 ("Lake Body"), the inlet waters of the lake ("Inlet") and some representative soil samples from talus soils above the lake ("Talus Soil"). The Weighted Unifrac calculated similarities of the (A) 16S rDNA and (B) 18S rDNA community. 16S communities of the inlet (white circles) waters clustered separately from those in the rest of the lake (black diamonds), and also clustered separately from the talus soils (black triangles). However, inlet communities were less different from talus communities than lake communities were, especially for the 18S communities (PERMANOVA, $P < 0.001$ for 16S and for 18S).

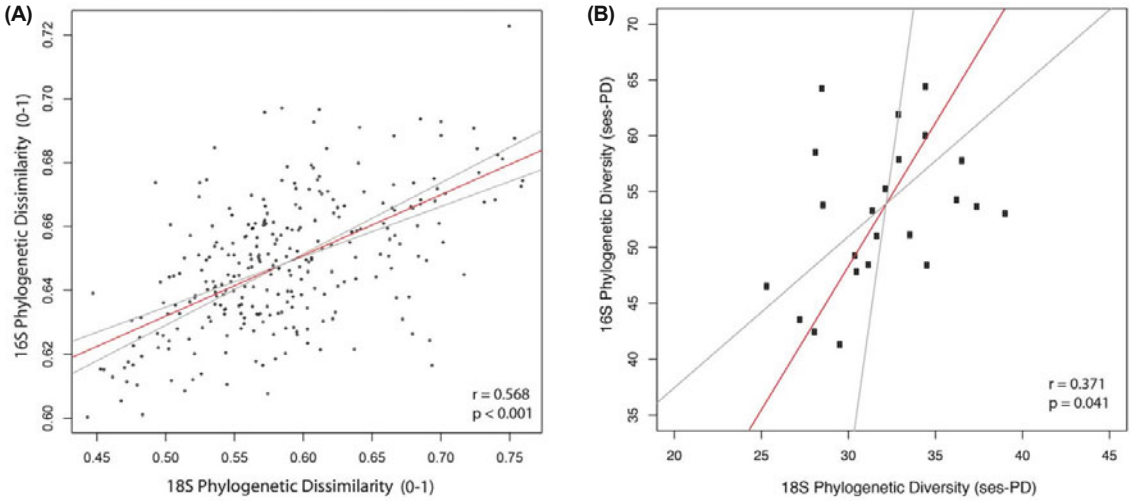

Fig. 5. Linear regression of 16S and 18S community diversity and composition. Model II linear regressions of the (A) pairwise beta-diversity correlation between 16S rDNA and 18S rDNA communities, and of the (B) correlation between 16S and 18S phylogenetic diversity (ses-PD). Pairwise beta-diversity comparisons were made using an unweighted Unifrac community dissimilarity metric, and phylogenetic diversity comparisons were made using a standardized-effect-size phylogenetic diversity metric. Correlations were tested using a model II linear regression because we do not know which variable is dependent on the other.

the inlet and lake body communities (Fig. 4) ($P < 0.001$). We also investigated the compositional overlap of these communities by comparing the total number of shared identical rDNA sequences in the soils versus the lake (Supplementary data Fig. S2A). After rarefaction (to 216,000 16S rDNA sequences), 31.5% of the talus soil 16S sequences were observed in the inlet, and only 2.01% of talus soil sequences were observed in the lake body. Out of the 338,000 18S rDNA talus soil sequences (after rarefaction), 6.21% of the 18S sequences were observed in the inlet, and only 0.71% of talus soil sequences were observed in the lake body.

To further explore the connections between the terrestrial and aquatic systems we compared the co-occurrence patterns of the soil and lake microbial communities and found that terrestrial autotrophic communities did not show significant co-occurrence with lake heterotrophic communities nor did the terrestrial heterotrophic communities significantly co-occur with lake autotrophic communities (Supplementary data Fig. S2), suggesting that terrestrial communities do not play a major role in structuring the communities in the body

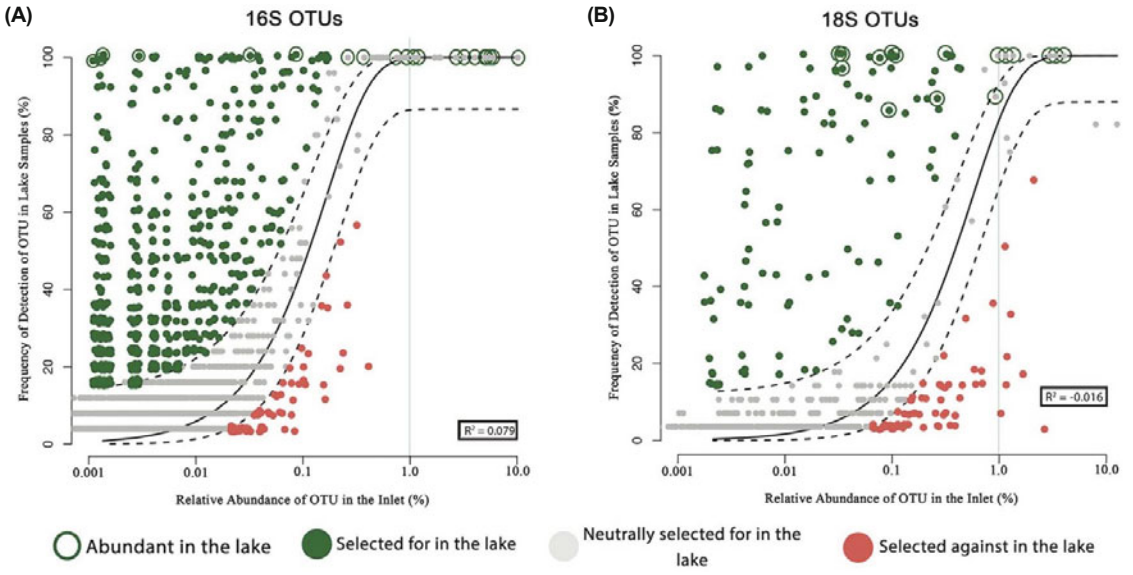

Fig. 6. Fit of (A) 16S (B) 18S data sets to the neutral community model. The predicted neutral model (solid black line) is surrounded by the 95% confidence intervals (dashed lines). The frequency of detection of 16S and 18S OTUs in the lake (target community, y-axis) are plotted against said OTU's abundance in the inlet (source community, x-axis) to indicate that the OTU is most likely neutrally dispersed into the lake (gray points), selected for (green points), or selected against (red points). Circled points indicate OTUs that were abundant in the lake (> 1% relative abundance).

of the lake. In contrast, co-occurrence analysis did reveal strong co-occurrence patterns between aquatic photoautrophs and aquatic heterotrophs. Lake autotrophic phylotypes were dominated by a few genera of algae, and one diatom OTU. Several heterotrophic phylotypes such as the *Polynucleobacter* and Cytophagaceae that strongly correlated with lake autotrophs have been previously found to be associated with blooms of algae, diatoms, and cyanobacteria (Riemann and Winding, 2001; Wu and Hahn, 2006).

We used Sloan's neutral community model (Sloan *et al.*, 2006; Venkataraman *et al.*, 2015) to investigate whether each OTU was selected for or against in the lake and which were neutrally selected (Fig. 6A). We found that the 16S community in GL4 was poorly fit by the neutral model ($R^2 = 0.079$) when the inlet was considered the source community. A high percentage of taxa did fall within the 95% confidence interval of the neutral model (grey dots in Fig. 6), but 96% of these were OTUs with low frequency (in < 20% of samples) in the body of the lake. Among the dominant (> 1% of the sequences) members of the lake community, 50% were OTUs that were relatively rare (< 0.1%) in the inlet community indicating positive selection of those OTUs in the lake.

The 18S community in GL4 was also poorly fit by the neutral model when the inlet was considered as the source community ($R^2 = -0.016$; Fig. 6B). As seen with the 16S community many OTUs fell within the 95% confidence interval of the neutral model, but 93% of these were rare members of the lake community (in < 20% of samples; grey dots in Fig. 6B). In contrast, 58% of the abundant OTUs (OTUs > 1% of sequences, circled dots) in the lake were selected for in the lake (green dots), while the remaining 42% of OTUs abundant in the lake were neutrally selected for in the lake, but these were all taxa that were abundant in the inlet. Lastly, 42% of the OTUs abundant in the inlet (> 1% of sequences) were selected against in the lake (red dots in Fig. 6B).

To determine the relative contribution of homogenous selection, variable selection, homogenizing dispersal, dispersal limitation, and undominated processes to community assembly processes between the inlet and the lake we implemented the approach of Stegen *et al.* (2013). The dominate processes predicted to govern assembly of both the 16S and 18S communities between the inlet to the lake were variable selection and dispersal limitation (Supplementary data Fig. S3). These results support the results reported above by indicating that selection (variable selection by abiotic or biotic pressures) plays a large role in the community differences between the lake and inlet and that dispersal limitation between the inlet and the main body of the lake may also play a role in the community differences.

Discussion

Although microbiological studies of high-elevation lakes have been done, most have focused on one domain of life: either Bacteria, Archaea, or Eukarya (Wang *et al.*, 2014; Kammerlander *et al.*, 2015; Zhong *et al.*, 2016; Hu *et al.*, 2018). In the present study, we used a spatially explicit sampling scheme to obtain the first high-resolution, comprehensive view of the bacterial and eukaryotic (16S rDNA and 18S rDNA) communities of an alpine lake in the Colorado Rocky Mountains. The 16S community found in GL4 was generally consistent with previous reports of bacterial phyla found in the microbial communities of other high-elevation lakes (Wu *et al.*, 2006; Liu *et al.*, 2013; Vila-Costa *et al.*, 2013). Microbial eukaryotic (18S) diversity in freshwater lakes is not as well documented as bacterial diversity; however, the eukaryotic communities we observed in GL4 were also similar to eukaryotic communities of other freshwater ecosystems (Šlapeta *et al.*, 2005; Lefèvre *et al.*, 2008; Lou *et al.*, 2011; Triadó-Margarit and Casamayor, 2012), even sharing some genera of diatoms (e.g. *Fragilaria* and *Craticula*) with high elevation lakes in the Altiplano of Chile (Angel *et al.*, 2016).

Our main goal was to determine what factors control the microbial community assembly (Nemergut *et al.*, 2013) in the main body of Green Lake 4 (GL4) and our results provide evidence that the microbial community of the lake is not structured by uphill microbial communities or the microbes entering the lake at the inlet. That is, both 16S and 18S communities in the inlet are significantly different from communities found in the main body of the lake (Figs. 2–4). The higher diversity in the inlet reflects the uphill sources of inoculum, including talus soils and streams as has been seen in some other studies of high-elevation and high-latitude lakes (Nelson *et al.*, 2009; Crump *et al.*, 2012; Adams *et al.*, 2014; Ruiz-Gonzalez *et al.*, 2015). However, due to known hydrological connectivity between the uphill talus soils and the valley's surface waters we expected to see more overlap between the talus and the lake communities (Liu *et al.*, 2004; Molotch *et al.*, 2008).

The strong signal of uphill environments detected in the inlet was mostly lost in the lake body (Supplementary data Fig. S2). The poor fit of the neutral model (Sloan *et al.*, 2006; Venkataraman *et al.*, 2015) to our lake data when the inlet was considered the source community (Fig. 6), further emphasized selection in the lake. For example, 58% of the 18S OTUs with greater than 1% relative abundance in the lake were OTUs that were rare (< 0.1%) or very rare (< 0.01%) in the inlet water (Fig. 6B). Conversely, 42% of the 18S OTUs that were abundant (> 1% of sequences) in the inlet were selected against in the lake and 93% of the inlet sequences that were neutrally selected in the lake were found in less than 20% of the lake samples according to the analyses shown in Fig. 6B. These analyses also indicated that the dominant 18S phylotype (in terms of relative abundance) in the inlet (*Hydrurus*), was at low relative abundance in the rest of the lake (Supplementary data Table S2), but was still found in 80% of the lake samples, indicating neutral dispersal (and perhaps dilution) in the lake. This same *Hydrurus* OTU was also the dominant 18S phylotype (range of 0.02–48.3% per sample) in the talus soils above GL4 and may be a snow alga (Freeman *et al.*, 2009; Naff *et al.*, 2013) as it is closely related to snow algae in Antarctica and the Arctic (Remias *et al.*, 2013; Darcy *et al.*, 2017), and to algae in cold glacial streams throughout the cryosphere (Rott *et al.*, 2006). *Hydrurus* has motile zoospores that may increase dispersal across the lake (Remias *et al.*, 2013). Even though *Hydrurus* was shown to be neutrally selected for in the lake, it is still perhaps the best indicator OTU for uphill inputs to the lake because of its high relative abundance in soil and stream habitats.

The bacterial community was found to fit the neutral model only slightly better ($R^2 = 0.079$) than the poor fit of the 18S community. Freshwater bacterial communities can display greater niche breadth and were found to be less influenced by environmental filters in some freshwater habitats following the 'size-plasticity' hypothesis (Farjalla *et al.*, 2012; Wu *et al.*, 2017), and we did observe that the percentage of the bacterial community selected against in the lake was less than that of the eukaryotic community (Fig. 6). However, we still observed that a significant proportion of the lake bacterial community was classified as selected for in the lake indicating that while some of bacterial taxa were being selected against in the lake, there were still 50% of the most abundant OTUs selected for (Fig. 6). Selection in the lake is also evident from an examination of which bacteria were dominant in the inlet compared to the lake body. The dominant bacterial phylotype in the inlet was most closely related to the ACK-M1 (or AcI) group of the Actinomycetales (Newton *et al.,* 2011; Comte *et al.*, 2015), whereas the rest of the lake (across all depths) was dominated by a *Flavobacterium* (Bacteroidetes) and a Verrucomicrobium phylotype in the Cerasicoccaceae (Opitutae; Supplementary data Table S1). The dominance of the ACK-M1 in the inlet is consistent with the work of Comte *et al.* (2015) who found that this group dominated (25% of total sequences) microbial communities in northern Canadian permafrost thaw ponds. Therefore, the presence of this group in the inlet waters may be indicative of the permafrost melting that is occurring in the Green Lakes Valley (Caine, 2010; Barnes *et al.*, 2014). These same bacteria are not found in significant abundance in the lake body (Supplementary data Table S1), again indicating that there are strong selective forces at work in the lake.

Our use of the approach of Stegen *et al.* (2013) shows that the two main determinants of 16S and 18S community assembly when comparing the inlet to the lake are "variable selection" and "dispersal limitation", with a relatively small influence from "undominated" processes (Supplementary data Fig. S3). These results further support the results from our implementation of Sloan's neutral community model (Fig. 6) because undominated processes are relatively stochastic, e.g. "weak selection, weak dispersal and/or random chance events" (Whitman *et al*., 2018), meaning we would expect to see a higher proportion of neutrally selected abundant OTUs in an undominated system (Hubbell, 2001; Chave, 2004), which is not what we found. The relationship between variable selection, dispersal limitation and the biotic and abiotic variables we measured are discussed in the following paragraphs.

Given the community shift between the inlet and the lake, and that "variable selection" (Whitman *et al.*, 2018) is one of the main forces driving this shift (Supplementary data Fig. S3), we used dbRDA to determine which abiotic and biotic variables show the strongest correlations with 18S and 16S community structure in the lake. One of the main variables correlating strongly with both communities was daily water discharge (Q) (Supplementary data Table S4). Lower daily discharge rates later in the growing season result in longer hydraulic residence times (HRT) in the lake. For example, HRTs vary from 8 days (27,000 m^3/Day) during peak runoff to 40 days (5,300 m^3/Day) during the late summer and fall in GL4 (Waters, 1999; Miller and McKnight, 2015), and during our study period we saw a near doubling of HRTs from 11.8–19.9 days. Therefore, later in the growing season microbial communities of GL4 are likely more dispersal limited and thereby would be subjected to longer periods during which biotic and abiotic forces could select for unique lake communities (i.e. "variable selection") that would differ from the upstream sources of inoculum and the inlet.

The dbRDA correlation between community structure and Q may also help us understand why the approach of Stegen *et al.* (2013) identified "dispersal limitation" as the other dominant process structuring the transition from the inlet to the lake community (Supplementary data Fig. S3). Longer HRTs during the summer months mean lower flow rates of water through the lake which would slow down convective dispersal in the summer and fall compared to periods of higher run off. In other words, the flux of cells into the lake at the inlet would not be high enough to overcome simple dilution effects as the inlet community mixes with the lake. Future work should be done to test this idea by sampling during periods of high and low discharge rates and by quantifying actual cell numbers in the inlet and the lake using flow cytometry or other methods. However, dispersal limitation does not fully explain the differences observed between the inlet and the body of GL4, and it has been shown *in vitro* that dispersal can homogenize bacterial communities even at low immigration rates (1:5 ratio of migrating to growing cells, Fodelianakis *et al.*, 2019). Despite this fact we did not observe a contribution of homogenizing dispersal in determining community composition between the inlet and lake body (Supplementary data Fig. S3). Future studies should incorporate the observed densities of these populations combined with *in situ* growth rates to address the role of immigration in structuring microbial communities in alpine lakes.

Other work in high-elevation and high latitude lakes has shown a similar pattern of microbial turnover between the inlet and lake body (e.g. Nelson *et al.*, 2009; Adams *et al.*, 2014). These studies attributed this turnover in the microbial community to strong selection or what is often called "species-sorting" effects (Crump *et al.*, 2012), where environmental conditions determine community composition through the selection of metacommunity members. "Species-sorting" can be driven by either abiotic ("bottom-up") or biotic ("top down") factors causing differential selection among members of the microbial community (Grossart *et al.*, 2008; Logue *et al.*, 2012; Peura *et al.*, 2012; Pommier *et al.*, 2012) and we used distance-based Redundancy Analysis (dbRDA: Supplementary data Table S4) and other approaches to explore which biotic and abiotic variables best explain the diversity patterns in GL4. In addition to discharge (discussed above), abiotic variables such as sulfate (SO_4^{2-}), and inorganic nitrogen (IN), explained the most variation in the data. Previous work supports these findings; for example, Miller and McKnight (2015) discussed that historical and present-day shifts in N deposition are responsible for shifts in the phytoplankton communities of GL4, and SO_4^{2-} has been shown to be correlated with drought-associated phytoplankton taxa (Flanagan *et al*., 2009). However, no previous work has examined the influence of abiotic factors on the entire microbial community in GL4. Previous research in other oligotrophic lakes

indicated that bacterial richness is largely controlled by nutrient availability (Lindström *et al.*, 2010; Logue *et al.*, 2012), but more work, such as nutrient addition experiments (Gardner *et al.*, 2008), is needed to better identify the effects of inorganic nutrients on the microbial communities of GL4.

Another possible explanation for the strong selection in GL4 is top-down pressures of predation by higher trophic groups, as has been observed in some (Medina-Sánchez *et al.*, 2004), but not all (Hinder *et al.*, 1999) high-mountain lakes. When we analyzed microbial beta-diversity in relation to biotic variables we found that zooplankton diversity in the lake was the strongest biotic variable explaining 16S community beta-diversity (Supplementary data Tables S3 and S4) perhaps indicating grazing by zooplankton is structuring the composition of the 16S community. In support of this idea, Maxillopoda sequences (e.g. Copepods) were among the dominant 18S phylotypes (and were also visually observed) in GL4 (Supplementary data Table S2), but not in the inlet. Copepods generally don't graze on bacteria but have been shown to supplement their diet with cyanobacteria (Schmidt and Jónasdóttir, 1997), and are readily colonized by bacteria (Tang *et al.*, 2010). In addition, bacteria are probably stimulated to grow by the turnover of nutrients caused by zooplankton grazing. In addition, smaller eukaryotic predators also can play an important role in the top-down structuring of bacterial communities in freshwater habitats (Newton *et al.*, 2011), and there are such predators in our 18S dataset (Supplementary data Table S2). Out of the top 50 most abundant 18S OTUs we observed several bacteriovore taxa such as members of the genus *Gymnodinium* (Löder *et al.*, 2014). Although the presence of these and other predators in GL4 does not prove that they are the main agent of species sorting in this lake, it is likely that they are a factor filtering out members of the inlet microbial community and structuring the microbial community of the lake, especially during periods with longer water residence times.

The potential influence of biotic filters was further highlighted by the observation that autotrophic and heterotrophic taxa showed strong co-occurrence patterns within the lake, but not between the talus soil and the lake (Supplementary data Fig. S2B). For example, the aquatic autotrophic community was primarily made up of the algae *Synura*, *Plagioselmis*, and *Chrysochromulina*, along with unidentified members of the Chrysophyceae, which all showed significant co-occurrence with heterotrophic bacteria such as Flavobacteria and Cytophagaceae. Other work has shown that these bacterial groups co-occur with phytoplankton blooms, and show a preferential use of algal derived (autochthonous) carbon (Riemann and Winding, 2001; Williams *et al.*, 2013). Furthermore, we saw a strong correlation of these autotrophs with a β-proteobacterium OTU belonging to *Polynucleobacter*, which has also been shown to be strongly associated with annual phytoplankton blooms in freshwater lakes (Wu and Hahn, 2006). Therefore, at least the 16S community is very likely being structured by algal derived carbon in GL4 and other lakes.

Finally, the high level of turnover of the microbial community between the inlet and the main body of the lake may have biogeochemical implications for the functioning of this and other lakes that show similar patterns. Microbes entering the lake contain all of the nutrients needed for life, and if they are being turned over by predation they are contributing to the growth of the lake communities through the aquatic "microbial loop" (Callieri *et al.*, 1999; Medina-Sánchez *et al.*, 2004). Release of nutrients due to seasonal turnover of soil microbial communities is one of the major sources of available nutrients in oligotrophic, high-elevation soils (Schadt *et al.*, 2003; Schmidt *et al.*, 2007), but, to our knowledge, no work has been done to assess the release of nutrients from the turnover of allochthonous microorganisms entering a lake. This potentially large source of nutrients could be an important missing piece of the nutrient balance of high-elevation (and other) watersheds (Ley *et al.*, 2004; Mladenov *et al.*, 2012) and a missing component of the landscape continuum model (Seastedt *et al.*, 2004). Of course, community turnover as measured by high-throughput sequencing does not constitute a quantitative method for determining mass or nutrient stoichiometry of the microbial biomass, so future work will be needed to determine these parameters for the inoculum entering the lake. Such work could lead to a new understanding of the biogeochemical and biological connectivity of terrestrial and aquatic ecosystems and allow us to better assess the controls on both microbial community assembly and its contributions to ecosystem dynamics.

Acknowledgements

We thank Robert Spencer, Sarah Power, Brian Straight, and Jessica Henley for field and laboratory assistance, and Diane McKnight and Pacifica Sommers for helpful discussions. Logistical support and meta-data were provided by the NSF supported Niwot Ridge Long-Term Ecological Research project and the University of Colorado Mountain Research Station. Funding was provided by the Niwot Ridge LTER program (NSF DEB-1637686) and grants to study global change effects on high-elevation microbial and plant communities (NSF DEB-1457827 and 1656978).

References

Adams, H.E., Crump, B.C., and Kling, G.W. 2014. Metacommunity dynamics of bacteria in an arctic lake: The impact of species sorting and mass effects on bacterial production and biogeography. *Front. Microbiol.* **5**, 82.

Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., and Huse, S.M. 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.

Angel, A., Vila, I., and Herrera, V. 2016. Extremophiles: photosynthetic systems in a high-altitude saline basin (Altiplano, Chile). *Inter. Aquat. Res.* **8**, 91–108.

Barnes, R.T., Williams, M.W., Parman, J.N., Hill, K., and Caine, N. 2014. Thawing glacial and permafrost features contribute to nitrogen export from Green Lakes Valley, Colorado Front Range, USA. *Biogeochemistry* **117**, 413–430.

Bueno de Mesquita, C.P., Sartwell, S.A., Ordemann, E.V., Porazinska, D.L., Farrer, E.C., King, A.J., Spasojevic, M.J., Smith, J.G., Suding, K.N., and Schmidt, S.K. 2018. Patterns of root colonization by arbuscular mycorrhizal fungi and dark septate endophytes across a mostly-unvegetated, high-elevation landscape. *Fungal Ecol.* **36**,

63–74.
Caine, N. 2010. Recent hydrologic change in a Colorado alpine basin: An indicator of permafrost thaw? *Ann. Glaciol.* **51**, 130–134.
Callieri, C., Pugnetti, A., and Manca, M. 1999. Carbon partitioning in the food web of a high mountain lake: from bacteria to zooplankton. *J. Limnol.* **58**, 144–151.
Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., *et al.* 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336.
Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., *et al.* 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624.
Chase, J.M., Kraft, N., Smith, K.G., and Vellend, M. 2011. Using null models to disentangle variation in community dis-similarity from variation in α-diversity. *Ecosphere* **2**, 1–11.
Chave, J. 2004. Neutral theory and community ecology. *Ecol. Lett.* **7**, 241–253.
Comte, J., Lovejoy, C., Crevecoeur, S., and Vincent, W.F. 2015. Co-occurrence patterns in aquatic bacterial communities across changing permafrost landscapes. *Biogeosciences* **13**, 175–190.
Crump, B.C., Amaral-Zettler, L.A., and Kling, G.W. 2012. Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J.* **6**, 1629–1639.
Darcy, J.L., King, A.J., Gendron, E.M.S., and Schmidt, S.K. 2017. Spatial autocorrelation of microbial communities atop a debris-covered glacier is evidence of a supraglacial chronosequence. *FEMS Microbiol. Ecol.* **93**, 1–11.
Edgar, R.C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461.
Faith, D.P. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10.
Farjalla, V.F., Srivastava, D.S., Marino, N.A.C., Azevedo, F.D., Dib, V., Lopes, P.M., Rosado, A.S., Bozelli, R.L., and Esteves, F.A. 2012. Ecological determinism increases with organism size. *Ecology* **93**, 1752–1759.
Flanagan, C.M., McKnight, D.M., Liptzin, D., Williams, M.W., and Miller, M.P. 2009. Response of the phytoplankton community in an Alpine Lake to drought conditions: Colorado Rocky Mountain Front Range, USA. *Arct. Antarct. Alp. Res.* **41**, 191–203.
Fodelianakis, S., Lorz, A., Valenzuela-Cuevas, A., Barozzi, A., Booth, J.M., and Daffonchio, D. 2019. Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. *Nat. Commun.* **10**, 1314.
Foley, B., Jones, I.D., Maberly, S.C., and Rippey, B. 2012. Long-term changes in oxygen depletion in a small temperate lake: Effects of climate change and eutrophication. *Freshwater Biol.* **57**, 278–289.
Freeman, K.R., Pescador, M.Y., Reed, S.C., Costello, E.K., Robeson, M.S., and Schmidt, S.K. 2009. Soil CO_2 flux and photoautotrophic community composition in high-elevation, 'barren' soil. *Environ. Microbiol.* **11**, 674–686.
Gardner, E.M., McKnight, D.M., Lewis, W.M., and Miller, M.P. 2008. Effects of nutrient enrichment on phytoplankton in an alpine lake, Colorado, USA. *Arct. Antarct. Alp. Res.* **40**, 55–64.
Grossart, H.P., Jezbera, J., Horňák, K., Hutalle, K.M.L., Buck, U., and Šimek, K. 2008. Top-down and bottom-up induced shifts in bacterial abundance, production and community composition in an experimentally divided humic lake. *Environ. Microbiol.* **10**, 635–652.
Hinder, B., Baur, I., Hanselmann, K., and Schanz, F. 1999. Microbial food web in an oligotrophic high mountain lake (Jöri Lake III, Switzerland). *J. Limnol.* **58**, 162–168.
Hu, Y., Cai, J., Bai, C., Shao, K., Tang, X., and Gao, G. 2018. Contrasting patterns of the bacterial and archaeal communities in a high-elevation river in northwestern China. *J. Microbiol.* **56**, 104–112.
Hubbell, S.P. 2001. The unified neutral theory of biodiversity and biogeography, p. 29. Princeton University Press, Princeton, NJ, USA.
Kammerlander, B., Breiner, H., Filker, S., Sommaruga, R., Sonntag, B., and Stoeck, T. 2015. High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalayan Mountains. *FEMS Microbiol. Ecol.* **91**, 1–10.
Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., and Webb, C.O. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464.
King, A.J., Farrer, E.C., Suding, K.N., and Schmidt, S.K. 2012. Co-occurrence patterns of plants and soil bacteria in the high-alpine subnival zone track environmental harshness. *Front. Microbiol.* **3**, 347.
King, A.J., Freeman, K.R., McCormick, K.F., Lynch, R.C., Lozupone, C., Knight, R., and Schmidt, S.K. 2010. Biogeography and habitat modelling of high-alpine bacteria. *Nat. Commun.* **1**, 53.
Langenheder, S. and Ragnarsson, H. 2007. The role of environmental and spatial factors for the composition of aquatic bacterial communities. *Ecology* **88**, 2154–2161.
Lefèvre, E., Roussel, B., Amblard, C., and Sime-Ngando, T. 2008. The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton. *PLoS One* **3**, e2324.
Legendre, P. 2014. Package 'lmodel2.' Model II Regression. R package version 1.7-2.
Legendre, P. and Anderson, M.J. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* **69**, 1–24.
Ley, R.E., Lipson, D.A., and Schmidt, S.K. 2001. Microbial biomass levels in barren and vegetated high-altitude talus soils. *Soil Sci. Soc. Am. J.* **65**, 111–117.
Ley, R.E., Williams, M.W., and Schmidt, S.K. 2004. Microbial population dynamics in an extreme environment: controlling factors in talus soils at 3750 m in the Colorado Rocky Mountains. *Biogeochemistry* **68**, 297–311.
Lindström, E.S., Feng, X.M., Granéli, W., and Kritzberg, E.S. 2010. The interplay between bacterial community composition and the environment determining function of inland water bacteria. *Limnol. Oceanogr.* **55**, 2052–2060.
Liu, F., Williams, M.W., and Caine, N. 2004. Source waters and flow paths in an alpine catchment, Colorado Front Range, United States. *Water Resour. Res.* **40**, 1–16.
Liu, Y., Yao, T., Jiao, N., Liu, X., Kang, S., and Luo, T. 2013. Seasonal dynamics of the bacterial community in Lake Namco, the largest Tibetan lake. *Geomicrobiol. J.* **30**, 17–28.
Löder, M.G.J., Boersma, M., Kraberg, A.C., Aberle, N., and Wiltshire, K.H. 2014. Microbial predators promote their competitors: commensalism within an intra-guild predation system in microzooplankton. *Ecosphere* **5**, 1–23.
Logue, J.B., Langenheder, S., Andersson, A.F., Bertilsson, S., Drakare, S., Lanzén, A., and Lindström, E.S. 2012. Freshwater bacterioplankton richness in oligotrophic lakes depends on nutrient availability rather than on species-area relationships. *ISME J.* **6**, 1127–1136.
Loria, K. 2019. Stream and lake water chemistry data for Green Lakes Valley, 1998–ongoing. Environmental Data Initiative. https://doi.org/10.6073/pasta/974f109b0e0e658ebb77da1c62c8f4bc (Accessed 27 March 2019).
Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J., and Knight, R. 2011. UniFrac: an effective distance metric for microbial community comparison. *ISME J.* **5**, 169–172.
McKnight, D.M. 2016. Stream and lake water chemistry data for Green Lake 1 Green Lake 3 Green Lake 4 Green Lake 5 from 1998–ongoing, weekly. http://niwot.colorado.edu
Medina-Sánchez, J.M., Villar-Argaiz, M., and Carrillo, P. 2004. Nei-

ther with nor without you: A complex algal control on bacterioplankton in a high mountain lake. *Limnol. Oceanogr.* **49**, 1722–1733.

Miller, M.P. and McKnight, D.M. 2010. Comparison of seasonal changes in fluorescent dissolved organic matter among aquatic lake and stream sites in the Green Lakes Valley. *J. Geophys. Res. Biogeosci.* **115**, G00F12.

Miller, M.P. and McKnight, D.M. 2015. Limnology of the Green Lakes Valley: Phytoplankton ecology and dissolved organic matter biogeochemistry at a long-term ecological research site. *Plant Ecol. Divers.* **8**, 689–702.

Miller, M.P., McKnight, D.M., Cory, R.M., Williams, M.W., and Runkel, R.L. 2006. Hyporheic exchange and fulvic acid redox reactions in an alpine stream/wetland ecosystem, Colorado Front Range. *Environ. Sci. Technol.* **40**, 5943–5949.

Mladenov, N., Williams, M.W., Schmidt, S.K., and Cawley, K. 2012. Atmospheric deposition as a source of carbon and nutrients to an alpine catchment of the Colorado Rocky Mountains. *Biogeosciences* **9**, 3337–3355.

Molotch, N.P., Meixner, T., and Williams, M.W. 2008. Estimating stream chemistry during the snowmelt pulse using a spatially distributed, coupled snowmelt and hydrochemical modeling approach. *Water Resour. Res.* **44**, 1–14.

Naff, C.S., Darcy, J.L., and Schmidt, S.K. 2013. Phylogeny and biogeography of an uncultured clade of snow chytrids. *Environ. Microbiol.* **15**, 2672–2680.

Nelson, C.E., Sadro, S., and Melack, J.M. 2009. Contrasting the influences of stream inputs and landscape position on bacterioplankton community structure and dissolved organic matter composition in high-elevation lake chains. *Limnol. Oceanogr.* **54**, 1292–1305.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P., *et al.* 2013. Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* **77**, 342–356.

Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., and Bertilsson, S. 2011. A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 14–49.

Peura, S., Eiler, A., Hiltunen, M., Nykanen, H., Tiirola, M., and Jones, R.I. 2012. Bacterial and phytoplankton responses to nutrient amendments in a Boreal lake differ according to season and to taxonomic resolution. *PLoS One* **7**, e38552.

Pommier, T., Douzery, E.J.P., and Mouillot, D. 2012. Environment drives high phylogenetic turnover among oceanic bacterial communities. *Biol. Lett.* **8**, 562–566.

Porazinska, D.L., Farrer, E.C., Spasojevic, M.J., Bueno de Mesquita, C.P., Sartwell, S.A., Smith, J.G., White, C.T., King, A.J., Suding, K.N., and Schmidt, S.K. 2018. Plant diversity and density predict belowground diversity and function in an early successional alpine ecosystem. *Ecology* **99**, 1942–1952.

Preston, D.L., Caine, N., McKnight, D.M., Williams, M.W., Hell, K., Miller, M.P., Hart, S.J., and Johnson, P.T.J. 2016. Climate regulates alpine lake ice cover phenology and aquatic ecosystem structure. *Geophys. Res. Lett.* **43**, 5353–5360.

Price, M.N., Dehal, P.S., and Arkin, A.P. 2010. FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490.

R Core Team. 2016. R: A language and environment for statistical computing. Vienna, Austria: R foundation for statistical computing. http://www.R-project.org/ (Accessed 22 November 2018).

Remias, D., Jost, S., Boenigk, J., Wastian, J., and Lütz, C. 2013. *Hydrurus*-related golden algae (Chrysophyceae) cause yellow snow in polar summer snowfields. *Phycol. Res.* **61**, 277–285.

Riemann, L. and Winding, A. 2001. Community dynamics of free-living and particle-associated bacterial assemblages during a freshwater phytoplankton bloom. *Microb. Ecol.* **42**, 274–285.

Rott, E., Cantonati, M., Füreder, L., and Pfister, P. 2006. Benthic algae in high altitude streams of the Alps - a neglected component of the aquatic biota. *Hydrobiologia* **562**, 195–216.

Ruiz-Gonzalez, C., Nino-Garcia, J.P., and Giorgio, P.A. 2015. Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol. Lett.* **18**, 1198–1206.

Schadt, C., Martin, A.P., Lipson, D.A., and Schmidt, S.K. 2003. Seasonal dynamics of previously unknown fungal lineages in tundra soils. *Science* **301**, 1359–1361.

Schmidt, S.K., Costello, E.K., Nemergut, D.R., Cleveland, C.C., Reed, S.C., Weintraub, M.N., Meyer, A.F., and Martin, A.M. 2007. Biogeochemical consequences of rapid microbial turnover and seasonal succession in soil. *Ecology* **88**, 1379–1385.

Schmidt, K. and Jónasdóttir, S.H. 1997. Nutritional quality of two cyanobacteria: How rich is 'poor' food? *Mar. Ecol. Prog. Ser.* **151**, 1–10.

Seastedt, T.R., Bowman, W.D., Caine, N., McKnight, D.M., Townsend, A., and Williams, M.W. 2004. The landscape continuum: a model for high-elevation ecosystems. *BioScience* **54**, 111–121.

Šlapeta, J., Moreira, D., and López-García, P. 2005. The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proc. Roy. Soc. B* **272**, 2073–2081.

Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., and Curtis, T.P. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.* **8**, 732–740.

Stanish, L.F., O'Neill, S.P., Gonzalez, A., Legg, T.M., Knelman, J., McKnight, D.M., Spaulding, S., and Nemergut, D.R. 2013. Bacteria and diatom co-occurrence patterns in microbial mats from polar desert streams. *Environ. Microbiol.* **15**, 1115–1131.

Stegen, J.C., Lin, X., Fredrickson, J.K., Chen, X., Kennedy, D.W., Murray, C.J., Rockhold, M.L., and Konopka, A. 2013. Quantifying community assembly processes and identifying features that impose them. *ISME J.* **7**, 2069–2079.

Stegen, J.C., Lin, X., Fredrickson, J.K., and Konopka, A.E. 2015. Estimating and mapping ecological processes influencing microbial community assembly. *Front. Microbiol.* **6**, 1–15.

Tang, K.W., Turk, V., and Grossart, H.P. 2010. Linkage between crustacean zooplankton and aquatic bacteria. *Aquat. Microb. Ecol.* **61**, 261–277.

Triadó-Margarit, X. and Casamayor, E.O. 2012. Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environ. Microbiol.* **14**, 2445–2456.

Venkataraman, A., Bassis, C.M., Beck, J.M., Young, V.B., Curtis, J.L., Huffnagle, G.B., and Schmidt, T.M. 2015. Application of a neutral community model to assess structuring of the human lung microbiome. *mBio* **6**, 1–9.

Vestheim, H. and Jarman, S.N. 2008. Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. *Front. Zool.* **5**, 1.

Vila-Costa, M., Barberan, A., Auguet, J.C., Sharma, S., Moran, M.A., and Casamayor, E.O. 2013. Bacterial and archaeal community structure in the surface microlayer of high mountain lakes examined under two atmospheric aerosol loading scenarios. *FEMS Microbiol. Ecol.* **84**, 387–397.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., *et al.* 2016. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* **1**, e00009-15.

Wang, J., Wang, F., Chu, L., Wang, H., Zhong, Z., Liu, Z., Gao, J., and Duan, H. 2014. High genetic diversity and novelty in eukaryotic plankton assemblages inhabiting saline lakes in the Qaidam Basin. *PLoS One* **9**, e112812.

Waters, S.B. 1999. Master's thesis. Responses of algal communities to environmental change in an alpine lake, Green Lakes Valley, Colorado. University of Colorado, Boulder, USA.

Webb, C.O., Ackerly, D.D., and Kembel, S.W. 2008. Phylocom: soft-

ware for the analysis of phy-logenetic community structure and trait evolution. Version 4.0.1. http://www.phylodiversity.net/phylocom/

Whitman, T., Neurath, R., Perera, A., Chu-Jacoby, I., Ning, D., Zhou, J., Nico, P., Pett-Ridge, J., and Firestone, M. 2018. Microbial community assembly differs across minerals in a rhizosphere microcosm. *Environ. Microbiol.* **20**, 4444–4460.

Williams, M.W., Hood, E., Molotch, N., Caine, N., Cowie, R., and Liu, F. 2015. The 'teflon basin' myth: hydrology and hydrochemistry of a seasonally snow-covered catchment. *Plant Ecol. Divers.* **8**, 639–661.

Williams, M.W., Losleben, M., Caine, N., and Greenland, D. 1996. Changes in climate and hydro-chemical responses in a high-elevation catchment in the Rocky Mountains, USA. *Limnol. Oceanogr.* **41**, 939–946.

Williams, T.J., Wilkins, D., Long, E., Evans, F., Demaere, M.Z., Raftery, M.J., and Cavicchioli, R. 2013. The role of planktonic *Flavobacteria* in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environ. Microbiol.* **15**, 1302–1317.

Wu, Q.L. and Hahn, M.W. 2006. High predictability of the seasonal dynamics of a species-like *Polynucleobacter* population in a freshwater lake. *Environ. Microbiol.* **8**, 1660–1666.

Wu, W., Lu, H.P., Sastri, A., Yeh, Y.C., Gong, G.C., Chou, W.C., and Hsieh, C.H. 2017. Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME J.* **12**, 485–494.

Wu, Q.L., Zwart, G., Schauer, M., Kamst-van Agterveld, M.P., and Hahn, M.W. 2006. Bacterioplankton community composition along a salinity gradient of sixteen High-Mountain Lakes located on the Tibetan Plateau, China. *Appl. Environ. Microbiol.* **72**, 5478–5485.

Zapala, M.A. and Schork, N.J. 2006. Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. *Proc. Natl. Acad. Sci. USA* **103**, 19430–19435.

Zhong, Z.P., Liu, Y., Miao, L.L., Wang, F., Chu, L.M., Wang, J.L., and Liu, Z.P. 2016. Prokaryotic community structure driven by salinity and ionic concentrations in plateau lakes of the Tibetan Plateau. *Appl. Environ. Microbiol.* **82**, 1846–1858.