

**Original citation:**

Ray, Jessica L., Althammer, Julia, Skaar, Katrine S., Simonelli, Paolo, Larsen, Aud, Stoecker, Diane, Sazhin, Andrey, Ijaz, Umer Z., Quince, Christopher, Nejstgaard, Jens C., Frischer, Marc, Pohnert, Georg and Troedsson, Christofer. (2016) Metabarcoding and metabolome analyses of copepod grazing reveal feeding preference and linkage to metabolite classes in dynamic microbial plankton communities. *Molecular Ecology*, 25 (21). pp. 5585-5602.

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/85060>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

This is the pre-peer reviewed version of the following article: Ray, J. L., Althammer, J., Skaar, K. S., Simonelli, P., Larsen, A., Stoecker, D., Sazhin, A., Ijaz, U. Z., Quince, C., Nejstgaard, J. C., Frischer, M., Pohnert, G. and Troedsson, C. (2016), Metabarcoding and metabolome analyses of copepod grazing reveal feeding preference and linkage to metabolite classes in dynamic microbial plankton communities. *Mol Ecol*, 25: 5585–5602.

doi:10.1111/mec.13844, which has been published in final form at

<http://dx.doi.org/10.1111/mec.13844>. This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

**A note on versions:**

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP url' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

1 **Metabarcoding and metabolome analysis of copepod grazing reveals feeding preference**  
2 **and linkage to metabolite classes in dynamic microbial plankton communities**

3 Jessica L. Ray<sup>1\*</sup>, Julia Althammer<sup>2□</sup>, Katrine S. Skaar<sup>1</sup>, Paolo Simonelli<sup>3</sup>, Aud Larsen<sup>1</sup>, Diane  
4 Stoecker<sup>4</sup>, Andrey Sazhin<sup>5</sup>, Umer Z. Ijaz<sup>6</sup>, Christopher Quince<sup>7</sup>, Jens C. Nejstgaard<sup>8</sup>, Marc  
5 Frischer<sup>1,9</sup>, Georg Pohnert<sup>2</sup> and Christofer Troedsson<sup>1</sup>

6 <sup>1</sup>Hjort Centre for Marine Ecosystem Dynamics, Uni Research Environment, Uni Research AS, Nygårdsgaten 112,  
7 N-5008 Bergen, Norway

8 <sup>2</sup>Institute for Inorganic and Analytical Chemistry, Friedrich Schiller University Jena, Lessingstr. 8, 07443 Jena,  
9 Germany

10 <sup>3</sup>Department of Biology, University of Bergen, Thormøhlensgt 53A, 5006 Bergen, Norway

11 <sup>4</sup>University of Maryland, Center of Environmental Science, Horn Point Lab, Cambridge, Massachusetts, 21613,  
12 USA

13 <sup>5</sup>P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Laboratory of Ecology of Plankton  
14 Organisms, Nakhimovsky Prospect 36, Moscow, Russia

15 <sup>6</sup>School of Engineering, University of Glasgow, Glasgow G12 8QQ, United Kingdom

16 <sup>7</sup>WMS - Microbiology and Infection, University of Warwick Medical School, Coventry, CV4 7AL, United  
17 Kingdom

18 <sup>8</sup>Leibniz Institute of Freshwater Ecology and Inland Fisheries, Dep. 3, Experimental Limnology/Alte Fischerhütte 2,  
19 16775 Stechlin, OT Neuglobsow, Germany

20 <sup>9</sup>Skidaway Institute of Oceanography, 10 Science Circle, Savannah, Georgia 31411, USA  
21

22 Keywords: *Calanus* grazing, *Phaeocystis pouchetii*, *Skeletonema marinoi*, seawater mesocosms,  
23 eukaryote V7 SSU, metabolome profiling, metabarcoding, symbiosis

24 Running Head: Multi-omic analysis of *Calanus* feeding

25 \*Corresponding author: E-mail [jessicalouiseray@gmail.com](mailto:jessicalouiseray@gmail.com), Telephone (+47) 55 58 26 73, Fax  
26 (+47) 55 58 96 74

27 □ Current address: Jenabios GmbH, Orlaweg 2, 07743 Jena, Germany.  
28  
29  
30  
31

32 **Abstract**

33 In order to characterize copepod feeding in relation to microbial plankton community dynamics,  
34 we combined metabarcoding and metabolome analyses during a 22-day seawater mesocosm  
35 experiment. Nutrient amendment of mesocosms promoted the development of haptophyte-

36 (*Phaeocystis pouchetii*) and diatom- (*Skeletonema marinoi*) dominated plankton communities in  
37 mesocosms, in which *Calanus* sp. copepods were incubated for 24-hours in flow-through  
38 chambers to allow access to prey particles (< 500  $\mu\text{m}$ ). Copepods and mesocosm water sampled  
39 six times spanning the experiment were analyzed using metabarcoding, while intracellular  
40 metabolite profiles of mesocosm plankton communities were generated for all experimental  
41 days. Taxon-specific metabarcoding ratios (ratio of consumed prey to available prey in the  
42 surrounding seawater) revealed diverse and dynamic copepod feeding selection, with positive  
43 selection on large diatoms, heterotrophic nanoflagellates and fungi, while smaller phytoplankton,  
44 including *P. pouchetii*, were passively consumed or even negatively selected according to our  
45 indicator. Our analysis of the relationship between *Calanus* grazing ratios and intracellular  
46 metabolite profiles indicates the importance of carbohydrates and lipids in plankton succession  
47 and copepod-prey interactions. This molecular characterization of *Calanus* sp. grazing therefore  
48 provides new evidence for selective feeding in mixed plankton assemblages and corroborates  
49 previous findings that copepod grazing may be coupled to the developmental and metabolic  
50 stage of the entire prey community rather than to individual prey abundances.

51

## 52 **Introduction**

53 The trophic efficiency of the marine food web depends upon the pathway of carbon flow from  
54 primary production to predatory fish - either through the classical food web (diatoms to  
55 mesozooplankton), the microbial food web (flagellates/bacteria to ciliates to mesozooplankton)  
56 (Landry 2002; Pepin *et al.* 2011), or via the recently discussed nutritunneling to bypass  
57 phytoplankton pathways (Pitta *et al.* 2016). Copepods are among the most abundant  
58 mesozooplankton in the global ocean and have long been assumed to be key regulators of carbon

59 transfer through their selective feeding in the marine food web (Sherr & Sherr 1988; Kleppel  
60 1993; Sanders & Wickham 1993). Understanding the interplay of copepod grazing as a top-down  
61 regulatory force, in concert with the bottom-up regulation of marine plankton assemblages, is  
62 essential for accurately predicting the flow of carbon and nutrients from primary production to  
63 fisheries (Sherr & Sherr 1988; Calbet & Saiz 2005).

64

65 Copepods can be selective grazers (Kiørboe *et al.* 1996; Kiørboe 2011 and references therein)  
66 dependent upon prey abundance, size, motility or chemical cues (Nejstgaard *et al.* 2008 and  
67 references therein). Elucidating copepod prey selection in natural environments presents a  
68 persistent methodological challenge, as full characterization of copepod feeding requires  
69 quantitative knowledge of the potential prey community as well as knowledge of the prey  
70 organisms that are actually consumed (reviewed in Pompanon *et al.* 2011). Chlorophyll *a* or  
71 pigment measurements of copepod gut content cannot be used to determine prey selection in  
72 natural prey assemblages as pigments show variable breakdown rates and do not reveal non-  
73 pigmented prey that may frequently be the most selected prey *in situ* (see data for *Calanus* in  
74 Nejstgaard *et al.* 1997; 2001b, and further discussion on methods in Nejstgaard *et al.* 2008).  
75 Molecular analysis of phylogenetic markers (metabarcoding; Taberlet *et al.* 2012) provides a  
76 promising alternative due to the universality of genomic DNA among cellular organisms and  
77 tunable precision of phylogenetic resolution without previous knowledge of community  
78 composition. In addition, tools exist to refine the output of molecular investigation toward a  
79 more prey-oriented analysis through the use of DHPLC-PCR (Troedsson *et al.* 2008a,b; Olsen *et*  
80 *al.* 2014), restriction endonucleases (Maloy *et al.* 2013) and/or blocking oligonucleotides  
81 (Troedsson *et al.* 2008b; Vestheim & Jarman 2008; Maloy *et al.* 2013) to selectively inhibit

82 amplification of predator sequences.

83

84 Ample evidence exists to suggest that chemical cues likely play an equitable role with prey  
85 availability in copepod feeding behavior (Poulet & Marsot 1978; Cowles *et al.* 1988). The  
86 combination of DNA-based metabarcoding methods with intracellular metabolite profiling may  
87 thus facilitate a deeper investigation of copepod feeding behavior (Woodson *et al.* 2007) that  
88 takes into account both taxonomy and chemical ecology (Barofsky *et al.* 2010; Kuhlisch &  
89 Pohnert 2015). We employed a multi-omic approach to characterize microbial plankton  
90 communities in seawater mesocosms dominated by the haptophyte *Phaeocystis pouchetii* or the  
91 diatom *Skeletonema marinoi*, and compared seawater microplankton communities to the  
92 “community” of prey organisms in the gut content of *Calanus* sp. copepods. The aim of this  
93 study was to utilize molecular proxies for feeding selection generated by metabarcoding analysis  
94 to investigate whether changes in *Calanus* feeding selection could be linked to changes in the  
95 chemical profile of co-occurring microbial planktonic communities.

96

## 97 **Materials and Methods**

### 98 **Mesocosm experiment**

99 A seawater mesocosm experiment was performed during a 22-day period from 8-30 March 2012  
100 at the Espeyrend Marine Biological Station at the University of Bergen, Norway. A detailed  
101 description of the experimental set-up is provided elsewhere (Nejstgaard *et al.* 2006; Stoecker *et*  
102 *al.* 2015; Ray *et al.* 2016). Briefly, triplicate 11 m<sup>3</sup> reinforced transparent polyurethane  
103 mesocosms bags were either left unamended (Control), amended with 16 μM NO<sub>3</sub><sup>-</sup> and 1 μM  
104 PO<sub>4</sub><sup>3-</sup> (NP) to selectively promote *P. pouchetii* growth, or amended with 16 μM NO<sub>3</sub><sup>-</sup>, 1 μM

105  $\text{PO}_4^{3-}$  and  $5 \mu\text{M SiO}_4^{2-}$  (NPSi) to selectively promote diatom growth. For reference,  
106 unmanipulated samples were also taken from Raunefjorden (Raunefjorden) directly adjacent to  
107 the mesocosm raft. A detailed description of mesocosm bloom development is available  
108 elsewhere (Stoecker *et al.* 2015; Ray *et al.* 2016). Briefly, exponential growth of the diatom *S.*  
109 *marinoi* occurred in all three mesocosm treatments (Control, NP and NPSi), peaking during 18-23  
110 March. Highest abundances of *S. marinoi* occurred in the NPSi treatment. Exponential growth of  
111 *P. pouchetii* blooms commenced after 25 March in both the NP and NPSi mesocosms, although  
112 highest abundances of *P. pouchetii* occurred in the NP mesocosms (Ray *et al.* 2016). The  
113 initially high nutrient levels in the Raunefjorden, from which mesocosm bags were filled,  
114 resulted in similar successive blooms in the different mesocosm treatments despite differential  
115 nutrient amendment (Stoecker *et al.* 2015).

116

### 117 **Sampling for metabarcoding analysis**

118 Samples for metabarcoding were collected on 11, 17, 21, 24, 28 and 30 March 2012 as described  
119 previously (Ray *et al.* 2016, Table 1). Briefly, triplicate 50-200 ml seawater samples from all  
120 mesocosms and Raunefjorden were filtered by gentle vacuum onto  $0.2 \mu\text{m}$  SUPOR filters (Pall  
121 Corporation). Filters were aseptically transferred to 2.0 ml tubes containing  $280 \mu\text{l}$  of  $56^\circ\text{C}$   
122 Buffer ATL and  $20 \mu\text{l}$  Proteinase K ( $20 \text{ mg ml}^{-1}$ ) (QIAGEN DNeasy Blood & Tissue kit). For  
123 feeding chamber incubations, *Calanus* sp. were first collected by net tows from Raunefjorden  
124 ( $60^\circ 16' 18''\text{N}$ ,  $5^\circ 10' 26''\text{E}$ ), and 20 individuals each of adult female or stage V copepodites were  
125 manually sorted into  $\sim 1.8 \text{ L}$  volume grazing chambers containing  $0.2 \mu\text{m}$ -filtered seawater. The  
126 detailed chamber construction is described in Ray *et al.* (2016). Sorted copepods inside grazing  
127 chambers were kept in the dark at  $8^\circ\text{C}$  until deployment inside mesocosms on the following  
128 morning. Three replicate grazing chambers containing copepods were deployed inside each

129 mesocosm. The 500  $\mu\text{m}$  nylon mesh covering both openings of the grazing chambers allowed  
130 constant vertical circulation of mesocosm water containing *in situ* microbial assemblages  
131 through the chambers. After a 24-hour incubation, mesocosm chambers were recovered one at a  
132 time, and copepods were rinsed with three successive washes of 0.2  $\mu\text{m}$ -filtered seawater prior to  
133 a final immersion in an anaesthetic seawater solution containing 0.37  $\text{mg ml}^{-1}$  ethyl 3-  
134 aminobenzoate methanesulfonate (MS222) (Sigma-Aldrich, Norway) (Simonelli *et al.* 2009).  
135 Pools of five copepods from each chamber were sorted into 1.5 ml tubes containing 180  $\mu\text{l}$  of  
136 Buffer ATL preheated to 56°C and 20  $\mu\text{l}$  (20  $\text{mg ml}^{-1}$ ) Proteinase K (QIAGEN DNeasy Blood &  
137 Tissue kit). In summary, we collected three biological replicate copepod samples (five copepod  
138 individuals per sample) and three biological replicate seawater samples (filters) per mesocosm  
139 and from Raunefjorden on each of the six sampling days. Because whole copepods were lysed,  
140 the copepod samples may also contain any symbionts on copepod surfaces or in tissues. For  
141 simplicity, however, we refer to these samples as “copepod gut content” samples throughout the  
142 manuscript. All filter and copepod samples for DNA extraction were lysed at 56°C overnight  
143 then frozen at -20°C until processing. DNA extraction was performed according to the  
144 manufacturer’s protocol, except that two rounds of elution with 100  $\mu\text{l}$  of 56°C Buffer AE  
145 (QIAGEN) were used instead of the recommended single elution step using 200  $\mu\text{l}$  of room  
146 temperature Buffer AE.

147

#### 148 **Amplicon library preparation**

149 The universal primers F-1183mod and R-1443mod (Table 1) targeting the V7 region of the small  
150 subunit ribosomal RNA (SSU rRNA) gene were used in polymerase chain reactions (PCR) to  
151 amplify microbial eukaryotes. Universality of primers was checked using the TestPrime function  
152 on the Arb-Silva website (<http://www.arb-silva.de/search/testprime/>) (Table S1). Primer F-

1183mod was modified with the Roche GS-FLX Lib-L Adapter B sequence (5' - CCT ATC  
CCC TGT GTG CCT TGG CAG TC - TCAG - primer - 3') and used in all PCR reactions.  
Unidirectional sequencing was performed from primer R-1443mod, which was modified with  
Adapter A (5' - CCA TCT CAT CCC TGC GTG TCT CCG AC -TCAG - barcode - primer - 3'),  
where a unique 10-mer barcode multiplex identifier (MID) was included for sample  
identification during demultiplexing. Primers were HPLC-purified to ensure uniform length and  
to eliminate contaminating DNA from lyophilized primer preparations. In order to block *Calanus*  
amplification from copepod samples, a combination of the *Calanus*-specific blocking oligos,  
Cal-SpcC3-block and Cal-PNA-block (Table 1 and Supporting Information) was used in 50 µl  
PCR reactions containing 5 µl of template DNA, 1X HF buffer (New England Biolabs, Ipswich,  
Massachusetts), 0.4 U Phusion DNA polymerase (New England Biolabs), 50 µM each dNTP,  
250 nM each primer, 2 µM Cal-SpcC3-block and 1 µM Cal-PNA-block. For amplification from  
seawater samples, no blocking oligos were included but otherwise PCR conditions were  
identical. Amplification was performed using a C-1000 thermal cycler (Bio-Rad) with one cycle  
of 95°C for 5 min, 30 cycles of 94°C for 20 sec/70.2°C for 10 sec/60.2°C for 20 sec/72°C for 25  
sec, and a final elongation of 72°C for 2 min. PCR products were pooled by sample, purified  
using 0.8 vol magnetic beads (Agencourt Ampure XP, Beckman-Coulter Inc., Indianapolis,  
Indiana, USA), and quantified using PicoGreen (Quant-It PicoGreen dsDNA Assay kit, Life  
Technologies Ltd, Paisley, UK) on a NanoDrop 3300 fluorospectrophotometer (Thermo Fisher  
Scientific, Waltham, Massachusetts). Equimolar amounts of each sample pool were combined  
and vacuum-concentrated to generate one amplicon library from copepod samples and one from  
seawater samples. A detailed description of amplicon library generation for metabarcoding may  
be found in the Supporting Information. Amplicon libraries were sequenced on ½ plate each



176 using Roche GS-FLX Titanium chemistry at the Norwegian Sequencing Centre (University of  
177 Oslo, Norway).

178

### 179 **Sequence analysis**

180 We used two different approaches for analysis of metabarcoding results - operational taxonomic  
181 unit (OTU) clustering and taxonomic classification. In order to identify taxa and OTUs common  
182 to both libraries, sequence data for the two data sets (copepod gut content and seawater amplicon  
183 libraries) were collated prior to taxonomic classification or to OTU clustering using 98%  
184 sequence similarity cut-off. OTU clustering with 98% cut-off was performed using  
185 AmpliconNoise v.1.29 (Quince *et al.* 2009; 2011). Alternatively, forward primer trimming and  
186 quality filtering using trim.seqs in mothur v.1.33 (Schloss *et al.* 2009) was performed with the  
187 following parameters: qaverage=25, maxambig=0, maxhomop=6, minlength=200, flip=T,  
188 pdiffs=1 with raw sequence fasta file and qfile as input. Taxonomic classification of these  
189 quality-filtered reads was performed using the CREST classifier with the SilvaMod database as  
190 reference set (Lanzén *et al.* 2012). CREST taxonomic classification results for individual  
191 samples files were collated using custom awk scripts. Because the SilvaMod database uses the  
192 Silva taxonomy without a strict taxonomic ranking, and because the CREST classifier uses a  
193 lowest common ancestor (LCA) algorithm, we refer to taxonomic assignments according to rank  
194 rather than according to standard taxonomic hierarchy.

195

### 196 **Metabolite profile analysis**

197 Phytoplankton samples for intracellular metabolic profiling were collected every day from 8 to  
198 30 March from all mesocosms and from Raunefjorden and analyzed using a modified protocol  
199 for metabolomic analysis of *Skeletonema marinoi* cultures (Vidoudez & Pohnert 2012). Briefly,

200 cells from 9 L of seawater were concentrated onto GF/C-filters (1.2  $\mu\text{m}$  nominal pore size) with  
201 gentle vacuum ( $\sim 400$  mbar). Depending on chlorophyll *a* (*chl a*) concentration and microscopy  
202 counts, the daily sampling volume was gradually decreased to 1 L during the experiment.  
203 Samples were extracted, derivatized and analysed as described previously (Vidoudez & Pohnert  
204 2012). A detailed description of methods used for metabolomic profiling is provided in the  
205 Supporting Information.

206

### 207 **Statistical analysis**

208 All statistical analyses, unless otherwise noted, were conducted in the R statistical computing  
209 environment (R Core Team, 2015). Alpha- and beta-diversity estimates were performed using  
210 the *rarecurve* and *estimateR* (richness) and *diversity* functions in the “vegan” package v.2-0.10  
211 (Oksanen *et al.* 2013). Non-metric multidimensional scaling (NDMS) ordination of OTU and  
212 taxonomy tables was performed using the *metaMDS* function in “vegan”. Data visualization was  
213 performed using the “*ggplot2*” package (Wickham 2009). Permutational analysis of variance  
214 using distance matrices (PERMANOVA) was performed using the *adonis* function in “vegan”.  
215 Constrained correspondence analysis (CCA) was performed using the *cca* function in “vegan”  
216 with *ordistep* for forward selection to identify significant environmental variables. Correlation  
217 analysis of metabarcoding ratios to individual metabolites was performed using *cor.test*  
218 (*method*="kendall") and custom R scripts. The data matrix from metabolomic runs was analysed  
219 using canonical analysis of principle coordinates by linear discriminant analysis (CAPdiscrim;  
220 Anderson & Willis 2003). CAPdiscrim was performed using the Windows-executable program  
221 CAP12 (Anderson 2004) using the following parameters: choice of transformation=none, choice  
222 of standardisation=none, choice of distance measure=Bray-Curtis, type of analysis=discriminant  
223 analysis, number of principal coordinate axes chosen by the program, number of random

224 permutations=999. Treatment of metabolomics runs to achieve the data matrix is described in the  
225 Supporting Information.

226

## 227 **Results**

### 228 **Metabarcoding analysis**

229 Unidirectional pyrosequencing of the seawater amplicon libraries resulted in 437 522 sequence  
230 reads, while the copepod gut content libraries generated 353 459 sequence reads (Table S2).

231 Read coverage for each sample was relatively even for the seawater library, with a maximum  
232 approximate two-fold variation in number of reads per sample from highest to lowest coverage  
233 and a relatively low standard deviation for both pipelines (Table S2). The *Calanus* gut content  
234 sample coverage was less even, with a nearly nine-fold difference in the number of sequence  
235 reads per sample between samples with lowest and highest sequence coverage (Table S2).

236 Denoising, quality filtering and OTU clustering and chimera removal using AmpliconNoise  
237 generated 3115 OTUs in total for both amplicon libraries, while taxonomic assignment of  
238 denoised, quality-filtered reads using the CREST classifier and the SilvaMod database as  
239 reference resulted in the identification of 1032 unique taxonomic rankings (herein referred to as  
240 taxa) in the collated seawater and *Calanus* gut content sequence data (Table S2). Initial  
241 inspection of OTU clustering and taxonomy results for the sequence data showed that many  
242 reads (2.9 - 36.6 %) present in the *Calanus* gut content amplicon libraries had highest similarity  
243 to *Calanus* copepod SSU sequences (black bars in Fig. S1). *Calanus*-like sequences were  
244 assumed to be copepod sequences that were amplified despite the blocking-PCR strategy used,  
245 and were therefore removed from the copepod gut content OTU and taxonomy tables prior to  
246 diversity or statistical analysis (Table S2, footnote 2). It should be noted that amplification  
247 strategy used is unable to distinguish between (1) *Calanus* sp. sequences that are the result of

248 incomplete blocking of predator DNA amplification and (2) *Calanus* sp. sequences that originate  
249 from ingestion of other calanoid species, nauplii and/or eggs. Assuming the former and thus  
250 removing these sequences prior to downstream analysis, we have therefore not assessed the  
251 contribution of cannibalism or predation on closely-related *Calanus* taxa.

252

253 Rarefaction analysis of AmpliconNoise OTUs (Fig. S2A, B) and CREST taxonomic assignments  
254 (Fig. S2C, D) demonstrated clear undersampling for all *Calanus* gut content samples (Fig. S2A,  
255 C), and in particular for those samples for which relatively few reads were obtained (Table S3).  
256 Seawater samples, however, were sampled to near saturation (Figures S2B, D), with the  
257 exception of seawater samples from Raunefjorden on 17 and 21 March 2012. Closer inspection  
258 of sequence data from these two seawater samples showed a larger number of singletons (data  
259 not shown) and higher biodiversity (Table S3) relative to the other seawater samples. Beta  
260 diversity analysis of abundance-normalized OTU and taxonomy tables based on the inverse  
261 Simpson's diversity index ( $1/D$ ) for Jaccard ( $1/D_J$ ) and Bray-Curtis ( $1/D_{BC}$ ) distance matrices  
262 suggested low variation in diversity between seawater samples regardless of analysis pipeline or  
263 distance metric used (Table S3, Mesocosm seawater). For *Calanus* gut content samples (from  
264 which *Calanus*-like reads were removed), however, we observed higher variation in beta  
265 diversity when Bray-Curtis was used as distance metric, but not when Jaccard distances were  
266 used (Table S3, *Calanus* gut content). In addition, we observed a trend of increasing beta  
267 diversity toward the end of the mesocosm experiment for *Calanus* gut content samples for all  
268 treatments, regardless of analysis pipeline (Table S3, *Calanus* gut content).

269

270 The diversity revealed by metabarcoding of seawater microbial plankton includes all major  
271 marine plankton groups, including rhizarians, diatoms, haptophytes, ciliates, dinoflagellates,

272 chlorophytes, cercozoans, choanoflagellates, cryptophytes, fungi, bivalves, gastropods, tunicates,  
273 crustaceans, land plants and others, indicating that the PCR primers used amplify DNA from a  
274 broad diversity of eukaryotic microorganisms present in the Raunefjorden ecosystem (Fig. 1).  
275 The taxonomic diversity observed was highly uneven, with the 20 most abundant OTUs (Fig.  
276 1A) or taxa (fifth-rank taxonomic assignments according to the Silva taxonomy) (Fig. 1B)  
277 comprising 20-95% of reads in both seawater and copepod gut content sequence libraries.  
278 Inspection of taxonomic assignments for the most abundant OTUs/taxa observed in copepod gut  
279 content revealed that these reads may represent organisms living in a symbiotic relationship  
280 (*sensu* Leung and Polin 2008) with *Calanus*, including Syndiniales (Dinophyceae),  
281 Oligohymenophorea (Ciliophora), *Amoebophyra* (Ciliophora), *Paradinium* (Rhizaria). Lacking  
282 proof, however, that these reads represent symbioses rather than ingested prey, and because these  
283 organisms would likely generate a chemical signal that would be detected by the metabolite  
284 profiling analysis, we chose to retain them in the sequence data for downstream statistical  
285 analyses.

286

287 Non-metric multidimensional scaling (NMDS) analysis of OTU or taxonomic classification  
288 results demonstrated that the strongest effect on sample diversity was sample type, that is,  
289 whether the sample originated from seawater or from *Calanus* gut content (Fig. 2). This  
290 significance was confirmed by PERMANOVA analysis of OTU (AmpliconNoise) and  
291 taxonomic (CREST) diversity, for which  $Pr(>F) = 0.001$  for both pipelines independent of  
292 distance metric (Table 2). Indeed, sample type explained 21-32.1% of the variation in OTU  
293 diversity or 34-47% of the variation in the taxonomic diversity (Table 2) observed.

294 PERMANOVA also identified sampling date as a significant explanatory variable for both  
295 AmpliconNoise OTU (6.8 - 8.8%) and CREST taxonomic (7.6 - 9.6%) diversity (Table 2).

296 Experiment type (mesocosm or fjord) explained a significant fraction of diversity for  
297 AmpliconNoise OTUs (3.8 - 4.6% of OTU diversity) but not for CREST taxonomic diversity (<  
298 1.5%). Treatment was not found to be significant regardless of pipeline or distance metric (Table  
299 2). Kruskal-Wallis rank sum tests identified several taxa (Fig. S3) that were significantly  
300 different in their relative abundance between seawater and *Calanus* gut content samples. In  
301 general, copepod samples were distinguishable by their significantly higher abundances of  
302 protozoans, primarily apostome ciliates and the rhizarian *Paradinium*. For two representative  
303 taxa, *Phaeocystis pouchetii* and *Skeletonema marinoi*, which exhibited exponential growth  
304 during the mesocosm experiment, metabarcoding signals were clearly stronger in seawater  
305 samples than in the corresponding copepod samples (Fig. 1 and Fig. S3). Seawater samples also  
306 contained significantly higher numbers of dinoflagellates, cercozoans, nanophytoplankton  
307 (Phaeocystales, Bacillariophyta, Prymnesiales, Pelagophyta, Chlorophyta), heterotrophic  
308 nanoflagellates and radiolarians relative to the corresponding copepod samples (Fig. S3).  
309 Copepod gut content samples, on the other hand, contained relatively higher numbers of ciliates,  
310 fungi, arthropods, and kinetoplastid protozoans (Fig. S3).

311

### 312 **Molecular proxy for *Calanus* feeding selection**

313 The quantitative power of the metabarcoding data was examined by comparison with previous  
314 results from qPCR and microscopy analysis of *P. pouchetii* and *S. marinoi* performed on samples  
315 from the same mesocosm experiment (Ray *et al.* 2016) (Fig. 3). Metabarcoding signal dynamics  
316 in seawater were generally consistent with both microscopic enumeration and qPCR  
317 quantification of these taxa in each treatment, although the decreasing microscopy and qPCR  
318 signals for *S. marinoi* at the end of the experiment in the NP and NPSi treatments were not  
319 evident in the metabarcoding data (Fig. 3B). In concert, however, these results support a semi-

320 quantitative interpretation of metabarcoding data to evaluate the dynamics of *Calanus* grazing  
321 responses to individual prey taxa as a function of specific prey abundance (D'Amore *et al.* 2016;  
322 Lanzén *et al.* 2016). The ratios of individual OTUs or taxa abundance in copepod gut content to  
323 their abundance in surrounding seawater were therefore calculated as a proxy, or selectivity  
324 index (Irigoien *et al.* 2000), for *Calanus* grazing (Ray *et al.* 2016) on individual OTUs and taxa  
325 (Table S4). Measurable grazing ratios (independent of treatment or sampling date) varied over  
326 seven orders of magnitude (Fig. 4), indicating a large range in apparent copepod feeding  
327 selection. High ratios were interpreted as suggestive of positive feeding selection, while low  
328 ratios were suggestive of negative feeding selection. Highest grazing ratios in this study were  
329 observed for large pennate and centric diatoms, fungi, kinetoplastid protozoans, marine  
330 invertebrates, heterotrophic nanoflagellates, hypotrich ciliates and unknown eukaryotes, while  
331 lowest grazing ratios were observed for oligotrich and choreotrich ciliates, dinoflagellates,  
332 autotrophic flagellates, haptophytes including *P. pouchetii* and small diatoms such as  
333 *Skeletonema* (Table S4, Part C). Grazing ratios for *P. pouchetii* and *S. marinoi* were almost  
334 always low ( $> 1$ ), potentially indicative of negative selection of these food particles by *Calanus*  
335 copepod incubated in mesocosm chambers despite their high relative abundance in the NP and  
336 NPSi mesocosms (Fig. 3 and Ray *et al.* 2016). The temporal dynamics of molecular grazing  
337 ratios calculated from AmpliconNoise OTU- and CREST taxonomic assignments for each  
338 treatment are shown in Fig. S4.

339

#### 340 **Correlation of intracellular metabolites with metabarcoding results**

341 In total, 274 features were detected in the intracellular metabolite profiles from all seawater  
342 samples (Table S5). Canonical correspondence analysis (CCA) with forward selection of  
343 peaksum-normalized metabolite profiles from all sampling days and treatments identified

344 Treatment ( $\text{Pr}( > F ) = 0.005$ ,  $F = 2.0215$ ,  $df = 3$ ) and Date ( $\text{Pr}( > F ) = 0.005$ ,  $F = 2.9512$ ,  $df = 1$ ) as  
345 significant structuring variables. Visual inspection of hierarchical clustering patterns in  
346 metabolite profiles based on chl *a*, phaeophytin and *P. pouchetii* colonial cell abundance  
347 dynamics reported previously (Ray *et al.* 2016) for the NP or NPSi mesocosms revealed three  
348 distinct stages of the mesocosm succession (Fig. 5). The “early” stage of succession, from 11-17  
349 March, was characterized by the exponential growth of the diatom *S. marinoi*, as indicated by an  
350 exponential increase in 0.2  $\mu\text{m}$ -filterable chl *a* (Ray *et al.* 2016) as well as microscopy counts of  
351 *S. marinoi* (Fig. 3). The subsequent plateau of chl *a* and concomitant increase in phaeophytin  
352 were indicators for the “middle” succession phase, which occurred from 18-23 March. Increasing  
353 numbers of *P. pouchetii* colonial cells (an indicator of exponential *P. pouchetii* growth)  
354 (FlowCAM measurements, Ray *et al.* 2016) delimited the “late” phase of mesocosms succession,  
355 from 24-30 March.

356

357 Succession stage was shown to be a highly significant grouping variable for ordination of the NP  
358 and NPSi intracellular metabolite profiles by CAPdiscrim analysis ( $P = 0.001$  with 1000  
359 permutations), explaining 94-95% of the variation in the principal components of the metabolite  
360 data (Fig. 6). Succession phase could also explain 49% of the variation in biomass estimates for  
361 microbial eukaryote taxonomic groups based on microscopy counts (available on  
362 <http://datadryad.org>) from NP and NPSi mesocosm seawater (PERMANOVA,  $df = 2$ ,  $SS =$   
363  $4.0224$ ,  $F = 20.027$ ,  $\text{Pr}( > F ) = 0.001$ ). CAPdiscrim analysis furthermore allowed us to identify  
364 intracellular metabolites, in particular carbohydrates and carbohydrate derivatives, which could  
365 discriminate between the *S. marinoi*-dominated “middle” succession phase and the *P. poucheti*-  
366 dominated “late” succession phase during the experiment (Fig. 5). For example, inositol isomers



367 (e.g. Metabolites 159, 173, 174) as well as saccharides (Metabolite 161) correlated significantly  
368 with the *S. marinoi* dominated succession phase, while Metabolites 134 and 207 were found to  
369 be saccharide-like metabolites whose high concentrations co-occurred significantly with *P.*  
370 *pouchetii* exponential growth. Furthermore, we identified several lipids that were either  
371 significantly correlated to the *S. marinoi*-dominated “middle” mesocosm phase, which included  
372 fatty acids (e.g. Metabolite 194), or to the *P. pouchetii*-dominated “late” mesocosm phase, which  
373 included a terpene (Metabolite 214) and sterols (Metabolites 260, 264, 267). Metabolites  
374 associated with the CAPdiscrim-identified mesocosm succession stages are shown in (Fig. 5).  
375 Based on microscopy biomass estimates (Table S6), several taxonomic groups were found to be  
376 significantly correlated to specific metabolites (Table S7). Weak positive correlation to  
377 Metabolites 173 (inositol isomer) and 205 (glucose derivative), for example, were identified for  
378 *Skeletonema* biomass estimates in NP mesocosms. Metabolites 203 (1-octadecanol), 207  
379 (unknown) and 209 (unknown) were found to be positively correlated to *Phaeocystis* biomass,  
380 while metabolite 262 (cholesterol) was found to be negatively correlated to *Phaeocystis* biomass,  
381 in NP mesocosms.

382

383 To further test whether *Calanus* grazing ratios on individual prey taxa correlated significantly to  
384 metabolites from all mesocosm treatments and Raunefjorden, we performed pairwise correlation  
385 analysis of the 274 detected metabolites to *Calanus* grazing ratios calculated from  
386 AmpliconNoise OTUs that were collated by fifth-rank taxonomic assignments (Table S8). Due  
387 to limited metabarcoding sample numbers and to missing data points for grazing ratios calculated  
388 from the metabarcoding data, it was not possible to identify statistically significant correlations  
389 between grazing ratio dynamics for individual OTUs/taxa and individual metabolites. Closer  
390 inspection of these non-significant results, however, revealed consistent relative

391 “associations” of grazing ratios with sugar derivatives, saccharides, amino acids and their  
392 derivatives, and fatty acids, thus implicating these metabolite classes in *Calanus* grazing  
393 selection (Table S8). Metabolites 159, 173 and 194, for example, which were significantly  
394 correlated with the *S. marinoi*-dominated “middle” mesocosm metabolic phase, were found to be  
395 negatively associated with the grazing ratios for Bacillariophyta in the diatom-dominated NPSi  
396 mesocosm treatment (Kendall's  $\tau = -0.733, -0.87$  and  $-0.733$ , respectively) (Table S8). For the  
397 NP mesocosms, which experienced strongest dominance by *P. pouchetii* during the “late”  
398 mesocosm metabolic phase, metabolites 134, 214, 260, 264 and 267 were found to be positively  
399 associated with *Calanus* grazing ratios on Phaeocystales (Kendall's  $\tau = 0.966, 0.690, 0.828,$   
400  $0.552$  and  $0.552$  respectively). Comprehensive results from correlation analysis of grazing ratios  
401 with metabolites can be found in Table S8.

402

## 403 **Discussion**

### 404 **Selective feeding by *Calanus* copepods**

405 Our molecular characterization of *Calanus* sp. grazing during a seawater mesocosm experiment  
406 provides new evidence for dynamic and discriminate feeding selection by this copepod in mixed  
407 natural microbial plankton assemblages. Although *Calanus* sp. copepods may generate feeding  
408 currents for passive grazing, our findings support numerous empirical studies demonstrating  
409 clear selectivity in copepod feeding behavior (Huntley 1988; Meyer-Harms *et al.* 1999;  
410 Nejstgaard *et al.* 1997; 2008; Barofsky *et al.* 2010, and references therein) in mixed natural  
411 assemblages of microbial plankton. Nutrient manipulation of mesocosms to promote a *P.*  
412 *pouchetii* bloom further indicated that the increase in prey particle size range caused by the  
413 formation and growth of *P. pouchetii* colonies during bloom development does not increase the  
414 ingestion of *P. pouchetii* by *Calanus*, negating size selection as a significant determinant of

415 copepod feeding (Nejstgaard *et al.* 2007). On the contrary, our molecular diet analysis provides  
416 direct evidence that *P. pouchetii* does not contribute significantly to *Calanus* copepod diet in a  
417 mesocosm setting (Nejstgaard *et al.* 2006; Ray *et al.* 2016). The timing of the bloom-like growth  
418 of *P. pouchetii* in this experiment was such that peak abundance occurred around the time when  
419 the experiment was stopped. The complete sample set does therefore not include any samples  
420 from a *P. pouchetii* “bloom decline” stage, precluding our ability to test the hypothesis of Estep  
421 *et al.* (1990) that ageing or senescent *P. pouchetii* colonies, either by merit of physical  
422 degradation or changes in chemical properties, are more readily consumed by *Calanus* copepods  
423 than younger colonies.

424

425 Our molecular grazing ratio results are in accord with idealized food web models and  
426 experimental observations in which mesozooplankton (copepods) have been shown to prefer  
427 ciliates and large diatoms over small autotrophs as a food source (e.g. Kleppel *et al.* 1991;  
428 Stoecker & Capuzzo 1990; Ohman & Runge 1994; Nejstgaard *et al.* 1997, 2001b; Calbet & Saiz  
429 2005; Yang *et al.* 2009; Fileman *et al.* 2010). The low grazing selection ratios that we observed  
430 for dinoflagellate taxa including *Gyrodinium*, the most abundant genus observed in our  
431 experiment (Stoecker *et al.* 2015, Ray *et al.* 2016), were more surprising as previous studies have  
432 shown high ingestion rates of dinoflagellates by calanoid copepods (Levinsen *et al.* 2000; Batten  
433 *et al.* 2001; Olson *et al.* 2006). For example, Fileman *et al.* (2010) observed high clearance rates  
434 by *C. helgolandicus* on *Gyrodinium fusiforme/spirale* in June in the English Channel; at this time  
435 large heterotrophic dinoflagellates dominated the microzooplankton biomass and large  
436 choreotrich ciliates were scarce. However, in most grazing experiments, choreotrich ciliates  
437 accounted for a larger proportion of the microzooplankton carbon ingested than did heterotrophic

438 dinoflagellates (Fileman *et al.* 2010). Based on our grazing proxy, our results suggest that  
439 *Gyrodinium* spp. dinoflagellates were not preferred prey during the mesocosm experiment. One  
440 explanation for this finding is active avoidance of *Calanus* predation by dinoflagellates (Granéli  
441 *et al.* 1993; Nejstgaard *et al.* 1997; 2001a; Verity 2010). Alternatively, ingested dinoflagellates  
442 might be rapidly digested in the *Calanus* digestive tract (e.g. Sullivan 2011), however we did not  
443 observe consistently low grazing ratios for all naked autotroph and protist taxa present in the  
444 grazing ratio data (Table S4), nor are we aware of studies demonstrating taxon-dependent  
445 differential prey DNA digestion rates. An additional surprising finding of our grazing ratio proxy  
446 is the apparent positive selection by *Calanus* copepods on fungi. Dikarya were highly grazed in  
447 all treatments and Raunefjorden throughout the experiment (Fig. S4). Inspection of taxonomic  
448 classification of these taxa revealed primarily peizizomycotina (Ascomycota) and agaricomycetes  
449 (Basidiomycota). The high relative presence of these fungal taxa in *Calanus* gut content may  
450 suggest grazing on fungal hyphae or spores, but may also suggest the consumption of detrital  
451 material by copepods (although see Paffenhöffer & Strickland 1970), as these fungal groups are  
452 known saprophytes of marine macroalgae (Richards *et al.* 2012). It should be further noted that  
453 the apparent gut content of harvested copepods likely reflects prey particles consumed only in  
454 the time period (< 1 hour) immediately prior to sampling, as prey DNA digestion (Troedsson *et*  
455 *al.* 2009) and rapid copepod gut passage (Nejstgaard *et al.* 2003) would certainly hinder  
456 detection of prey particles cumulatively consumed during the entire mesocosm incubation  
457 period. The metabarcoding data from copepod gut content samples thus represents a snapshot of  
458 grazing activity by *Calanus* copepods incubated in mesocosm grazing chambers. In order to  
459 ensure that the seawater microbial plankton diversity would provide a concurrent diversity  
460 “scaffold” upon which to assess prey selection by feeding copepods, seawater samples for

461 metabarcoding analysis were collected from mesocosms and Raunefjorden immediately prior to  
462 copepod harvesting.

463

464 The interpretation of metabarcoding results as an indication of positive or negative grazing  
465 selection by *Calanus* on specific OTUs or taxa requires the implicit assumption that  
466 metabarcoding abundances correspond to biologically meaningful quantities, i.e. prey  
467 abundances. The potential sources of PCR and target gene copy number bias that can limit the  
468 quantitative power of metabarcoding approaches in microbial ecology have been described in  
469 other publications (Richards & Bass 2005; Potvin & Lovejoy 2009; Amend *et al.* 2010; Stoeck *et*  
470 *al.* 2010; Deagle *et al.* 2010; 2013; Gong *et al.* 2013). Cognizant of these cautions, we have  
471 utilized grazing ratios rather than absolute abundances, thus normalizing read abundance data for  
472 some of the bias inherent in PCR-based sequencing library preparation. We demonstrate  
473 correspondence between metabarcoding analysis, microscopy counts and qPCR estimations for  
474 key mesocosm taxa, namely *P. pouchetii* and *S. marinoi*. The semi-quantitative interpretation of  
475 metabarcoding results for individual OTUs or taxa (D'Amore *et al.* 2016; Lanzén *et al.* 2016) has  
476 thus generated a numerical proxy by which grazing selection may be evaluated, and which has  
477 provided a useful tool for correlative analysis of copepod feeding selection with plankton  
478 metabolic succession. Contemporary calls for an integrated approach to marine biodiversity  
479 assessment and baseline establishment include the contribution of molecular analysis (Duffy *et*  
480 *al.* 2013), thus critical evaluations of the quantitative power of, for example, metabarcoding data  
481 to investigate trophic interactions (King *et al.* 2008), are both timely and relevant.

482

483 **Metabarcoding analysis marker selection**

484 Sequence read length limitations of earlier generation sequencing technologies restricted the  
485 choice of SSU rRNA hypervariable target region to those regions (typically V9 for eukaryotes,  
486 ~200 bp) representing the best compromise between phylogenetic information density and  
487 shortest possible target fragment length (Amaral-Zettler *et al.* 2009; Brown *et al.* 2009; Stoeck *et*  
488 *al.* 2009; Edgcomb *et al.* 2011). Later improvements in sequencing chemistry, however, have  
489 facilitated sequencing of longer target regions, thus increasing the possibility for exploration of  
490 microbial diversity using other hypervariable regions within the SSU rRNA gene (Chariton *et al.*  
491 2010; Stoeck *et al.* 2010; Behnke *et al.* 2011; Monchy *et al.* 2011; Lanzén *et al.* 2016).  
492 Molecular analysis of trophic interactions, however, must also assume considerable prey DNA  
493 degradation inside the host gut, which reduces the informational advantage obtained with longer  
494 sequence reads (Troedsson *et al.* 2009). The diversity of prey organisms observed in copepod gut  
495 content samples in this study (Fig. 1) confirms that the V7 region of the SSU rRNA gene  
496 targeted in this study (Hadziavdic *et al.* 2014) provides a satisfactory compromise between  
497 fragment length (260-360 bp) and phylogenetic/taxonomic resolution (Table S1) for broad  
498 exploration of microbial eukaryote plankton diversity in the context of copepod grazing  
499 selection.

500

#### 501 **Putative protist symbionts of *Calanus***

502 The most abundant taxa identified among classified reads and OTUs in the copepod gut content  
503 sequence data had highest similarity to the taxonomically ambiguous *Paradinium* (Shields 1994;  
504 Carman & Dobbs 1997; Skovgaard & Daugbjerg 2008), the dinoflagellate groups Blastodinia  
505 and Syndinia (Shields 1994), Apicomplexa (Rueckert *et al.* 2011), and the apostome ciliate clade  
506 Oligohymenophorea (Carman and Dobbs 1997; Prokopowicz *et al.* 2010; Guo *et al.* 2012;

507 Chantangsi *et al.* 2013). Symbioses between *Calanus* and commensals and/or parasites would be  
508 characterized by our proxy as having high grazing ratios despite the fact that *Calanus* had not  
509 consumed these organisms *per se*. We have therefore not drawn conclusions about *Calanus*  
510 grazing activity on these taxa. Furthermore, the nature of the biological relationship of these taxa  
511 with the *Calanus* copepods in which they were detected falls outside the scope of this study, yet  
512 raises questions about the incidence of *Calanus* parasitism and the effect of parasitism on  
513 copepod health, reproduction and feeding behavior (Skovgaard & Saiz 2006; Cirtwill *et al.* 2015;  
514 Worden *et al.* 2015). Future studies of *Calanus* feeding using state-of-the-art high-throughput  
515 sequencing technology might increase the sequencing depth of *Calanus* gut content sample  
516 coverage to improve the sensitivity of detection of prey organisms even in the presence of  
517 symbioses. Alternatively, prey enrichment strategies such as blocking PCR oligos (Troedsson *et*  
518 *al.* 2008b; Hu *et al.* 2014) against dominant suspected symbiont sequences, DHPLC-PCR  
519 (Troedsson *et al.* 2008a,b; Olsen *et al.* 2012; 2014) and/or restriction enzyme treatment (Maloy  
520 *et al.* 2013) might also be employed to overcome symbiont phylogenetic signal. Amplicon  
521 libraries were not generated from starved copepods in this study, which would have enabled a  
522 more conclusive assessment of prey versus symbiont. The sequencing strategy employed in this  
523 study to investigate *Calanus* grazing choice has nonetheless uncovered a high diversity of  
524 *Calanus*-associated microbial eukaryotes that suggests a considerable degree of endemic  
525 symbioses in copepods in accord with previous observations (Skovgaard & Saiz 2006; Rueckert  
526 *et al.* 2011; Bickel *et al.* 2012; Lima-Mendez *et al.* 2015).

527

## 528 **Role of Metabolites in Grazing Selection by *Calanus***

529 The importance of carbohydrates and lipids was a feature common to all correlation analyses

530 (metabolite succession phase, biomass of taxonomic groups, molecular grazing ratios) performed  
531 in this study, implicating the importance of these metabolite classes in the trophodynamics of  
532 *Calanus* feeding. For example, inositol-related peaks (Metabolites 159, 173 and 174) were  
533 positively correlated with the *S. marinoi*-dominated “middle” metabolite succession phase as  
534 well as *Skeletonema* biomass estimates, but negatively associated to *Calanus* grazing ratios on  
535 diatoms, in the NPSi treatment. Using a qPCR-based molecular grazing proxy, we have  
536 previously shown that *Calanus* grazing selection on *S. marinoi* was low during exponential  
537 growth of *S. marinoi*, but increased with diatom bloom senescence (Ray *et al.* 2016). Indeed,  
538 grazing ratios for diatoms, albeit generally low, increased toward the end of the mesocosm  
539 experiment in the NP and NPSi treatments in this study (Fig. S5, “Bacillariophyta”), suggesting  
540 that copepod grazing on diatoms only increased as the diatom bloom declined during the “late”  
541 metabolite succession phase. These results further suggest that exponential-phase *S. marinoi* is  
542 not preferred prey for *Calanus*, and that *Calanus* grazing selection may rather be linked to the  
543 developmental status of the greater prey community that is reflected in its chemical composition,  
544 as previously demonstrated (Barofsky *et al.* 2010; Ray *et al.* 2016).

545

546 The complexity of plankton succession in natural assemblages presents an ongoing challenge of  
547 our ability to identify taxon-associated metabolites and their role(s) in grazing selection by  
548 copepods. For example, Metabolite 260, a steroid, was significantly positively correlated to the  
549 “late” metabolome succession phase of the NP and NPSi treatments, however it was not  
550 significantly correlated to microscopy-based biomass estimates for *P. pouchetii* (Table S7),  
551 which was a dominant phytoplankton taxon during this phase. This suggests correspondence of  
552 this metabolite to the “late” succession phase itself, rather than to one specific taxon. Our



553 additional observation that this metabolite was positively associated with *Calanus* grazing ratios  
554 on *P. pouchetii* in the NP treatment (Table S8) is therefore likely anecdotal, i.e. that *Calanus*  
555 simultaneously consumed *P. pouchetii* colonies or colony fragments that were physically  
556 associated with other prey particles such as diatoms (Smetacek *et al.* 2002), whose lipids and  
557 carbohydrates may have become more bioavailable to *Calanus* due to diatom bloom senescence  
558 during the “late” succession phase. Indeed, Metabolite 207 (galactosylglycerol) was weakly  
559 positively correlated with *P. pouchetii* biomass estimates (Table S7), but also positively  
560 associated with molecular grazing ratios for diatoms (Table S8). The positive association  
561 between Metabolite 207 and *P. pouchetii* biomass may therefore indicate that the underlying  
562 biological interaction was in fact due to the co-occurring *S. marinoi* bloom decline rather than to  
563 the increase in *P. pouchetii* biomass *per se*. In concert, these results highlight the biological and  
564 chemical complexity of the seawater mesocosm plankton communities.

565

## 566 **Conclusions**

567 In order to generate new knowledge about copepod grazing behavior, we have implemented a  
568 proxy for *Calanus* feeding selection based on metabarcoding analysis of eukaryotic microbial  
569 plankton communities. This proxy has revealed a diverse prey landscape, and allowed us to  
570 identify associations between copepod feeding selection dynamics and metabolic signatures of  
571 co-occurring microbial plankton communities. Our correlative approach has provided consistent  
572 indication of the importance of carbohydrates and lipids during *Calanus* feeding selection. In line  
573 with the findings of previous studies, the combined metabarcoding and metabolomic analyses  
574 therefore suggest that prey selection by copepods is determined by the developmental status of  
575 the diverse prey community, rather than by relative abundances of individual prey taxa. We have

576 demonstrated that the combined analysis of high-resolution biological and chemical factors can  
577 provide new information about copepod feeding selecting in dynamic and complex prey  
578 assemblages. This knowledge improves our understanding of the efficiency of the marine food  
579 web and increases our ability to predict responses to perturbation and changing climate on a  
580 temporal and spatial scale relevant for evolutionary forces on plankton.

581

## 582 **Acknowledgments**

583 This study was funded by the Research Council of Norway (RCN) research project “A novel  
584 cross-disciplinary approach to solve an old enigma: the food-web transfer of the mass-blooming  
585 phytoplankter *Phaeocystis pouchetii*” (Phaeonigma, project number 204479/F20). Additional  
586 support was received from the European Research Council Advanced Grant ERC-AG-LS8  
587 “Microbial Network Organisation” (MINOS, project number 250254), the German Research  
588 Foundation within the framework of the CRC1127 ChemBioSys and the RCN project “Processes  
589 and players in Arctic marine pelagic food webs - biogeochemistry, environment and climate  
590 change” (MicroPolar, project number 225956/E10). JLR received Short Term Scientific Mission  
591 travel support through the COST-EU funding program ES1103-Microbial ecology & the earth  
592 system: collaborating for insight and success with the new generation sequencing tools (COST-  
593 STSM-ES1103-16732 and COST □ STSM □ ES1103 □ 21370).

594

## 595 **References**

- 596 Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying  
597 protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-  
598 subunit ribosomal RNA genes. PLoS One 4: e6372
- 599 Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454  
600 pyrosequencing: does read abundance count? Mol Ecol 19: 5555-5565

- 601 Anderson MJ, Willis TJ (2003) Canonical analysis of principal coordinates: a useful method of  
602 constrained ordination for ecology. *Ecol* 84: 511-525
- 603 Anderson MJ (2004) CAP: a FORTRAN computer program for canonical analysis of principal  
604 coordinates. Department of Statistics, University of Auckland, New Zealand
- 605 Ausloos P, Clifton CL, Lias SG, Mikaya AI, Stein SE, Tchekhovskoi DV, Sparkman OD, Zaikin  
606 V, Zhu D (1999) The critical evaluation of a comprehensive mass spectral library. *J Am Soc*  
607 *Mass Spec* 10: 287-299
- 608 Barofsky A, Simonelli P, Vidoudez C, Troedsson C, Nejstgaard JC, Jakobsen HH, Pohnert G  
609 (2010) Growth phase of the diatom *Skeletonema marinoi* influences the metabolic profile of the  
610 cells and the selective feeding of the copepod *Calanus* spp. *J Plankt Res* 32: 263-272
- 611 Batten SD, Fileman ES, Halvorsen E (2001) The contribution of microzooplankton to the diet of  
612 mesozooplankton in an upwelling filament off the north west coast of Spain. *Prog Oceanogr* 51:  
613 385-398
- 614 Behnke A, Engel M, Christen R, Nebel M, Klein RR, Stoeck T (2011) Depicting more accurate  
615 pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA  
616 gene regions. *Environ Microbiol* 13: 340-349
- 617 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful  
618 approach to multiple testing. *J Roy Stat Soc B* 57: 289-300
- 619 Bickel SL, Tang KW, Grossart H-P (2012) Ciliate epibionts associated with crustacean  
620 zooplankton in German lakes: Distribution, motility and bacterivory. *Front Microbiol* 3: 27-54
- 621 Brown MV, Philip GK, Bunge JA, Smith MC, Bissett A, Lauro FM, Fuhrman JA, Donachie SP  
622 (2009) Microbial community structure in the North Pacific ocean. *ISME J* 3: 1374-1386
- 623 Calbet A, Saiz E (2005) The ciliate-copepod link in marine ecosystems. *Aquat Microb Ecol* 38:  
624 157-167
- 625 Carman KR, Dobbs FC (1997) Epibiotic microorganisms on copepods and other marine  
626 crustaceans. *Microscop Res Tech* 37: 116-135
- 627 Chantangsi C, Lynn DH, Rueckert S, Prokopowicz AJ, Panha S, Leander BS (2013) *Fusiforma*  
628 *themisticola* n. gen., n. sp., a new genus and species of apostome ciliate infecting the hyperiid  
629 amphipod *Themisto libellula* in the Canadian Beaufort Sea (Arctic Ocean), and establishment of  
630 the Pseudocolliniidae (Ciliophora, Apostomatia). *Protist* 164: 793-810
- 631 Chariton AA, Court LN, Hartley DM, Colloff MJ, Hardy CM (2010) Ecological assessment of  
632 estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Front Ecol Environ* 8: 233-

- 633 238
- 634 Cirtwill AR, Stouffer DB (2015) Concomitant predation on parasites is highly variable but  
635 constrains the ways in which parasites contribute to food web structure. *J Anim Ecol* 84: 734-744
- 636 Cowles TJ, Olson RJ, Chisholm SW (1988) Food selection by copepods: discrimination on the  
637 basis of food quality. *Mar Biol* 100: 41-49
- 638 D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M,  
639 Quince C, Hall N (2016) A comprehensive benchmarking study of protocols and sequencing  
640 platforms for 16S rRNA community profiling. *BMC Genomics* 17: 55
- 641 Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to  
642 determine diet of little penguins: is what goes in what comes out? *Conserv Genet* 11: 2039-2048
- 643 Deagle BE, Thomas AC, Shaffer AK, Trites AW, Jarman SN (2013) Quantifying sequence  
644 proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: which counts  
645 count? *Mol Ecol Res* 13: 620-633
- 646 Duffy JE, Amaral-Zettler LA, Fautin DG, Paulay G, Rynearson TA, Sosik HM, Stachowicz JJ  
647 (2013) Envisioning a Marine Biodiversity Observation Network. *63*: 350-361.
- 648 Edgar RC (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads.  
649 *Nature Meth* 10: 996-998
- 650 Edgcomb V, Orsi W, Bunge J, Jeon S, Christen R, Leslin C, Holder M, Taylor GT, Suarez P,  
651 Varela R, Epstein S (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I.  
652 Pyrosequencing vs Sanger insights into species richness. *ISME J* 5: 1344-1356
- 653 Estep KW, Nejstgaard JC, Skjoldal HR, Rey F (1990) Predation by copepods upon natural  
654 populations of *Phaeocystis pouchetii* as a function of the physiological state of the prey. *Mar*  
655 *Ecol Prog Ser* 67: 235-249
- 656 Fileman E, Petropavlovsky A, Harris R (2010) Grazing by the copepods *Calanus helgolandicus*  
657 and *Acartia clausi* on the protozooplankton community at station L4 in the Western English  
658 Channel. *J Plankt Res* 32: 709-724
- 659 Gong J, Dong J, Liu X, Massana R (2013) Extremely high copy numbers and polymorphisms of  
660 the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist*  
661 164: 369-379
- 662 Granéli E, Olsson P, Carlsson P, Granéli W, Christer N (1993) Weak 'top-down' control of  
663 dinoflagellate growth in the coastal Skagerrak. *J Plankt Res* 15: 213-237

- 664 Guo Z, Liu S, Hu S, Li T, Huang Y, Liu G, Zhang H, Lin S (2012) Prevalent ciliate symbiosis on  
665 copepods: high genetic diversity and wide distribution detected using small subunit ribosomal  
666 RNA gene. PLoS One 7: e44847
- 667 Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C (2014)  
668 Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. PLoS  
669 One 9: e87624
- 670 Hu S, Guo Z, Li T, Carpenter EJ, Liu S, Lin S (2014) Detecting *in situ* copepod diet diversity  
671 using molecular technique: development of a copepod/symbiotic ciliate-excluding eukaryote-  
672 inclusive PCR protocol. PLoS One 9: e103528.
- 673 Huntley M (1988) Feeding biology of *Calanus*: a new perspective. Hydrobiol 167: 83-99
- 674 Irigoien X (1998) Gut clearance rate constant, temperature and initial gut contents: a review. J  
675 Plankt Res 20: 997-1003
- 676 King RA, Read DS, Traugott M, Symondson WO (2008) Molecular analysis of predation: a  
677 review of best practice for DNA-based approaches. Mol Ecol 17: 947-963
- 678 Kiørboe T, Saiz E, Viitasalo M (1996) Prey switching behaviour in the planktonic copepod  
679 *Acartia tonsa*. Mar Ecol Prog Ser 143: 65-75
- 680 Kiørboe T (2011) How zooplankton feed: mechanisms, traits and trade-offs. Biol Rev 86: 311-  
681 340
- 682 Kleppel GS, Holliday DV, Pieper RE (1991) Trophic interactions between copepods and  
683 microplankton: a question about the role of diatoms. Limnol Oceanogr 36: 172-178
- 684 Kleppel GS (1993) On the diets of calanoid copepods. Mar Ecol Prog Ser 99: 183-195
- 685 Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2012)  
686 Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation  
687 sequencing-based diversity studies. Nucl Acids Res gks808
- 688 Kuhlisch C, Pohnert G (2015) Metabolomics in chemical ecology. Nat Prod Rep 32: 937-955
- 689 Landry MR (2002) Integrating classical and microbial food web concepts: evolving views from  
690 the open-ocean tropical Pacific. Hydrobiologia 480: 29-39
- 691 Lanzén A, Jørgensen SL, Huson DH, Gorfer M, Grindhaug SH, Jonassen I, Øvreås L, Urich T  
692 (2012) CREST—Classification resources for environmental sequence tags. PLoS One 7: e49334
- 693 Lanzén A, Lekang K, Jonassen I, Thompson EM, Troedsson C (2016) High-throughput

- 694 metabarcoding of eukaryotic diversity for environmental monitoring of offshore oil drilling  
695 activities. *Mol Ecol (in press)*: 10.1111/mec.13761
- 696 Leung TLF, Polin R (2008) Parasitism, commensalism, and mutualism: Exploring the many  
697 shades of symbioses. *Life & Environ* 58: 107-115.
- 698 Levinsen H, Turner JT, Nielsen TG, Hansen BW (2000) On the trophic coupling between  
699 protists and copepods in arctic marine ecosystems. *Mar Ecol Prog Ser* 204: 65-77
- 700 Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinosa  
701 JC, Roux S, Vincent F, Bittner L, Darzi Y, Wang J, Audic S, Berline L, Bontempi G, Cabello  
702 AM, Coppola L, Cornejo-Castillo FM, d'Ovidio F, De Meester L, Ferrera I, Garet-Delmas M-J,  
703 Guidi L, Lara E, Pesant S, Royo-Llonch M, Salazar G, Sánchez P, Sebastian M, Souffreau C,  
704 Dimier C, Picheral M, Searson S, Kandels-Lewis S, Tara Oceans coordinators, Gorsky G, Not F,  
705 Ogata H, Speich S, Stemmann L, Weissenbach J, Wincker P, Acinas SG, Sunagawa S, Bork P,  
706 Sullivan MB, Karsenti E, Bowler C, de Vargas C, Raes J (2015) Determinants of community  
707 structure in the global plankton interactome. *Science* 348: 1262073
- 708 Maloy AP, Culloty SC, Slater JW (2013) Dietary analysis of small planktonic consumers: a case  
709 study with marine bivalve larvae. *J Plankt Res*: fbt027
- 710 Meyer-Harms B, Irigoien X, Head R, Harris R (1999) Selective feeding on natural phytoplankton  
711 by *Calanus finmarchicus* before, during, and after the 1997 spring bloom in the Norwegian Sea.  
712 *Limnol Oceanogr* 44: 154-165
- 713 Monchy S, Sancier G, Jobard M, Rasconi S, Gerphagnon M, Chabé M, Cian A, Meloni D, Niquil  
714 N, Christaki U, Viscogliosi E (2011) Exploring and quantifying fungal diversity in freshwater  
715 lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. *Environ*  
716 *Microbiol* 13: 1433-1453
- 717 Nejstgaard JC, Gismervik I, Solberg PT (1997) Feeding and reproduction by *Calanus*  
718 *finmarchicus*, and microzooplankton grazing during mesocosm blooms of diatoms and the  
719 coccolithophore *Emiliana huxleyi*. *Mar Ecol Prog Ser* 147: 197-217
- 720 Nejstgaard JC, Naustvoll L-J, Sazhin A (2001a) Correcting for underestimation of  
721 microzooplankton grazing in bottle incubation experiments with mesozooplankton. *Mar Ecol*  
722 *Prog Ser* 221: 59-75
- 723 Nejstgaard JC, Hygym B, Naustoll L-J, Båmstedt U (2001b) Zooplankton growth, diet and  
724 reproductive success compared in simultaneous diatom- and flagellate-microzooplankton  
725 dominated plankton blooms. *Mar Ecol Prog Ser* 221: 77-91
- 726 Nejstgaard JC, Frischer ME, Raule CL, Gruebel R, Kohlberg KE, Verity PG (2003) Molecular

- 727 detection of algal prey in copepod guts and faecal pellets. *Limnol Oceanogr Meth* 1: 29-38
- 728 Nejstgaard JC, Frischer ME, Verity PG, Anderson JT, Jacobsen A, Zirbel MJ, Larsen A,  
729 Martínez-Martínez J, Sazhin AF, Walters T, Bronk DA (2006) Plankton development and  
730 trophic transfer in seawater enclosures with nutrients and *Phaeocystis pouchetii* added. *Mar Ecol*  
731 *Prog Ser* 321: 99-121
- 732 Nejstgaard JC, Tang KW, Steinke M, Dutz J, Koski M, Antajan E, Long JD (2007) Zooplankton  
733 grazing on *Phaeocystis*: a quantitative review and future challenges. *Biogeochem* 83: 147-172
- 734 Nejstgaard JC, Frischer ME, Simonelli P, Troedsson C, Brakel M, Adiyaman F, Sazhin AF,  
735 Artigas LF (2008) Quantitative PCR to estimate copepod feeding. *Mar Biol* 153: 565-577
- 736 Ohman MD, Runge JA (1994) Sustained fecundity when phytoplankton resources are in short  
737 supply: Omnivory by *Calanus finmarchicus* in the Gulf of St. Lawrence. *Limnol Oceanogr* 39:  
738 21-36
- 739 Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos  
740 P, Stevens MHH, Wagner H (2015) vegan: Community Ecology Package. R package version  
741 2.3-0. <http://CRAN.R-project.org/package=vegan>
- 742 Olsen BR, Dahlgren K, Schander C, Båmstedt U, Rapp HT, Troedsson C (2012) PCR–DHPLC  
743 assay for the identification of predator–prey interactions. *J Plankt Res* 27: fbr110
- 744 Olsen BR, Troedsson C, Hadziavdic K, Pedersen RB, Rapp HT (2014) A molecular gut content  
745 study of *Themisto abyssorum* (Amphipoda) from Arctic hydrothermal vent and cold seep  
746 systems. *Mol Ecol* 23: 3877-3889
- 747 Olson MB, Lessard EJ, Wong CHJ, Bernhardt MJ (2006) Copepod feeding selectivity on  
748 microplankton, including the toxigenic diatoms *Pseudo-nitzschia* spp. in the coastal northwest  
749 Pacific. *Mar Ecol Prog Ser* 326: 207-220
- 750 Pepin P, Colbourne E, Maillet G. (2011) Seasonal patterns in zooplankton community structure  
751 on the Newfoundland and Labrador Shelf. *Prog Oceanogr* 91: 274-285
- 752 Pitta P, Nejstgaard JC, Tsagaraki TM, Zervoudaki S, Egge JK, Frangoulis C, Lagaria A,  
753 Magiopoulos I, Psarra S, Sandaa R-A, Skjoldal EF, Tanaka T, Thyrhaug R, Thingstad TF (2016)  
754 Confirming the “Rapid phosphorous transfer from microorganisms to mesozooplankton in the  
755 Eastern Mediterranean Sea” scenario through a mesocosm experiment. *J Plankt Res, in press*
- 756 Pompanon F, Deagle BE, Symondson WO, Brown DS, Jarman SN, Taberlet P (2012) Who is  
757 eating what: diet assessment using next generation sequencing. *Mol Ecol* 21: 1931-1950
- 758 Potvin M, Lovejoy C (2009) PCR-based diversity estimates of artificial and environmental 18S

- 759 rRNA gene libraries. *J Euk Microbiol* 56: 174-181
- 760 Poulet SA, Marsot P (1978) Chemosensory grazing by marine calanoid copepods (Arthropoda:  
761 Crustacea). *Science* 200: 1403-1405.
- 762 Prokopowicz AJ, Rueckert S, Leander BS, Michaud J, Fortier L (2010) Parasitic infection of the  
763 hyperiid amphipod *Themisto libellula* in the Canadian Beaufort Sea (Arctic Ocean), with a  
764 description of *Ganymedes themistos* sp. n. (Apicomplexa, Eugregarinorida). *Polar Biol* 33: 1339-  
765 1350
- 766 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2012) The  
767 SILVA ribosomal RNA gene database project: improved data processing and web-based tools.  
768 *Nucl Acids Res* 41: D590-D596
- 769 Quince C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF, Sloan WT (2009)  
770 Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Meth* 6:  
771 639-641
- 772 Quince C, Lanzén A, Davenport RJ, Turnbaugh PJ (2011). Removing noise from pyrosequenced  
773 amplicons. *BMC Bioinform* 12: 38
- 774 Paffenhöfer GA, Strickland JDH (1970) A note on the feeding of *Calanus helgolandicus* on  
775 detritus. *Mar Biol* 5: 97-99
- 776 R Development Core Team (2015) *R: A Language and Environment for Statistical Computing*. R  
777 Foundation for Statistical Computing, Vienna, Austria
- 778 Ray JL, Skaar KS, Simonelli PS, Larsen A, Sazhin A, Jakobsen HH, Nejstgaard JC, Troedsson C  
779 (2016) Molecular gut content analysis demonstrates that *Calanus* grazing on *Phaeocystis*  
780 *pouchetii* and *Skeletonema marinoi* is sensitive to bloom phase but not prey density. *Mar Ecol*  
781 *Prog Ser* 542: 63-77
- 782 Richards TA, Bass D (2005) Molecular screening of free-living microbial eukaryotes: diversity  
783 and distribution using a meta-analysis. *Curr Op Microbiol* 8: 240-252
- 784 Richards TA, Jones MDM, Leonard G, Bass D (2012) Marine fungi: their ecology and molecular  
785 diversity. *Annu Rev Mar Sci* 4: 495-522
- 786 Rueckert S, Simdyanov TG, Aleoshin VV, Leander BS (2011) Identification of a divergent  
787 environmental DNA sequence clade using the phylogeny of gregarine parasites (Apicomplexa)  
788 from crustacean hosts. *PLoS One* 6: e18163d quality and population control. *Aquat Microb Ecol*  
789 7: 197-223
- 790 Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,



- 791 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF  
792 (2009) Introducing mothur: Open-source, platform-independent, community-supported software  
793 for describing and comparing microbial communities. *Appl Environ Microbiol* 75: 7537-7541
- 794 Sherr E, Sherr B (1988) Role of microbes in pelagic food webs: A revised concept. *Limnol*  
795 *Oceanogr* 33: 1225-1227
- 796 Shields JD (1994). The parasitic dinoflagellates of marine crustaceans. *Annu Rev Fish Diseases*  
797 4: 241-271
- 798 Simonelli P, Troedsson C, Nejstgaard JC, Zech K, Larsen JB, Frischer ME (2009) Evaluation of  
799 DNA extraction and handling procedures for PCR-based copepod feeding studies. *J Plankt Res*  
800 31: 1465-1474
- 801 Skovgaard A, Saiz E (2006) Seasonal occurrence and role of protistan parasites in coastal marine  
802 zooplankton. *Mar Ecol Prog Ser* 327: 37-49
- 803 Skovgaard A, Daugbjerg N (2008) Identity and systematic position of *Paradinium poucheti* and  
804 other *Paradinium*-like parasites of marine copepods based on morphology and nuclear-encoded  
805 SSU rDNA. *Protist* 159: 401-413
- 806 Smetacek V, Klaas C, Menden-Deuer S, Rynearson TA (2002) Mesoscale distribution of  
807 dominant diatom species relative to the hydrographical field along the Antarctic Polar Front.  
808 *Deep Sea Res II* 49: 3835-3848
- 809 Stoeck T, Behnke A, Christen R, Amaral-Zettler L, Rodriguez-Mora MJ, Chistoserdov A, Orsi  
810 W, Edgcomb VP (2009) Massively parallel tag sequencing reveals the complexity of anaerobic  
811 marine protistan communities. *BMC Biol* 7: 72
- 812 Stoeck T, Bass D, Nebel M, Christen R, Jones MD, Breiner HW, Richards TA (2010) Multiple  
813 marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic  
814 community in marine anoxic water. *Mol Ecol* 19: 21-31
- 815 Stoecker DK, Capuzzo JM (1990) Predation on protozoa: its importance to zooplankton. *J Plankt*  
816 *Res* 12: 891-908
- 817 Stoecker DK, Nejstgaard JC, Madhusoodhanan R, Pohnert G, Wolfram S, Jakobsen HH, Šulčius  
818 S, Larsen A (2015) Underestimation of microzooplankton grazing in dilution experiments due to  
819 inhibition of phytoplankton growth. *Limnol Oceanogr* 60: 1426-1438
- 820 Sullivan LJ (2010) Gut evacuation of larval *Mnemiopsis leidyi* A. Agassiz (Ctenophora, Lobata).  
821 *J Plankt Res* 32: 69-74
- 822 Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation

- 823 biodiversity assessment using DNA metabarcoding. *Mol Ecol* 21: 2045-2050
- 824 Troedsson C, Lee RF, Stokes V, Walters TL, Simonelli P, Frischer ME (2008a) Development of  
825 a denaturing high-performance liquid chromatography method for detection of protist parasites  
826 of metazoans. *Appl Environ Microbiol* 74: 4336-4345
- 827 Troedsson C, Lee RF, Walters T, Stokes V, Brinkley K, Naegele V, Frischer ME (2008b)  
828 Detection and discovery of crustacean parasites in blue crabs (*Callinectes sapidus*) by using 18S  
829 rRNA gene-targeted denaturing high-performance liquid chromatography. *Appl Environ*  
830 *Microbiol* 74: 4346-4353
- 831 Troedsson C, Simonelli P, Naegele V, Nejstgaard JC, Frischer ME (2009) Quantification of  
832 copepod gut content by differential length amplification quantitative PCR (dla-qPCR). *Mar Biol*  
833 156: 253-259
- 834 Verity PG (2010) Expansion of potentially harmful algal taxa in a Georgia Estuary (USA).  
835 *Harmful Algae* 9: 144-152
- 836 Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare  
837 sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. *Front Zool*  
838 5: 12
- 839 Vidoudez C, Pohnert G (2012) Comparative metabolomics of the diatom *Skeletonema marinoi* in  
840 different growth phases. *Metabolomics* 8: 654-669
- 841 Wagner C, Sefkow M, Kopka J (2003) Construction and application of a mass spectral and  
842 retention time index database generated from plant GC/EI-TOF-MS metabolite profiles.  
843 *Phytochem* 62: 887-900
- 844 Wickham H (2009) *ggplot2: elegant graphics for data analysis*. Springer, New York
- 845 Woodson CB, Webster DR, Weissburg MJ, Yen J (2007) Cue hierarchy and foraging in calanoid  
846 copepods: ecological implications of oceanographic structure. *Mar Ecol Prog Ser* 330: 163-177
- 847 Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ (2015)  
848 Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science*  
849 347: 736-745
- 850 Yang EU, Kang H-K, Yoo S, Hyun J-H (2009) Contribution of auto- and heterotrophic protozoa  
851 to the diet of copepods in the Ulleung Basin, East Sea/Japan Sea. *J Plankt Res* 31: 647-659

852

853 **Data Accessibility**

854 Pyrosequencing flowgram files (\*.sff) have been submitted to the NCBI Sequence Read Project  
855 with accession number SRP076974. The AmpliconNoise OTU table, the CREST taxonomic  
856 assignment table are available from the Dryad Digital Repository  
857 <http://dx.doi.org/10.5061/dryad.6tn7c>.

858

### 859 **Author Contributions**

860 JLR, JA, PS, JCN, AL, DS, MF, GP and CT designed the research. AL and JCN coordinated the  
861 mesocosm experiment. JLR, JA, PS, DS, AS, GP, MF and CT collected samples. JLR, JA, KSS,  
862 DS and AS performed assay optimization and laboratory processing of samples. JLR, JA, UI and  
863 CQ analyzed the data. JLR, JA, PS, JCN, AL, DS, MF, GP and CT wrote the manuscript.

864

865

### 866 **Table and Figure Captions**

867

868 **Table 1.** Primers and blocking oligonucleotide probes used in this study.

869 **Table 2.** PERMANOVA analysis of metabarcoding diversity from OTU clustering

870 (AmpliconNoise) or taxonomic classification (CREST). Significant *P*-values are indicated in

871 bold. Explanatory variable levels tested were: Sample Type - seawater vs. copepod gut content;

872 Experiment - mesocosm vs. fjord; Treatment - Control, NP, NPSi or Raunefjorden; Date - 11, 17,

873 21, 24, 28 or 30 March 2012.

874 **Table S1.** *In silico* coverage of Silva SSU r121 RefNR sequence collection using TestPrime v2.0

875 (Klindworth *et al.* 2012) with primers F-1183mod and R-1443mod and allowing 0, 1, or 2

876 mismatches between primers and database sequence. Numbers in parenthesis indicate percentage

877 of the r121 RefNR sequence collection detected *in silico* for each number of allowed  
878 mismatches.

879 **Table S2.** Summary metrics from analysis of metabarcoding results from mesocosm seawater (*n*  
880 = 24) and copepod samples (*n* = 24) analyzed using AmpliconNoise for OTU clustering or  
881 CREST for taxonomic assignment.

882 **Table S3.** Read distribution per sample and beta diversity estimates from metabarcoding  
883 analysis.

884 **Table S4.** Grazing selection by *Calanus* as identified through metabarcoding analysis of  
885 copepod gut content and co-occurring seawater plankton assemblages. Molecular grazing ratios  
886 (ratio of each OTU/taxon in copepod gut content to their abundance in seawater) were calculated  
887 for (A) individual AmpliconNoise OTUs, (B) AmpliconNoise OTUs grouped by fifth-level (or  
888 highest available) CREST taxonomic assignments, (C) Highest-level CREST taxonomic  
889 assignment ranks and (D) CREST taxonomic assignments grouped by fifth-level (or highest  
890 available) taxonomic ranks. Only finite, non-zero values are shown.

891 **Table S5.** Metabolome profiles. (A) GC-MS peak identifications. Metabolite, peak number; Ion,  
892 ionization energy in *eV*; RT, retention time; Identification, certainty is expressed as the reverse  
893 match factor of each metabolite when searched against the National Institute of Standards and  
894 Technology chemistry database: from 750 to 850 (?), 650 to 750 (??), or 550 to 650 (???). (B)  
895 Peaksum-normalized metabolite concentrations in each mesocosm on all sampling days. Top  
896 column names indicate *<treatment\_sampling date>*. Bottom column names indicate mesocosm  
897 number. (C) Mean metabolite concentration per treatment on each sampling date. Samples  
898 (rownames) are indicated as *<treatment\_sampling date>*.

899 **Table S6.** Biomass estimates for major eukaryotic plankton taxonomic groups in Control, NP  
900 and NPSi mesocosms. Biomass estimates were calculated from microscopy counts according to  
901 previously described methods (Ray *et al.* 2016).

902 **Table S7.** Metabolites with significant correlation to biomass estimates for major microbial  
903 eukaryote plankton taxonomic groups present in the NP and NPSi seawater mesocosms. Only  
904 significant correlations are shown. Adjusted *P*-value indicates significance correction for  
905 multiple treatment comparisons according to (Benjamini & Hochberg 1995). Significance, \*\*\* =  
906 0.001, \*\* = 0.01, \* = 0.05.

907 **Table S8.** Correlation results (Kendall's *tau*) for intracellular metabolites to *Calanus* grazing  
908 ratios calculated from AmpliconNoise OTUs grouped into lower-level taxonomic ranks.  
909 Metabolite - feature identification number; Correlation - Kendall's *tau* correlation test statistic;  
910 Adj. P-value - significance adjusted for multiple treatment comparisons. Question marks  
911 preceding metabolite identifications indicate degree of uncertainty, and represent the reverse  
912 match factor of each metabolite when searched against the National Institute of Standards and  
913 Technology chemistry database: from 750 to 850 (?), 650 to 750 (??), or 550 to 650 (???)

914 **Figure 1.** The 20 most abundant taxa present in *Calanus* gut content (“Copepod”) and mesocosm  
915 seawater (“Water”) samples and their relative abundance in the metabarcoding data (excluding  
916 *Calanus* sequences). (A) AmpliconNoise OTUs, (B) CREST fifth-rank taxonomic assignments.  
917 Taxonomic assignments of OTUs and of sequence reads were performed using the CREST  
918 classifier with SilvaMod reference database.

919 **Figure 2.** Non-metric multidimensional scaling (NMDS) ordination of (A) AmpliconNoise OTU  
920 diversity and (B) CREST taxonomic diversity. *Calanus*-like reads were removed from the

921 sequence data prior to ordination. Dashed ellipses indicate 95% confidence intervals.

922 **Figure 3.** Comparison of methods used for quantification of (A) *Phaeocystis pouchetii* and (B)  
923 *Skeletonema marinoi* in mesocosm water. Methods used were direct taxonomic classification of  
924 pyrosequencing amplicons (Metabarcoding), cell counts per ml (Microscopy), and qPCR  
925 (qPCR). Units shown are reads per ml, cells per ml, or SSU rRNA gene copies per ml,  
926 respectively, for Metabarcoding, Microscopy and qPCR. Note logarithmic y-axis. Microscopy  
927 and qPCR counts are from (Ray *et al.* 2016).

928 **Figure 4.** Numerical distribution of (A) AmpliconNoise OTU or (B) CREST taxonomic  
929 assignment non-zero, finite grazing ratios (copepod gut content / seawater). The distribution of  
930 grazing ratio values for all treatments and Raunefjorden into six arbitrary numerical categories,  
931 indicated in the legend at the right, indicate various degrees of putatively “positive” (>1) or  
932 “negative” (<1) grazing selection by *Calanus*. Full taxonomic content of AmpliconNoise OTU  
933 and CREST grazing ratio tables can be found in the Supporting Information (Table S4).

934 **Figure 5.** Heatmap of metabolites, which are significant (correlation  $R > 0.235$  of CAPdiscrim  
935 of the peaksum normalized data set, Pearson’s correlation test used for determination of  
936 significance level) for the separation (“early”, “middle”, “late”) in both NP and NPSi treatments.  
937 Black=highest concentration, white=lowest concentration, Met=number of metabolite,  
938 Ion=iontrace used for quantification, RT=retention time, tag “?”, “??”, or “???”=reverse match  
939 factor of NIST database between 750 and 850, 650 and 750, or 550 and 650.compound classes:  
940 CH=carbohydrates and derivatives, LP=lipides, TP=terpenoides, ST=steroles.

941 **Figure 6.** Clustering of intracellular metabolite profiles according to mesocosm succession phase  
942 using canonical analysis of principle components using linear discriminant analysis

943 (CAPdiscrim). (A) NP mesocosms, Eigenvalues=0.95 and 0.40, Mis-classification error=25.7 %,   
944 Permutation test:  $P = 0.001$  with 1000 permutations (B) NPSi mesocosms, Eigenvalues=0.96 and   
945 0.40, Mis-classification error=22.5 %, Permutation test:  $P = 0.001$  with 1000 permutations.

946 **Figure S1.** Contribution of copepod and putative symbiont reads to sequence datasets. Height of   
947 bars represents the total number of filtered 454 reads used for OTU clustering or taxonomic   
948 identification. The number of copepod OTUs/reads in each sample are indicated by black bars,   
949 while putative symbiont OTUs/reads are shown as grey bars. Percentage of reads remaining after   
950 exclusion of copepod and putative symbiont OTUs/reads is shown above each column. (A)   
951 AmpliconNoise OTUs, (B) CREST taxonomic classifications.

952 **Figure S2.** Rarefaction analysis of 454 sequence data from *Calanus* gut content (“COP”, A and   
953 C) and seawater (“FIL”, B and D) samples. OTU clustering (A and B) was performed with a   
954 98% similarity cut-off using AmpliconNoise v.1.29 with otherwise default parameters.   
955 Taxonomic assignments were performed using the CREST classifier and the SilvaMod database   
956 as reference taxonomy (C and D). Rarefaction is based on non-normalized read counts and a   
957 sampling frequency of 100 (A and B) or 10 (C and D). Treatment labels: blank, Control   
958 mesocosm; NP, NP mesocosm; NPSi, NPSi mesocosm; fjord, Raunefjorden. Numbers in sample   
959 labels indicate date of sampling in March 2012.

960 **Figure S3.** Microbial eukaryotes that distinguish *Calanus* gut content diversity from mesocosm   
961 seawater diversity when abundance normalized fifth-rank CREST taxonomic assignments were   
962 analyzed using Kruskal-Wallis rank sum tests ( $\alpha = 0.001$ ) with Sample Type (copepod or   
963 seawater) as grouping variable and significance correction for multiple comparisons (q-value)   
964 (Benjamini & Hochberg 2005). Copepod sequences were removed prior to analysis.

965 **Figure S4.** Dynamics of *Calanus* sp. feeding selection over time in three mesocosm treatments  
966 (Control, NP and NPSi) and Raunefjorden as assessed by metabarcoding analysis. The height of  
967 bars (vertically centered at grazing ratio = 1, dashed black line) indicates ratios of the relative  
968 abundance of taxonomic groups in copepod gut content divided by their relative abundance in  
969 co-occurring plankton communities. Grazing ratios  $> 1$  indicate higher abundance of a taxon in  
970 *Calanus* gut content relative to taxon abundance in co-occurring plankton communities, while  
971 grazing ratios  $< 1$  indicate lower abundance in gut content relative to co-occurring plankton  
972 communities. Sampling date (March 2012) is shown on the  $x$ -axis. (A) AmpliconNoise OTUs  
973 grouped by fifth-rank taxonomic classification, (B) Fifth-rank taxonomic classifications.



**Table 1.**

<b>Name</b>	<b>Sequence (5'-3')<sup>1</sup></b>	<b>T<sub>an2</sub></b>	<b>Position<sup>3</sup></b>	<b>Reference</b>
F-1183mod	AATTTGACTCAACRCGGG	60.2	1183-1200	Hadziavdic <i>et al.</i> 2014
R-1443mod	GRGCATCACAGACCTG		1443-1428	Hadziavdic <i>et al.</i> 2014
Cal-SpcC3-block	CTGTTATTGCTCAATCTY GTGCGAC[SpcC3]	70.2	1430-1406	This study
Cal-PNA-block	[NH <sub>2</sub> ]-CTAAGAGTCGCCA GTCCC-[COOH]	70.2	1406-1389	This study

<sup>1</sup> [SpcC3], -C-C-C-OH; [NH<sub>2</sub>], N-terminus of peptide backbone; [COOH], C-terminus of peptide backbone

<sup>2</sup> Annealing temperature used for PCR, in degrees Celsius

<sup>3</sup> Numbering based on reference alignment described in Hadziavdic *et al.* 2014

Table 2.

Pipeline	Distance metric	Explanatory variable	F-value	P-value	R <sup>2</sup>
Amplicon Noise	Jaccard	Sample Type	13.279	<b>0.001</b>	0.210
		Experiment	2.420	<b>0.017</b>	0.038
		Treatment	0.549	0.996	0.017
		Date	4.385	<b>0.001</b>	0.069
	Bray-Curtis	Sample Type	22.592	<b>0.001</b>	0.300
		Experiment	3.454	<b>0.003</b>	0.046
		Treatment	0.443	0.996	0.012
		Date	6.279	<b>0.001</b>	0.084
CREST taxonomic assignment	Jaccard	Sample Type	30.930	<b>0.001</b>	0.362
		Experiment	1.170	0.267	0.014
		Treatment	1.754	0.080	0.041
		Date	7.919	<b>0.001</b>	0.093
	Bray-Curtis	Sample Type	51.043	<b>0.001</b>	0.470
		Experiment	1.151	0.286	0.011
		Treatment	1.986	0.071	0.037
		Date	10.465	<b>0.002</b>	0.096

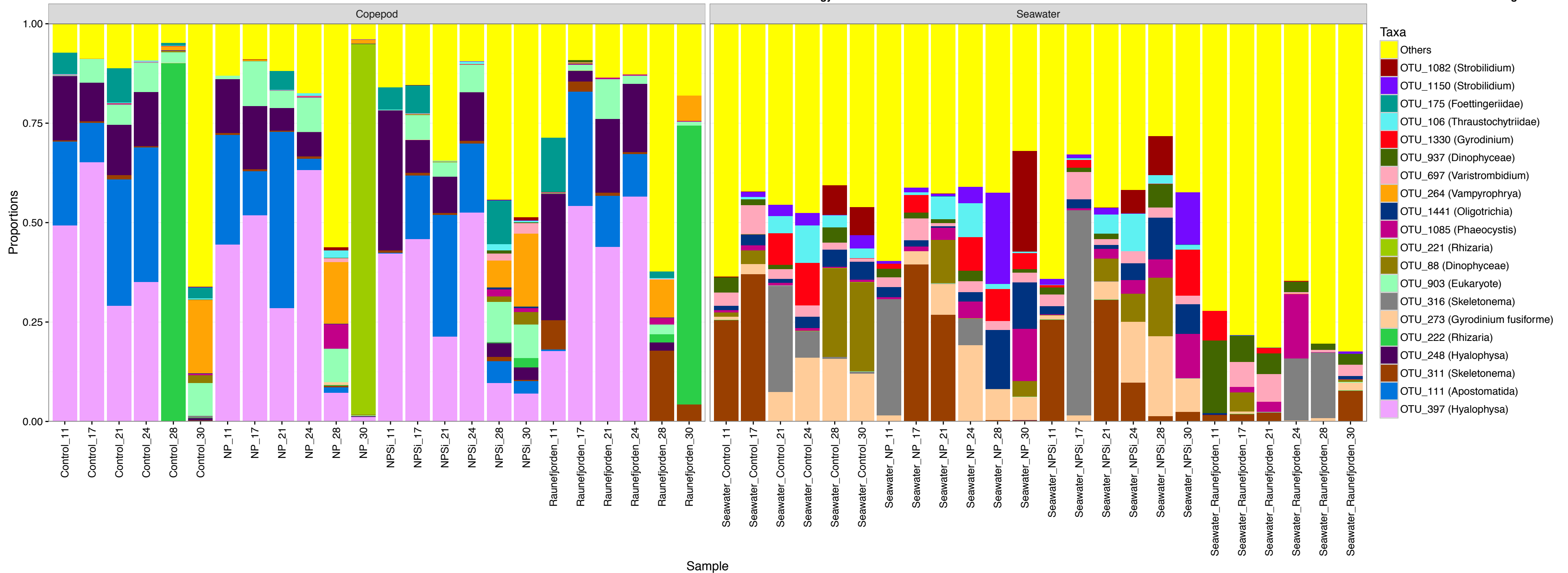
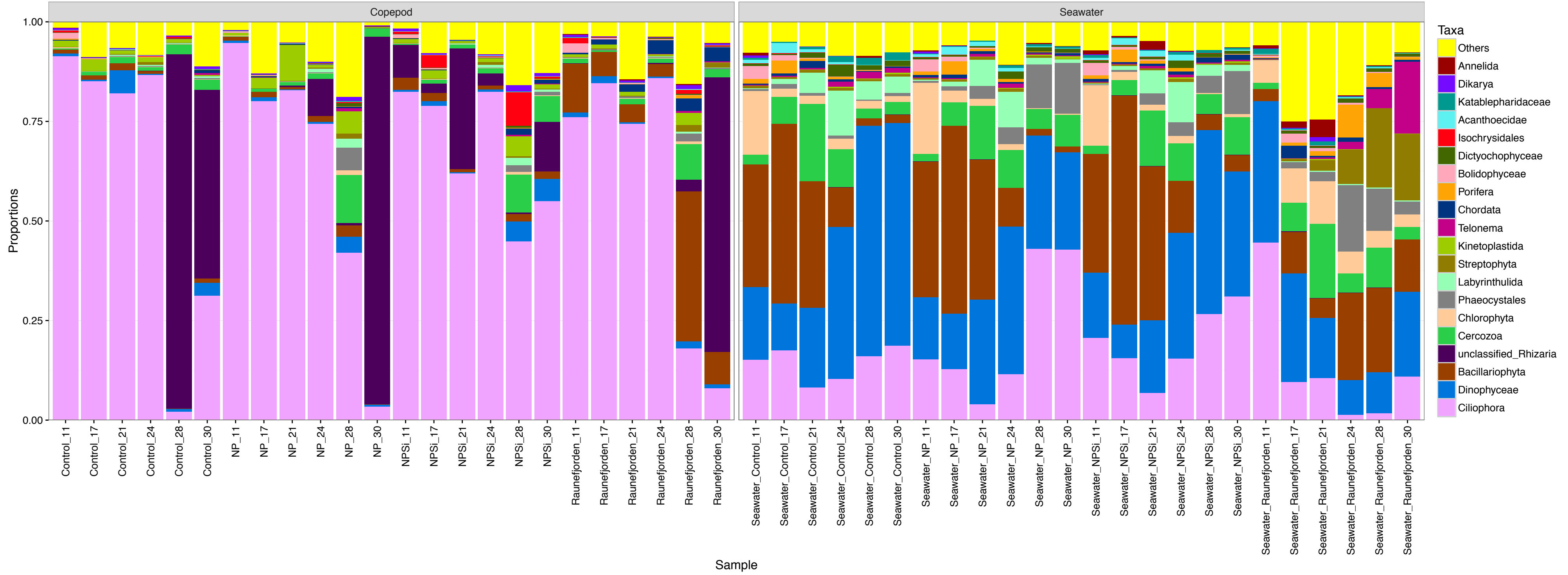
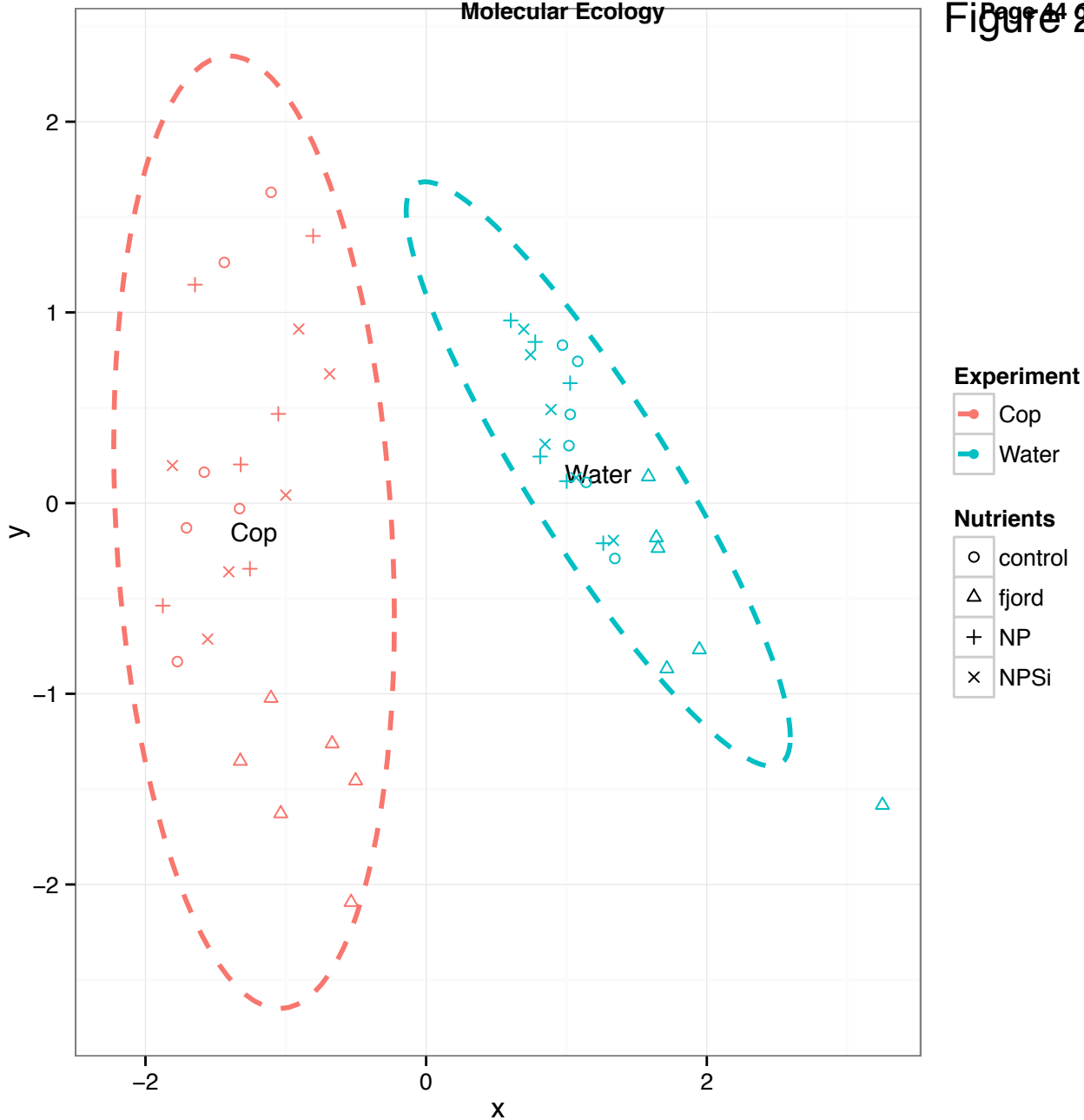
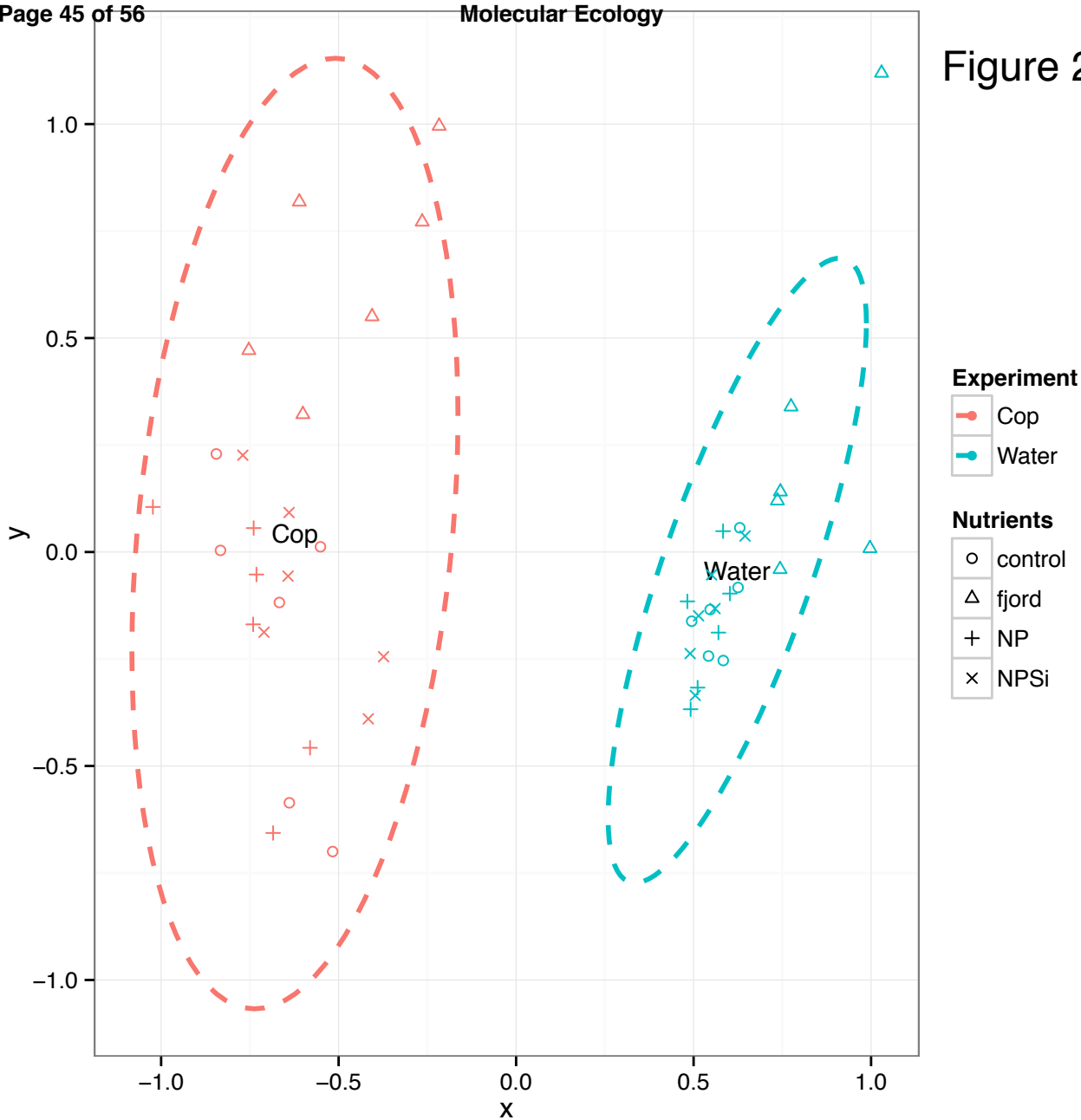


Figure 1B  
Molecular Ecology

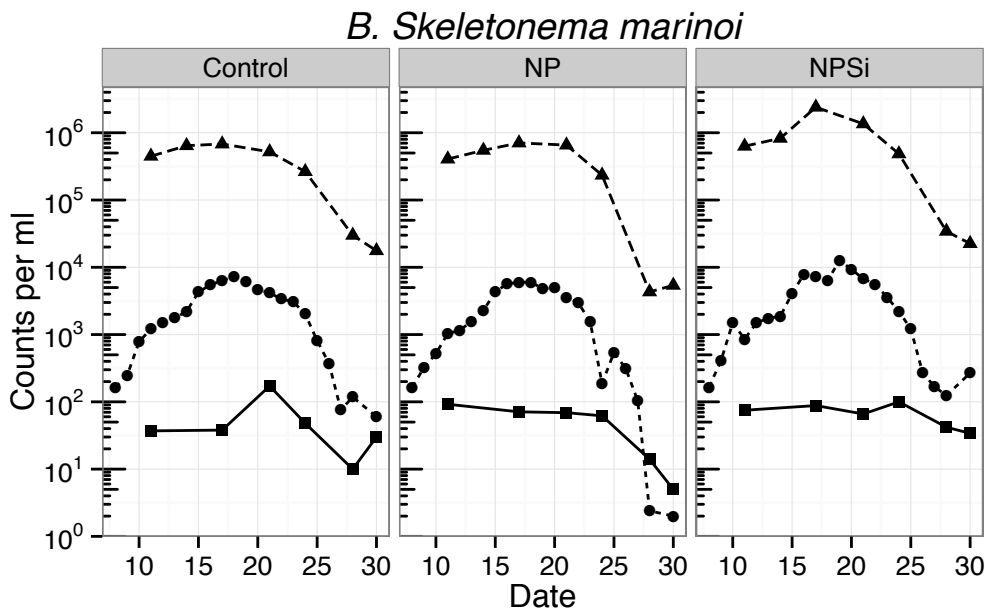
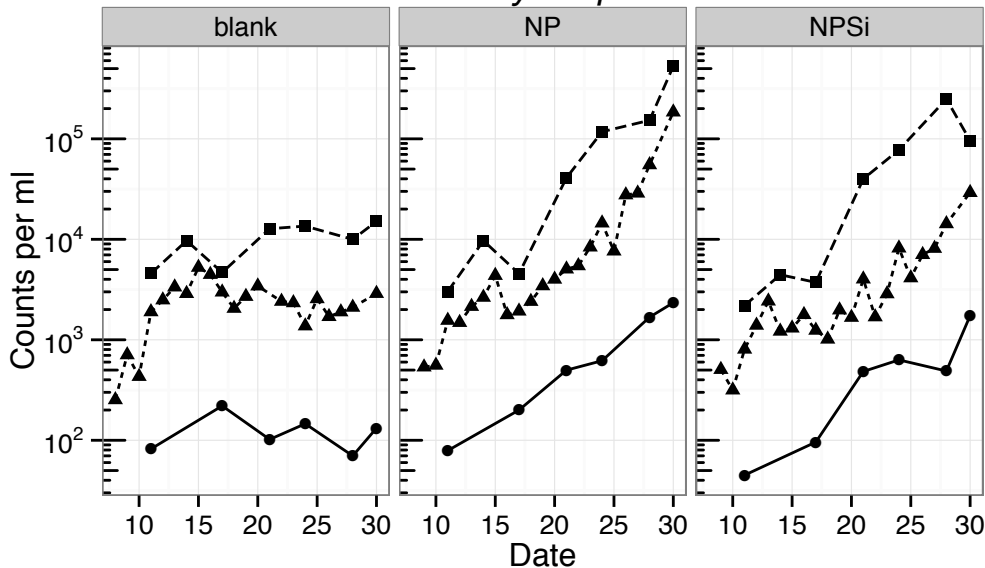






Molecular Ecology  
*A. Phaeocystis pouchetii*

Figure 3



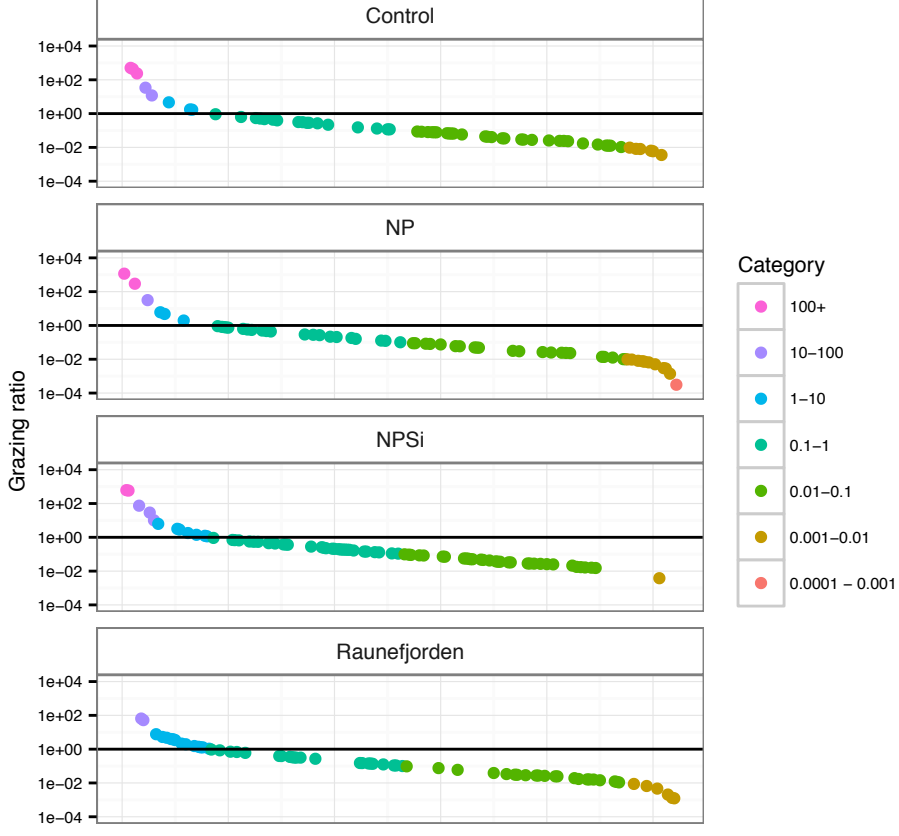


Figure 4B. CREST taxonomy ratios

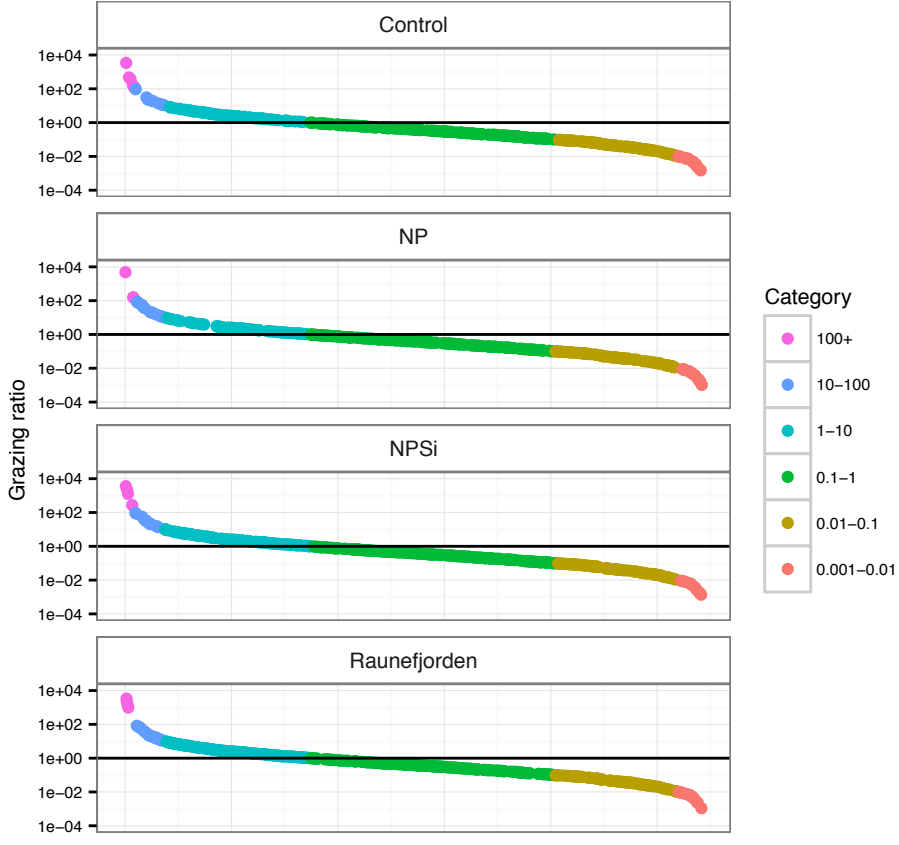






Figure 6

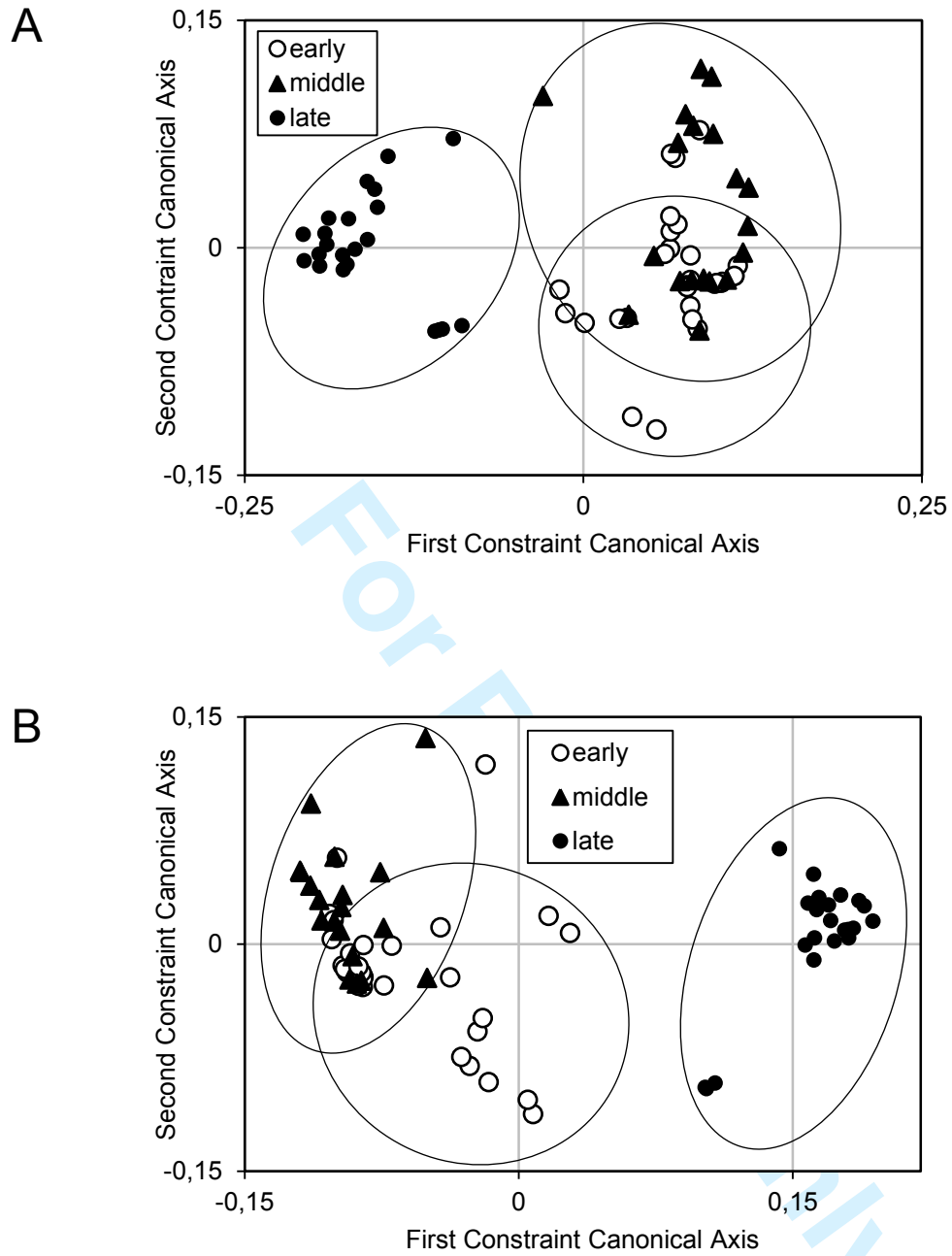


Figure S1A

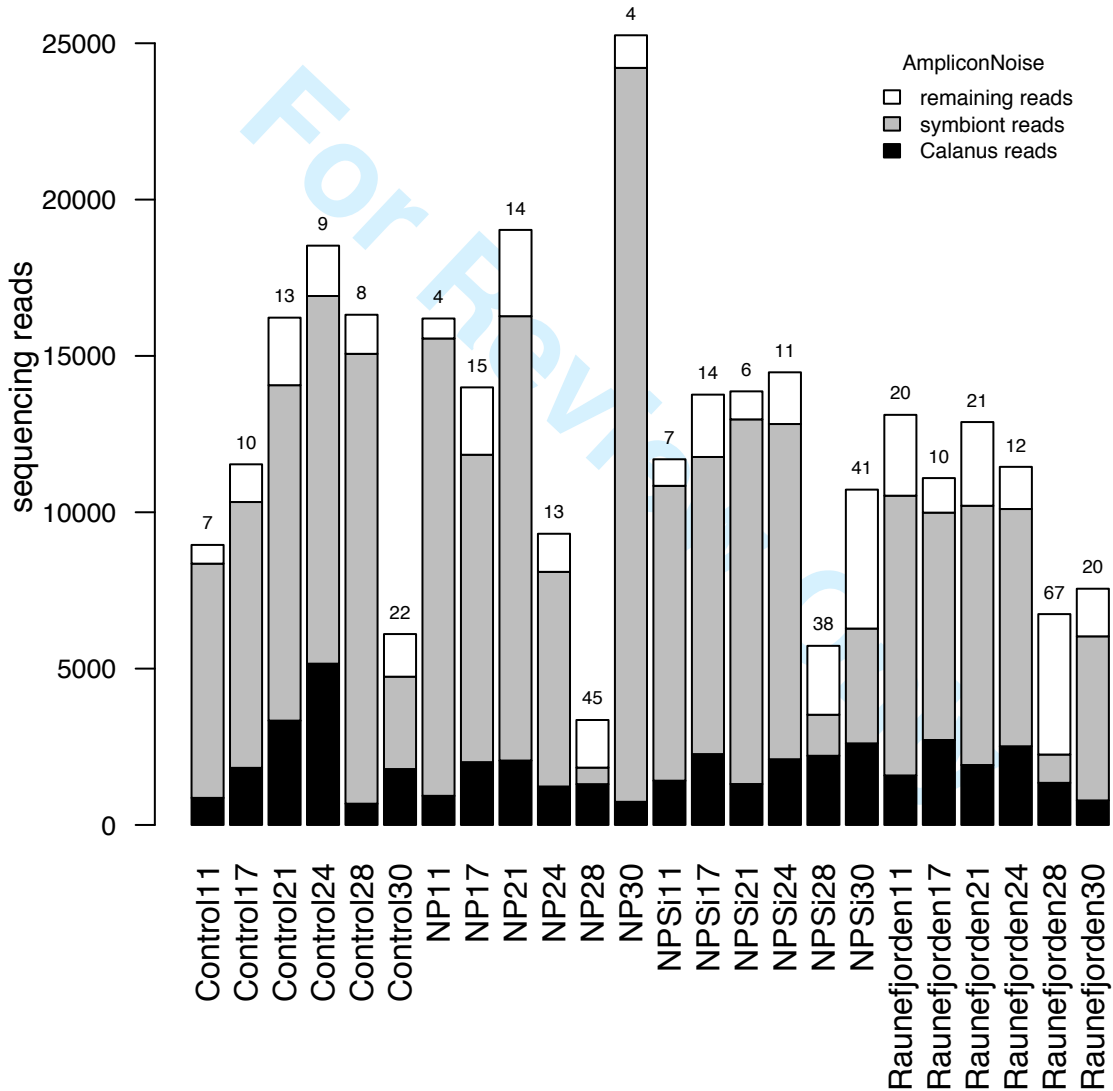


Figure S1B

