

Northumbria Research Link

Citation: Aguirre de Cárcer, Daniel, Pedrós-Alió, Carlos, Pearce, David and Alcamí, Antonio (2016) Composition and interactions among bacterial, microeukaryotic and T4-like viral assemblages in lakes from both polar zones. *Frontiers in Microbiology*, 7 (337). ISSN 1664-302X

Published by: Frontiers

URL: <http://doi.org/10.3389/fmicb.2016.00337> <<http://doi.org/10.3389/fmicb.2016.00337>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/26307/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

www.northumbria.ac.uk/nrl



Composition and interactions among bacterial, microeukaryotic and T4-like viral assemblages in lakes from both polar zones

Daniel Aguirre De Cárcer¹, Carlos Pedrós-Alió², David A. Pearce³, Antonio Alcami^{1*}

¹Centro de Biología Molecular Severo Ochoa, CSIC., Spain, ²Institut de Ciències del Mar, CSIC, Spain, ³British Antarctic Survey, Natural Environment Research Council, United Kingdom

Submitted to Journal:
Frontiers in Microbiology

Specialty Section:
Aquatic Microbiology

ISSN:
1664-302X

Article type:
Original Research Article

Received on:
21 Dec 2015

Accepted on:
03 Mar 2016

Provisional PDF published on:
03 Mar 2016

Frontiers website link:
www.frontiersin.org

Citation:

Aguirre_de_cárcer D, Pedrós-alió C, Pearce DA and Alcami A(2016) Composition and interactions among bacterial, microeukaryotic and T4-like viral assemblages in lakes from both polar zones. *Front. Microbiol.* 7:337. doi:10.3389/fmicb.2016.00337

Copyright statement:

© 2016 Aguirre_de_cárcer, Pedrós-alió, Pearce and Alcami. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Provisional

1 **Composition and interactions among bacterial, microeukaryotic**
2 **and T4-like viral assemblages in lakes from both polar zones**

3
4
5
6
7 **Daniel Aguirre de Cárcer¹, Carlos Pedrós-Alió², David A. Pearce^{3,4,5} & Antonio**
8 **Alcami^{1,*}**

9
10 ¹Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones
11 Científicas (CSIC)–Universidad Autónoma de Madrid, Madrid, Spain.

12 ²Institut de Ciències del Mar, CSIC, Barcelona, Spain.

13 ³British Antarctic Survey, Natural Environment Research Council, Cambridge, UK.

14 ⁴Faculty of Health and Life Sciences, University of Northumbria, Newcastle Upon Tyne,
15 UK.

16 ⁵University Center in Svalbard, Loneyarbyen, Norway.

17
18
19 ***Correspondence:**

20 Antonio Alcamí

21 aalcami@cbm.csic.es

22
23
24
Provisional

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

ABSTRACT

In this study we assess global biogeography and correlation patterns among three components of microbial life: bacteria, microeukaryotes, and T4-like myoviruses. In addition to environmental and biogeographical considerations, we have focused our study on samples from high-latitude pristine lakes from both poles, since these simple island-like ecosystems represent ideal ecological models to probe the relationships among microbial components and with the environment. Bacterial assemblages were dominated by members of the same groups found to dominate freshwater ecosystems elsewhere, and microeukaryotic assemblages were dominated by photosynthetic microalgae. Despite inter-lake variations in community composition, the overall percentages of OTUs shared among sites was remarkable, indicating that many microeukaryotic, bacterial, and viral OTUs are globally-distributed. We observed an intriguing negative correlation between bacterial and microeukaryotic diversity values. Notably, our analyses show significant global correlations between bacterial and microeukaryotic community structures, and between the phylogenetic compositions of bacterial and T4-like virus assemblages. Overall, environmental filtering emerged as the main factor driving community structures.

Keywords

Bacteria, Microeukaryotes, T4 phages, Pyrosequencing, Diversity, Lakes, Polar

Provisional

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94

1. INTRODUCTION

The functioning of aquatic ecosystems is heavily dependent on the microbial food web (Azam et al., 1983), which consists of several components with different cellular (bacteria, archaea, and eukaryotes) or non-cellular (viruses) organizations. Polar freshwater environments embody some of the least human-impacted habitats on the planet (Convey and Stevens, 2007), and normally represent simple ecosystems with truncated food webs dominated by microorganisms. Arctic and Antarctic freshwater environments share common features, such as extreme annual cycles of temperature, sunlight, and ice phenology, and yet they are separated by long geographical distances. These characteristics make these environments a unique model to shed light on fundamental questions of aquatic microbial ecology, such as how communities vary across spatial scales and environmental gradients, or the association among different components of microbial life.

Previous studies assessing the links between different components of microbial life have focused on local patterns, such as the microbial inventory of a single coral reef ecosystem (McCliment et al., 2012), or the time-series analysis of a specific marine site (Jones et al., 2013; Chow et al., 2014). All these studies reported co-occurrence patterns among members of the different components of microbial life. Another significant question in aquatic microbial ecology is the importance of dispersal limitations (Martiny et al., 2011). However, most studies tackling this issue have focused on marine samples, presumably showing enhanced connectivity related to global oceanic circulation (Sul et al., 2011; Ghiglione et al., 2012). On the other hand, the question of whether microbial species in freshwater bodies from both poles are the same or different is also of interest from the viewpoint of biogeography and genetic exchange (Bano and Hollibaugh, 2002).

The aim of this study is two-fold; first, to assess if similar microbial communities populate freshwater bodies from both polar zones. Second, to test whether the previously observed co-occurrence patterns among components of microbial life appear at a global scale, or instead represent local phenomena. To do so, we have studied the resident microbial communities from four Arctic lakes located in the Svalbard archipelago, and nine Antarctic lakes sampled across a latitudinal transect along the Antarctic Peninsula. Using a massive parallel sequencing approach targeting phylogenetic marker genes, we have studied the community structures of three components of microbial life: bacteria, microeukaryotes, and viruses. Since the latter lack a common phylogenetic marker gene, we have focused on the T4-like myovirus group. This group represents a diverse and abundant (Filee et al., 2005; Williamson et al., 2008) group of bacteria-infecting viruses (Wichels et al., 1998; Clokie et al., 2010) amenable to the proposed approach based on their g23 major capsid gene, which has been shown to serve as a phylogenetic proxy (Comeau and Krisch, 2008). After obtaining the community profiles, we carried out several analyses, both at the OTU level and based on overall community phylogenetic compositions, to study the relationships among the three components of microbial life studied, and with respect to biogeographical and limnological (both physical and chemical) factors.

2. MATERIALS AND METHODS

2.1. Sampling and community DNA extraction

95

96 Planktonic samples were taken from freshwater bodies in the Arctic (Spitsbergen Island,
97 Svalbard archipelago, Norway), and along the Antarctic Peninsula (Table 1, Suppl. Fig. 1).
98 Lake Nordammen (SvL1) was completely frozen (samples represent a combination of
99 existing melted top ice from three different sites), while Lake Tenndammen (SvL2) is a
100 shallow lake with frozen surface at the time of sampling. All the other lakes had open
101 waters at the time of sampling and representative samples of the water column were taken
102 at different depths, except for Limnopolar lake where water was taken from 4 m depth. All
103 the lakes were sampled around summer. A portable probe was deployed *in situ* before
104 sampling the Antarctic sites to measure several limnological variables (Temperature,
105 Conductivity, pH, Chlorophyll; Suppl. Table 1). Lake Green's measurement was deemed
106 unreliable due to technical issues and hence data from this lake were removed from those
107 analyses including limnological data.

108 Single ninety liter samples from each water body were filtered through a 30 μm nylon
109 mesh. Subsequent filtration by 0.45 μm tangential flow filtration (TFF) using a Centramate
110 holder (Pall) separated the free viral community (defined as the <0.45 μm fraction) from
111 the cellular (and associated virus) fraction (defined as the fraction between 0.45 and 30
112 μm). Viral fractions were subsequently concentrated 100 times by 70-kDa TFF as
113 described (López-Bueno et al. 2009; Aguirre de Cárcer et al. 2015). All samples were
114 preserved at -20°C/-80°C prior to DNA extraction. Cellular fraction community DNA was
115 extracted using PowerSoil DNA isolation kit (MoBio, Carlsbad, CA) according to the
116 manufacturer's instructions. Viral DNA was obtained from frozen stocks thawed at 4°C,
117 and passed through a 25% sucrose cushion by centrifugation for 16 h at 60 000 g and 4°C.
118 The pellets were re-suspended in 10 mM Tris pH 8, 1 mM EDTA, and filtered using a 0.45
119 μm syringe filter. Viral concentrates were then treated with DNase I (500 U.ml⁻¹),
120 Nuclease S7 (420 U.ml⁻¹), RNase A (100 $\mu\text{g.ml}^{-1}$) and RNase H (2 U per reaction) for 30
121 min at room temperature to remove free nucleic acids. Nuclease reactions were stopped
122 with 12 mM EDTA/2 mM EGTA, and viral capsids and envelopes were then disrupted
123 with SDS (0.5%) and proteinase K (200 $\mu\text{g.ml}^{-1}$) treatment. Finally, viral DNA was
124 extracted with phenol-chloroform and ethanol precipitated.

125

126 **2.2. PCR amplification of community marker genes, and massive parallel sequencing**

127

128 Bacterial 16S rRNA marker genes were amplified from the cellular fractions using primers
129 8F15B (5'-[RocheAdapterB]AGAGTTTGATCCTGG-3') and 515R14AM (5'-
130 [RocheAdapterA]-bc-TTACCGCGGCTGCT-3') (Aguirre de Cárcer et al., 2011b). All
131 PCR reactions were carried out using 1 μl of template DNA, 0.5 μl Phusion High-fidelity
132 polymerase (NEB), 20 nmol dNTPs, 20 pmol of each primer, 1.5 μl DMSO, 0.4 mM
133 MgCl₂, in a final volume of 50 μl . Reaction conditions included an initial denaturation step
134 of 30s at 98°C, followed by 25 cycles of 10s at 98°C, 30s at 53°C, 30s at 72°C, and a final
135 elongation step of 5 min at 72°C. For the analysis of the microeukaryotic and T4-like viral
136 assemblages we followed a validated two-step barcoding strategy (Aguirre de Cárcer et al.,
137 2011b) that allows a universal set of bar-coded sequencing primers to be appended to an
138 amplified PCR product without introducing discernible biases. In the first step, one of the
139 target-specific primers is modified to include a linker sequence. After amplification, a
140 second primer consisting of the bar code and linker is used to tag the amplicon. The
141 eukaryotic community in the cellular fractions was analyzed using primer sequences
142 targeting the 18S rRNA gene (Bates et al., 2013) (F515; 5'-
143 GTGCCAGCMGCCGCGGTAA-3' and R1119 5'-GGTGCCCTTCCGTCA-3'). The T4-
144 like virus communities in both cellular and free viral fractions were assessed using primer

145 sequences (Filee et al., 2005) targeting the g23 major capsid protein gene (MZIA1bis 5'-
146 GATATTTGIGGIGTTTCAGCCCIATGA-3' and MZIA6 5'-
147 CGCGGTTGATTTCCAGCATGATTTTC-3'). In this case, we performed an initial
148 amplification using the unmodified (no linker or 454-adaptor sequences) primers (45
149 cycles, melting temperature of 50°C), followed by agarose gel extraction of DNA bands of
150 appropriate size. The resulting products were re-amplified using the modified primers (10
151 cycles), at this point linking with the two-step protocol. The primer pairs employed in this
152 study have been reported to target most known sequences within their target groups. Final
153 concentrations of PCR products were measured using a PicoGreen dsDNA Assay Kit (Life
154 Tech.), equal amounts for each sample pooled, agarose gel-extracted using the QIAquick
155 Gel Extraction Kit (QIAGEN), and sequenced using a Roche 454 FLX sequencer with
156 titanium chemistry. Bacterial and Eukaryotic profiles from each lake were generated in
157 triplicate (hence three different barcodes/reactions per sample) to mitigate potential
158 reaction-level PCR biases (Bates et al., 2013). All sequences have been deposited at ENA
159 under id PRJEB10639.

160

161 **2.3. Sequence processing and data analysis**

162

163 The sequences obtained were denoised with ACACIA (Bragg et al., 2012) then processed
164 using QIIME (Caporaso et al., 2010). Sequences were first assigned to each sample using
165 their respective barcodes. They were next filtered for correct length and quality values
166 (Maximum number of ambiguous bases; 5. Mean quality score; >25. Maximum
167 homopolymer length; 6. Maximum mismatches in primer; 0. Chimera removal with
168 *usearch*). Later, all sequences were grouped into operational taxonomic units (OTUs) at
169 0.97 distance thresholds using the Uclust algorithm, and OTUs not appearing in at least
170 two replicates across each dataset were discarded to eliminate noise and possible artifacts.
171 At this point, the results from the per-sample technical triplicates were pooled to obtain
172 per-sample community profiles. The resulting sample-by-OTUs matrices were subsampled
173 to the minimum number of sequences at any given site (independently for each marker
174 gene dataset) to normalize sampling efforts, and singletons were removed (Aguirre de
175 Cárcer et al., 2011a). Additionally, representative sequences from each bacterial and
176 eukaryotic OTU (the most abundant sequence of that OTU across the dataset) were
177 confronted against the Greengenes (DeSantis et al., 2006)/Silva (Quast et al., 2013)
178 reference alignments (for 16S and 18S sequences respectively) for taxonomic affiliation.
179 Shannon diversity indices, Chao1 richness estimates, and percentage of shared OTUs
180 between samples were obtained using dedicated QIIME scripts.

181

182 Statistical and analytical procedures were carried out in *R* (R Core Team 2013).

183 Overall differences between poles or sampling fractions (only for T4-like viruses) in terms
184 of richness, diversity, percentage of shared OTUs, and relative abundance of major
185 eukaryotic taxa were assessed using t tests (*t.test* function).

186

187 The relationships between diversity values and latitude, limnological variables, or the
188 relative abundance of major taxonomic groups were explored through linear regression.
189 Bootstrapping was conducted to test whether or not the slope of the regression was
190 significantly different from zero (*boot* function within package of the same name (Davison
191 and Hinkley, 1997)), and alternatively, performing a Wald test for multiple coefficients
192 (*f.robftest* function from package *sfsmisc*).

193 The existence of correlation between percentages of shared OTUs and geographical
194 distance or limnological profiles was assessed through Mantel tests (*mantel.randtest*,
195 package *ade4*).

196 The initial exploration of community structure data was carried-out through double
197 principal coordinates analysis (DPCoA) using functions from the *ade4* package. This is an
198 ordination method that takes into account phylogenetic (genetic distance) relatedness
199 between OTUs when explaining variation in the data, hence quantifying community
200 dissimilarity based on phylogenetic relatedness. Statistical significance of *a priori*
201 community groupings (e.g. pole of origin) was tested by between class analysis (BCA) and
202 constrained double principal coordinates analysis (cDPCoA) (phylogenetic-aware) (Dray et
203 al., 2015) available in *ade4*. Identification of limnological variables correlated with
204 community structure was undertaken by analysis with instrumental variables (IV) (Baty et
205 al., 2006) using the *pcaiv* function of the same package. Finally, associations among
206 community structures derived from different marker genes, as well as with existing
207 environmental data (limnological profiles, geographic distance, latitude) were assessed
208 with Mantel tests using Bray-Curtis and Rao (phylogenetic-aware) distance matrices
209 derived from the community profiles.

210

211 **3. RESULTS**

212

213 **3.1. OTUs richness and diversity**

214

215 Samples from four water bodies in Spitsbergen Island (78°N, Svalbard, Norway) and nine
216 water bodies along the Antarctic Peninsula (62°-67°S) were collected in three different
217 years (Table 1, Suppl. Fig. 1). From these samples, community DNA was prepared using
218 both cellular and free virus fractions, and bacterial 16S rRNA, eukaryotic 18S rRNA and
219 T4-like virus g23 major capsid protein genes were PCR-amplified and sequenced using
220 Roche 454 technology. The strategy employed produced 78000 (16S), 23000 (18S), and
221 38000 (g23) sequences passing the initial quality control. Subsequent filtering resulted in
222 5302 ± 1040 (16S), 1805 ± 440 (18S), and 1514 ± 496 (g23) sequences per sample (Suppl.
223 Table 1). However, we were unable to produce reliable g23 sequence data for a small
224 subset of the sites and fractions (4 out of 26), and hence these data points were removed
225 from some analyses. Clustering of sequences at 97% similarity produced 1548 (16S), 704
226 (18S), and 176 (g23) OTUs in total. The number of bacterial OTUs per sample ranged
227 between 126 in Horseshoe Lake and 864 in Lake SvL1 (Suppl. Table 1). The number of
228 singletons was moderate in all cases, resulting in Chao 1 estimates that were only about
229 twice as large as the retrieved OTUs (ranging between 121 and 1466). Corresponding
230 values were lower for eukaryotes with values ranging between 87 OTUs in Lakes Biscoe
231 and SvL2 and 418 in Lake IR2 (average 197), and even lower for viruses. Interestingly,
232 viral richness was higher in the cell-associated (average 33 OTUs) than in the free viral
233 fractions (average 21 OTUs, Suppl. Table 1) although difference did not reach statistical
234 significance ($p > 0.05$).

235

236 The microbial communities studied shared a noticeable percentage of OTUs: averaging 14,
237 10.2 and 5.8 % for bacteria, eukarya and viruses respectively (Table 2). Communities from
238 the same polar zone shared a significantly higher percentage of OTUs. The only exception
239 was the Arctic bacterial assemblages, which only shared approximately 10% of their OTUs
240 among themselves. In the case of T4-like viral assemblages, a higher percentage of shared
241 OTUs between the free viral and cellular fractions arising from the same sample was
242 observed when compared to the inter-lake average (31.4% vs. overall 5.8%, $p < 0.0005$,

243 respectively). For the Antarctic bacterial and microeukaryotic data sets, we found no
244 significant correlation between the percentage of shared OTUs and either geographical
245 distance or limnological profiles based on temperature, conductivity, pH, and Chlorophyll
246 (Suppl. Table 2).

247

248 Shannon's entropy index was used to study community diversity (Table 3). No significant
249 differences were found when comparing diversity values between the two polar zones,
250 either for the bacterial (Antarctic; 4.7 ± 1.1 , Arctic; 4.7 ± 0.8) or microeukaryotic
251 assemblages (Antarctic; 3.9 ± 1.1 , Arctic; 4.0 ± 1.4). Also, for the Antarctic datasets, no
252 correlation was observed between community diversity values and latitude along the
253 studied transect. Moreover, observed diversity values and recorded limnological variables
254 (Suppl. Table 2) only correlated in the case of 16S-based bacterial community diversity
255 and Chlorophyll *a* content, where a positive correlation was found (Slope 0.85, $R^2=0.3$),
256 although the observed correlation only reached significance ($p<0.05$) with one of the two
257 statistical tests applied (Figure 1A). Strikingly, a negative correlation (Slope -0.54, $R^2=$
258 0.21) was observed between bacterial and microeukaryotic diversity values (Figure 1B),
259 with the negative association reaching a noticeable degree of significance with both
260 statistical tests applied (linear regression bootstrapping $p<0.05$; Wald test $p = 0.078$).
261 Subsequent analyses failed to show significant correlations between diversity indices of
262 bacterial and microeukaryotic assemblages and the relative abundances of any major
263 taxonomic groups.

264 In the case of the T4-like viral assemblages (Table 3) no significant differences in diversity
265 estimates were observed with regards to pole of origin. 16S and 18S-based values obtained
266 for the same samples gave no significant correlations when compared to T4-like virus
267 values. Moreover, geographical (distance, latitude) and limnological parameters available
268 for the Antarctic sites did not show an influence on T4-like virus community diversity.
269 Lastly, the cellular fractions were more diverse than their free viral fraction counterparts,
270 yet such differences did not reach statistical significance ($p= 0.11$). One peculiar case was
271 that of the free viral fraction of Lake Biscoe that consisted of a single OTU. This OTU was
272 also detected at 11% relative abundance in the cellular fraction of Lake SvL1 in the Arctic
273 suggesting that it was not an artifact.

274

275 **3.2. Taxonomic compositions**

276

277 Bacterial community profiles (Figure 2A) were dominated by sequences classified as order
278 Burkholderiales (39%), containing mainly representatives of families Oxalobacteriaceae
279 (13%) and Comamonadaceae (22%). Bacteroidetes-related sequences were also very
280 abundant (35%), dominated by Flavobacteriaceae (29%) and Sphingobacteriales (5%).
281 Actinobacteria-related sequences were abundant (8%), mainly corresponding to the
282 Microbacteriaceae (3%) and ACK-M1 (4%) clades. Finally, Cyanobacteria-related
283 sequences averaged 8%, yet a closer inspection revealed that, although a few sequences
284 belonged to the Synechococcophycidae and Nostocaceae clades, the great majority (94%)
285 corresponded to algal chloroplast-related sequences. When the taxonomic composition was
286 examined separately for each lake, substantial differences emerged (Figure 3A). The
287 predominance of Betaproteobacteria and Bacteroidetes was apparent, but some lakes were
288 dominated by the former, such as Cierva, Horsheshoe, SvL2, and Sv11, while others were
289 dominated by Bacteroidetes, like Domo, Pourquois-Pas, and IR2. Chloroplasts appeared
290 only in low-latitude Antarctic lakes. Finally, Actinobacteria were abundant in several
291 Antarctic lakes but only in one Arctic lake. Interestingly, within the Betaproteobacteria,

292 family Oxaloacetaceae was the most abundant in two Arctic lakes (SvL1 and SvL2), while
293 Comamonadacea was the most abundant family in most other lakes.

294

295 Globally, microeukaryotic assemblages (Figure 2B) were dominated by sequences
296 classified as Stramenopiles (26%), Viridiplantae (19%), and Alveolata (15%). A major
297 difference between the two polar zones became evident when the composition was
298 examined for the different lakes separately (Figure 3B): Arctic lakes were clearly
299 dominated by Alveolata while the Antarctic lakes had mostly Viridiplantae and
300 Stramenopiles, with differences reaching statistical significance ($p < 0.05$). The
301 Stramenopiles belonged mostly to the Chrysophyceae (12%) and Ochromonadaceae (6%).
302 The Viridiplantae were mostly Chlorophyta, and the Alveolata mostly Dinophyceae (8%).
303 Some Antarctic lakes were dominated by Viridiplantae while others were dominated by
304 Stramenopiles. Other groups were present in some lakes only, and in smaller proportions.
305 Cercozoa were one such group. They were abundant in two Arctic and one Antarctic lake
306 and present in a few more lakes. There were also sequences related to fungi. In particular,
307 one extremely abundant (52.4%) OTU in Lake SvL2 related to the Basidiomycota, and
308 Chytridiomycota-affiliated sequences in lake IR2 (12.5%). Another interesting case was
309 that of a *Telonema* OTU that was very abundant in Lake IR1. *Telonema* is a widely
310 distributed marine heterotrophic flagellate that forms a deep branch in the tree of eukaryal
311 life. Finally, an average 13% of eukaryotic sequences could not be properly assigned, and
312 17% were related to uncultured eukaryotes.

313

314 Overall, bacterial community compositions for the studied environments were dominated
315 by members of the Burkholderiales, Bacteroidetes, and Actinobacteria, the same groups
316 found to dominate both Polar and non-Polar freshwater ecosystems (Newton et al.,
317 2011; Logares et al., 2013; Barberan and Casamayor, 2014; Vick-Majors et al., 2014). The
318 same was true for the microeukaryotic assemblages, where the dominance of
319 Stramenopiles, Chlorophyta, and Alveolata, is in line with previous knowledge that most
320 eukaryotes in other Antarctic lakes were photosynthetic microalgae (Wilkins et al.,
321 2013; Vick-Majors et al., 2014). Phototrophic microeukaryotes accounted for about 60% of
322 the eukaryotic sequences. On the other hand, 16S sequences provided very few sequences
323 of free-living cyanobacteria and a substantial amount of chloroplast sequences. This
324 indicates that the primary producers in the studied pelagic ecosystems were eukaryotes,
325 which is in contrast to the dominant role of Cyanobacteria in the benthos of Antarctic
326 freshwater systems (Wharton et al., 1983; Vincent, 2000). However, our results are in
327 agreement with the fact that viral metagenomes from polar freshwater environments were
328 found to be dominated by likely microeukaryote-infecting viruses (Lopez-Bueno et al.,
329 2009; Aguirre de Carcer et al., 2015), and with results from microscopy observations
330 (Izaguirre et al., 1998).

331

332 **3.3 Community structure**

333

334 Double principal coordinates analyses were used to explore the relationships between
335 phylogenetic community compositions at each site. The major source of variation in the
336 bacterial data set (Figure 4A) related to the relative proportions of sequences affiliated as
337 Flavobacteria and Betaproteobacteria. Lakes Avian, Domo, IR2 and Pourquois-Pas showed
338 increased abundances of Flavobacteria and reduced abundances of Betaproteobacteria,
339 whereas Cierva, SvL1, SvL2, and Green exhibited the opposite pattern. On the other hand,
340 variation in the microeukaryotic data set was principally related to the relative proportions
341 of sequences affiliated to the Stramenopiles, Alveolata, and Viridiplantae clades (Figure

342 4B). As already mentioned Antarctic communities had significantly ($p < 0.05$) lower relative
343 abundances of Alveolata-affiliated sequences and higher proportions of those affiliated to
344 the Stramenopiles and Viridiplantae clades, when compared to the studied Arctic sites.
345 Furthermore, the results confirmed two opposing clusters of Antarctic sites along a
346 Stramenopiles-Viridiplantae axis.

347
348 For the remaining exploration of the datasets, we carried-out two types of analyses (Table
349 4): those based solely on OTUs relative abundances (Mantel tests of Bray-Curtis distances,
350 Between Class Analysis), and analyses based on both OTUs relative abundances and
351 phylogenetic relatedness between OTUs (Mantel tests of Rao distances, constrained
352 Double Principal Coordinates Analysis). First, we examined the relationship between the
353 composition of the cellular components and both geographical distance and limnological
354 parameters (Table 4A). We found no significant association between either component and
355 geographic distance or latitude. On the other hand, we observed a significant correlation
356 between bacterial community structure and the limnological parameters recorded in the
357 Antarctic sites, and marginal association ($p = 0.068$) between the same parameters and the
358 microeukaryotic assemblages. In both cases, associations seemed to be related to
359 temperature and conductivity (analysis with instrumental variables based on Bray-Curtis
360 distances; $p < 0.05$). Nevertheless, comparisons became statistically not significant when
361 taking into account the overall phylogenetic composition of the communities.

362
363 Next, we analyzed the correlations among the different microbial components (Table 4B).
364 Bacteria and microeukaryotic assemblages showed a significant correlation employing
365 Bray-Curtis distances, yet statistical significance disappeared when using Rao distances
366 (phylogenetic-aware). On the contrary, bacterial and viral cellular fractions showed a
367 significant correlation with Rao distances but not with Bray-Curtis distances. As expected,
368 cellular and free virus fraction T4-like virus assemblages from the same lake were more
369 similar to each other than to those of the other lakes, also when accounting for the genetic
370 similarities among sequences. Interestingly, we observed a significant discrimination
371 between virus assemblages arising from either cellular or free viral fractions with Bray-
372 Curtis distances, although such differences disappeared when Rao distances were
373 considered.

374
375 Finally, we examined potential differences between the two polar zones (Table 4C).
376 Bacterial assemblages from the same pole were more similar to each other than to those of
377 the other pole, yet the overall genetic diversity of both poles was undistinguishable. In the
378 cases of the microeukaryotic and T4-like virus assemblages, on the other hand, both poles
379 were shown to harbor distinct community structures, also when accounting for the genetic
380 diversity of the OTUs.

381 382 **4. DISCUSSION**

383 384 **4.1. Composition of polar freshwater microbial communities**

385 The phylogenetic analysis clustered the lakes on the basis of their microeukaryotic
386 composition, with Arctic lakes bearing higher proportions of Alveolata/Dinophyceae, and
387 Antarctic lakes partitioned along a Viridiplantae (Avian, Cierva, Pourquois-Pas and
388 Refugio) vs. Stramenopiles (Limnopolar, Biscoe, Horseshoe, Green, and Domo) axis. This
389 second axis roughly matches the classification of the lakes according to their higher or
390 lower Chlorophyll concentration (Suppl. Table 2). Interestingly, this partitioning is also in
391 line with that reported by Izaguirre et al. (Izaguirre et al., 1998) for lakes in Hope Bay

392 (Antarctica), where trophic status was shown to impact microeukaryotic assemblages:
393 more oligotrophic lakes were dominated by Chrysophyceae (Stramenopiles), and more
394 eutrophic lakes were dominated by Chlorophyceae (Viridiplantae). These authors reported
395 that trophic status of the freshwater bodies in Hope Bay depended strongly on their
396 proximity to bird colonies. In the present study most lakes did not appear to have a strong
397 influence of marine birds and mammals. However, partitioning along the abovementioned
398 axis correlated with estimated zoogenic input (Table 1). Moreover, Lake Refugio was
399 surrounded by substantial numbers of elephant seals, likely explaining the very high value
400 of Chlorophyll in this lake and the dominance of Viridiplantae.

401

402 We detected significant differences between T4-like virus assemblages arising from the
403 cellular and free virus fractions. However, such differences appeared only at the level of
404 OTUs, and not when accounting for the overall phylogenetic compositions of the fractions.
405 These observations may indicate that different phylogenetically-related viral OTUs have a
406 preference to localize in one or the other fraction. This fact may have important
407 methodological consequences, since most protocols used to study overall viral assemblages
408 (e.g. through viral shot-gun metagenomics) rely on the study of the free viral fraction
409 exclusively, in order to reduce the massive proportion of undesired bacterial DNA in the
410 resulting datasets (Lopez-Bueno et al., 2009; Fancello et al., 2013; Brum and Sullivan,
411 2015). Also, the free virus fraction was less diverse than the cellular fraction (e.g. intra-
412 cellular and membrane-attached viruses). Moreover, a significant correlation was found
413 between bacterial and T4-like viral assemblages present in the cellular fractions.
414 Altogether, these results indicate that studying solely free viral fractions may result in a
415 rather biased picture of the ecology of viruses.

416

417 **4.2. Limnological parameters drive Antarctic community structures**

418

419 We could not find evidence for latitude or geographical distance having a significant
420 influence on the community composition of Antarctic sites (Table 4A). This was true for
421 both the bacterial and the microeukaryotic components. On the other hand, limnological
422 parameters correlated with both bacterial and microeukaryotic community structure (Table
423 4A). The same weak or inexistent short to mid-range biogeographical patterns in microbial
424 communities, combined with strong environmental filtering have previously been observed
425 in similar ecosystems, such as Antarctic lakes (Logares et al., 2013), or sub-arctic and
426 Arctic marine environments (Winter et al., 2013). Subsequent analyses revealed that both
427 conductivity and temperature co-varied with community structure. However, it is important
428 to note that in the present dataset temperature and conductivity were correlated (Suppl.
429 Table 1). The influence of temperature was not unexpected, since it had previously been
430 shown to affect bacterial community structure in northern European lakes (Lindström et
431 al., 2005). The effects of temperature and conductivity were significant exclusively on the
432 Bray-Curtis distances, but they did not affect overall phylogenetic compositions (Table
433 4A). The fact that the latter remained unaffected by environmental variables likely
434 indicates the existence of ecologically-redundant, phylogenetically-related, OTUs fit to
435 slightly different limnological scenarios. We had anticipated significant pH effects, as they
436 had previously been reported in studies of high altitude (Barberan and Casamayor, 2014)
437 (pH range 4.5-9), northern Europe (Lindström et al., 2005) (pH 5.5-8.7), and Wisconsin
438 lakes (Yannarell and Triplett, 2005) (pH 5.4-8.6. It is possible that we did not detect pH-
439 related effects due to the reduced dataset size. On the other hand, pH range along our sites
440 did not reach two units (5.8-7.7), while in all abovementioned studies the range exceeded 3
441 units.

442
443
444

4.3. Bipolar patterns in community structures

446

447 The overall percentages of shared OTUs were considerable (between 6 and 19%), further
448 substantiating the idea that while each lake has a particular community structure, many
449 lakes can harbor the same OTUs (Newton et al., 2011). Nevertheless, the percentages of
450 shared bacterial OTUs between lakes from different poles were noticeably lower than those
451 observed between marine Arctic and Antarctic bacterial assemblages (*ca.* 20-30%)
452 (Ghiglione et al., 2012), suggesting a stronger effect of water-borne bacterial dispersal
453 and(or) a greater variability in the physicochemical environment between lakes versus
454 ocean samples. In particular, Ghiglione et al. (2012) found indications that deep water
455 circulation might be responsible for dispersal of OTUs from one polar ocean to the other.
456 Lakes are not communicated by oceanic circulation and, thus, dispersal can only occur
457 through the air or through migrating Arctic terns. In this regard, it is not surprising that the
458 number of shared OTUs is lower in freshwater bodies than in marine polar areas.
459 Additional examples of similar or even identical sequences retrieved from both polar
460 marine areas exist for example for *Nitrosospira* {Hollibaugh, 2002 #375} or for ammonia-
461 oxidizing Chrenarchaeota {Kalanetra, 2009 #377}. In fact one of the latter sequences was
462 found to be dominant during the dark winter period in the Arctic Ocean {Alonso-Saez,
463 2012 #378}. Another interesting example is that of cyanobacteria isolated from mats in
464 ponds from both Polar zones, Nadeau et al. found that psychrophilic strains had almost
465 identical sequences in their 16S rDNA, while psychrotolerant strains were more different.

466

467 There were no differences in diversity estimates between poles for the three components of
468 microbial life. Nevertheless, the results based on community structure, phylogenetic
469 composition, and shared OTUs, all point to a consistent biogeographical pattern
470 segregating both poles. The only exception was that the bacterial assemblages were not
471 significantly different in terms of phylogenetic content. This pattern had already been
472 observed for the bacterial assemblages of Antarctic vs. Scandinavian lakes (Logares et al.,
473 2013), as well as for the bacterial and microeukaryotic assemblages present in cryoconite
474 holes from both poles (Cameron et al., 2012).

475

4.4. Relationships among the three microbial components

476

477 We found a significant correlation between bacterial and microeukaryotic community
478 structures. However, such link disappeared when accounting for the overall phylogenetic
479 content of the communities. These results indicate that while the overall phylogenetic
480 compositions of each of these microbial groups are not coordinated, there are particular
481 associations between several of their OTUs. Similar findings have recently been reported
482 for a single marine site (Chow et al., 2014). While these signals may simply reflect the
483 associated OTUs' preferences for similar environmental parameters, they could also
484 indicate true biotic interactions.

485

486 Our analyses also showed a significant correlation among bacterial and T4-like virus
487 assemblages present in the cellular fractions, although in this case only when accounting
488 for the phylogenetic relatedness among OTUs. Links between viral and bacterial
489 community compositions had previously been detected in a large scale study of sub-polar
490 and Arctic marine sites (Winter et al., 2013). The results from the latter study, however,
491

492 were not conclusive, since observed links were regionally restricted and failed to show
493 such pattern for the entire dataset. The circumstance that we did not observe such pattern
494 using the free virus dataset may relate to increased noise, since this community may be
495 dominated by viral strains recently arising from exponential outbursts. Nevertheless, the
496 fact that such pattern was detected when accounting for phylogenetic relatedness among
497 OTUs is in line with the idea that many viral strains may infect more likely
498 phylogenetically-related hosts, although this predominant idea has not yet been properly
499 addressed (Flores et al., 2011).

500

501 **Conclusion**

502

503 This study represents, to our knowledge, the first massive parallel sequencing-based
504 assessment of three components of microbial life; bacteria, microeukaryotes and T4-like
505 viruses. In this regard, we believe this study offers an important complement to recent
506 efforts to monitor these relationships following time-series analysis of a single marine
507 ecosystem employing Automated Ribosomal Intergenic Spacer Analysis (ARISA) and
508 terminal restriction fragment length polymorphism (T-RFLP) of phylogenetic marker
509 genes (Needham et al., 2013;Chow et al., 2014). These reports were able to show particular
510 temporal dynamics among some members of the studied communities. Here we have
511 assessed the role of biogeographical and limnological patterns in driving community
512 composition, as well as global co-variation among the different components of microbial
513 life. Overall, the results derived from this study support previous reports on the
514 composition of polar freshwater microbial communities. They are also in agreement with
515 environmental filtering representing a predominant factor driving community structure,
516 with geographical distance effects appearing only at large scales. . Also, the noticeable
517 fraction of shared OTUs among sites seemingly indicates reduced dispersal limitations for
518 microorganisms in these ecosystems (Logares et al., 2013). The availability of deep-
519 sequencing phylogenetic marker genes data has allowed the study of community structure
520 no only based on OTUs' relative abundance, but also on overall phylogenetic content. The
521 instances when these two complementary approaches have disagreed have provided
522 important information regarding the ecology of the ecosystems. For example,
523 compositional shifts along the narrow limnological space sampled at the Antarctic sites
524 were only detected based on OTUs' relative abundances, but not when taking into account
525 their phylogenetic relationships. Since ecological coherence and gene conservation have
526 been shown to be negatively correlated with phylogenetic distance (Philippot et al.,
527 2010;Zaneveld et al., 2010), a plausible explanation is that small differences in
528 limnological parameters shifted OTU-level community structure, yet the overall link
529 between ecological function and phylogenetic structure remained stable.

530

531 **Author Contributions**

532 AA, DAC, and DP conceived the study. DAC undertook all wet-lab procedures. DAC and
533 CPA analyzed the data and wrote the manuscript, with input from the other contributors.

534

535 **Acknowledgements**

536 This work was funded by the Spanish Polar Programme and the Spanish Ministry of
537 Science and Innovation grants CTM2008-05134-E/ANT, CTM2009-08644-E and
538 CTM2011-15091-E/ANT. We thank Roche for their support to the Antarctic expedition. D.
539 A. C. was supported by the Marie Curie IIF grant PIIF-GA-2012-328287. We thank the
540 logistic support from the Maritime Technology Unit (CSIC) and *Las Palmas* crew (Spanish
541 Navy). We thank the Mapping and Geographic Information Centre and D. Hodgson

542 (British Antarctic Survey, UK) for their help in identifying suitable water bodies along the
543 Antarctic Peninsula. We thank A. López-Bueno, J. M. Alonso-Lobo, M. L. Avila, L. Little
544 and the students of the University Center in Svalbard Master Course AB-327 on Arctic
545 Microbiology (2011-2012) for their help collecting and processing the freshwater samples.
546 The authors declare no conflict of interests.
547
548

Provisional

550 **REFERENCES**

- 551 Aguirre de Cárcer, D., Denman, S.E., McSweeney, C., and Morrison, M. (2011a). Evaluation of
552 Subsampling-Based Normalization Strategies for Tagged High-Throughput Sequencing
553 Data Sets from Gut Microbiomes. *Applied and Environmental Microbiology* 77, 8795-8798.
554 doi: 10.1128/aem.05491-11.
- 555 Aguirre de Cárcer, D., Denman, S.E., McSweeney, C., and Morrison, M. (2011b). Strategy for
556 Modular Tagged High-Throughput Amplicon Sequencing. *Applied and Environmental*
557 *Microbiology* 77, 6310-6312. doi: 10.1128/aem.05146-11.
- 558 Aguirre de Cárcer, D., Lopez-Bueno, A., Pearce, D.A., and Alcamí, A. (2015). Biodiversity and
559 Distribution of Polar Freshwater DNA Viruses. *Science Advances* 1, e1400127.
- 560 Amaral-Zettler, L., Artigas, F., Baross, J., Bharathi, L., Boetius, A., Chandramohan, D., Herndl, G.,
561 Kogure, K., Neal, P., Pedrós-Alió, C., Ramette, A., Schouten, S., Stal, L., Thessen, J.L., and
562 Sogin, M. (2010). "A global census of marine microbes," in *Life in the World's Oceans.*
563 *Diversity, Distribution, and Abundance. Census of Marine Life*, ed. A.g.c.o.m. microbes.
564 (Oxford, U.K. : Wiley-Blackwell), Oxford, U.K. .
- 565 Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983). The ecological role
566 of water-column microbes in the sea. *Marine ecology progress series. Oldendorf* 10, 257-
567 263.
- 568 Barberan, A., and Casamayor, E.O. (2014). A phylogenetic perspective on species diversity, beta-
569 diversity and biogeography for the microbial world. *Mol Ecol* 23, 5868-5876.
- 570 Bates, S.T., Clemente, J.C., Flores, G.E., Walters, W.A., Parfrey, L.W., Knight, R., and Fierer, N.
571 (2013). Global biogeography of highly diverse protistan communities in soil. *Isme J* 7, 652-
572 659. doi: 10.1038/ismej.2012.147
- 573 Baty, F., Facompre, M., Wiegand, J., Schwager, J., and Brutsche, M. (2006). Analysis with respect
574 to instrumental variables for the exploration of microarray data structures. *BMC*
575 *Bioinformatics* 7, 422.
- 576 Becking, L.G.M.B. (1934). *Geobiologie of inleiding tot de milieukunde*. W.P. Van Stockum & Zoon.
- 577 Bragg, L., Stone, G., Imelfort, M., Hugenholtz, P., and Tyson, G.W. (2012). Fast, accurate error-
578 correction of amplicon pyrosequences using Acacia. *Nat Meth* 9, 425-426.
- 579 Brum, J.R., and Sullivan, M.B. (2015). Rising to the challenge: accelerated pace of discovery
580 transforms marine virology. *Nat Rev Micro* 13, 147-159. doi: 10.1038/nrmicro3404.
- 581 Cameron, K.A., Hodson, A.J., and Osborn, A.M. (2012). Structure and diversity of bacterial,
582 eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the
583 Antarctic. *FEMS Microbiol Ecol* 82, 254-267.
- 584 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N.,
585 Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E.,
586 Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky,
587 J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight,
588 R. (2010). *QIIME allows analysis of high-throughput community sequencing data*. *Nat*
589 *Methods*. 2010 May;7(5):335-6. doi: 10.1038/nmeth.f.303.
- 590 Clokie, M.R., Millard, A.D., and Mann, N.H. (2010). T4 genes in the marine ecosystem: studies of
591 the T4-like cyanophages and their role in marine ecology. *Virology* 7, 7-291.
- 592 Comeau, A.M., and Krisch, H.M. (2008). The capsid of the T4 phage superfamily: the evolution,
593 diversity, and structure of some of the most prevalent proteins in the biosphere. *Mol Biol*
594 *Evol* 25, 1321-1332.
- 595 Convey, P., and Stevens, M.I. (2007). Antarctic Biodiversity. *Science* 317, 1877-1878. doi:
596 10.1126/science.1147261.

- 597 Chow, C.E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., and Fuhrman, J.A. (2014). Top-down controls on
598 bacterial community structure: microbial network analysis of bacteria, T4-like viruses and
599 protists. *ISME J* 8, 816-829. doi: 10.1038/ismej.2013.199.
- 600 Davison, A.C., and Hinkley, D.V. (1997). *Bootstrap Methods and Their Applications*. Cambridge:
601 Cambridge University Press.
- 602 DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D.,
603 Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene
604 database and workbench compatible with ARB. *Appl Environ Microbiol* 72, 5069-5072.
605 doi: 10.1128/AEM.03006-05.
- 606 Dray, S., Pavoine, S., and Aguirre de Cárcer, D. (2015). Considering external information to
607 improve the phylogenetic comparison of microbial communities: a new approach based
608 on constrained Double Principal Coordinates Analysis (cDPCoA). *Molecular Ecology*
609 *Resources* 15, 242-249. doi: 10.1111/1755-0998.12300.
- 610 Fancello, L., Trape, S., Robert, C., Boyer, M., Popgeorgiev, N., Raoult, D., and Desnues, C. (2013).
611 Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds
612 of the Mauritanian Sahara. *ISME J* 7, 359-369.
- 613 Filee, J., Tetart, F., Suttle, C.A., and Krisch, H.M. (2005). Marine T4-type bacteriophages, a
614 ubiquitous component of the dark matter of the biosphere. *Proc Natl Acad Sci U S A* 102,
615 12471-12476.
- 616 Flores, C.O., Meyer, J.R., Valverde, S., Farr, L., and Weitz, J.S. (2011). Statistical structure of host-
617 phage interactions. *Proceedings of the National Academy of Sciences* 108, E288-E297. doi:
618 10.1073/pnas.1101595108.
- 619 Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M.V., Green, J.L., and Brown, J.H.
620 (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proc Natl Acad Sci U*
621 *S A* 105, 7774-7778.
- 622 Ghiglione, J.F., Galand, P.E., Pommier, T., Pedros-Alio, C., Maas, E.W., Bakker, K., Bertilson, S.,
623 Kirchman, D.L., Lovejoy, C., Yager, P.L., and Murray, A.E. (2012). Pole-to-pole
624 biogeography of surface and deep marine bacterial communities. *Proc Natl Acad Sci U S A*
625 109, 17633-17638. doi: 10.1073/pnas.1208160109
- 626 Hawkins, B.A. (2001). Ecology's oldest pattern? *Endeavour* 25, 133.
- 627 Izaguirre, I., Vinocur, A., Mataloni, G., and Pose, M. (1998). Phytoplankton communities in relation
628 to trophic status in lakes from Hope Bay (Antarctic Peninsula). *Hydrobiologia* 369-370, 73-
629 87. doi: 10.1023/a:1017070415024.
- 630 Lindström, E.S., Kamst-Van Agterveld, M.P., and Zwart, G. (2005). Distribution of Typical
631 Freshwater Bacterial Groups Is Associated with pH, Temperature, and Lake Water
632 Retention Time. *Applied and Environmental Microbiology* 71, 8201-8206. doi:
633 10.1128/aem.71.12.8201-8206.2005.
- 634 Logares, R., Lindstrom, E.S., Langenheder, S., Logue, J.B., Paterson, H., Laybourn-Parry, J.,
635 Rengefors, K., Tranvik, L., and Bertilsson, S. (2013). Biogeography of bacterial communities
636 exposed to progressive long-term environmental change. *ISME J* 7, 937-948.
- 637 Lopez-Bueno, A., Tamames, J., Velazquez, D., Moya, A., Quesada, A., and Alcami, A. (2009). High
638 diversity of the viral community from an Antarctic lake. *Science* 326, 858-861. doi:
639 10.1126/science.1179287
- 640 Lliros, M., Inceoglu, O., Garcia-Armisen, T., Anzil, A., Leporcq, B., Pigneur, L.M., Viroux, L.,
641 Darchambeau, F., Descy, J.P., and Servais, P. (2014). Bacterial community composition in
642 three freshwater reservoirs of different alkalinity and trophic status. *PLoS One* 9.
- 643 Martiny, J.B., Eisen, J.A., Penn, K., Allison, S.D., and Horner-Devine, M.C. (2011). Drivers of
644 bacterial beta-diversity depend on spatial scale. *Proc Natl Acad Sci U S A* 108, 7850-7854.

- 645 Needham, D.M., Chow, C.E., Cram, J.A., Sachdeva, R., Parada, A., and Fuhrman, J.A. (2013). Short-
646 term observations of marine bacterial and viral communities: patterns, connections and
647 resilience. *ISME J* 7, 1274-1285.
- 648 Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., and Bertilsson, S. (2011). A guide to the natural
649 history of freshwater lake bacteria. *Microbiol Mol Biol Rev* 75, 14-49.
- 650 Peura, S., Eiler, A., Hiltunen, M., Nykanen, H., Tiirola, M., and Jones, R.I. (2012). Bacterial and
651 phytoplankton responses to nutrient amendments in a boreal lake differ according to
652 season and to taxonomic resolution. *PLoS One* 7, 8.
- 653 Philippot, L., Andersson, S.G., Battin, T.J., Prosser, J.I., Schimel, J.P., Whitman, W.B., and Hallin, S.
654 (2010). The ecological coherence of high bacterial taxonomic ranks. *Nat Rev Microbiol* 8,
655 523-529. doi: 10.1038/nrmicro2367
- 656 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glockner, F.O.
657 (2013). The SILVA ribosomal RNA gene database project: improved data processing and
658 web-based tools. *Nucleic Acids Res* 41, D590-596. doi: 10.1093/nar/gks1219
- 659 Ren, L., Jeppesen, E., He, D., Wang, J., Liboriussen, L., Xing, P., and Wu, Q.L. (2015). pH influences
660 the importance of niche-related versus neutral processes in lacustrine bacterioplankton
661 assembly. *Applied and Environmental Microbiology*. doi: 10.1128/aem.04042-14.
- 662 Sul, W.J., Cole, J.R., Jesus Eda, C., Wang, Q., Farris, R.J., Fish, J.A., and Tiedje, J.M. (2011). Bacterial
663 community comparisons by taxonomy-supervised analysis independent of sequence
664 alignment and clustering. *Proc Natl Acad Sci U S A* 108, 14637-14642.
- 665 Sul, W.J., Oliver, T.A., Ducklow, H.W., Amaral-Zettler, L.A., and Sogin, M.L. (2013). Marine bacteria
666 exhibit a bipolar distribution. *Proc Natl Acad Sci U S A* 110, 2342-2347.
- 667 Vick-Majors, T.J., Priscu, J.C., and Amaral-Zettler, L.A. (2014). Modular community structure
668 suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic
669 lakes. *ISME J* 8, 778-789.
- 670 Vincent, W.F. (2000). Evolutionary origins of Antarctic microbiota: invasion, selection and
671 endemism. *Antarctic Science* 12, 374-385. doi: doi:10.1017/S0954102000000420.
- 672 Wharton, R., Parker, B., and Simmons, G. (1983). Distribution, species composition and
673 morphology of algal mats in Antarctic dry valley lakes. *Phycologia* 22, 355-365. doi:
674 doi:10.2216/i0031-8884-22-4-355.1.
- 675 Wichels, A., Biel, S.S., Gelderblom, H.R., Brinkhoff, T., Muyzer, G., and Schutt, C. (1998).
676 Bacteriophage diversity in the North Sea. *Appl Environ Microbiol* 64, 4128-4133.
- 677 Wilkins, D., Yau, S., Williams, T.J., Allen, M.A., Brown, M.V., DeMaere, M.Z., Lauro, F.M., and
678 Cavicchioli, R. (2013). Key microbial drivers in Antarctic aquatic environments. *FEMS*
679 *Microbiol Rev* 37, 303-335.
- 680 Williamson, S.J., Rusch, D.B., Yooshef, S., Halpern, A.L., Heidelberg, K.B., Glass, J.I., Andrews-
681 Pfannkoch, C., Fadrosch, D., Miller, C.S., Sutton, G., Frazier, M., and Venter, J.C. (2008). The
682 Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses
683 within aquatic microbial samples. *PLoS One* 3, 0001456.
- 684 Winter, C., Matthews, B., and Suttle, C.A. (2013). Effects of environmental variation and spatial
685 distance on Bacteria, Archaea and viruses in sub-polar and arctic waters. *ISME J* 7, 1507-
686 1518. doi: 10.1038/ismej.2013.56.
- 687 Yannarell, A.C., and Triplett, E.W. (2005). Geographic and Environmental Sources of Variation in
688 Lake Bacterial Community Composition. *Applied and Environmental Microbiology* 71, 227-
689 239. doi: 10.1128/aem.71.1.227-239.2005.
- 690 Zaneveld, J.R., Lozupone, C., Gordon, J.I., and Knight, R. (2010). Ribosomal RNA diversity predicts
691 genome diversity in gut bacteria and their relatives. *Nucleic Acids Res* 38, 3869-3879.

692

693

694 **FIGURE LEGENDS**

695

696 **Figure 1. Exploration of diversity trends.** A) Relationship between bacterial diversity
697 (Y-axis) and chlorophyll *a* concentration (X-axis). B) Relationship between bacterial (Y-
698 axis) and microeukaryotic diversity (X-axis). Diversity is represented by Shannon's
699 indexes and Chlorophyll *a* is in µg/L.

700

701 **Figure 2. Krona graphs depicting averaged relative abundances of taxonomic groups**
702 **in the studied communities.** A; Bacteria. B; Eukarya. **u:** uncultured. **euk:** eukaryote.

703

704 **Figure 3. Per-lake relative abundances of taxonomic groups in the studies**
705 **communities.** A) Bacteria. B) Eukarya.

706

707 **Figure 4. Phylogenetic composition similarities between polar freshwater**
708 **assemblages.** The figures depict results from Double Principal Coordinates Analyses
709 (DPCoA) based on 16S (A) and 18S (B) marker gene profiles. The analysis ordinales the
710 sites (white boxes) based on the phylogenetic distances and per-site relative abundances of
711 OTUs. The position of each point (OTU) represents its association with each site. Points
712 are colored and collectively labelled based on their taxonomic affiliation.

713

714

Provisional

714

715

Table 1. Location, sampling date and characteristics of the sampled environments.

	Pole	Position ¹	Date	Zoogenic input ²
Lake Domo*	Antarctic	-62.64 -60.97	03/02/2010	Very low
Lake Refugio*	Antarctic	-62.65 -61.00	03/02/2010	Next to sea elephant colony
Lake Limnopolar*	Antarctic	-62.66 -61.10	01/02/2010	Very low
Caleta Cierva	Antarctic	-64.16 -61.01	21/01/2010	Nearby seabird colonies
Green Island	Antarctic	-65.31 -64.15	23/01/2010	Accessible to seabirds
Biscoe Point	Antarctic	-65.43 -65.48	29/01/2011	Very low
Pourquoi-Pas Island	Antarctic	-67.66 -67.25	27/01/2011	Very low
Avian Island	Antarctic	-67.76 -68.88	25/01/2010	Nearby seabird colonies
Horseshoe Island	Antarctic	-67.84 -67.19	25/01/2011	Very low
IR2	Arctic	78.04 13.69	27/09/2011	Frequent seabird activity
Lake Tunsjøen(IR1)	Arctic	78.05 13.65	26/09/2011	Frequent seabird activity
Lake Nordammen(Sv11)	Arctic	78.63 16.63	05/06/2012	Accessible to seabirds
Lake Tenndamen(Sv12)	Arctic	78.10 15.03	06/06/2012	Accessible to seabirds

¹Lat Long. ²*in situ* observations. *Lakes in Byers Peninsula, Livingston Island.

716

Table 2. Shared OTUs. Numbers represent averaged values (%) of geometric means of the pairwise comparisons.

	16SS	18S	g23
Overall	14.0±4	10.2±2.4	5.8±11.6
Intra-Antarctic	19.4±4***	13.7±2.2***	5.7±2.2*
Intra-Arctic	10.1±1.1	11.7±2.8*	10.9±9.3*

*p<0.05;***p<0.0005. Paired t-tests. Comparison with respect to the overall shared OTUs.

717

Table 3. Shannon's diversity indices.

	Bacteria	Eukarya	virus ^f	virus ^c
Domo	4.28	2.95	NA	2.25
Refugio	6.50	4.46	NA	NA
Limnopolar	4.95	3.99	3.01	2.84
Cierva	3.37	5.61	NA	1.52
Green	6.45	2.31	2.63	2.39
Biscoe	4.72	2.68	0	2.16
Pourquoi-pas	4.10	4.76	0.39	1.94
Avian	4.57	5.04	2.28	2.29
Horseshoe	3.51	4.12	1.36	1.84
IR2	4.00	5.67	1.95	1.34
Tunsjøen IR1	5.29	4.80	1.32	1.45
Tenndamen SvL2	5.62	2.26	1.83	1.99
Nordammen SvL1	4.08	3.58	1.84	4.17

718

^fFree virus fraction. ^cCellular fraction.

Table 4. Community structure analyses.

Comparison	Test	Sign	Comparison	Test	Sign
A. Antarctic vs. Environment:			B1. Correlation between assemblages:		
Bacteria Vs. Geo. Distance	Mantel(BC)	na	Bacteria Vs. Microeukaryotes	Mantel(BC)	***
	Mantel(Rao)	na		Mantel(Rao)	na
Bacteria Vs. Latitude	Mantel(BC)	na	Bacteria Vs. Virus ^F	Mantel(BC)	na
	Mantel(Rao)	na		Mantel(Rao)	na
Bacteria Vs. Physicochem.	Mantel(BC)	**	Bacteria Vs. Virus ^C	Mantel(BC)	na
	Mantel(Rao)	na		Mantel(Rao)	**
Microeukaryotes Vs. Geo.Distance	Mantel(BC)	na	Microeukaryotes Vs. Virus ^F	Mantel(BC)	na
	Mantel(Rao)	na		Mantel(Rao)	na
Microeukaryotes Vs. Latitude	Mantel(BC)	na	Microeukaryotes Vs. Virus ^C	Mantel(BC)	na
	Mantel(Rao)	na		Mantel(Rao)	na
Microeukaryotes Vs. Physicochem.	Mantel(BC)	*	C. Between-poles:		
	Mantel(Rao)	na	Bacteria	BCA	**
B2. Virus (other):				cDPCoA	na
Between lakes	BCA	**	Microeukaryotes	BCA	**
	cDPCoA	**		cDPCoA	**
Virus ^F Vs. Virus ^C	BCA	**	Virus	BCA	**
	cDPCoA	na		cDPCoA	**

Tests: Mantel(BC)[Mantel test based on Bray-Curtis distances]; Mantel(Rao) [Mantel test based on Rao distances {Phylogenetic}]; BCA [Between class analysis]; cDPCoA [Constrained Double Principal Coordinates Analysis {Phylogenetic}]. **Sign:** na (p>0.1), *** (p<0.001), ** (p<0.05), * (p<0.1). ^F free virus fraction, ^C cellular fraction.

Figure 1.TIF

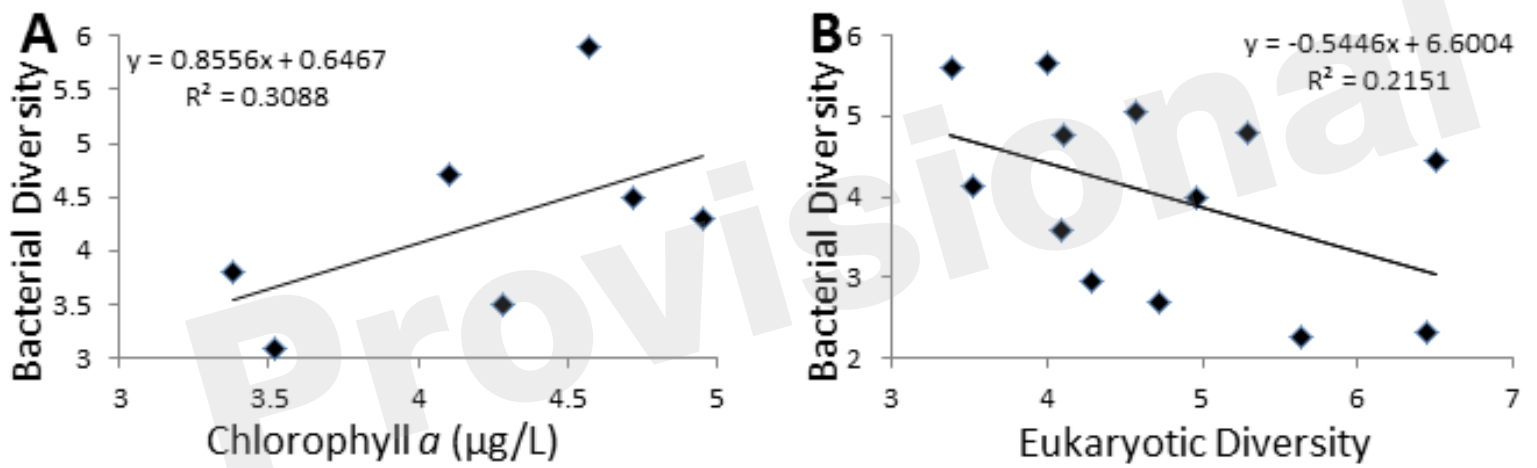


Figure 2.TIF

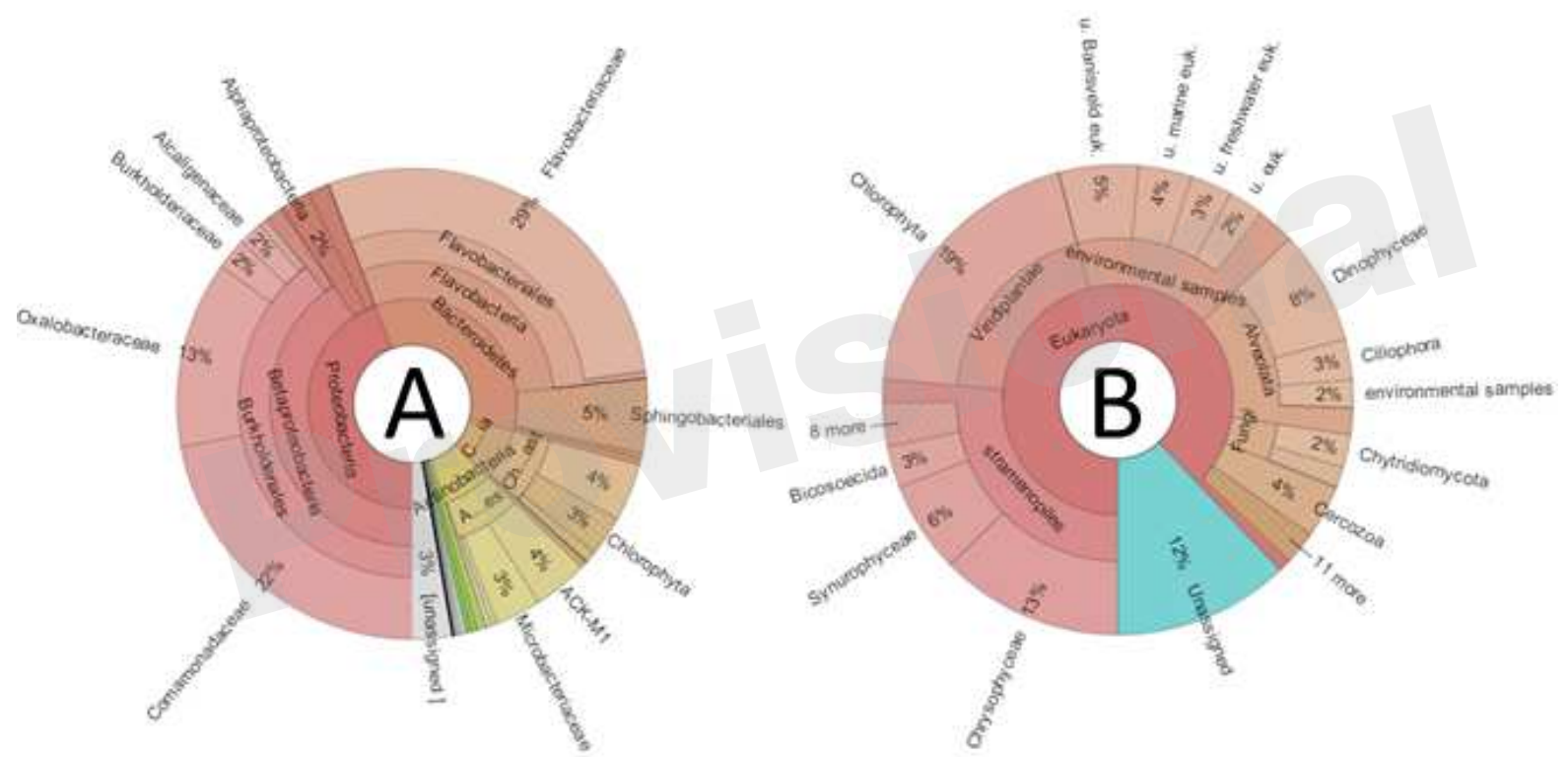


Figure 4.TIF

