The bioreactor microbiome of mass cultivated marine diatoms for industrial carbon capture and utilization

Nerea Johanna Aalto[1,2], Ingeborg Hulda Giæver[1], Gunilla Kristina Eriksen[1], Linn Israelsen[1], Stina Krsmanovic[1], Sebastian Petters[1], Hans C. Bernstein[1,2]*

Affiliations:

1. Norwegian College of Fisheries Sciences, UiT - The Arctic University of Norway, Tromsø, Norway

2. The Arctic Centre for Sustainable Energy, UiT - The Arctic University of Norway, Tromsø, Norway

***Correspondence**: Hans C. Bernstein, UiT-The Arctic University of Norway, BFE-fak., Postboks 6050 Langnes, 9037 Tromsø, NORWAY*; Phone: +47 776 46 114; Email:* Hans.C.Bernstein@uit.no

Running Title: Industrial microbial communities for decarbonization

Keywords: *Algae, Mass cultivation, Phycosphere, Community Succession, Porosira glacialis*

ABSTRACT

Marine microalgae are a promising innovation platform for carbon capture and utilization (CCU) biotechnologies to mitigate industrial greenhouse gas emissions. However, industrial-scale cultivation of algal mono-cultures is challenging and often unscalable. Non-axenic microalgae in large semi-open CCU photobioreactors leads to the co-cultivation of diverse microbial communities. There is limited knowledge about the "bioreactor ecology" involving microalgae interacting with the microbiome and its subsequent impact on stability and productivity. In this study, we describe the semi-continuous industrial mass cultivation of the cold-adapted marine diatom, *Porosira glacialis* UiT201, by investigating the prokaryotic and microeukaryotic communities. Data were collected in two consecutive time series experiments, representing the initiation and operation of an preindustrial scale CCU photobioreactor (300,000 liters). The first experiment experienced a culture "crash" of the focal strain after 39 days, while the second culture remained stable and "healthy" for 60 days. The results highlight that this mass cultivation system represents a unique industrial marine microbial ecosystem. The succession of the prokaryotic community was primarily driven by species replacement, indicating turnover due to selective bioreactor conditions and/or biological interactions. Nonetheless, the bioreactor consistently harbors a recurring and abundant core microbiome, suggesting that the closely associated bacterial community is influenced by microalgae-specific properties and can endure a dynamic and variable environment. The observed culture collapse of *P. glacialis* coincided with changes in the core microbiome structure and different environmental growth conditions compared to the stable and "healthy" experiment. These findings imply that cohabiting microbial taxa within industrial microalgae cultivation likely play a critical role in stabilizing the conversion of industrial CO_2 into marine biomass, and changes in community structure serve as an indicator of process stability.

1 Introduction

Growing levels of atmospheric carbon have an impact on biodiversity, human health, the economy, and the global geopolitical landscape. As the world's climate nears new tipping points [1], it is imperative that we not only reduce emissions but also take new steps toward carbon capture and utilization (CCU). The state-of-the-art for CCU, however, is insufficient and constrained by finances, human capability, and limitations imposed by the requirement to replace outdated industial installations. Innovative biotechnologies that capitalize on photosynthetic carbon fixation, which naturally controls global carbon cycles, will contribute significantly to the solution [2]. Marine microalgae are specifically promising because they have a much higher carbon fixation capacity than plants. Living microalgae make up just 1-2% of global biomass, but they account for almost 50% of the world's primary production [3, 4]. The utility of microalgae for CCU is being demonstrated at an industrial installation in (Arctic) northern Norway, that harnesses cold-adapted marine diatoms to capture CO_2 at the source from an operational ferrosilicon smelter plant [5]. This is a prominent example of how innovative algal biotechnology can be deployed for CCU by taking an advantage of locally isolated diatom strains adapted to the seasonality and seawater resources available to localized large scale industry.

Economically feasible mass cultivation of microalgae for CCU must, however, overcome several barriers that include managing the microbial popluations within large industrial scale photobioreactors. Axenic mass cultivation of microalgae is both technically challenging and cost ineffective. Hence, large photobioreactors are often populated by multikingdom microbial communities and function as their own unique ecosystems [6, 7]. A better understanding of how these communities assemble and behave under stable and unstable – i.e., reactor crash – conditions is needed to scale bioreactors up to volumes capable of sufficient CO_2 fixation. Open airlift column bioreactors are effective for scaled-up industrial operation, land space utilization,

76 and improved CO_2 gas-to-liquid transfer [5]. However, the challenge linked to open-systems is
77 a constant exposure to the environment and thereby capacity of recruiting new microbial taxa
78 to the system [7]. The extent to which a bioreactor recruits microbial taxa from the environment
79 depends on how the inputs (e.g., water, nutrients, innocula) and control volume are exposed to
80 the environment. This presents a unique intersection between bioprocess engineering for CCU
81 and microbial ecology.
82
83 Coinhabiting microbial taxa within industrial photobioreactors are not necessarly harmful.
84 Infact, recent studies have demonstrated that the majority of the interspecific interactions
85 between microalgae and bacteria are mutualistic [8-10]. The beneficial associations are based
86 on exchange of synthesized compounds enhancing growth of both partners [11, 12]. The
87 invasion of other microeukaryotes instead can cause a rapid culture collapse of the focal strain
88 [7]. Some mass cultivation attempts used euryhaline and high pH tolerate focal strain to prevent
89 the contamination of competitive or predative species [13, 14]. However, extreme conditions
90 reduce the number of potential strains and might complicate the operational processes. To
91 uncover the potential and risks associated with the coexisting microbiome, a shift in focus from
92 focal strain-centered investigations to ones that encompass the entire biological component of
93 the production system is necessary [15].
94
95 Major knowledge gaps exist in our understanding of how to control or even measure the
96 complex interactions between microalgae and bacteria. Only few previous studies have
97 measured the microbial diversity of photobioreactors and none (to our knolwege) have
98 investigated the effects on the stability of CCU-specific operations. In order to address some of
99 these knowledge gaps, mass cultivation of the marine centric diatom *Porosira glacialis* UiT201
100 in a preindustrial scale CCU operation was combined with an investigation of the community

101 composition and temporal dynamics of the bioreactor microbiome. Specifically, we addressed
102 three questions: i) What are the most abundant prokaryotic and eukaryotic taxa mass cultivated
103 in cooccurrence with the focal strain? ii) What are the major succession patterns among
104 prokaryotic and microeukaryotic members of the bioreactor microbiome and how do the
105 identified temporal patterns of taxon-specific blooms relate to a variable bioreactor
106 environment? iii) How did the two time series experiments – one ending to an unplanned culture
107 crash after 39 days and one staying stable and healthy for 60 days – differ from each other and
108 what were the similarities with respect to microbiome structure and bioreactor growth
109 conditions?
110

111 **2 Materials and Methods**

112 *2.1 Algae inoculum*

113 The focal strain used to inoculate bioreactors was the centric diatom *Porosira glacialis* (UiT-
114 201). It was originally single-cell isolated from a sediment sample collected in the Barents Sea
115 (N 76° 27.54′, E 033° 03.54′) in 2014 [16]. The monoculture was maintained in a temperature
116 regulated incubator (Termaks KBP 6151) at 8 °C with a scalar irradiance of 65 µmol m^{-2} s^{-1}
117 and diluted weekly with filtered (0.45 µm) and autoclaved seawater containing inorganic
118 marco- and micronutrients from a plant nutrient mix YaraTera Kristalon Purple (Yara, Norway)
119 and natriummetasilikat $5H_2O$ (Permakem, Norway).
120

121 *2.2 Photobioreactor operation*

122 Mass cultivation of *P. glacialis* took place in northern Norway (69.22°N; 18.08°E) in two
123 consecutive time series experiments: 19.08-27.09 and 19.10-17.12 in 2021. Hereafter the first
124 time series recording is referred to as "T1-crash" due to unplanned discontinuation of
125 cultivation caused by contamination and collapse of *P. glacialis* culture and the second as "T2-

126 healthy" in which the focal strain was maintained at targeted abundance and growth rates until
127 the end of time series. The cultivation platform is integrated into an industrial ferrosilicon
128 smelting plant, Finnfjord AS, in collaboration with UiT – The Arctic University of Norway.
129 The platform contains four vertical column airlift photobioreactors with volumes of 6 m^3, 6 m^3,
130 14 m^3, and 300 m^3 [5]. This study deployed the largest 300 m^3 reactor operated with culture
131 volume of 150 m^3 under semicontinuous conditions. Seawater for cultivation medium was
132 obtained from an adjacent fjord at the depth of 25 m and filtered through a multichannel
133 filtration unit with 0.45 µm as a smallest cartridge pore size and through a UV-sterilization unit.
134 Compressed air and flue gas from the factory with CO_2 content of 3-7% were added into the
135 bottom of the reactor using a rotating gas dispersing device [5]. Flue gas was sparged
136 continuously during the cultivation periods except for three dates: 28.10., 08.12. and
137 14.12.2021. Illumination was provided for the reactors by ambient sunlight plus submerged
138 LED lights. Additional artificial illumination was provided by a collection of 200 x 50W LED
139 light units, placed above the culture surface of the bioreactor. Inorganic nutrients (nitrite, nitrate
140 and phosphate) a plant nutrient mix YaraTera Kristalon Purple (Yara, Norway), and
141 $Na_2SiO_3 \cdot 5H2O$ (Permakem, Norway) were added manually to the culture from the start of T1-
142 crash and following 9-days (19.10.-29.10.2021). Thereafter, the nutrients were automatically
143 added using a dosing pump to ensure continuous addition of nutrients.
144

145 *2.3 Inoculation of the 300 m^3 photobioreactor*

146 Prior to each time series experiment, the inoculation volume of *P. glacialis* was increased in
147 two smaller (6 m^3 and 14 m^3) bioreactors. The 300 m^3 bioreactor was pre-filled with ~45 000 L
148 of filtered seawater before the inoculum was introduced. Thereafter, the volume was increased
149 for the next two days until a volume of 150 m^3 was reached as this was the maximum operation
150 capacity for the light delivery system. Semicontinuous operation was maintained by

151 continuously adding filtered seawater for four days a week (Monday – Friday morning) with a
152 mean inflow of 2100 L h^{-1}. The culture surplus was extracted from overflows (150 m^3) and
153 algae was harvested.
154
155 *2.4 Sample collection and growth condition monitoring*
156 Daily measurements were obtained during scheduled work shifts (Monday-Friday) from the
157 300 m^3 bioreactor for: temperature (C°; Endress-Hauser sensor iTHERMA ModuLine TM131
158 w) and dissolved oxygen (DO mg L^{-1}; WTWMulti 360 meter with CellOx 325 sensor, Xylem
159 Anlaytics, Welheim, Germany) from subsurface at ~1 m depth. Liquid samples for cell counts
160 were collected daily and for analysis of inorganic nutrients every second day. Samples were
161 collected twice weekly (Monday and Friday) for genomic DNA isolation and dissolved organic
162 carbon (DOC mg L^{-1}). These samples were also collected from the 300 m^3 bioreactor pre-filled
163 with filtered intake water (hereafter referred as "SW bioreactor"), from inoculum (sample from
164 6 m^3 and 14 m^3 was combined), and from intake water used to semicontinuously dilute the
165 reactor.
166
167 *2.5 Sample processing and analyses*
168 Concentration of inorganic nutrients of nitrite-nitrate (NO_2^-+NO_3^-), phosphate (PO_4^{3-}), and
169 silicate ($Si(OH)_4$) were analyzed immediately after filtration of sample through a GF/C
170 microfiber filter (Whatman). Nutrient specific Spectroquant kit (Merck Millipore, US) was used
171 and the concentration was determined spectrophotometrically using a 96-well plate with a
172 microplate reader (FilterMax F5, Molecular Devices, CA) in which the absorbance was
173 measured at 540 nm. DOC samples were filtered through 0.2 µm cellulose syringe filters
174 (Corning®) into acid washed and pre-combusted 60 mL glass vials and stored in dark at +4 °C
175 for 12 months before acidification with 250 µl of 2N HCl to determine the concentration with

176 TOC-VPH analyzer (Shimadzu, Japan). Modified Utermöhl method was used to obtain an
177 estimation of *P. glacialis* cell density. Cells from two separate transects were counted from a
178 4-well dish with inverted light microscope (Zeiss PrimoVert) after 2 h of settling (Nunc,
179 Thermo Fisher Scientific, US). Depending on the culture density up to 1 L of sample was
180 filtered through each (n = 5 per sampling day) 0.22 µm Sterivex polyethersulfone membrane
181 filter unit (Millipore) to obtain biomass for DNA extraction and following amplicon sequencing
182 of 16S and 18S small subunit ribosomal RNA (SSU rRNA) gene. The filters were immediately
183 stored at -80 °C for 3 months before DNA extraction took place at the UiT – The Arctic
184 University of Norway.
185
186 *2.6 Microbiome – DNA extraction and Illumina amplicon sequencing*
187 DNeasy Power Water Kit (Qiagen, Carlsbad, CA) was used for the purification of genomic
188 DNA. DNA was extracted from the collected biomass following the manufacturer's protocol
189 and stored at -80 °C until sequencing. DNA quality and quantity were assessed by the
190 NanoDrop™ Spectrophotometer and Qubit fluorometer (Invitrogen TM). Bacterial
191 communities were identified following the Earth Microbiome Project (EMP) [17] 16S Standard
192 Illumina library preparation protocol, using the forward-barcoded 515f (5′-3′
193 GTGYCAGCMGCCGCGGTAA) [18] and 806r (5′-3′ GGACTACNVGGGTWTCTAAT)
194 [19] primers targeting the V4 region of 16S SSU rRNA. Targeted gene sequencing of
195 microeukaryotes was determined following the EMP 18S Illumina Amplicon Protocol using
196 the universal Euk_1391f (5′-3′ AATGATACGGCGACCACCGAGATCTACAC) and the
197 barcoded EukBr (5′-3′ TGATCCTTCTGCAGGTTCACCTAC) primers targeting the V9
198 region of the 18S SSU rRNA [20, 21]. The PCR products were confirmed by gel electrophoresis
199 in 1% agarose gels and visualized with GelRed® staining. Sequencing was performed on an

200 Illumina MiSeq instrument (Illumina, San Diego, CA) at the Environmental Sample
201 Preparation and Sequencing Facility at Argonne National Laboratory (Lemont, IL, USA).
202
203 *2.7 Microbiome – Bioinformatic pipeline*
204 Forward and reverse Illumina reads were imported into QIIME2 and thereafter demultiplexed
205 using the complementary barcode file, for each sequencing run, respectively. These steps were
206 executed using the EMP paired-end flag. DADA2 V2021.0 algorithm within QIIME2 was used
207 for denoising and merging the reads [22]. During the denoising steps of the DADA2 package,
208 the reads were pooled, filtered and chimera-checked. Additionally, the replicated reads were
209 discarded. After removing the chimeric sequences and singletons, the reads of the 16S runs and
210 18S runs were merged, respectively. Subsequently, the DADA2 package was used for inferring
211 and filtering amplicon sequence variants (ASVs). Taxonomy was assigned by aligning these
212 ASV sequences against the SILVA V138.1 database, using a 16S and 18S rRNA self-trained
213 classifier [23, 24]. Classifiers were trained using RESCRIPt [25]. Although the 16S and 18S
214 based taxonomic classification within SILVA V138.1 database is a standardized approach, it
215 has limitations regarding classification and taxonomic resolution. It is noted that manual
216 correction of taxonomic annotation was beyond the scope of this study. Nevertheless, order
217 *Enterobacterales* was manually changed to order *Alteromonadales* as the main classified
218 families obtained within *Enterobacterales* in our 16S dataset were *Colwelliaceae*,
219 *Alteromonadaceae* and *Pseudoalteromonadaea* that are known to be members of
220 *Alteromonadales* [26]. The phylogenetic rooted tree, generated with QIIME2 using MAFFT
221 algorithm, was used for downstream diversity analysis.
222

223 *2.8 Microbiome – Diversity analysis*

224 All downstream analyses were completed using R [27]. The microbiome time-series data was
225 analyzed using the "microeco" package (v16.0) [28] unless another package is specifically
226 mentioned. Prior to community analysis, all reads that belonged to mitochondria, chloroplast
227 and kingdom *Eukaryota* or were unassigned (no assigned kingdom) were removed from the
228 16S dataset. Filtering from the 18S dataset removed ASVs assigned to the kingdoms of
229 *Bacteria* and *Archae*a and phyla of *Vertebrata*, *Arthropoda*, *Nematozoa*, *Annelida*, *Cnidaria*,
230 *Mollusca*, *Tunicate*, *Phragmoplastophyt*a and *Basidiomycota* as well as unassigned reads were.
231 Both datasets, 16S and 18S, contained a high number of ASVs with low number of reads. For
232 example, 2074 ASVs found just in 300 m^3 bioreactor comprised only 0.6% of the total reads
233 within 16S dataset. As we are only considering taxa with higher abundance, an abundance-filter
234 was used to remove rare and possibly ambiguous ASVs from the datasets. The filter-threshold
235 included only ASVs which mean relative abundance within the complete 16S and 18S dataset
236 was > 0.01%, respectively. As a result of this the number of ASVs decreased from 7736 to 430
237 and from 1577 to 58 within the 16S and 18S datasets, respectively. Counts of unique ASVs per
238 sample were used to measure richness – i.e., alpha-diversity [29]. The "multicomp" R package
239 was used to perform *post hoc* multiple comparison test with Tukey's HSD [30] on alpha-
240 diversity of different sample types between and within T1-crash and T2-healthy using
241 generalized linear models with Poisson distribution. Potential core prokaryotic taxa associated
242 with the focal strain from inoculum to the end of time series were identified applying an
243 occurrence filter with 75% presence treshold separately on T1-crash and T2-healthy 16S
244 dataset.

245

246 The compositional differences in prokaryotic and microeukaryotic communities – i.e., beta
247 diversity was measured with unweighted Jaccard distance metric to emphasize

248 presence/absence of community members. Three sample days, both from the beginning and the
249 end of 300 m^3 bioreactor time series, were included to infer how the bioreactor microbiome
250 composition in the start and end differs and how they relate to the communities in inoculum,
251 intake water and SW bioreactor (pre-filled 300 m^3 reactor with filtrated seawater). The
252 community differences were visualized in Principal Coordinate Analysis (PCoA) and the
253 pairwise permutational multivariate analysis of variance (PERMANOVA) was performed to
254 test the (dis)similarity associated to each sample type and chosen time series fraction within
255 and between T1-crash and T2-healthy. Beta diversity was also measured via Jaccard distances
256 between T1-crash and T2-healthy bioreactor time series samples, respectively. The Jaccard
257 distances were partitioned to account for the contributions of "nestedness" (species loss) and
258 "turnover" (species gain) towards the measure of total dissimilarity – i.e., $\beta_{Jac} = \beta_{Jne} + \beta_{Jtu}$ [31,
259 32]. This was obtained using the R package "betapart" [33]. The associated changes within total
260 beta diversity, turnover, and nestedness, respectively, by sampling day were tested using
261 "adonis" function (PERMANOVA) with 999 permutations from the R package "vegan" [34].
262

263 *2.9 Microbiome – Clustering around temporal patterns*

264 The K-medoids clustering method was depolyed following the previously described approach
265 applied to microbiome time series data [35, 36]. It was used to investigate shifts in microbiome
266 structure by examining community members that share similar succession patterns across the
267 crash and healthy time series. Prior to clustering the ASV counts were preprocessed through
268 normalization steps including variance stabilizing transformation (R package "DESeq2") [37],
269 detrending (R package "pracma") [38], and scaling via z-transformation to convert values on z-
270 score scale to ensure intercomparability between ASVs. As the K-medoids is a distance-based
271 clustering method pairwise Euclidean distances were calculated before each ASV was
272 partitioned into K clusters according to similarity criterion using the R package "cluster" [39].

273 The clustering efficacy – a balance between number of clusters and data overfitting – was
274 assessed via the Calinski-Harabasz index (Supplementary Figure S1) [40, 41].
275
276 *2.10 Growth conditions of 300 m^3 bioreactor – Statistics*
277 The Mann-Whitney U test (also known as Wilcoxon rank sum test) was performed to reveal
278 differences in bioreactor physical and biogeochemical conditions between T1-crash and T2-
279 healthy. The potential association between the measured bioreactor conditions and their
280 relationship with identified temporal patterns (cluster medoids) was obtained using Spearman's
281 rank correlation analysis. Non-parametric analyses were conducted as half of the included
282 factors indicated non-normal distribution based on Shapiro-Wilk normality test (Supplementary
283 Table S1).
284
285 *2.11 Data repository and reproducible analyses*
286 The unprocessed genetic sequencing data for both SSU rRNA 16S and 18S gene amplicons is
287 available in ENA repository under project PRJEB66331. All data including ASV, taxonomy
288 and relative abundance tables, and bioreactor measurements used for analysis and graphing
289 along with R markdown script are available on the Open Science Framework (osf.io) as part of
290 this project: https://osf.io/k2egy/. All Supplementary Data referred to in this study is available
291 and found on https://osf.io/k2egy/.
292
293 **3 Results**
294 *3.1 The bioreactor´s microbiome*
295 Open mass cultivation of marine diatoms within the industrial CCU setting resulted in a
296 complex microbiome that accompanied growth of the target strain *P. glacialis*. The prokaryotic
297 and microeukaryotic community members were diverse and relatively dynamic during both
298 crashed and healthy cultivation periods. This was observed across two time series experiments

299 with lengths of 39 (crashed) and 60 days (healthy), respectively. The first time series – T1-crash
300 – ended with the loss of dominating abundance of the focal strain, *P. glacialis*, as determined
301 by decrease in growth and microscopy observations revealed severe increases of dead and lysed
302 cells. During the second time series from a mass cultivation run – T2-healthy – *P. glacialis* kept
303 its growth and abundance until the operation was ended and biomass was harvested after 60
304 days. Despite the major differences in conditions and outcomes, the observed richness,
305 especially for the prokaryotic component of the microbiome, across both time series
306 experiments shared major similarities (Figure 1). The total number of prokaryotic ASVs in T1-
307 crash was 339 and in T2-healthy 368 and the corresponding microeukaryotic richness was 47
308 and 56 ASVs, respectively. The prokaryotic richness measured in 300 m^3 bioreactor did not
309 differ between the T1-crash and T2-health as in total 291 and 256 unique ASVs were observed
310 across time (Tukey's HSD results given in Supplementary Table S2). However, only 54-128
311 ASVs in T1-crash and 49-113 in T2-healthy were present in each sample day which
312 corresponds to the diversity measured within each inoculum, but not within the SW bioreactor
313 (pre-filled 300 m^3 reactor with filtrated seawater) or intake water samples which comprised a
314 higher number of ASVs (Figure 1a, b). The microeukaryotic portion of the bioreactor
315 microbiome (measured via 18S rRNA gene amplicons) showed a contrasting result as compared
316 to the prokaryotic community (Figure 1c, d). The SW reactor and intake water samples
317 comprised less or corresponding number of unique ASVs than the inoculum and per day 300
318 m^3 bioreactor samples. Also, the mean number of observed ASVs in the bioreactor across time
319 was significantly different between T1-crash and T2-healthy (Tukey's HSD, $p < 0.001$;
320 Supplementary Table S3). In both time series richness and changes that occurred over time
321 were somewhat similar until day 32. Thereafter the number of observed ASVs stayed constant
322 in T2-healthy whereas in T1-crash a strong decrease was observed towards the end (Figure 1c,
323 d).

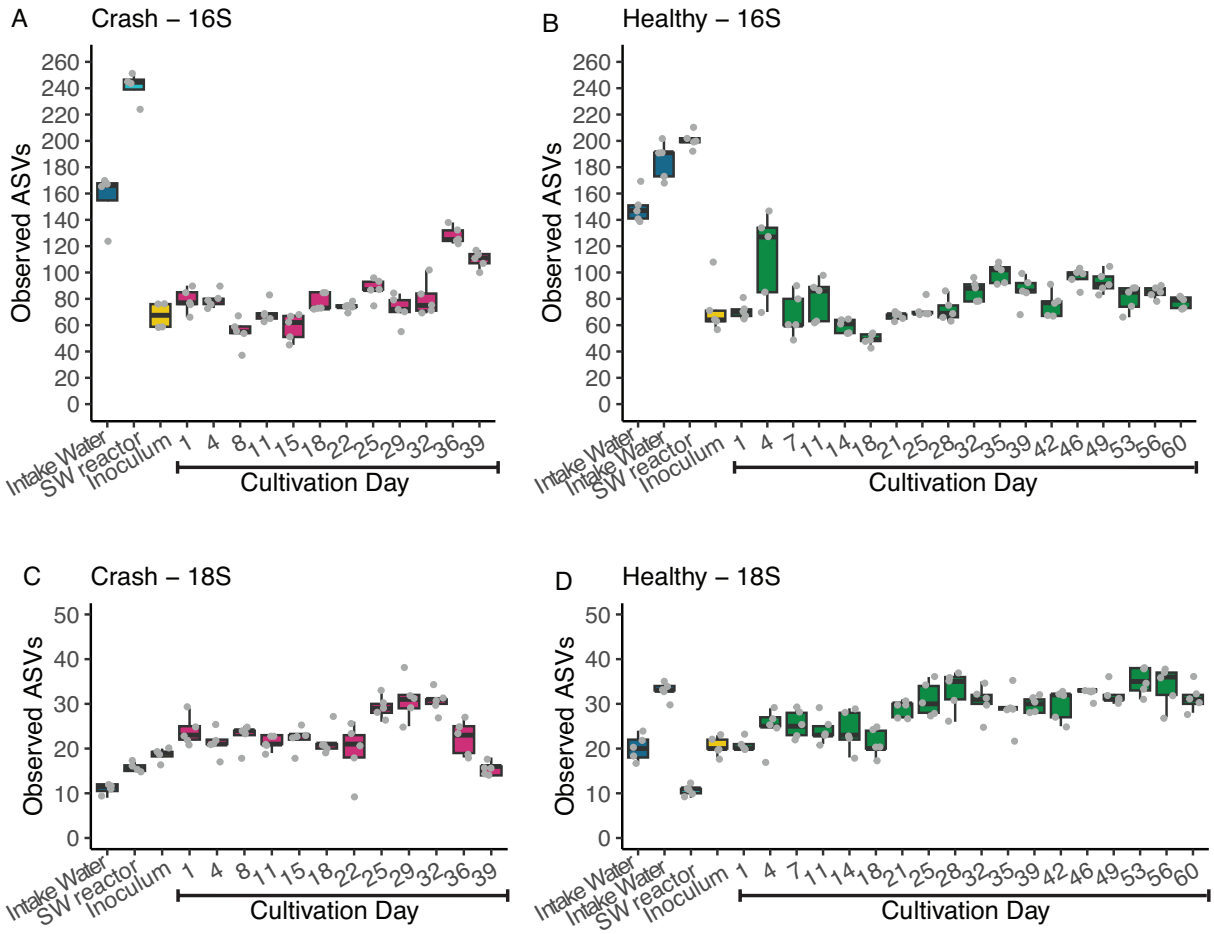


Figure 1. Alpha diversity measured as observed number of unique ASVs in each sample type and across bioreactor time series cultivations. Prokaryotic richness (16S) in time series experiment (A) crash and (B) healthy. Microeukaryotic richnes (18S) in time series experiment (C) crash and (D) healthy. The values of biological replicates (n = 4 or n = 5) are shown over boxplots. Note the different y-axis range between 16S and 18S datasets.

3.2 The reactor's prokaryotic community structure was distinct between crash versus healthy production

When the 16S ASVs were allocated to the taxonomic level order, the prokaryotic community in both time series studies exhibited a dominance of three groups (Figure 2a, b). Specifically, in the T1-crash, the community was dominated by taxa classified within the order *Rhodobacterales* (22%-75% of the relative abundance) until sample day-11, comprising mainly ASVs assigned to genus *Planktotalea* (Figure 2a). In contrast, T2-healthy, *Rhodobacterales* became the most abundant class in the end of the time series, after sample day 46 (28%-54%

340 of relative abundance) and was dominated by ASVs not assigned to a lower taxonomic level
341 (Figure 2b). Taxa indicative to the order *Flavobacteriales* were predominant, notably ASVs
342 assigned to genera NS9 marine group and *Winogradskyella*, for the first half of the T2-healthy
343 time series with exception of sample day-14, while in T1-crash the predominance of the same
344 genera was shorter and limited in the middle of the time series (Figure 2a, b). The third
345 predominant prokaryotic order was different between the time series experiments as taxa
346 classified within order *Chitinophagales*, mainly ASVs assigned to the family *Saprospiraceae*,
347 became the most abundant in the end of T1-crash (25%-64% of relative abundance) (Figure
348 2a). In contrast to that in T2-healthy taxa classified as *Opitutales*, comprised almost solely by
349 genus indicative to *Lentimonas*, had a high relative abundance in most of the sample days and
350 was the most abundant on sample day-14 (49% of the relative abundance) and in the middle of
351 the second half of the time series (Figure 2b).
352
353 *3.3 Mass cultivation has an abundant core microbiome*
354 Nearly the same number of ASVs and taxonomic assignments within prokaryotes were found
355 when ASVs were filtered on occurrence (i.e., presence) at a threshold of 75% among
356 the inoculum and 300 m^3 bioreactor samples. These were 27 and 21 in T1-crash and 25 and 22
357 in T2-healthy, respectively, that were present throughout the mass cultivation runs. The core
358 ASVs accounted for 20-50% of the unique ASVs observed each sample day in bioreactor of
359 both time series runs. Fourteen of the core ASVs, categorized into nine different bacterial
360 orders, were discovered in the bioreactor and inoculum of the two time series periods. Between
361 44% and 91% of the prokaryotic community's daily relative abundance in T1-crash and,
362 correspondingly, between 33% and 91% in T2-healthy, were the taxa that the core ASVs were
363 classified into (core and other ASV(s), henceforth referred to as core taxa). Supplementary

364 Table S4 contains a list of all the core ASVs together with the corresponding taxonomic
365 assignments.
366
367 *3.4 The bioreactor's microeukaryotic community structure was similar until the culture*
368 *collapse in T1-crash*
369 The focal strain, *P. glacialis*, is currently classified within class *Mediophyceae* [42], although
370 this classification was not given by the SILVA v138 database as 19 different ASVs indicative
371 to class *Mediophyceae* were assigned under the same annotation (*Mediophyceae*) down to
372 genus level. Among these ASVs, ASV *032943128192fa7f0a17b2d2991a72f4*, comprised 2-3
373 magnitude higher number of reads compared to the other ASVs. It is therefore considered that
374 most of the relative abundance within class *Mediophyceae* is composed by the focal strain, *P.*
375 *glacialis* in the 18S microeukaryotic communities.
376
377 The community structure of microeukaryotes at the taxonomic level class was similar between
378 T1-crash and T2-healthy samples until a rapid turn from *P. glacialis* dominance to
379 predominance of ASVs classified within the class *Bacillariophyceae* after sample day 29 (59%-
380 89% of relative abundance) in T1-crash (Figure 2c, d). It is notable that members of
381 *Bacillariophyceae* were present from the start in both time series with similar abundances until
382 strong occupation took place in T1-crash whereas in T2-healthy the relative abundance stayed
383 constant until the end of experiment. This contamination was caused by small pennate diatoms,
384 as confirmed by preliminary microscopic detection (Supplementary Figure S2). The most
385 abundant ASV within *Bacillariophyceae* was *f030aa085cd7c8b4bb6538425a95161c*. ASVs
386 assigned to the class *Bacillariophyceae* were also present in the community in the SW
387 bioreactor and intake water samples in the T1-crash, but to a lower degree in the T2-healthy
388 run, where taxa representative of the class *Chrysophyceae* were prevalent, particularly in the

intake water. However, *Chrysophyceae* did not show increasing occupation across time in the T2-healthy microbiome (Figure 2c, d).

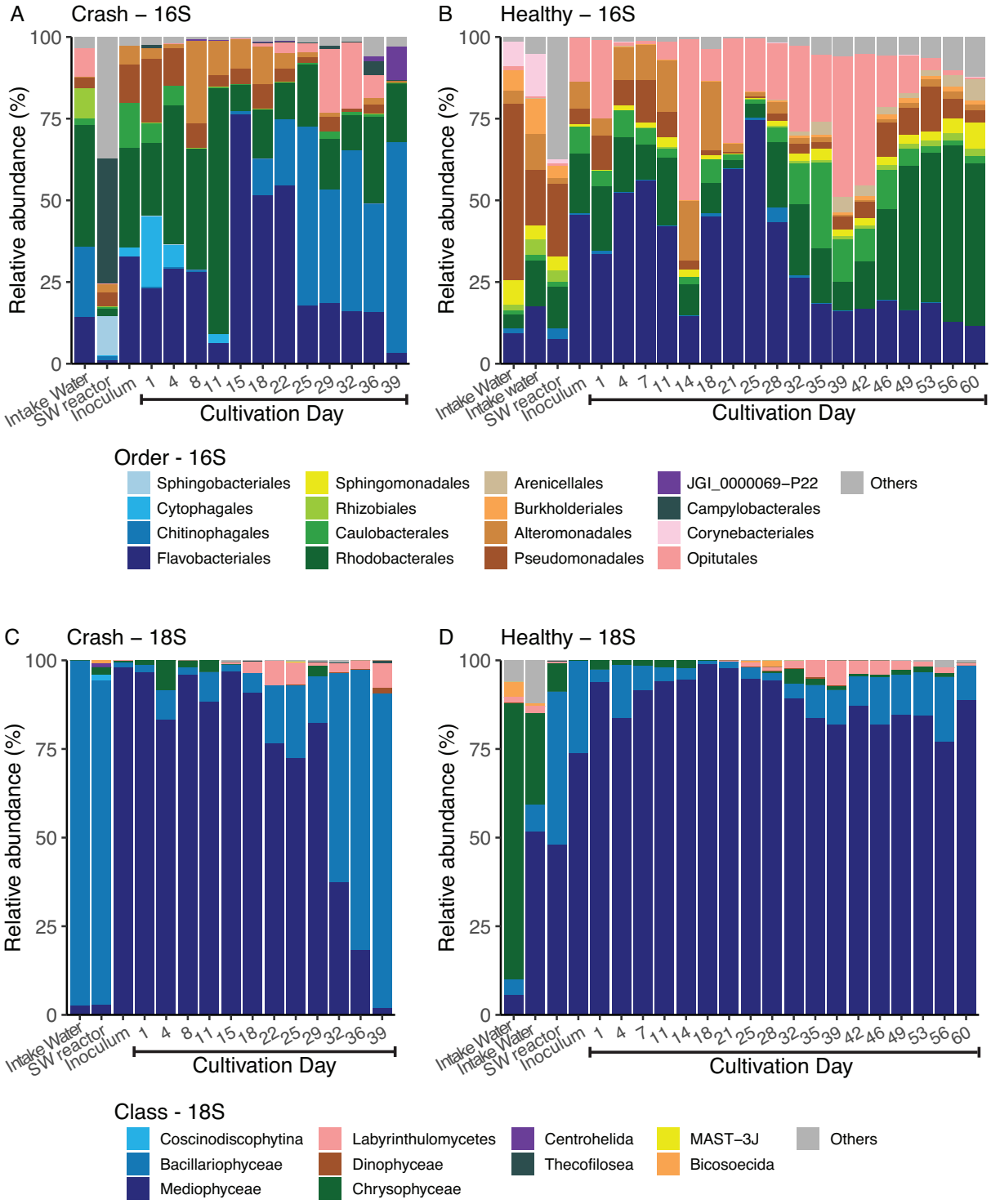


Figure 2. Comparison of community structure in each sample type and across bioreactor time series cultivations. The most common prokaryotic taxa classified within taxonomic level order in (A) crash and (B) healthy time series. The most common microeukaryotic taxa classified within taxonomic level class in (C) crash and (D) healthy time series.

397 *3.5 Beta diversity – microbiome community composition between sample types and across*
398 *time*

399 The community composition of prokaryotes and microeukaryotes changed from inoculation to
400 the end-point in both time series experiments as inferred from pairwise PERMANOVA
401 comparison of the three first (start) and last (end) sample days (T1-crash: 16S, $R^2 = 0.50$ and p
402 < 0.01 & 18S, $R^2 = 0.61$ and $p < 0.01$; T2-healthy: 16S, $R^2 = 0.61$, $p < 0.01$ & 18S, $R^2 = 0.63$,
403 $p < 0.01$). The communities within start- and end-time fractions were significantly dissimilar
404 to those obtained in corresponding inoculum, intake water, and SW bioreactor samples and
405 these sample types were also dissimilar in relation to each other (Supplementary Table S5 and
406 S6). The Jaccard dissimilarity distances within 16S and 18S datasets, respectively, were
407 visualized by ordination with PCoA which showed that 39.7% and 44.0% of the total variation
408 among all included prokaryotic and microeukaryotic samples, respectively, were explained by
409 the two first principal coordinate axis (Figure 3). Closer inspection of the ordination results
410 indicated that in both time series experiments, the community composition in the early time
411 series were more similar with the parallel inoculum community, except for microeukaryotes in
412 T1-crash, than in the end as the communities exhibited strong differentiation across time
413 (Figure 3). The communities within SW bioreactor and intake water were clearly distinct from
414 those in inoculum and 300 m^3 bioreactor samples but also from each other, although the 300
415 m^3 bioreactor was pre-filled with intake water (Figure 3).

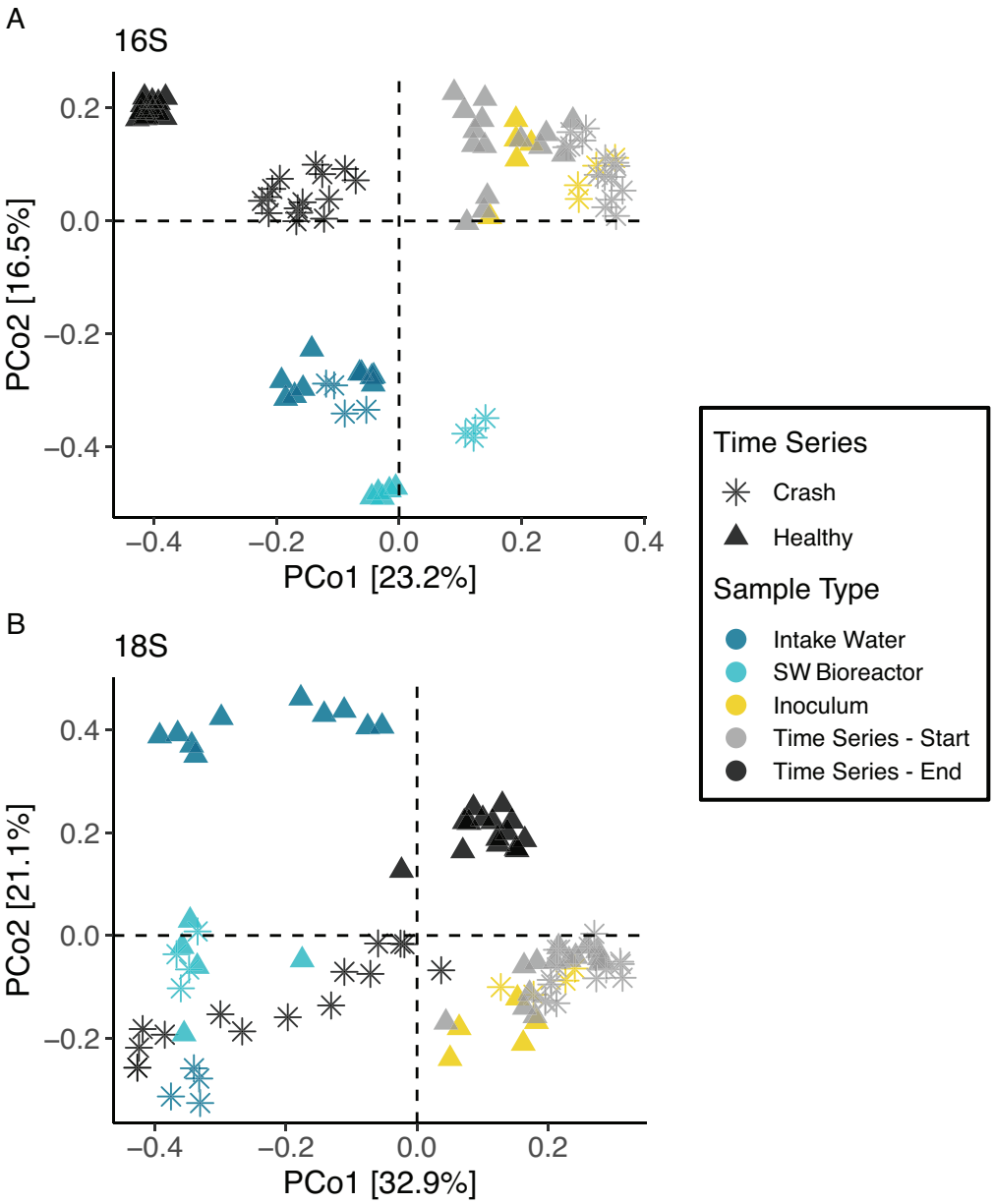


Figure 3. Beta diversity in prokaryotic and microeukaryotic commmunity composition between sample types and bioreactor time series start- and end-time fractions measured with unweighted Jaccard distances and visualized via Principal coordinate analysis. Dissimilarity within (A) prokaryotic and (B) microeukaryotic communities. Time series start-time fraction comprises cultivation days 1, 4 and 8 in crash and 1, 4 and 7 in healthy experiment. The corresponding end-time fractions are 32, 36 and 39 for crash and 53, 56 and 60 for healthy time series.

3.6 Beta diversity – contribution of species-loss and replacement to the bioreactor's community succession patterns

Temporal variation in the community composition of the bacterial fraction of the microbiome was driven to a greater degree by turnover (species replacement) than nestedness (species loss)

428 in both time series (Figure 4). This inference was obtained by analysis that disentangles the
429 additive contributions of these two antithetic processes to the total beta diversity (as measured
430 by the Jaccard distance) between all-time points across T1-crash and T2-healthy, respectively
431 [31]. The variance attributed to the sampling day was much lower from nestedness than from
432 turnover and total Jaccard in both time series as these explained over 84% and 78% of the
433 observed temporal variation associated with the time points, respectively, in both time series
434 runs (Figure 4 and Supplementary Table S7). Strong turnover is also indicated by the observed
435 prokaryotic richness which stayed rather constant across T1-crash and T2-healthy as under
436 nestedness some level of decrease would be expected through temproal species loss (Figure 1).

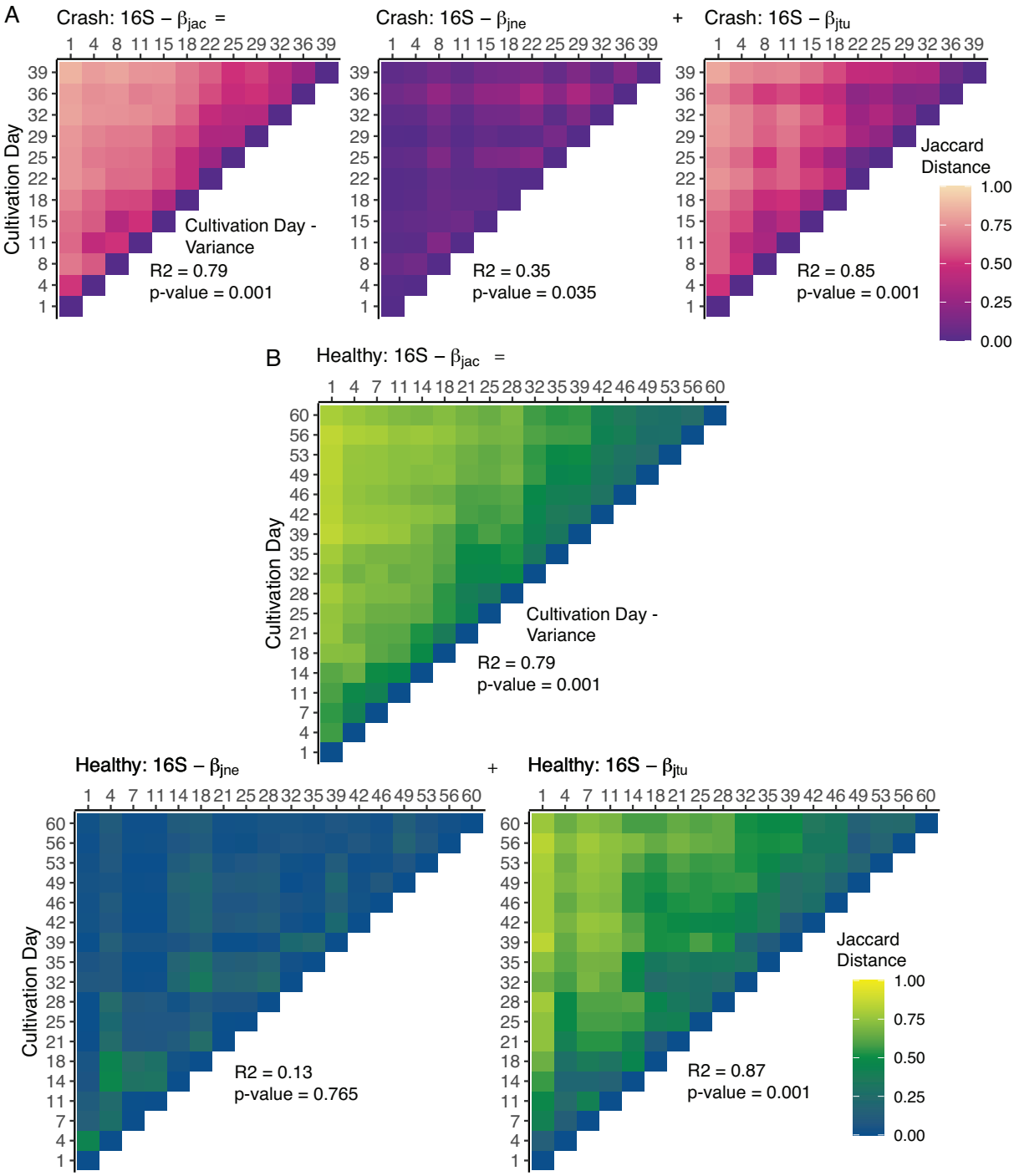
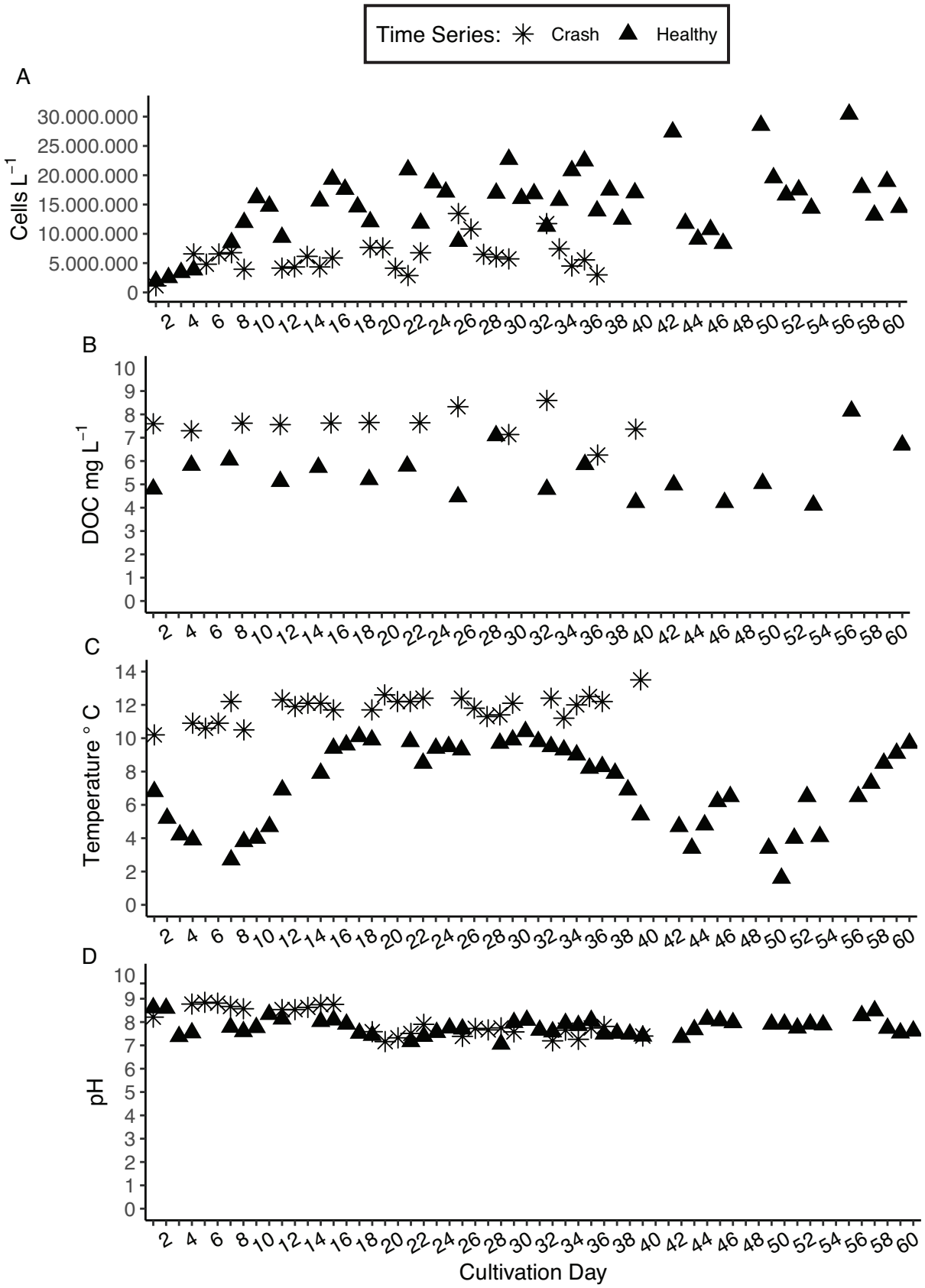


Figure 4. Changes in prokaryotic communities across time. Unweighted Jaccard distance measured between cultivation days in (A) crash and (B) healthy time series experiments. The total Jaccard distance (β_{jac}) is partiotioned by contributions from nestedness (β_{jne}; species loss) and turnover (β_{jtu}; species replacement) i.e., $\beta_{jac} = \beta_{jne} + \beta_{jtu}$. Lighter color represents cultivation days with high dissimilarity. Variance and correspondig significance associated with cultivation day are shown for each processes.

445 *3.7 Temporal succession patterns and cooccurring taxa in distinct bioreactor growth*
446 *environments*

447 The physical and chemical environment and cell density of *P. glacialis* in the large 300 m^3
448 photobioreactor was distinct with respect to measurements obtained during the T1-crash and
449 T2-healthy mass cultivation runs. This was established using the Mann-Whitney U test that
450 showed statistically significant differences between *P. glacialis* cell number and raw
451 fluorescence, concentration of inorganic nutrients, DOC, and temperature measured across both
452 time series experiments – i.e., most of the monitored factors except pH level (Supplementary
453 Table S9). The measured values of these factors were higher in T2-healthy compared to the T1-
454 crash apart from DOC concentration and bioreactor temperature which were lower (Figure 5
455 and Supplementary Figure S3).



456

457 Figure 5. Variation in bioreactor environment across time in crash and healthy cultivation.
458 Daily or twice a week measurements of (A) *P. glacialis* cell density, (B) DOC concentration,
459 (C) bioreactor temperature and (D) pH level. Note that the time series crash was ended after
460 cultivation day-39.

461

462 The potential association of community members to the variable bioreactor environment that
463 prevailed across the time series experiments was evidenced through correlation with temporal
464 patterns. A K-medoid clustering algorithm was used to reduce the temporal complexity within
465 the bioreactor microbiome by identifying the major temporal patterns and revealing the related
466 community members within each time series (T1-crash and T2-healthy) and dataset (16S and
467 18S) driving these distinct patterns. The ASVs within both time series and datasets were divided
468 into three clusters based on assessment of the Calinski-Harabasz index which was used as a
469 quality metric to quantify the ratio for inter-intra-cluster variance for a different number of
470 clusters (Supplementary Figure S1) [41]. The observation of succession patterns among
471 prokaryotes as specified and represented by a medoid taxon, revealed that parts of the
472 community members from both time series datasets, specifically 61 ASVs in T1-crash (cluster
473 2) and 75 ASVs in T2-healthy (cluster 2), exhibited similar temporal dynamics with increase in
474 abundance in the middle of the time series (Figure 6a, b). The medoid of cluster 2 in both time
475 series indicated a significant positive relationship to one of the two factors measuring *P.*
476 *glacialis* cell density. Yet, the association to the other growth condition factors was not
477 consistent (Figure 6c, d). Also, most of the prokaryotic taxa within cluster 2 (T1-crash and T2-
478 healthy) belonged to the same orders *Flavobacteriales*, *Rhodobacterales*, *Opitutales* and
479 *Enterobacterales* (Figure 6e, f). Together these orders comprised the highest mean relative
480 abundance in both time series although cluster 2 contained only half of the number of ASVs
481 compared to the two other clusters in T1-crash and was not the richest cluster in T2-healthy
482 either. The K-medoid clustering revealed that a high number of prokaryotes within cluster 1
483 and 3 first decreased and thereafter increased in their abundance towards the culture crash of *P.*
484 *glacialis* in the end of T1-crash (Figure 6a). Especially in cluster 1 the temporal pattern of
485 medoid showed a strong inverse relationship with *P. glacialis* cell density and DOC

concentration and indicated an intensive proliferation of ASVs classified within *Rhodobacterales* and *Chitinophagales* in the last two sample days as these orders comprised the major taxonomic groups within cluster 1 (Figure 6c, d).

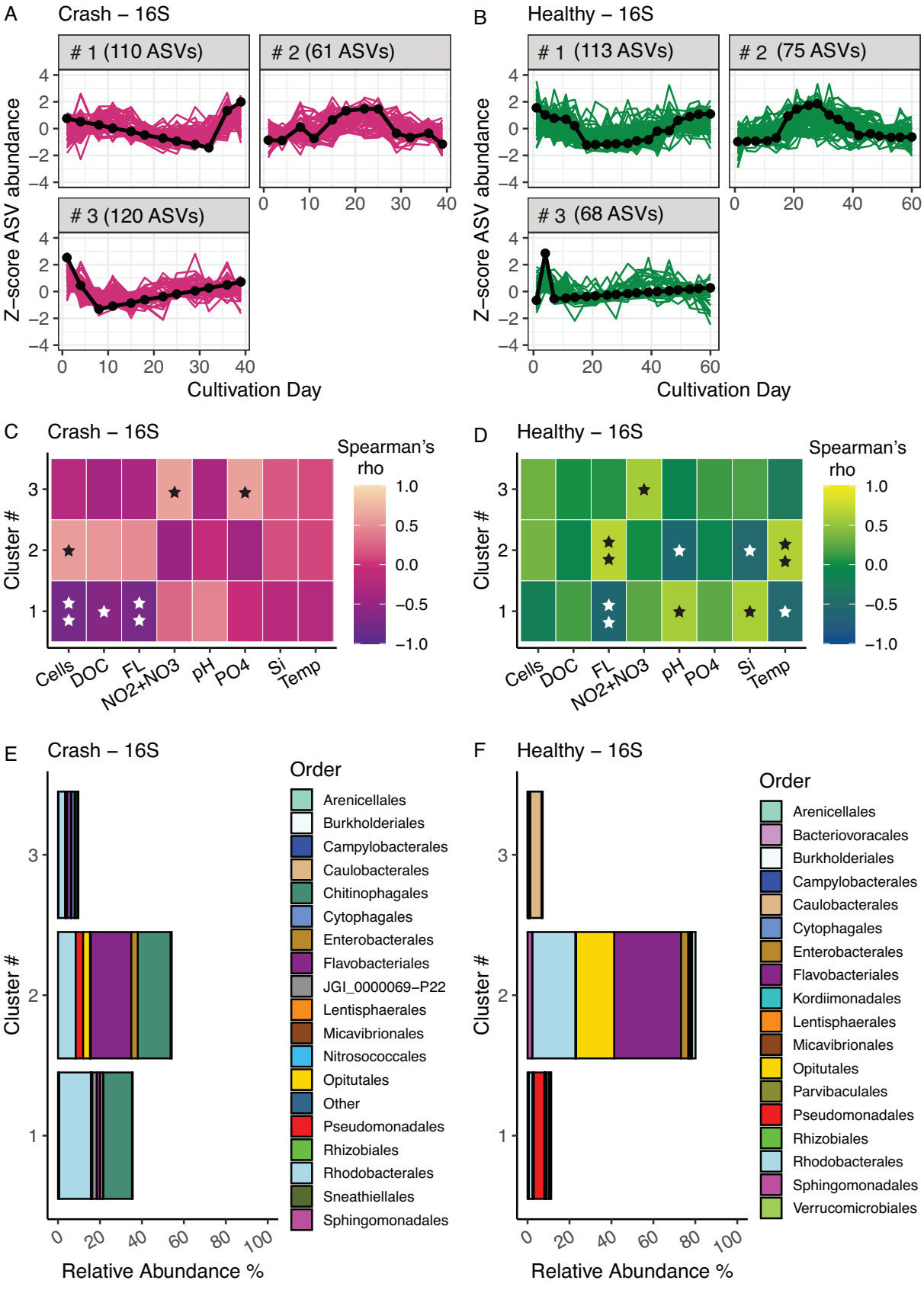


Figure 6. The major temporal succession patterns within prokaryotic communities as assessed via K-medoid clustering analysis with z-transformation and detrending of 16S ASVs. Temproal dynamics of prokaryotic taxa and identified number of ASVs in bioreactor (A) crash and (B)

493 healthy time series experiemnts. The black line indicates cluster specific medoid taxon, a
494 representative for cluster's temporal pattern. ASV specific z-score is a y-axis and a value 0
495 denotes the mean abundance. Heatmap of Spearman's correaltion between temporal dynamics
496 of cluster specific medoid taxon (z-scores) and measured bioreactor factors in time series (C)
497 crash and (D) healthy. The color indicates Spearmna's rank correlation coefficient (rho) and
498 the significance is denoted with ★ (★, $p \leq 0.05$; ★ ★, $p \leq 0.01$; ★ ★ ★, $p \leq 0.001$). Cells, *P.*
499 *glacialis* cell density L^{-1}; DOC, DOC mg L^{-1}; FL, raw fluorescence as a proxy for
500 increase/decrease of *P. glacialis* cell density; NO2-NO3, $NO_2^- + NO_3^-$ µM; pH, pH level; PO4,
501 PO_4^{3-} µM; Si, $Si(OH)_4$ µM, Temp, bioreactor temperature ºC. The most common prokaryotic
502 taxa at the taxonomic level order within each cluster and their mean relative abundance in a
503 whole community in time series (E) crash and (F) healthy.
504

505 The cluster medoids identified within 18S microeukaryotic community revealed a strong
506 fluctuation in relative abundances and weak correlation to bioreactor conditions across time in
507 both time series (Figure 7a-d). ASVs classified within class *Mediophyceae* and
508 *Bacillariophyceae* were the main groups driving distinct temporal patterns in T1-crash and T2-
509 healthy. In both time series experiments the class *Bacillariophyceae* belonged to cluster 3 in
510 which the cluster medoid exhibited strong decrease in relative abundance in the middle of time
511 series (Figure 7e, f).

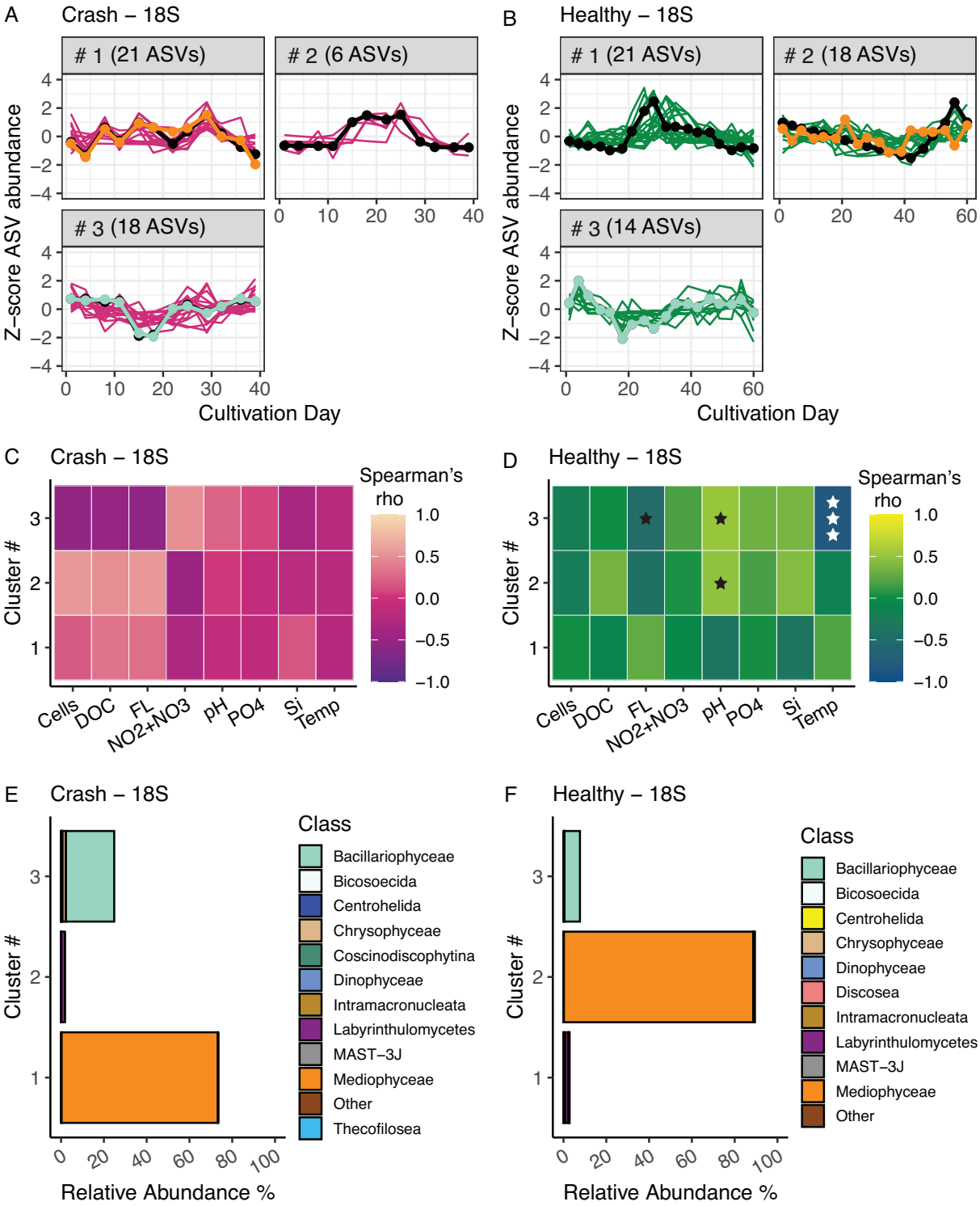


Figure 7. The major temporal succession patterns within microeukaryotic communities as assessed via K-medoid clustering analysis with z-transformation and detrending of 16S ASVs. Temproal dynamics of prokaryotic taxa and identified number of ASVs in bioreactor (A) crash and (B) healthy cultivation experiemnts. The black line indicates cluster specific medoid taxon, a representative for cluster's temporal pattern. ASV specific z-score is a y-axis and a value 0 denotes the mean abundance. The orange line specifies temporal dynamics of the most abundant ASV classified as *Mediophyceae* representing an expected focal strain, and the light green line specifies the most abundant ASVs classified as *Bacillariophyceae*, representing a potential pennate contaminant. Heatmap of Spearman's correaltion between temporal dynamics of

522 cluster specific medoid taxon (z-scores) and measured bioreactor factors in time series (C) crash
523 and (D) healthy. The color indicates Spearmna's rank correlation coefficient (rho) and the
524 significance is denoted with ★ (★, $p \leq 0.05$; ★★, $p \leq 0.01$; ★★★, $p \leq 0.001$). Cells, *P. glacialis*
525 cell density L^{-1}; DOC, DOC mg L^{-1}; FL, raw fluorescence as a proxy for increase/decrease of
526 *P. glacialis* cell density; NO2-NO3, NO_2^- + NO_3^- µM; pH, pH level; PO4, PO_4^{3-} µM; Si,
527 $Si(OH)_4$ µM, Temp, bioreactor temperature ºC. The most common microeukaryotic taxa at the
528 taxonomic level class within each cluster and their mean relative abundance in a whole
529 community in time series (E) crash and (F) healthy.
530

531 **4 Discussion**

532 This study was designed to measure and compare the microbial community succession patterns
533 of an industrial scale CCU photobioreactor during both successful and crashed operations.
534 While contamination control is important, it is not usually feasible to operate large scale, open
535 airlift photobioreactors as axenic cultures of microalgae. Hence, industrial scale production of
536 microalgal biomass in open systems is also mass cultivation of the associated microbial
537 community members. This is not necessarily a hindrance to the goals of algal CCU where pure
538 cultures are not required for success, if the bioreactor environment can effectively select for a
539 stable microbial community that promotes growth and abundant proliferation of a focal strain
540 capable of converting CO_2-enriched flue gas into biomass and bioproducts**.** This study
541 highlights an example of how a unique core microbiome, associated with cold-water marine
542 environments, stabilizes within a photobioreactor used for mass cultivation of diatoms and how
543 these types of microbiomes represent an important signature and possibly a controlling variable
544 associated with successful operation.
545

546 *4.1 CCU bioreactors are unique industrial marine microbial ecosystems*

547 This study characterized a diverse microbial ecosystem within an industrial scale
548 photobioreactor that was distinct from sea water inflow and inoculum (Figure 2 and 3). This is
549 despite the fact that the operation has been designed to promote growth of large marine diatoms
550 – i.e., *P. glacialis.* The microbial community structure is also sensitive to unique temporal
551 succession patterns that serve as signatures of success or crash. New taxa are recruited to the

552 system via inputs from the semicontinuous inflow of sea water from the adjacent costal bay
553 [43] and likely the air-water exchanges within the industrial environment. The majority of
554 taxonomic recruitment was observed within the prokaryotic portion of the microbiome but
555 observations were also made with respect to microeukaryotic taxa. Changes were observed on
556 the time scale of days, in a similar fashion to previous observations from temperate and coastal
557 cold-water microbial ecosystems [36, 44]. It is expected that the diverse community comprises
558 members with different ecological properties, which potentially have a strong impact on the
559 function of the production system in relation to CCU [15]. In addition, most of the major
560 bacteria groups *Flavobacteriales*, *Rhodobacterales* and orders within *Gammaproteobacteria*
561 found in the 300 m^3 bioreactor have frequently been reported to associate with diatom driven
562 phytoplankton blooms in natural ecosystems or in diatom monoculture experiments (Figure 2a,
563 b) [10, 29, 45, 46].
564
565 Bacteria that are commonly associated with marine diatoms are known to utilize different alga-
566 derived organic substrates – *i.e.,* DOC that is released across different bloom or growth stages
567 or generated through temporal substrate degradation [44, 47]. These dynamic processes at least
568 partially underpin the successional blooming patterns of the prokaryotic groups that were
569 identified in this study and other investigations [36, 45]. The orders *Opitutales*
570 (*Verrumicrobiota*) and *Chitinophagales* which also belonged to the most abundant community
571 ASVs from the current study are not considered either uncommon or specific to our production
572 system as they have also been documented in environmental and experimental studies.
573 Members within *Opitutales* are likely part of free-living community thriving in eutrophic
574 environments such as wastewater related microalgae cultivations, with a capacity to degrade
575 complex polysaccharides [48]. Members within the *Saprospiraceae* family (*Chitinophagales*)
576 have been reported to interact with diatoms [49].

577

578 *4.2 Importance of a strong core microbiome in mass cultivation*

579 The relatively long and consecutively repeated temporal coverage on microbial community
580 composition provided us an excellent opportunity to infer reoccurring community succession
581 processes. The bioreactor harbors a prominent core microbiome although the taxonomic
582 composition underwent effective replacement of prokaryotes (strong turnover) between time
583 points (Figure 4). We found that the core ASVs comprised a high portion of the total number
584 of unique ASVs observed on each sample day in both time series. Many prokaryotic taxa that
585 are introduced by intake water and the pre-filled bioreactor at the beginning are replaced after
586 a while as they do not integrate into the core community, as evidenced by the strong
587 contribution of turnover towards total community dissimilarity. Consequently, the bioreactor
588 functions as a selective environment because of the very different physical and biogeochemical
589 conditions that these taxa suddenly encounter in comparison to the coastal environment, such
590 as low pH and elevated levels of light, nutrients, and DOC, or because of the core taxa's
591 competition and antagonistic relationships, which inhibit the growth of the other taxa. It is also
592 expected that *P. glacialis* imposes a selective force that shapes core microbiome as partially
593 evinced by the observed changes in the prokaryotic community when small pennate diatom
594 contaminates supplanted the *P. glacialis* focal strain. The relative strength of diatom-driven
595 selection compared to the physical bioreactor environment remains unclear, although it is
596 understood that the closely associated community is structured by microalgae-specific manner
597 and can persist through environmental variation [11, 50, 51]. The high similarity in the core
598 microbiome between the two cultivation experiments indicates for a certain level of specificity
599 with the production system as it is designed for mass cultivation of *P. glacialis*. It is noted that
600 at least part of the infiltered taxa may reside in the system below detection or were excluded
601 from the analysis along with the applied abundance filter (0.01%). The core taxa (including

602 core and other ASVs classified within the same taxonomic assignment) accounted for an
603 average of 69% and 77% of the relative abundance within bioreactor samples taken from T1-
604 crash and T2-healthy, respectively. Nevertheless, on the majority of sample days, the high
605 abundance was accounted for by a small number of core taxa, whose predominance fluctuated
606 and was only permanent for a short while. Consequently, the temporal change in the prokaryotic
607 community structure was significantly influenced by a small subset of core taxa. The effects of
608 strong temporal variations in community structure and the interactions of the (core) microbiome
609 on the productivity, stability, and/or controllability of the production system are not well
610 understood. Increasing complexity, or the number of interactions within cooccurring
611 relationships, might positively affect stability and, in turn, have an impact on the growth and
612 biomass concentration of the focal strain. [52]. On the other hand, a previous study on the
613 microbiome of marine sponges inferred that a dense core microbiome with few abundant taxa
614 had a stronger positive impact on community level stability than more diverse assemblies of
615 sparsely abundant taxa. [53].
616
617 *4.3 P. glacialis is a promising candidate for mass cultivation*
618 Despite the relatively diversified coinhabiting microbiome, *P. glacialis* grows robustly in the
619 mass cultivation system from inoculation through semicontinuous operation of the 300 m^3
620 photobioreactor. While this work is the first to look at the whole community of the bioreactor,
621 mass cultivation of *P. glacialis* for CCU has been successfully demonstrated in industrial scale
622 trials for several years before [5]. Reactor crash events are rare and conversion of industrial
623 CO_2 to *P. glacialis* biomass is typically successful at scale, regardless of cooccurring
624 microeukaryotes identified in this study. This relatively novel marine diatom is a viable agent
625 for large scale algal CCU and production of high-value biomass [16].
626

627 *4.4 Biological factors associated with reactor crash*

628 Even with a well controlled inoculation and a robust start during mass cultivation, the reactor
629 can occasionally crash. One of the main obstacles when expanding the size of a bioreactor or
630 switching from a closed to an open cultivation system is the high input water volume delivered
631 into the reactor in production systems with (semi)continuous operation, which makes them
632 susceptible to contamination. This is a well-recognized challenge for open cultivation systems
633 even though bioreactor crash events are rarely documented in peer reviewed scientific literature.
634 The research reports on biomass production collapse or culture crash often attribute it to
635 contaminants exhibiting predatory behavior or resource competition [7, 14, 54]. In the present
636 study, the culture collapse in T1-crash was presumably not caused by predation but rather
637 competition or antagonist interactions from another marine diatom or a member from the
638 prokaryotic community, however our study design does not allow for a conclusive mechanistic
639 understanding of the factors that lead to the culture crash. Our findings suggest two taxa that
640 showed temporal blooming patterns in relation to the observed crash. Particularly, toward the
641 end of the T1-crash, the *Saprospiraceae* family (order *Chitinophagales*) and the order
642 *Bacillariophyceae* demonstrated strongly rising dominance within prokaryotic and
643 microeukaryotic communities, respectively, after sample days-22 and 29 (Figure 6a, e and 7a,
644 e). The only highly prevalent taxonomic order that was not a component of the T2-healthy core
645 microbiome was *Chitinophagales*, albeit it was only present in the second half of the T1-crash
646 (Figure 2a). Algicidal activity has been connected on filamentous members of *Saprospiraceae*
647 [55]. In a previous study *Saprospir*a-like bacteria were shown to exhibit a direct attack with the
648 help of its gliding motility on the diatom *Cheatoceros ceratosporum* cell causing complete lysis
649 of diatom cells [49]. If algicidal activity occurred against *P. glacialis* one could ask why the
650 pennate diatoms within *Bacillariophyceae* proliferated and seemed undegraded and why the
651 abundance of *Saprospiraceae* increased after three weeks of successful cultivation? As the

652 presence of algicidal member of *Saprospiraceae* will remain highly tentative without further
653 investigation another potential explanation is that *P. glacialis* in T1-crash was not able to
654 withstand a competition over resources against members of *Bacillariophyceae* whose
655 increasing presence was likely also supported by infiltration of these taxa as the community
656 measured in intake water was dominated by *Bacillariophyceae* (Figure 2c).
657
658 *4.5 Variable environments and blooming patterns within the bioreactor*
659 The large scale mass cultivation system is also unique in relation to the prevailing environment
660 created in the bioreactor and therefore difficult to replicate. This is primarily a result of seasonal
661 variations in the coastal environment and constant production optimization. Although the
662 measured growth conditions varied and were distinct between the time series experiments, the
663 K-medoids clustering did not reveal a clear correlation between the temporal abundance
664 patterns of microeukaryotes and bioreactor factors. Instead, the impact on the prokaryotic
665 community structure might have been stronger (Figure 6). For example, the direct impact of
666 lower and more variable temperature in T2-healthy than in T1-crash on the growth of focal
667 strain was probably minor as *P. glacialis* is shown to obtain a strong growth across the
668 temperature range of 2-12 ºC [16]. Whereas the growth rate of marine bacteria is known to have
669 a pronounced positive relationship within the similar temperature range (White et al., 1991).
670 Moreover, major differences were observed on the cell density of *P. glacialis* and
671 biogeochemical environment. While the cell density of *P. glacialis* was clearly lower across
672 T1-crash than T2-healthy, the measured DOC concentration revealed the opposite (Figure 5).
673 The dataset does not provide a clear insight on why the DOC level was almost constantly 2.1
674 mg L^{-1} higher in T1-crash than in T2-healthy or how it potentially governed the cooccurring
675 microbial community. DOC excretion can reflect physiological state of microalgae as a
676 response to stress or it can be triggered by biological interactions [56-58]. The lower cell density

677 in this case does not directly relate to decreased growth (before culture crash) which in turn
678 could indicate stress, because the dilution rate was higher in T1-crash than T2-healthy.
679 However, the concentration of inorganic nutrients, especially silicate in the first half of the time
680 series, was clearly lower in T1-crash than in T2-healthy and therefore might have impacted the
681 cellular response of *P. glacialis* (Supplementary Figure S3).
682

683 **5 Conclusion**

684 This work shows how basic concepts of marine microbial ecology can be used to gain a better
685 understanding of, and perhaps even control over, a promising industrial biotechnology for CCU.
686 Nevertheless, it is difficult to determine the mechanistic influence of the microbiome on large-
687 scale biomass production in large industrial-scale photobioreactors since they are difficult and
688 expensive to modify and operate with high replication for only experimental purposes. Each
689 cultivation run of the large vertical photobioreactor is distinct in terms of the cooccurring
690 microbiome community dynamics and maintained growth conditions throughout the *P.*
691 *glacialis* production process because of its location within a highly seasonal coastal factory
692 setting. However, the reactor is home to a relatively diverse microbial ecosystem that
693 transforms into a core microbiome via turnover of taxa introduced by the inoculum and
694 seawater intake. Changes to the core microbiome are either affected by or causative crashing
695 populations of the *P. glacialis* strain. Hence, the reactors microbiome is a unique signature of
696 stable CCU operation that likely plays a critical role in stabilizing the conversion of industrial
697 CO_2 into marine biomass.
698

699 **Author contributions**

700 HCB and NJA conceptualized the study. IHG, GKE and LI carried out the cultivation and
701 sampling. SK and SP processed molecular samples and amplicon sequence data. NJA and HCB

conducted data analysis and wrote and the manuscript draft. All authors contributed to the final version.

Aknowledgements

This project was made possible through tight collaboration and technical support from Finnfjord AS, with specific acknolwdegements to Hans Christian Eilertsen, Jo Strømholt, John-Steinar Bergum and Geir-Henning Wintervoll.

Funding

This work was supported by the project AlgScaleUp, funded by the Research Councile of Norway as part of the Green Platform (NFR project number: 328654) and ABSORB – Arctic Carbon Storage from Biomes, which is a strategic funding inniative from UiT – The Arctic University of Norway (https://site.uit.no/absorb/).

Conflict of interest

The authors have no conflicts of interest to declare.

References

1. Masson-Delmotte, V., et al., *IPCC Special Report: Global Warming of 1.5° C.* Summary for Policymakers. https://www.ipcc.ch/sr15/chapter/spm/, 2018.
2. Schweitzer, H., et al., *Innovating carbon-capture biotechnologies through ecosystem-inspired solutions.* One Earth, 2021. **4**(1): p. 49-59.
3. Field, C.B., et al., *Primary production of the biosphere: integrating terrestrial and oceanic components.* science, 1998. **281**(5374): p. 237-240.
4. Falkowski, P.G., *The role of phytoplankton photosynthesis in global biogeochemical cycles.* Photosynthesis research, 1994. **39**: p. 235-258.
5. Eilertsen, H.C., et al., *Mass cultivation of microalgae: I. Experiences with vertical column airlift photobioreactors, diatoms and CO2 sequestration.* Applied Sciences, 2022. **12**(6): p. 3082.
6. Fulbright, S.P., et al., *Bacterial community changes in an industrial algae production system.* Algal research, 2018. **31**: p. 147-156.

7. Davies, F.K., et al., *Microbiota associated with the large-scale outdoor cultivation of the cyanobacterium Synechococcus sp. PCC 7002.* Algal Research, 2021. **58**: p. 102382.
8. Amin, S.A., M.S. Parker, and E.V. Armbrust, *Interactions between Diatoms and Bacteria.* Microbiology and Molecular Biology Reviews, 2012. **76**(3): p. 667-684.
9. Durham, B.P., et al., *Cryptic carbon and sulfur cycling between surface ocean plankton.* Proceedings of the National Academy of Sciences, 2015. **112**(2): p. 453-457.
10. Buchan, A., et al., *Master recyclers: features and functions of bacteria associated with phytoplankton blooms.* Nature Reviews Microbiology, 2014. **12**(10): p. 686-698.
11. Raina, J.-B., et al., *Chemotaxis shapes the microscale organization of the ocean's microbiome.* Nature, 2022. **605**(7908): p. 132-138.
12. Amin, S.A., et al., *Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria.* Nature, 2015. **522**(7554): p. 98-101.
13. Meseck, S.L., *Controlling the growth of a cyanobacterial contaminant, Synechoccus sp., in a culture of Tetraselmis chui (PLY429) by varying pH: Implications for outdoor aquaculture production.* Aquaculture, 2007. **273**(4): p. 566-572.
14. von Alvensleben, N., et al., *Salinity tolerance of Picochlorum atomus and the use of salinity for contamination control by the freshwater cyanobacterium Pseudanabaena limnetica.* PloS one, 2013. **8**(5): p. e63569.
15. Lian, J., et al., *The effect of the algal microbiome on industrial production of microalgae.* Microbial biotechnology, 2018. **11**(5): p. 806-818.
16. Svenning, J.B., et al., *Temperature dependent growth rate, lipid content and fatty acid composition of the marine cold-water diatom Porosira glacialis.* Algal Research, 2019. **37**: p. 11-16.
17. Thompson, L.R., et al., *A communal catalogue reveals Earth's multiscale microbial diversity.* Nature, 2017. **551**(7681): p. 457-463.
18. Parada, A.E., D.M. Needham, and J.A. Fuhrman, *Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples.* Environmental microbiology, 2016. **18**(5): p. 1403-1414.
19. Apprill, A., et al., *Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton.* Aquatic Microbial Ecology, 2015. **75**(2): p. 129-137.
20. Amaral-Zettler, L.A., et al., *A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes.* PloS one, 2009. **4**(7): p. e6372.
21. Stoeck, T., et al., *Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water.* Molecular ecology, 2010. **19**: p. 21-31.
22. Callahan, B.J., et al., *DADA2: high-resolution sample inference from Illumina amplicon data.* Nature methods, 2016. **13**(7): p. 581-583.
23. Quast, C., et al., *The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.* Nucleic acids research, 2012. **41**(D1): p. D590-D596.
24. Yilmaz, P., et al., *The SILVA and "all-species living tree project (LTP)" taxonomic frameworks.* Nucleic acids research, 2014. **42**(D1): p. D643-D648.
25. Robeson, M.S., et al., *RESCRIPt: Reproducible sequence taxonomy reference database management.* PLoS computational biology, 2021. **17**(11): p. e1009581.
26. Parte, A.C., et al., *List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ.* International Journal of Systematic and Evolutionary Microbiology, 2020. **70**(11): p. 5607.

27. R Core Team, *R: A language and environment for statistical computing, R Foundation for Statistical Computing.* 2021.
28. Liu, C., et al., *microeco: an R package for data mining in microbial community ecology.* FEMS Microbiology Ecology, 2021. **97**(2): p. fiaa255.
29. Aalto, N.J., et al., *Diversity and selection of surface marine microbiomes in the Atlantic-influenced Arctic.* Frontiers in Microbiology, 2022.
30. Hothorn, T., F. Bretz, and P. Westfall, *Simultaneous inference in general parametric models.* Biometrical Journal: Journal of Mathematical Methods in Biosciences, 2008. **50**(3): p. 346-363.
31. Baselga, A., *The relationship between species replacement, dissimilarity derived from nestedness, and nestedness.* Global Ecology and Biogeography, 2012. **21**(12): p. 1223-1232.
32. Brislawn, C.J., et al., *Forfeiting the priority effect: turnover defines biofilm community succession.* The ISME journal, 2019. **13**(7): p. 1865-1877.
33. Baselga, A. and C.D.L. Orme, *betapart: an R package for the study of beta diversity.* Methods in ecology and evolution, 2012. **3**(5): p. 808-812.
34. Oksanen, J., et al., *Community ecology package.* R package version, 2013. **2**(0).
35. Coenen, A.R., et al., *A primer for microbiome time-series analysis.* Frontiers in genetics, 2020. **11**: p. 310.
36. Aalto, N.J., et al., *Microbial Community Dynamics during a Harmful Chrysochromulina leadbeateri Bloom in Northern Norway.* Applied and Environmental Microbiology, 2023. **89**(1): p. e01895-22.
37. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome biology, 2014. **15**(12): p. 1-21.
38. Borchers, H.W. and M.H.W. Borchers, *Package 'pracma'.* accessed on, 2022. **4**.
39. Maechler, M., et al., *Package 'cluster'.* Dosegljivo na, 2013.
40. Lord, E., et al., *Using the stability of objects to determine the number of clusters in datasets.* Information Sciences, 2017. **393**: p. 29-46.
41. Caliński, T. and J. Harabasz, *A dendrite method for cluster analysis.* Communications in Statistics-theory and Methods, 1974. **3**(1): p. 1-27.
42. Guiry, M.D. and G.M. Guiry, *AlgaeBase. World-wide electronic publication, National University of Ireland, Galway.* http://www.algaebase.org, 2022.
43. Aalto, N.J., et al., *Drivers of atmosphere-ocean CO2 flux in northern Norwegian fjords.* Frontiers in Marine Science, 2021: p. 841.
44. Teeling, H., et al., *Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom.* Science, 2012. **336**(6081): p. 608-611.
45. Teeling, H., et al., *Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms.* elife, 2016. **5**: p. e11888.
46. Schäfer, H., et al., *Genetic diversity of 'satellite'bacteria present in cultures of marine diatoms.* FEMS microbiology ecology, 2002. **42**(1): p. 25-35.
47. Eigemann, F., et al., *Phytoplankton Producer Species and Transformation of Released Compounds over Time Define Bacterial Communities following Phytoplankton Dissolved Organic Matter Pulses.* Applied and Environmental Microbiology, 2023. **89**(7): p. e00539-23.
48. Orellana, L.H., et al., *Verrucomicrobiota are specialist consumers of sulfated methyl pentoses during diatom blooms.* The ISME Journal, 2022. **16**(3): p. 630-641.
49. Furusawa, G., et al., *Algicidal activity and gliding motility of Saprospira sp. SS98-5.* Canadian journal of microbiology, 2003. **49**(2): p. 92-100.

50. Mönnich, J., et al., *Niche-based assembly of bacterial consortia on the diatom Thalassiosira rotula is stable and reproducible.* The ISME Journal, 2020. **14**(6): p. 1614-1625.
51. Jackrel, S.L., et al., *Host specificity of microbiome assembly and its fitness effects in phytoplankton.* The ISME journal, 2021. **15**(3): p. 774-788.
52. Paddock, M.B., J.D. Fernández-Bayo, and J.S. VanderGheynst, *The effect of the microalgae-bacteria microbiome on wastewater treatment and biomass production.* Applied microbiology and biotechnology, 2020. **104**(2): p. 893-905.
53. Björk, J.R., et al., *The dynamic core microbiome: Structure, dynamics and stability.* BioRxiv, 2017: p. 137885.
54. Wang, Y., et al., *Identification of harmful protozoa in outdoor cultivation of Chlorella and the use of ultrasonication to control contamination.* Algal research, 2018. **31**: p. 298-310.
55. McIlroy, S.J. and P.H. Nielsen, *The family saprospiraceae.* The Prokaryotes: Other Major Lineages of Bacteria and The Archaea.[4th ed.], 2014: p. 863-889.
56. Mühlenbruch, M., et al., *Mini-review: Phytoplankton-derived polysaccharides in the marine environment and their interactions with heterotrophic bacteria.* Environmental microbiology, 2018. **20**(8): p. 2671-2685.
57. Gärdes, A., et al., *Effects of Marinobacter adhaerens HP15 on polymer exudation by Thalassiosira weissflogii at different N: P ratios.* Marine Ecology Progress Series, 2012. **461**: p. 1-14.
58. Hulatt, C.J. and D.N. Thomas, *Dissolved organic matter (DOM) in microalgal photobioreactors: a potential loss in solar energy conversion?* Bioresource technology, 2010. **101**(22): p. 8690-8697.