

# Cyanobacterial taxonomy, metabolism, and cyanophage association in dark and anoxic sediment

Elias Broman (✉ [elias.broman@su.se](mailto:elias.broman@su.se))

Stockholm University <https://orcid.org/0000-0001-9005-5168>

Karin Holmfeldt

Linneuniversitet

Stefano Bonaglia

Stockholms Universitet

Per OJ Hall

Goteborgs Universitet

Francisco JA Nascimento

Stockholms Universitet

---

## Research

**Keywords:** Cyanobacterial taxonomy, metabolism, cyanophage, dark environments, anoxic conditions

**Posted Date:** December 5th, 2019

**DOI:** <https://doi.org/10.21203/rs.2.18208/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Background Cyanobacteria are photosynthetic ancient bacteria ubiquitous in terrestrial and aquatic environments. Even though they carry a photosynthesis apparatus they are known to survive in dark environments. Cyanophages are viruses that infect and lyse cyanobacterial cells, adding bioavailable carbon and nutrients into the environment. Here we present the first study that investigate the metabolic spectrum of cyanobacteria in dark and anoxic environments, as well as their associated cyanophages. We sampled surface sediments during April 2018 located along a water depth gradient of 60–210 m—representing oxic, hypoxic and anoxic conditions—in the largest dead zone in the world. We combined metagenomic and total RNA sequencing to investigate cyanobacterial taxonomy, activity and their associated cyanophages.

Results Cyanobacteria were detected at all four stations (n = 3 per station) along the sampled gradient, including the anoxic sediment. Top active genera in the anoxic sediment included *Anabaena* (19% RNA data), *Synechococcus* (16%), and *Cyanobium* (5%). The mRNA data showed that cyanobacteria were surviving through i) anaerobic carbon metabolism indicated by glycolysis plus fatty acid biosynthesis, and ii) nitrogen (N<sub>2</sub>) fixation (likely by heterocystous *Anabaena*). Interestingly, cyanobacteria were also actively transcribing photosynthesis, phytochromes, and gas vesicle genes. Cyanophages were also detected at all stations, and compared to the oxic sediment had a different beta diversity in the hypoxic-anoxic sediment. Moreover, our results show that these cyanophages infect cyanobacteria affecting the photosystem and phosphate regulation of cyanobacteria.

Conclusions Cyanobacteria were found to transcribe genes for photosynthesis, phytochromes, and gas vesicles and this suggests that cyanobacteria are trying to ascend to the surface waters. The difference in cyanophage beta diversity between oxic and hypoxic-anoxic sediment suggests that anaerobic cyanobacteria select for specific cyanophages. Cyanobacteria are known to fuel oxygen depleted benthic ecosystems with phosphorous (so called internal loading), and our study suggests that cyanophage-controlled lysis of cyanobacteria likely provides a source of nitrogen. Photosynthetic cyanobacteria are commonly thought to have been essential in the great oxygenation event on Earth ca. 2.4 billion years ago. Our results suggest that active cyanobacteria might also provide nutrients (via N<sub>2</sub> fixation and viral lysis) in dark and anoxic environments.

## Background

Cyanobacteria are ubiquitous free-living bacteria in terrestrial and aquatic environments [1]. Many cyanobacterial species are capable of photosynthesis and N<sub>2</sub> fixation [1], and are therefore essential in ecosystems as primary producers channeling organic carbon and nitrogen to higher trophic levels [2–6]. During high nutrient load and warm periods cyanobacterial growth and reproduction increases and forms large algal blooms in aquatic systems [7]. Because a large portion of cyanobacteria are capable of N<sub>2</sub> fixation they are able to flourish even when inorganic nitrogen is depleted [8]. In addition to shading the euphotic zone and affecting food web dynamics [9], such cyanobacteria also release toxins into the water

[9, 10]. A portion of these blooms eventually decay and filamentous cyanobacteria sink down below the euphotic zone and are deposited on the seabed [7, 11, 12]. In the sediment surface the algal bloom material provides organic matter for benthic consumers and adds to the internal loading of phosphorus in the system [2, 6, 8, 13]. Eventually microbial degradation of this material depletes oxygen in the bottom water and so called hypoxic ( $< 62.5 \mu\text{mol L}^{-1} \text{O}_2$ ) or anoxic “dead zones” are formed [11, 14]. Interestingly, cyanobacteria are known to survive in darkness [15, 16] even under anoxic conditions [17] and it is therefore likely that sinking cyanobacteria survive in dead zone sediments. During dark and anoxic conditions cyanobacteria taxa such as *Anabaena*, *Nostoc*, and *Oscillatoria* are able to ferment the carbohydrates glucose, fructose, and sucrose, which yields production of  $\text{CO}_2$ ,  $\text{H}_2$ , lactate, ethanol, formate, and acetate [18]. Some fermentative cyanobacteria have also been shown to produce sulfide by using elemental sulfur as a terminal electron acceptor [19].

In addition to fermentation, certain cyanobacteria can survive in the sediment as dormant spore-like cells called akinetes under harsh conditions [20]. Upon favorable conditions (e.g. light and oxygen) akinetes turn into vegetative cells and via newly formed gas vesicles ascend to the upper water column [20]. Such mechanism has been seen to be important for the recruitment of cyanobacterial blooms in shallow lake and estuarine systems [21, 22]. However, in deeper water bodies cyanobacterial gas vesicles collapse at too high pressure (up to 0.98 MPa) and this has been estimated to be at 90–100 m water depth in the Baltic Sea, northern Europe [23]. This is significant, because i) the Baltic Sea is a severely eutrophied aquatic system with large cyanobacterial blooms during summer [24], and ii) the largest dead zone in the world is located in the Baltic Sea [25], which has anoxic waters below a permanent oxycline situated at 65–125 m depth [26]. This would indicate that sinking cyanobacteria (either as akinetes or filaments) in the Baltic Sea are trapped in the dark anoxic bottom water unable to ascend, and would have to adapt to these extreme conditions. Based on prokaryotic metagenome and metatranscriptome studies, cyanobacteria have been detected in Baltic Sea coastal anoxic sediment [27, 28], as well as in the deepest point of the Baltic Sea, the Landsort Deep which reaches a maximum depth of 466 m [29]. However, there is a paucity of studies focusing on investigating cyanobacterial taxonomy and metabolism, as well as their association with cyanophages in dead zone sediments.

Cyanophages are viruses that are known to infect cyanobacteria and after infection lytic cyanophages eventually lyse the host cell [30–32], and thus contribute to the release of carbon and nutrients into the environment [33, 34]. Furthermore, cyanophages might increase the decay of pelagic cyanobacteria and increase the amount sinking down towards the seabed through increased particle formation caused by viral lysis [35]. In addition to being attached to sinking cyanobacteria and possibly transported, viruses can also be active and persist in sediment for thousands of years [36, 37]. For example, Cai L, Jørgensen BB, Suttle CA, He M, Cragg BA, Jiao N and Zhang R [37] showed that cyanophages persist for centuries in the anoxic sediment at the deepest point of the Baltic Sea (down to 37 m below the seafloor), indicating that cyanophages already exist in the sediment before sinking cyanobacteria reaches the seabed. Cyanobacterial blooms in the Baltic Sea have been traced back to ca. 7000 years (based on pigment analysis in deep sediments) [38], and it is likely that the ancient cyanophages are remnants from these

sinking blooms. Considering that cyanophages directly affect the abundance of cyanobacteria and the release of bioavailable carbon and nutrients, cyanophages might have a key role in the food web of anoxic sediments. However, the significance of cyanophages and their association with cyanobacteria in benthic ecosystems has yet to be studied.

Here, we aimed to elucidate the role of cyanobacteria and cyanophages in both oxic and dark hypoxic-anoxic sediments. We did this by sampling the top 0–2 cm sediment surface from four stations along a water depth gradient (60–210 m water depth, from oxic to anoxic sediment; Table 1) in the largest dead zone in the world (Baltic Sea). Sediment was sequenced using the latest DNA and RNA sequencing technology (Illumina NovaSeq) to investigate cyanobacterial activity and associated cyanophages in the different sediments. We hypothesized that 1) bloom-forming cyanobacteria are alive and active in the anoxic sediment and survive by metabolic processes such as fermentation and N<sub>2</sub> fixation, and 2) the diversity of cyanophages are related to the community structure of cyanobacteria, and sediment with hypoxic-anoxic conditions might therefore contain different cyanophages compared to oxic ones.

## Results

### *Cyanobacterial taxonomy and diversity*

Oxygen concentrations measured in the 0–0.5 mm sediment surface decreased along the sampled gradient, with station A (330  $\mu\text{mol L}^{-1}$ ), D (8.8  $\mu\text{mol L}^{-1}$ ), E (1.8  $\mu\text{mol L}^{-1}$ ), and F (0  $\mu\text{mol L}^{-1}$ ) (Table 1). Cyanobacteria were detected in all metagenome and RNA-seq sediment samples at all stations ( $n = 3$  per station for both DNA and RNA data). The number of classified reads belonging to cyanobacteria was on average 3.3 % (range 2.5–4.5 %) in the DNA data, and showed a decreasing trend along the water depth (i.e. oxic station A to anoxic station F) in the RNA data from 7.9 to 4.2 % (Additional File 1: Fig. S1). Both the DNA and RNA datasets were dominated by cyanobacterial populations belonging to the oxygenic photosynthetic class *Oxyphotobacteria*, with especially two taxonomic orders: 1) *Synechococcales* (up to 65% in the RNA data) and 2) *Nostocales* (up to 44% in the RNA data; Fig. 1). In the DNA data there was an increasing pattern of reads classified to *Nostocales*, from 13 to 33 %, along the water depth gradient (i.e. station A to F; Fig. 1). Looking closer at the active cyanobacteria in the hypoxic-anoxic sediment (station D, E and F), most of the classified reads were attributed to 1) the genera *Cyanobium* and *Synechococcus* both belonging to the order *Synechococcales* (both in the DNA and RNA data), and 2) the genus *Anabaena* belonging to the order *Nostocales* (15–20%; mainly in the RNA data; Fig. 2). *Cyanobium* was attributed more DNA and RNA-seq classified reads at the oxic station A compared to the hypoxic-anoxic stations D, E and F, with DNA 51% and RNA 14% compared to DNA 14% and RNA 5%, respectively. *Synechococcus* showed the opposite pattern with more numbers of reads at station D, E, and F (DNA 19% and RNA 16% compared to station A with DNA 12% and RNA 11%; Fig. 2). These results show that cyanobacteria, even bloom forming oxygenic photosynthetic taxa, are present and active in the dead zone sediment. A full list of read counts and sequence classifications against cyanobacterial genomes in the RefSeq database is available in Additional File 2: Data S1.

Cyanobacterial alpha diversity (Shannon's H) in the sediment was significantly higher in the oxic station A (Shannon's H  $4.17 \pm 0.02$ , one standard deviation shown) compared to the hypoxic-anoxic stations D, E, and F (Shannon's H 3.92–3.94; One-Way ANOVA,  $F_{(3,8)} = 129.76$ ,  $P = 0.0000004$ ; and Tukey post hoc tests  $P < 0.00001$ ; Fig. 3).

### *Metabolism of cyanobacteria in the dead zone sediments*

The RNA transcripts data indicate that the metabolism of the cyanobacteria was higher at the oxic station A compared to the hypoxic-anoxic stations D, E, and F (Fig. 4A). RNA transcripts for the KEGG MODULES fatty acids biosynthesis and glycolysis were prevalent at all stations, indicating that anaerobic carbon metabolism such as fermentation was likely occurring. The reductive citrate cycle (CO<sub>2</sub> fixation), cystine amino acid synthesis, and the photosystems were also indicated to be active processes in the sediment at all stations (Fig. 4A). In contrast, nitrate assimilation was prominent only at station A (Fig. 4A). Because not all proteins could be classified into KEGG MODULES, we also decided to inspect the UniProtKB/Swiss-Prot classified protein data manually. A full list of proteins affiliated with Cyanobacteria is available in Additional File 2: Data S2, KEGG KOs and Modules in Additional File 2: Data S3, while a list of proteins with a significant difference between stations are available in Additional File 2: Data S4.

RNA polymerase was present at all stations in the cyanobacterial RNA data and indicated that active transcription of genes was occurring (Additional File 2: Data S2). The oxic station A had more RNA transcripts classified to ribosomal proteins (30S and 50S) when compared to the other stations in hypoxic and anoxic conditions (edgeR analysis,  $FDR < 0.05$ ). Furthermore, the number of significantly different proteins between station A and D, E, and F increased with water depth, with: A vs D 58 proteins, A vs E 69 proteins, and A vs F 90 proteins (Additional File 2: Data S4). This shows that with decreasing oxygen concentrations, and further away the oxic station A, the difference in cyanobacterial metabolism increases.

Cyanobacterial RNA transcripts for N<sub>2</sub> fixation (genes *nifDEHKNU*) were present at all stations (Fig. 4B). The number of *nifU* RNA transcripts was higher at station A compared to D and F ( $FDR < 0.05$ ) and *NifH* RNA transcripts was higher in station E compared to A ( $FDR < 0.05$ ). Genes related to assimilatory nitrate reduction and uptake (nitrate reductase *narB* and nitrate transporters *nrtABCDP*) had together more RNA transcripts at the oxic station A than in the hypoxic anoxic stations (20–30 GeTMM at station A compared to  $< 1$  GeTMM at the other stations; Fig. 4B). Of these, RNA transcripts attributed to genes *narB* and *nrtP* were significantly higher at station A ( $FDR < 0.05$ ) than in station D, E, and F (Additional File 2: Data S2).

RNA transcripts attributed to genes affiliated with the photosystems decreased with water depth and O<sub>2</sub> concentrations (photosystem I: *ycf3*, *ycf4*, *ycf12*, *btpA*, *psaABCDEFGHIJKLM*, and photosystem II: *psbA1A2A3BDHIJKLMOTWXYZ*, *psb27*, *psb28*; Fig. 4B). There was a variation between stations with e.g. *psbD* having more RNA transcripts at station A compared to the others ( $FDR < 0.05$ ), and *psaD* showing

the opposite pattern with lower RNA transcripts in station A ( $FDR < 0.05$ ). Interestingly, *psbA2* had the highest amount of RNA transcripts at the deepest anoxic station F ( $FDR < 0.05$ ). In addition to the photosystem, cyanobacterial phytochromes used to detect light were attributed RNA transcripts at all stations with no significant difference (*phyAB*; Fig. 4B).

The hypoxic-anoxic stations D, E, and F had a higher amount of RNA transcripts attributed to gas vesicle proteins (range 52–96 GeTMM) when compared to the oxic station A ( $< 3$  GeTMM, genes *gvpAJN*; Fig. 4B and Table 2), with RNA transcripts for *gvpA* being dominant of the *gvp* proteins and significantly higher at stations D, E and F ( $FDR < 0.05$ ; Additional File 2: Data S2).

In addition to these metabolic processes, compared to the oxic station A the deepest anoxic station F had more RNA transcripts attributed to proteins used in e.g. iron deficiency (Iron uptake protein A2), stress related protein Chaperone protein ClpB, Pyruvate kinase proteins used in glycolysis (PK1 and PK2), and ribosome hibernation to stop protein synthesis (dark-response and light-repressed hibernation promoting factors) (all  $FDR < 0.05$ ; Table 2).

### *Cyanophage community composition and diversity*

A total of 90 partial cyanophage genomes (i.e. contigs) were assembled from the co-assembly metagenome data with an average genome size of 5076 bp (range 2000–47,157 bp). The network analysis to identify cyanophages based on protein content (Additional File 2: Data S5) showed that three clusters of cyanophages were formed: x similar to *Siphoviridae* cyanophages, y similar to *Podoviridae* cyanophages, and z similar to *Myoviridae* cyanophages (Fig. 5). See Additional File 1: Fig. S2 for the whole vConTACT network analysis including all reference viruses. Of these 90 identified cyanophages 77 had a mapped read coverage of at least 75% and were therefore considered to be present in the sediment at the stations (Fig. 6 and Additional File 2: Data S6 for a full list of cyanophages). Cyanophage alpha diversity was  $2.37 \pm 0.11$  (station A, Shannon's H),  $3.07 \pm 0.11$  (D),  $3.51 \pm 0.43$  (E), and  $3.11 \pm 0.39$  (F). The alpha diversity was significantly higher at station E when compared to station A ( $P = 0.007$ , One-Way ANOVA with Tukey post hoc test, whole model  $F_{(3,8)} = 7.39$ ,  $P = 0.011$ ; Fig. S3). The community composition of the cyanophages was different in oxic sediment at Station A when compared to the hypoxic-anoxic sediment at stations D, E, and F (Bray-Curtis beta diversity, PERMANOVA (9999 permutations),  $F_{(3,8)} = 5.71$ ,  $P = 0.0006$ ; Additional File 1: Fig. S4). These differences in cyanophage alpha and beta diversity is also visible in Fig. 6, and interestingly most cyanophages with a high relative abundance in the oxic sediment at station A belonged to cluster x, while cyanophages in the hypoxic-anoxic sediment mainly belonged to cluster y and z (Fig. 6).

### *Cyanobacteria and associated cyanophages in the dead zone sediments*

Cyanophages in the sediment were found to largely cluster with stations D and E when comparing with the beta diversity of the cyanobacteria (RNA data; Fig. 7A). The cyanophages distributed in the sediments also showed an association with the cyanobacteria proteins. Here, cyanophages x16, x27, x57, z7, 79, z81 occurring at station A (Fig. 6) clustered towards the beta diversity of the UniProtKB/Swiss-Prot classified cyanobacteria proteins (mRNA data) in the sediment of station A (Fig. 7B). Majority of the cyanophages clustered with the cyanobacteria proteins in the hypoxic-anoxic stations (right side in Fig. 7B). BVSTEP analysis showed that the combination of the three cyanophages y2, x16, and z58 best explained the beta diversity of the cyanobacteria proteins ( $\rho = 0.80$ ). That the cyanophages had an effect on the cyanobacterial community in the hypoxic-anoxic sediment was also indicated by several cyanophages (z5, z10, z12, z13, z17, z19, z45, z46, y48, z58, z74, z76, z78, z82) correlating positively with the number of mapped RNA transcripts attributed to cyanobacterial protein CRISPR-associated endoribonuclease Cas2 3 (gene *cas2-3*;  $\rho = 0.60-0.86$ ,  $n = 12$  for each cyanophage, all  $P < 0.05$ ; Fig. 7C). Finally, the protein classification of the cyanophage genomes revealed that they were carrying a large number of gene copies related to hypothetical proteins, but also e.g. glycosyl transferase related genes, heat-shock protein lbpA, p-starvation inducible protein (gene *phoH*), and photosystem II D1 (gene *psbA*) (Additional File 2: Data S5).

## Discussion

Here we show that dead zone sediments contain active cyanobacteria as indicated by the RNA data. This was in accordance to previous studies demonstrating that cyanobacteria are able to survive in dark and anoxic environments either through spore-like akinetes [20] or fermentative metabolism [18]. In the Baltic Sea a large portion of cyanobacteria in the sediment derives from algal blooms in the surface waters [7, 11, 12]. This was also indicated in our dataset with the filamentous genus *Anabaena* being present. Interestingly, the relative proportion of classified reads attributed to *Anabaena* was higher in the RNA-seq compared to the DNA data, and this might be explained by that 1) *Anabaena* are known to produce akinetes [20], and 2) certain species of *Anabaena* are known to be heterotrophic and survive in darkness even under anaerobic conditions [18]. The presence of *Anabaena* in the hypoxic-anoxic sediment is also explained by that organic matter degradation is slower in oxygen deficit sediments [39, 40], and settled pelagic cyanobacterial material might therefore be preserved for a considerable time. However, mRNA transcript data directly affiliated with cyanobacteria suggest that these organisms are alive and metabolically active in dead zone sediments. *Anabaena* might therefore have survived by fermentative metabolism (endo- or exogenous [18]) and akinetes resting stages (with a low metabolic signal that was present in the RNA data). For example, the species *Anabaena cylindrica* has been observed to produce  $H_2$  by utilizing endogenous substrates [41]. Compared to *Anabaena* the other top genera detected in the hypoxic-anoxic sediment were *Cyanobium* and *Synechococcus*, and they are not known to form akinetes [42]. In addition, *Cyanobium* and *Synechococcus* have not been shown to be capable of fermentation of exogenous substrates, and might therefore survive by either fermentation of endogenous carbohydrates [18] or other means such as carbon fixation. Further studies such as controlled laboratory experiments are needed to fully elucidate their survival mechanism in dark and anoxic sediment.

The mRNA data from cyanobacteria in the hypoxic-anoxic sediment also shows that N<sub>2</sub> fixation is an important metabolic strategy for survival. The sampled stations have scarce or no nitrate and nitrite [43], and this explains why assimilatory nitrate reduction and uptake was only prevalent in the oxic sediment at station A. In the hypoxic-anoxic sediments, N<sub>2</sub> fixation would therefore be a viable strategy to convert N<sub>2</sub> to biologically available nitrogenous compounds [2]. Our data indicates that this is likely conducted partly by *Anabaena* in the sediment surface as N<sub>2</sub> fixation also occurred at the oxic station A. *Anabaena* is known to contain heterocysts that makes it possible to perform N<sub>2</sub> fixation even in the oxic sediment [1]. Furthermore, other non-heterocystous cyanobacteria carrying the nitrogenase enzyme could also fix N<sub>2</sub> [44] in the dead zone sediment as the environment is anoxic.

Interestingly, mRNA transcript classified to genes coding for photosynthesis and phytochrome proteins were detected at all stations. Accumulation of mRNA transcripts attributed to the photosystem in darkness has previously been observed for the strain *Synechocystis* sp. PCC 6803 [45], and an increase in photosystem II units for *Anabaena variabilis* ATCC 2941 [15]. Smart LB and McIntosh L [45] also showed in their laboratory experiment that the absence of glucose caused a lower accumulation of transcripts and larger degradation of mRNA. This can explain why RNA transcripts attributed to photosynthesis genes decreased along the sampled oxygen gradient in the sampled region (i.e. from station A to F). We also detected a decrease in reads classified to cyanobacteria along the oxygen gradient, which might further explain the decrease in photosynthesis RNA transcripts. Furthermore, dark-response ribosome hibernation that lowers potential protein synthesis [46] was higher in the hypoxic-anoxic sediment and is therefore likely a survival strategy in these environments. It is therefore surprising that a large portion of RNA transcripts were still attributed to photosynthesis. Our data indicates that heterotrophic cyanobacteria in dark anoxic sediments actively produce photosynthesis related mRNA and are therefore able to assemble photosynthesis proteins.

We also detected more RNA transcripts translating for proteins involved in gas vesicles in the hypoxic-anoxic sediment. Gas vesicles allow cyanobacteria to regulate buoyance and ascend to the upper water column [20]. Even though gas vesicles have been shown to be produced by cyanobacteria in darkness [47], the pressure at the depth of the hypoxic-anoxic stations would be too high (1.3–2.0 MPa) and gas vesicles would collapse according to previous studies looking at *Aphanizomenon* and *Nodularia* gas vesicles that collapsed at 0.98 MPa [23]. In addition to these two genera, filamentous *Anabaena* is also one of the major cyanobacteria contributing to the large summer blooms in the Baltic Sea [48]. This is likely why we detected a large portion of reads classified to *Anabaena*, and it is therefore unknown if these cyanobacteria are able to ascend to the surface waters in the Baltic Sea [48]. The high turgor pressure at the deep hypoxic-anoxic stations might have caused cyanobacteria to actively transcribe mRNA that translates for gas vesicle proteins (any formed gas vesicles might however have collapsed) as an attempt to escape unfavorable conditions. Further studies are needed to investigate the role of cyanobacterial gas vesicles in dark and anoxic sediment located hundreds of meters below the water surface.

In addition to the extreme environmental conditions in dead zone sediments, cyanophages also infect and directly affect the abundance of cyanobacteria [30–32]. Viruses in sediment are affected by water flow, environmental shifts due to seasonal changes, as well as microbial abundances and activity [49]. However, dead zones in the Baltic Sea are stable environments due to a strong halocline trapping the heavy saline and oxygen deficient bottom water [25]. In the sampled gradient we detected that cyanophages had a different beta diversity in the hypoxic-anoxic sediment when compared to the oxic station A. This could partly be explained by that deep anoxic water in the Baltic Sea has a higher abundance of viruses [50], which might explain why cyanophage community structure was higher at station D, E, and F compared to oxic station A. In addition, cyanophages were indicated to have an association to different cyanobacterial communities and their metabolic activity across the sampled gradient (as indicated in Fig. 7A and 7B). This indicates that cyanobacteria in the hypoxic-anoxic sediment have to cope with different cyanophages compared to that of the oxic station. Considering that cyanophages can persist for thousands of years in anoxic sediments [37], differences in the cyanophage community composition and diversity might also be an effect of accumulating viruses in dead zone sediments. That cyanobacteria were indeed expressing an immune response against cyanophage infection was indicated by CRISPR-associated endoribonuclease related RNA transcripts (attributed to gene *cas2–3*) correlating with the relative abundance of several cyanophages. On the bacterial genome CRISPR (clustered regularly interspaced short palindromic repeats) loci consists of virus-specific DNA repeats with adjacent *cas* genes. Upon viral infection the virus DNA is sampled and inserted into a CRISPR locus; transcribed into crRNA; which are then used to guide Cas endonucleases proteins to cleave invading virus DNA [51]. These findings strongly suggest that cyanobacteria in this study were being targeted by cyanophages. We have here shown that cyanophages have a different beta diversity in hypoxic-anoxic sediment compared to oxic, and are actively infecting cyanobacteria triggering the CRISPR/Cas system.

We observed cyanobacterial N<sub>2</sub> fixation in the hypoxic-anoxic sediment, and it is therefore likely that viral lysis of these cells contribute carbon, phosphorus, and nitrogen to other microorganisms [33, 34]. The bacterial uptake of lysed material has been shown to be especially important in oligotrophic environments where nutrients are scarce [52], and this further suggests that viruses might have an important role in the studied dead zone sediments as they have been shown to be low or depleted in nitrate [43]. In addition, cyanophages are known to carry genes for the photosystem II [53], as was also discovered in our study, and it is therefore possible that changes in cyanobacterial photosynthesis regulation could have been due to viral infection by cyanophages. Studies looking at prokaryotic photosystem genes or protein counts might therefore partly be affected by prokaryote-viral interactions. In addition, cyanophages were also found to carry genes for the phosphorous regulon *phoH*. Even though the function of the translated protein is still not fully understood [53], the transcription of the *phoH* gene has been shown to be induced by phosphate starvation [54]. Cyanophages might therefore have influenced the cycling of phosphate by infecting cyanobacteria in the sediment. Further studies are needed to investigate the impact of such nutrient fluxes in dark and anoxic sediments. It is not possible to determine if cyanophages influenced the community composition of the cyanobacteria (or vice versa), or

how many of the assembled cyanophage genomes in this study were derived from sinking pelagic cyanobacteria or were already residing in the sediment. However, the difference in cyanophage beta diversity between the oxic and hypoxic-anoxic sediment suggests that host associated cyanobacteria capable of surviving in oxygen deficient environments select for specific cyanophages.

## Conclusions

We have here shown that cyanobacteria detected in the largest dead zone in the world are alive and adapted to survive in dark and anoxic sediments. Methods of survival includes anaerobic glycolysis and fatty acid biosynthesis (likely fermentative metabolism),  $N_2$  fixation, and possibly carbon fixation. Top prevalent cyanobacteria detected include genera not capable of forming akinetes (*Cyanobium* and *Synechococcus*) and akinetes forming *Anabaena*, which also contain species known to be heterotrophs and grow in darkness. Furthermore, active transcription of photosynthesis, phytochromes, and gas vesicle related genes indicate that these cyanobacteria are trying to ascend to the surface waters. These findings indicate that a seed bank of viable cyanobacteria is present in the dead zone sediment, and this might increase the recruitment to the surface water algal blooms. However, further work is needed to investigate if such recruitment is possible and under which circumstances. Considering that large amounts of cyanobacteria sinking to dead zone sediments are known to fuel the benthic ecosystems with phosphorous (so called internal loading), our cyanophage data indicate viral lysis of cyanobacteria and this might provide the nitrate-depleted sediment with nitrogen. Furthermore, the results showed that the cyanobacteria have to cope with different cyanophages in the hypoxic-anoxic sediment when compared to the oxic station. To conclude, cyanobacteria are commonly thought to have evolved from a last common ancestor carrying an evolved photosynthesis apparatus, as well as being essential in the great oxygenation event on Earth ca. 2.4 billion years ago [55]. Our results suggest that active cyanobacteria might also provide nutrients (via  $N_2$  fixation and viral lysis) in dark and anoxic environments.

## Methods

### Study area and sampling

Sediment was collected using a modified box corer [56] during 23–26 April 2018 from four stations, located below the euphotic zone, following a water depth and oxygen gradient (stations A, D, E, and F) in the dead zone of Eastern Gotland basin, Baltic Sea (the stations and geochemical data have previously been presented in Marzocchi U, Bonaglia S, van de Velde S, Hall POJ, Schramm A, Risgaard-Petersen N and Meysman FJR [43]) (Table 1; 2018 oxygen data first presented in Broman E, Bonaglia S, Holovachov O, Marzocchi U, Hall POJ and Nascimento FJA [57]). Three PVC cylinders (5 cm diameter; 30 cm length) from each site were inserted into the box-core sediment and moved onto a sterile bench. The sediment was gently extruded and the top 0–2 top cm directly sliced into a 50 mL centrifuge tube. The sample was flash frozen in liquid  $N_2$  and stored at  $-80^\circ\text{C}$  until DNA and RNA extraction.

# Nucleic acids extraction

DNA and RNA were extracted from thawed and homogenized sediment with the DNeasy PowerMax Soil kit (QIAGEN) and RNeasy PowerSoil kit (QIAGEN), with 10 g and 2 g input material, respectively. Extracted RNA was DNase treated with the TURBO DNA-free kit (Invitrogen), and ribosomal depletion using the RiboMinus Transcriptome Isolation Kit (bacterial version, ThermoFisher Scientific). Multiplexed libraries were prepared with the ThruPLEX DNA-seq (Rubicon Genomics) and TruSeq RNA Library Prep v2 (Illumina) kits for DNA and RNA, respectively. Paired-end 2 × 150 bp sequencing was conducted at the Science for Life Laboratory, Stockholm on the Illumina NovaSeq platform (DNA on one lane NovaSeq6000 S2; and RNA on one lane Illumina NovaSeq6000 S4). The sequencing yielded on average 41 million (range 32–52) and 82 million (range 74–89) read-pairs for each DNA and RNA sample, respectively.

# Bioinformatics

Illumina adapters were removed from the raw sequence data by using SeqPrep 1.2 [58] on default settings targeting the adapter sequences. Phi-X174 control sequences were removed from the data by mapping the reads with bowtie2 2.3.4.3 [59] against the PhiX genome (NCBI Reference Sequence: NC\_001422.1). Reads were quality trimmed using Trimmomatic 0.36 [60] with the parameters: LEADING:20 TRAILING:20 MINLEN:50, yielding quality trimmed reads with a an average read length of 136 bp (range 135–139 bp).

Metagenome and RNA-seq quality trimmed sequences (trimmomatic: paired without unpaired, PwU) were classified against the NCBI RefSeq genome database (downloaded: 1 March 2019) using Kraken2 2.0.7 with a paired-end setup [61]. The DNA data consisted on average of 103,043 cyanobacteria counts (range 51,939–152,036), while the RNA data consisted on average of 3,860,948 cyanobacteria counts (range 2,647,420–5,345,457). The RNA samples still contained on average 89% rRNA sequences (range 87–92%) after ribosomal depletion and therefore contained a large amount of rRNA alongside mRNA, this would have helped to further identify the active cyanobacterial community in the sediment. This method made it possible to classify a large amount of reads and detect present and active cyanobacteria. However, because the reads classify against genes with different lengths in the RefSeq genomes it is not possible to directly compare minor differences in relative abundances of taxonomic groups (although it gives a good indication of the distribution of reads). The final taxonomy data was analyzed and visualized as relative proportion of classified reads (%) in the software Explicet 2.10.5 [62].

The DNA PwU quality trimmed reads were used to construct a metagenome co-assembly using MEGAHIT 1.1.2 on default settings [63]. The Prokka 1.12 software tool [64] was used for gene prediction with Prodigal 2.6.3 [65] and annotation using BLAST 2.6.0+ [66] against the UniProtKB/Swiss-Prot database (database downloaded: 31 January 2019), with the parameters:—proteins uniprot\_sprot.fasta—metagenome. DNA and RNA PwU quality trimmed reads were then mapped onto the co-assembly using

bowtie2 on default settings and the output.sam files were converted to .bam with samtools 1.9 [67]. Final sequence counts were estimated by using htseq-count from the HTSeq python package 0.9.1 [68] with the .bam files and PROKKA output.gff file as input, with the parameters: -s no -f bam -t CDS -i ID. Sequence counts were normalized among samples as Gene length corrected Trimmed Mean of M-values (GeTMM) [69]. Unique UniProtKB/Swiss-Prot identifiers were merged and identifiers affiliated with Cyanobacteria in the UniProtKB/Swiss-Prot database were extracted. The final data consisted on average of 175,134 mapped cyanobacteria DNA sequences per sample (range 143,600–209,247), and an average of 5080 mapped cyanobacteria RNA transcripts per sample (range 4075–6048). Proteins were categorized into KEGG MODULES by linking UniProtKB/Swiss-Prot identifiers to KEGG KO identifiers using the “Retrieve/ID mapping” function available on the official UniProt website, <https://www.uniprot.org/uploadlists/>.

Contigs longer than 2000 bp in the metagenome co-assembly were extracted using sqtek 1.3, and identified as viruses with VirSorter 1.0.3 [70] on default settings via the CyVerse Discovery Environment interface (<https://de.cyverse.org>). Contigs classified as “Pretty sure” and “Quite sure” by VirSorter were used for further analyses. Open reading frames (ORFs) were detected in these contigs by using prodigal 2.6.3 on default setting [65]. To predict potential cyanophages, network analyses to isolated, genome-sequenced bacterial and archaeal viruses was conducted using vConTACT2 with default settings (version 0.9.8, NCBI Bacterial and Archaeal Viral RefSeq V88; [71]). The ORFs were further classified using BLASTP 2.9.0+ against a database of all viral genomes from the NCBI RefSeq database (downloaded date: 12 July 2019) with the following parameters: -max\_target\_seqs 1 -outfmt 6 -evalue 0.001 [66] to verify the cyanophages. Bowtie2 was used to map DNA PwU reads on the contigs with a 90% identity alignment threshold. Considering the reads had an average length of 136 bp after quality trimming, using Bowtie2 parameters: “-score-min L,0,-0.6”, allows for 14 mismatches without including additional gap penalties. The output.sam files were converted to .bam and sorted by coordinates using samtools 1.9 [67]. BEDTools 2.27.1 [72] were used to calculate total read depth per contig (mapped bp) and coverage per contig (%) using parameters “bedtools genomecov -d” and “bedtools genomecov -pc”, respectively. The dataset was delimited to cyanophage identified contigs with a coverage of at least 75% to ensure the viral genome was present in the sample [73]. The amount of mapped bp per contig was normalized among samples by dividing the contig length (kb) and the total number of bases in the metagenome (Mb) [73]. A full list of bioinformatic statistics is available in Data S7, such as number reads obtained after sequencing, quality trimming, mapping against the metagenome co-assembly, and numbers of reads classified against the UniProtKB/Swiss-Prot database.

## Statistics

Shannon’s H alpha diversity index of cyanobacteria was calculated in the software Explicit based on the RNA data after subsampling to the lowest sample size (2,647,420 counts) and bootstrap × 100. The final alpha index was reported as the mean of the bootstraps. Shannon’s H alpha diversity was calculated on the normalized sequence counts for the cyanophages. Non-metric multidimensional scaling (NMDS) showing cyanophage Bray-Curtis beta diversity was based on the normalized counts and analyzed in the

software past 3.25 [74]. The best explanatory cyanophages for the cyanobacterial community structure was analyzed with the BVSTEP method [75] using R 3.6.0 [76] and the bvStep function with default settings in the sinkr R package [77]. The bvStep was run by using the distance methods bray-curtis for cyanobacteria proteins and euclidean distances for cyanophages. NMDS plots of cyanobacteria community structure or cyanobacteria proteins beta diversity with cyanophages as overlying triplot were constructed using the metaMDS function in the vegan R package [78]. Spearman correlations were conducted to find patterns between cyanophages and cyanobacteria proteins by using the function rcorr in the Hmisc R package [79]. Differences in alpha diversity between stations were tested with One-Way ANOVA tests and post hoc Tukey tests using IBM SPSS 26. Differences in the beta diversity between stations were tested with PERMANOVA (9999 permutations) using the software past. Differences in RNA transcripts for annotated proteins between stations were tested with the R package edgeR 3.24.3 [80] by using the "run\_DE\_analysis.pl" script supplied with Trinity 2.8.2 [81]. False discovery rates ( $FDR$ ) < 0.05 were used to indicate statistical significances.

## Declarations

### *Ethics approval and consent to participate*

Not applicable.

### *Consent for publication*

Not applicable.

### *Availability of data and materials*

The raw sequencing data supporting the conclusions of this article is available in the NCBI BioProject repository, PRJNA531756.

<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA531756>

### *Competing interest*

The authors declare that they have no competing interest.

### *Funding*

Individual financial support was provided by the Swedish Research Council Formas to SB (Grant no. 2017-01513); the Swedish Research Council, Vetenskapsrådet to KH (Grant no 2013-4554); the Swedish Research Council VR to POJH (Grant no. 2015-03717); the Stockholm University's strategic funds for Baltic Sea research to FN; and the Swedish Research Council Formas to FN (Grant no. 2016-00804). This material is based upon work supported by the National Science Foundation under Award Numbers DBI-0735191, DBI-1265383, and DBI-1743442. URL: [www.cyverse.org](http://www.cyverse.org).

### *Author contributions*

EB drafted the manuscript, conducted laboratory work, bioinformatics, and data analyses. KH conducted bioinformatics, data analyses, and gave feedback on the manuscript. SB sampled in the field, conducted sediment microprofiling, and gave feedback on the manuscript. POJH led the sea expedition and gave feedback on the manuscript. FJAN coordinated the study and helped draft the manuscript. The research was designed by EB, SB, and FJAN. All authors read and approved the final manuscript.

### *Acknowledgements*

The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. We thank the captain and crew of University of Gothenburg R/V Skagerak for skillful support at sea.

## **References**

- 1.Castenholz RW: *General Characteristics of the Cyanobacteria*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Ltd; 2015
- 2.Karlson AML, Duberg J, Motwani NH, Hogfors H, Klawonn I, Ploug H, Barthel Svedén J, Garbaras A, Sundelin B, Hajdu S, et al: *Nitrogen fixation by cyanobacteria stimulates production in Baltic food webs*. *Ambio* 2015, *44*:413–426.
- 3.Karlson AML, Nascimento FJA, Elmgren R: *Incorporation and burial of carbon from settling cyanobacterial blooms by deposit-feeding macrofauna*. *Limnology and Oceanography* 2008, *53*:2754–2758.
- 4.Brocke HJ, Wenzhoefer F, de Beer D, Mueller B, van Duyl FC, Nugues MM: *High dissolved organic carbon release by benthic cyanobacterial mats in a Caribbean reef ecosystem*. *Scientific Reports* 2015, *5*:8852.

- 5.O'Neil JM, Roman MR: *Ingestion of the cyanobacterium Trichodesmium spp. by pelagic harpacticoid copepods Macrosetella, Miracia and Oculoseitella. Hydrobiologia* 1994, 292:235–240.
- 6.Nascimento FJA, Karlson AM, Elmgren R: *Settling blooms of filamentous cyanobacteria as food for meiofauna assemblages. Limnology and Oceanography* 2008, 53:2636–2643.
- 7.Emeis KC, Struck U, Leipe T, Pollehne F, Kunzendorf H, Christiansen C: *Changes in the C, N, P burial rates in some Baltic Sea sediments over the last 150 years - relevance to P regeneration rates and the phosphorus cycle. Marine Geology* 2000, 167:43–59.
- 8.Vahtera E, Conley DJ, Gustafsson BG, Kuosa H, Pitkänen H, Savchuk OP, Tamminen T, Viitasalo M, Voss M, Wasmund N: *Internal ecosystem feedbacks enhance nitrogen-fixing cyanobacteria blooms and complicate management in the Baltic Sea. AMBIO: A journal of the Human Environment* 2007, 36:186–195.
- 9.Havens KE: *Cyanobacteria blooms: effects on aquatic ecosystems. In Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs. Edited by Hudnell HK. New York, NY: Springer New York; 2008: 733–747*
- 10.Bláha L, Babica P, Maršálek B: *Toxins produced in cyanobacterial water blooms - toxicity and risks. Interdisciplinary toxicology* 2009, 2:36–41.
- 11.Conley DJ: *Ecology: Save the Baltic Sea. Nature* 2012, 486:463–464.
- 12.Peinert R, Saure A, Stegmann P, Stienen C, Haardt H, Smetacek V: *Dynamics of primary production and sedimentation in a coastal ecosystem. Netherlands Journal of Sea Research* 1982, 16:276–289.
- 13.Nascimento FJA, Karlson AML, Näslund J, Gorokhova E: *Settling cyanobacterial blooms do not improve growth conditions for soft bottom meiofauna. Journal of Experimental Marine Biology and Ecology* 2009, 368:138–146.
- 14.Middelburg JJ, Meysman FJR: *Burial at Sea. Science* 2007, 316:1294–1295.
- 15.Mannan RM, Pakrasi HB: *Dark Heterotrophic Growth Conditions Result in an Increase in the Content of Photosystem II Units in the Filamentous Cyanobacterium Anabaena variabilis ATCC 29413. Plant Physiology* 1993, 103:971–977.
- 16.Pelroy RA, Bassham JA: *Photosynthetic and dark carbon metabolism in unicellular blue-green algae. Archiv für Mikrobiologie* 1972, 86:25–38.
- 17.Richardson LL, Castenholz RW: *Enhanced Survival of the Cyanobacterium Oscillatoria terebriformis in Darkness under Anaerobic Conditions. Applied and Environmental Microbiology* 1987, 53:2151–2158.

18. Stal LJ, Moezelaar R: *Fermentation in cyanobacteria*<sup>1</sup> Publication 2274 of the Centre of Estuarine and Coastal Ecology, Yerseke, The Netherlands. 1. *FEMS Microbiology Reviews* 1997, 21:179–211.
19. Moezelaar R, Bijvank SM, Stal LJ: *Fermentation and Sulfur Reduction in the Mat-Building Cyanobacterium Microcoleus chthonoplastes*. *Applied and Environmental Microbiology* 1996, 62:1752–1758.
20. Kaplan-Levy RN, Hadas O, Summers ML, Rücker J, Sukenik A: *Akinetes: dormant cells of cyanobacteria*. In *Dormancy and resistance in harsh environments*. Springer; 2010: 5–27
21. Myers JH, Beardall J, Allinson G, Salzman S, Robertson S, Gunthorpe L: *Potential triggers of akinete differentiation in Nodularia spumigena (Cyanobacteriaceae) isolated from Australia*. *Hydrobiologia* 2011, 671:165.
22. Huber AL: *Factors Affecting the Germination of Akinetes of Nodularia spumigena (Cyanobacteriaceae)*. *Applied and Environmental Microbiology* 1985, 49:73–78.
23. Walsby AE, Hayes PK, Boje R: *The gas vesicles, buoyancy and vertical distribution of cyanobacteria in the Baltic Sea*. *European Journal of Phycology* 1995, 30:87–94.
24. Voss M, Dippner JW, Humborg C, Hürdler J, Korth F, Neumann T, Schernewski G, Venohr M: *History and scenarios of future development of Baltic Sea eutrophication*. *Estuarine, Coastal and Shelf Science* 2011, 92:307–322.
25. Diaz RJ, Rosenberg R: *Spreading dead zones and consequences for marine ecosystems*. *Science* 2008, 321:926–929.
26. Bonaglia S, Klawonn I, De Brabandere L, Deutsch B, Thamdrup B, Brüchert V: *Denitrification and DNRA at the Baltic Sea oxic–anoxic interface: Substrate spectrum and kinetics*. *Limnology and Oceanography* 2016, 61:1900–1915.
27. Broman E, Sjöstedt J, Pinhassi J, Dopson M: *Shifts in coastal sediment oxygenation cause pronounced changes in microbial community composition and associated metabolism*. *Microbiome* 2017, 5.
28. Broman E, Sachpazidou V, Pinhassi J, Dopson M: *Oxygenation of hypoxic coastal Baltic Sea sediments impacts on chemistry, microbial community composition, and metabolism*. *Frontiers in Microbiology* 2017, 8.
29. Thureborn P, Franzetti A, Lundin D, Sjöling S: *Reconstructing ecosystem functions of the active microbial community of the Baltic Sea oxygen depleted sediments*. *PeerJ* 2016, 4:e1593.
30. Waterbury JB, Valois FW: *Resistance to co-occurring phages enables marine synechococcus communities to coexist with cyanophages abundant in seawater*. *Applied and environmental*

*microbiology* 1993, 59:3393–3399.

31. Suttle CA, Chan AM: *Marine cyanophages infecting oceanic and coastal strains of Synechococcus: abundance, morphology, cross-infectivity and growth characteristics. Marine Ecology-Progress Series* 1993, 92:99–99.

32. Sullivan MB, Waterbury JB, Chisholm SW: *Cyanophages infecting the oceanic cyanobacterium Prochlorococcus. Nature* 2003, 424:1047–1051.

33. Cairns J, Coloma S, Sivonen K, Hiltunen T: *Evolving interactions between diazotrophic cyanobacterium and phage mediate nitrogen release and host competitive ability. Royal Society open science* 2016, 3:160839–160839.

34. Holmfeldt K, Titelman J, Riemann L: *Virus Production and Lysate Recycling in Different Sub-basins of the Northern Baltic Sea. Microbial Ecology* 2010, 60:572–580.

35. Fuchsman CA, Palevsky HI, Widner B, Duffy M, Carlson MCG, Neibauer JA, Mulholland MR, Keil RG, Devol AH, Rocap G: *Cyanobacteria and cyanophage contributions to carbon and nitrogen cycling in an oligotrophic oxygen-deficient zone. The ISME Journal* 2019.

36. Kate R. H, N. John A, Martha R. J. C: *Recovery of viable cyanophages from the sediments of a eutrophic lake at decadal timescales. FEMS Microbiology Ecology* 2013, 83:450–456.

37. Cai L, Jørgensen BB, Suttle CA, He M, Cragg BA, Jiao N, Zhang R: *Active and diverse viruses persist in the deep sub-seafloor sediments over thousands of years. The ISME Journal* 2019, 13:1857–1864.

38. Bianchi TS, Engelhaupt E, Westman P, Andrén T, Rolff C, Elmgren R: *Cyanobacterial blooms in the Baltic Sea: Natural or human-induced? Limnology and Oceanography* 2000, 45:716–726.

39. Jessen GL, Lichtschlag A, Ramette A, Pantoja S, Rossel PE, Schubert CJ, Struck U, Boetius A: *Hypoxia causes preservation of labile organic matter and changes seafloor microbial community composition (Black Sea). Science Advances* 2017, 3:e1601897.

40. Hulthe G, Hulth S, Hall POJ: *Effect of oxygen on degradation rate of refractory and labile organic matter in continental margin sediments. Geochimica et Cosmochimica Acta* 1998, 62:1319–1328.

41. Hallenbeck PC, Kochian LV, Benemann JR: *Hydrogen evolution catalyzed by hydrogenase in cultures of cyanobacteria. Zeitschrift für Naturforschung C* 1981, 36:87–92.

42. Sukenik A, Rucker J, Maldener I: *Chapter 4 - Dormant Cells (Akinetes) of Filamentous Cyanobacteria Demonstrate a Great Variability in Morphology, Physiology, and Ecological Function. In Cyanobacteria. Edited by Mishra AK, Tiwari DN, Rai AN: Academic Press; 2019: 65–77*

43. Marzocchi U, Bonaglia S, van de Velde S, Hall POJ, Schramm A, Risgaard-Petersen N, Meysman FJR: *Transient bottom water oxygenation creates a niche for cable bacteria in long-term anoxic sediments of the Eastern Gotland Basin. Environmental Microbiology* 2018, *20*:3031–3041.
44. Bergman B, Gallon JR, Rai AN, Stal LJ: *N<sub>2</sub> Fixation by non-heterocystous cyanobacteria. FEMS Microbiology Reviews* 1997, *19*:139–185.
45. Smart LB, McIntosh L: *Expression of photosynthesis genes in the cyanobacterium Synechocystis sp. PCC 6803: psaA-psaB and psbA transcripts accumulate in dark-grown cells. Plant Molecular Biology* 1991, *17*:959–971.
46. Ueta M, Ohniwa RL, Yoshida H, Maki Y, Wada C, Wada A: *Role of HPF (Hibernation Promoting Factor) in Translational Activity in Escherichia coli. The Journal of Biochemistry* 2008, *143*:425–433.
47. Deacon C, Walsby AE: *Gas vesicle formation in the dark, and in light of different irradiances, by the cyanobacterium Microcystis sp. British Phycological Journal* 1990, *25*:133–139.
48. Stal LJ, Albertano P, Bergman B, Bröckel Kv, Gallon JR, Hayes PK, Sivonen K, Walsby AE: *BASIC: Baltic Sea cyanobacteria. An investigation of the structure and dynamics of water blooms of cyanobacteria in the Baltic Sea—responses to a changing environment. Continental Shelf Research* 2003, *23*:1695–1714.
49. Vandieken V, Sabelhaus L, Engelhardt T: *Virus Dynamics Are Influenced by Season, Tides and Advective Transport in Intertidal, Permeable Sediments. Frontiers in Microbiology* 2017, *8*.
50. Weinbauer MG, Brettar I, Höfle MG: *Lysogeny and virus-induced mortality of bacterioplankton in surface, deep, and anoxic marine waters. Limnology and Oceanography* 2003, *48*:1457–1465.
51. Barrangou R: *The roles of CRISPR–Cas systems in adaptive immunity and beyond. Current Opinion in Immunology* 2015, *32*:36–41.
52. Riemann L, Holmfeldt K, Titelman J: *Importance of Viral Lysis and Dissolved DNA for Bacterioplankton Activity in a P-Limited Estuary, Northern Baltic Sea. Microbial Ecology* 2009, *57*:286–294.
53. Gao EB, Huang Y, Ning D: *Metabolic Genes within Cyanophage Genomes: Implications for Diversity and Evolution. Genes* 2016, *7*:80.
54. Kim SK, Makino K, Amemura M, Shinagawa H, Nakata A: *Molecular analysis of the phoH gene, belonging to the phosphate regulon in Escherichia coli. Journal of Bacteriology* 1993, *175*:1316–1324.
55. Herrero A, Flores E, Flores FG: *The Cyanobacteria: Molecular Biology, Genomics, and Evolution. Caister Academic Press; 2008.*
56. Blomqvist S, Ekeröth N, Elmgren R, Hall PO: *Long overdue improvement of box corer sampling. Marine Ecology Progress Series* 2015, *538*:13–21.

57. Broman E, Bonaglia S, Holovachov O, Marzocchi U, Hall POJ, Nascimento FJA: *Uncovering diversity and metabolic spectrum of animals in dead zone sediments. Communications Biology* 2019. In review.
58. St John J: *SeqPrep*. 2014.
59. Langmead B, Salzberg SL: *Fast gapped-read alignment with Bowtie 2. Nature Methods* 2012, 9:357.
60. Bolger AM, Lohse M, Usadel B: *Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics* 2014, 30:2114–2120.
61. Wood DE, Salzberg SLJGB: *Kraken: ultrafast metagenomic sequence classification using exact alignments. 2014, 15:R46.*
62. Robertson CE, Harris JK, Wagner BD, Granger D, Browne K, Tatem B, Feazel LM, Park K, Pace NR, Frank DN: *Explicit: graphical user interface software for metadata-driven management, analysis and visualization of microbiome data. Bioinformatics* 2013, 29:3100–3101.
63. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW: *MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods* 2016, 102:3–11.
64. Seemann T: *Prokka: rapid prokaryotic genome annotation. Bioinformatics* 2014, 30:2068–2069.
65. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ: *Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics* 2010, 11:119.
66. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: *Basic local alignment search tool. Journal of Molecular Biology* 1990, 215:403–410.
67. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: *The sequence alignment/map format and SAMtools. Bioinformatics* 2009, 25:2078–2079.
68. Anders S, Pyl PT, Huber W: *HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics* 2015, 31:166–169.
69. Smid M, Coebergh van den Braak RRJ, van de Werken HJG, van Riet J, van Galen A, de Weerd V, van der Vlugt-Daane M, Bril SI, Lalmahomed ZS, Kloosterman WP, et al: *Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. 2018, 19:236.*
70. Roux S, Enault F, Hurwitz BL, Sullivan MB: *VirSorter: mining viral signal from microbial genomic data. PeerJ* 2015, 3:e985.
71. Bolduc B, Jang HB, Douclet G, You Z-Q, Roux S, Sullivan MB: *vConTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. PeerJ* 2017, 5:e3243.

- 72.Quinlan AR, Hall IM: *BEDTools: a flexible suite of utilities for comparing genomic features*. *Bioinformatics* 2010, 26:841–842.
- 73.Brum JR, Ignacio-Espinoza JC, Roux S, Doulier G, Acinas SG, Alberti A, Chaffron S, Cruaud C, De Vargas C, Gasol JM: *Patterns and ecological drivers of ocean viral communities*. *Science* 2015, 348:1261498.
- 74.Hammer Ø, Harper DAT, Ryan PD: *PAST: Paleontological statistics software package for education and data analysis*. *Palaeontologia Electronica* 2001, 4:9.
- 75.Clarke K, Warwick R: *Quantifying structural redundancy in ecological communities*. *Oecologia* 1998, 113:278–289.
- 76.R Core Team: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. 2013.
- 77.*sinkr*: *Collection of functions with emphasis in multivariate data analysis*. R package version 0.6.
- 78.Oksanen J, Blanchet FG, Kindt R, Legendre P, O'hara R, Simpson GL, Solymos P, Stevens MHH, Wagner HJhcr-poAe: *vegan: Community Ecology Package*. R package version 2.5–2. 2018.
- 79.Harrell Jr FE, Dupont C: *Hmisc: harrell miscellaneous*. R package version 2008, 3.
- 80.Robinson MD, McCarthy DJ, Smyth GK: *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. *Bioinformatics* 2010, 26:139–140.
- 81.Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al: *De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis*. *Nature Protocols* 2013, 8:1494–1512.

## Tables

**Table 1** Triplicate sediment samples (top 0-2 cm sediment surface) were collected during 2018 in the Eastern Gotland basin dead zone in the Baltic Sea. The stations were located along a water depth gradient (all aphotic) with different bottom water oxygen conditions. Oxygen concentrations shown were measured in the top 0-0.5 mm sediment surface, with station A being oxic, D hypoxic, E near-anoxic, and F anoxic.

Station	Water depth (m)	O <sub>2</sub> (μmol L <sup>-1</sup> )	Date	Longitude	Latitude
A	60	330	April 25	19°04'951	57°23'106
D	130	8.8	April 26	19°19'414	57°19'671
E	170	1.8	April 23	19°30'451	57°07'518
F	210	0	April 23	19°48'035	57°17'225

**Table 2** RNA transcripts attributed to proteins with a significant higher difference (edgeR analysis,  $FDR < 0.05$ ) in the hypoxic sediment of station D plus anoxic stations E and F when compared to station A. The values in the table shows the log fold change difference for stations D, E, and F compared to station A.

	Protein name	D	E	F
	3',5'-cyclic adenosine monophosphate phosphodiesterase CpdA		4.1	3.7
	30S ribosomal protein S18	5.9		5.6
	3-oxoacyl-[acyl-carrier-protein] reductase			5.2
	Acetyl-coenzyme A synthetase (AcCoA synthetase) (Acs)		3.0	
	Acyl carrier protein (ACP)		3.1	3.1
	Adenylate cyclase		1.9	
	ATP synthase gamma chain	5.5	4.5	4.4
	ATP-dependent DNA helicase RecG	2.0		
	Carbon dioxide-concentrating mechanism protein CcmK homolog 2			2.1
	Carbonic anhydrase	3.9	3.8	3.5
	Chaperone protein ClpB 1			4.4
	C-phycoerythrin class 1 subunit alpha			5.1
	Demethyl-4-deoxygadusol synthase (DDGS)	5.4	5.1	4.9
	DNA-binding protein HU			5.9
	Elongation factor Tu (EF-Tu)			5.4
	Gas vesicle structural protein (GVP), gvpA	6.2	5.4	5.9
	Glucokinase		2.6	
	Glutathione synthetase	1.8	1.9	
	HTH-type transcriptional activator CmpR		1.5	
	Iron uptake protein A2 (Iron deficiency-induced protein A) (PP2)			5.2
	Metalloprotease slr0863	3.6		3.9
	Nitrogenase iron protein		2.1	
	Photosystem I reaction center subunit II (PSI-D)	5.9	5.1	5.7
	Phytochrome-like protein cph1		1.6	
	Probable transposase for insertion sequence element IS701	1.4	1.4	1.4
	Protein GvpF/L	5.9		6.6
	Protein RcaC		7.3	
	Protein Smf	2.4	2.0	2.1
	Putative biopolymer transport protein ExbB-like 1	5.4		
	Putative dihydroflavonol 4-reductase (DFR)			2.0
	Putative pyruvate-flavodoxin oxidoreductase			1.4
	Putative RNA-binding protein RbpA	2.3	2.1	1.6
	Pyruvate kinase 1 (PK 1)	6.3		5.8
	Pyruvate kinase 2 (PK 2)		5.3	5.6
	Ribonuclease 3	1.7	2.7	2.4
	Ribosome hibernation promotion factor (HPF) (Dark-response protein 21)	2.8	2.8	2.2
	Ribosome hibernation promotion factor (HPF) (Light-repressed protein A)			2.7
	RNA 3'-terminal phosphate cyclase (RNA cyclase)			5.5
	SsrA-binding protein (Small protein B)			5.6
	Thiamine-phosphate synthase (TP synthase) (TPS)	3.8	3.9	3.5
	TPR repeat-containing protein SYNPC7002 A0425			5.8
	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	4.8	3.5	3.5
	Uncharacterized hydrolase sll0100	3.7		
	Uncharacterized methyltransferase slr0309			1.7
	Uncharacterized phycocyanin operon protein Z (ORF Z)			2.1
	Uncharacterized protein sll0925		6.4	
	Uncharacterized protein sll1119			5.4
	Uncharacterized protein sll1483	1.9	1.6	1.8
	Uncharacterized protein slr0039			1.9

# Figures

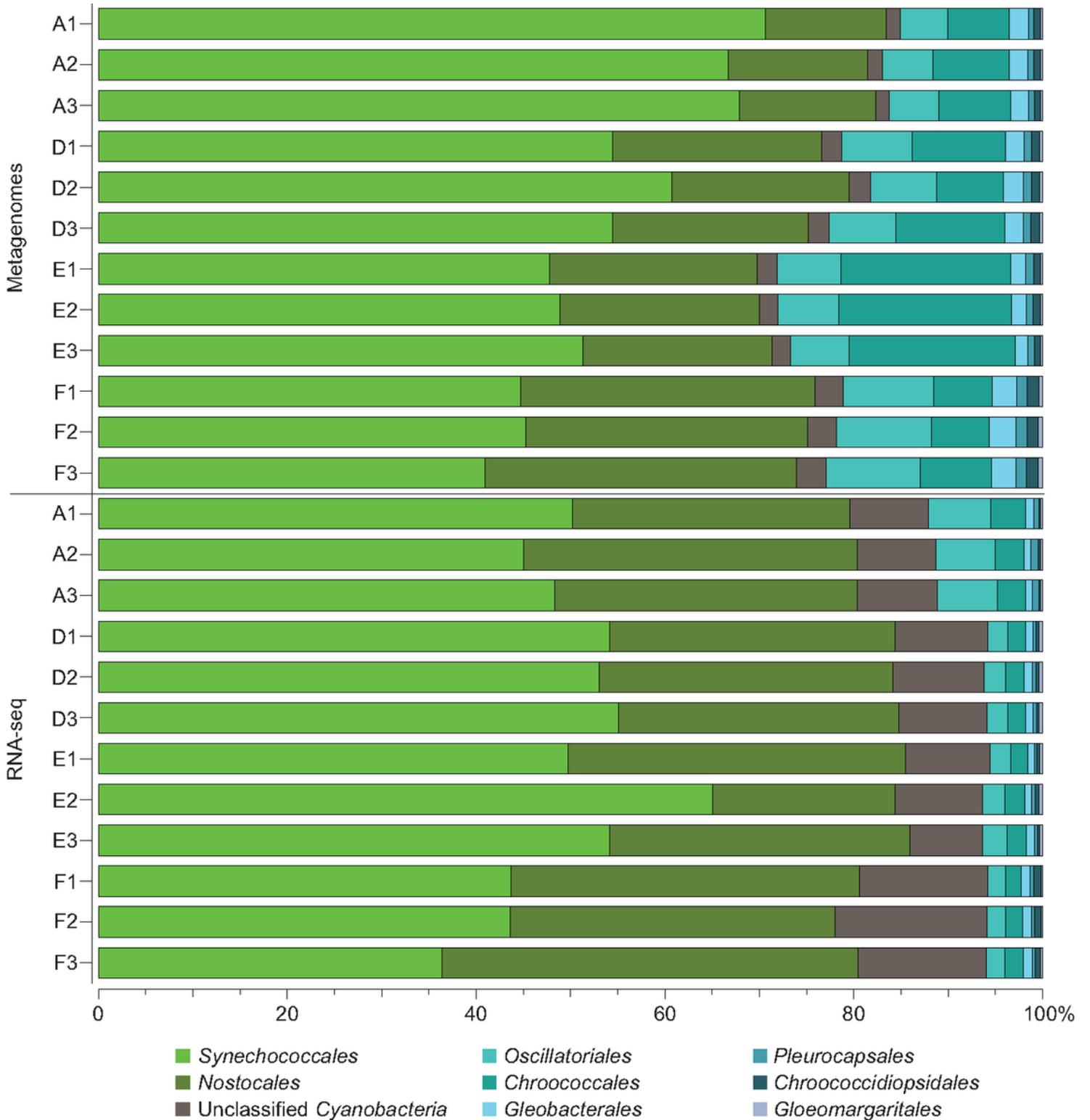
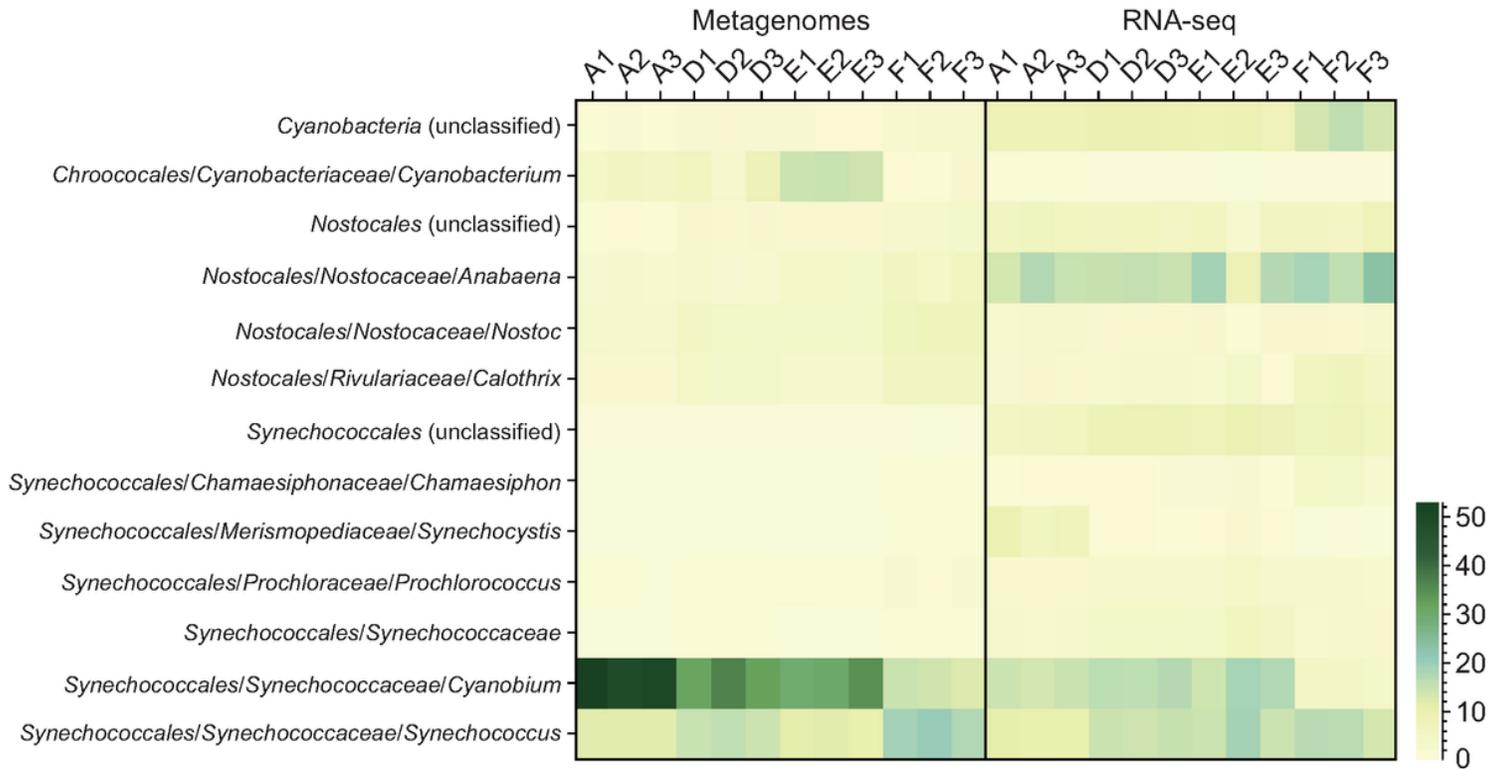


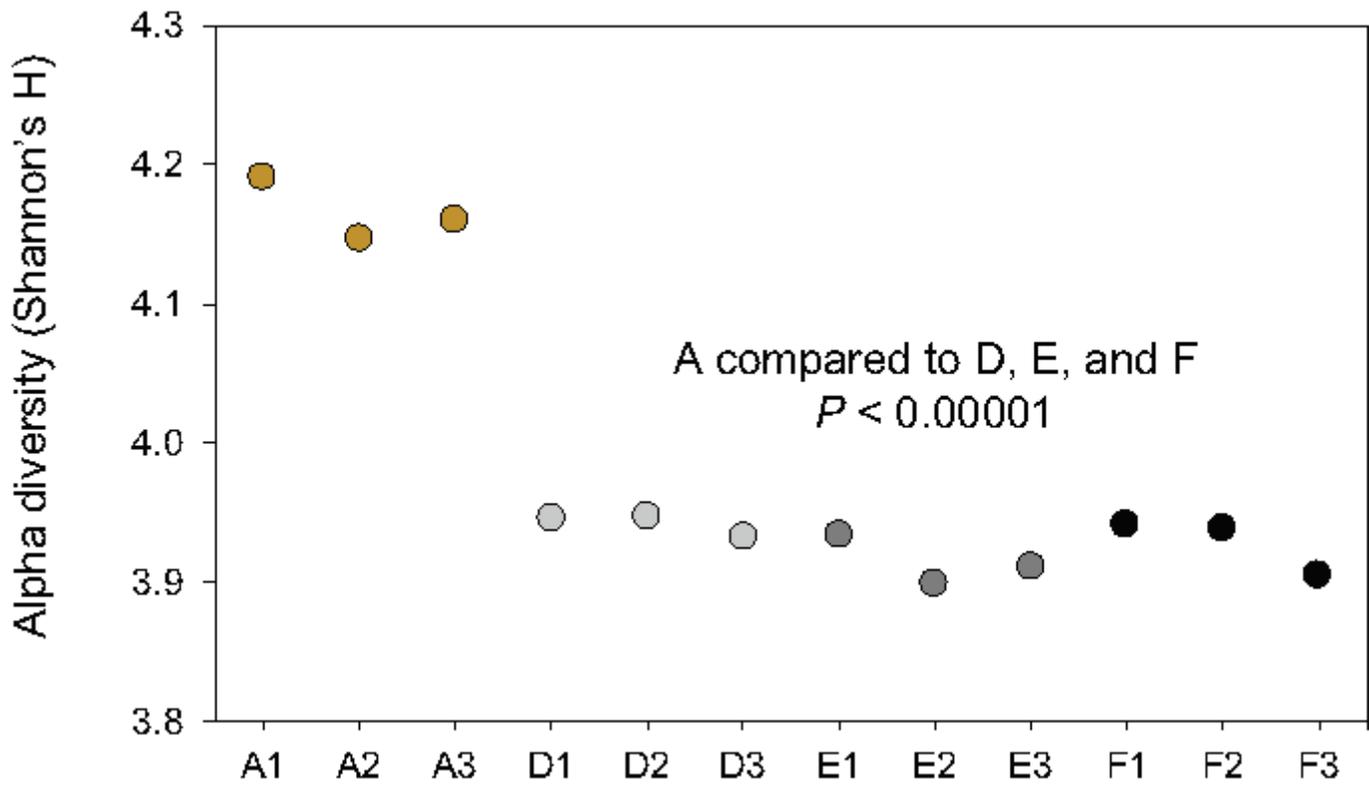
Figure 1

The stacked bars show the relative proportion of classified DNA and RNA sequences against cyanobacteria genomes in the RefSeq database. The data shows % of Cyanobacteria and the color legend shows taxonomic orders, while the y-axis shows the triplicate samples for each station (oxic sediment at station A, hypoxic D, and the anoxic sediment at E and F).



**Figure 2**

Heatmap showing the relative proportion of classified reads against cyanobacteria genomes in the RefSeq database. The data shows the lowest taxonomic classification with at least an average of >2% in the DNA or RNA samples.



**Figure 3**

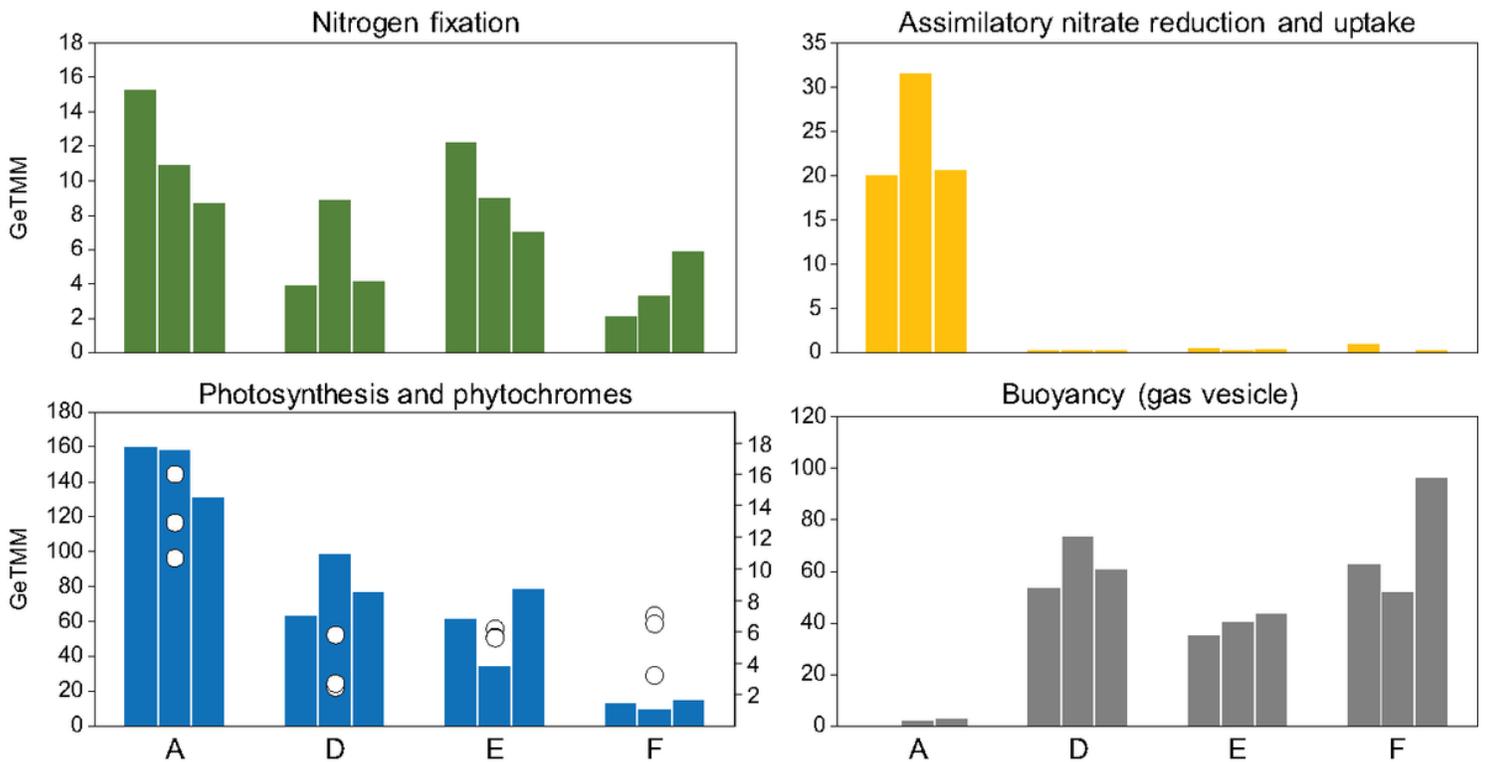
Shannon's H alpha diversity based on the lowest cyanobacteria taxonomic classifications in the RNA-seq data. B) NMDS based on the presence/absence Sørensen-Dice beta diversity of the Cyanobacteria in the RNA-Seq data.

**A**

KEGG MODULE	A1	A2	A3	D1	D2	D3	E1	E2	E3	F1	F2	F3
Cysteine biosynthesis	6.9	8.4	10.7	3.0	1.4	3.2	2.2	2.5	4.1	2.9	2.6	4.7
Fatty acid biosynthesis	9.0	11.6	6.9	4.1	4.0	5.1	6.7	2.7	7.2	6.5	10.6	6.8
Fatty acid biosynthesis, elongation	5.7	7.2	4.0	1.6	2.2	2.6	4.6	1.8	3.9	3.9	5.6	4.7
Fatty acid biosynthesis, initiation	3.3	5.2	2.4	1.6	2.2	2.3	2.7	2.0	2.6	3.3	1.0	1.6
F-type ATPase	4.7	9.7	5.1	4.0	6.6	5.9	2.5	3.5	4.9	3.2	3.6	8.3
Glycolysis (Embden-Meyerhof pathway)	6.1	5.7	2.7	2.5	4.8	1.5	5.6	4.0	2.9	3.7	2.1	3.5
Heme biosynthesis	4.9	7.5	4.9	4.0	5.2	6.0	12.8	4.0	5.2	9.0	3.2	7.1
NAD(P)H:quinone oxidoreductase	5.3	7.1	6.4	2.0	1.2	2.8	2.8	1.2	2.1	2.5	2.5	3.1
Nitrate assimilation	14.4	23.9	14.6	0.1	0.1	0.3	0.4	0.2	0.3	0.4	0.0	0.1
Pentose phosphate pathway (Pentose phosphate cycle)	5.8	6.8	3.4	1.3	2.1	1.6	3.0	1.7	2.0	3.4	2.3	1.8
Photosystem I	10.1	9.9	9.7	1.3	1.8	2.6	2.2	2.3	3.1	4.6	2.5	3.0
Photosystem II	136.9	140.4	110.8	55.2	90.5	68.4	51.5	28.0	71.0	3.4	1.7	5.1
Pimeloyl-ACP biosynthesis, BioC-BioH pathway	5.7	7.2	4.0	1.6	2.2	2.6	4.6	1.8	3.9	3.9	5.6	4.7
Reductive citrate cycle (Arnon-Buchanan cycle)	7.8	5.1	5.5	2.0	5.0	2.1	3.7	1.8	2.8	2.8	3.3	4.7
Reductive pentose phosphate cycle (Calvin cycle)	5.2	3.8	4.8	1.2	3.4	1.5	2.9	0.4	1.4	1.3	0.8	2.2
Siroheme biosynthesis	4.7	3.5	2.9	2.0	2.1	3.2	6.5	1.6	2.4	3.6	1.4	3.1
Tetrahydrofolate biosynthesis	4.2	4.5	2.5	1.1	1.8	1.1	2.1	1.5	2.1	2.7	3.6	2.7
Uridine monophosphate biosynthesis	12.8	7.4	15.1	3.2	2.0	2.8	2.0	0.6	0.8	1.0	1.0	0.9
Tetrahydrofolate biosynthesis, mediated by PTPS	3	4	2	1	2	1	2	2	2	3	4	3
3-Hydroxypropionate bi-cycle	3	5	2	1	2	2	2	2	3	3	1	2

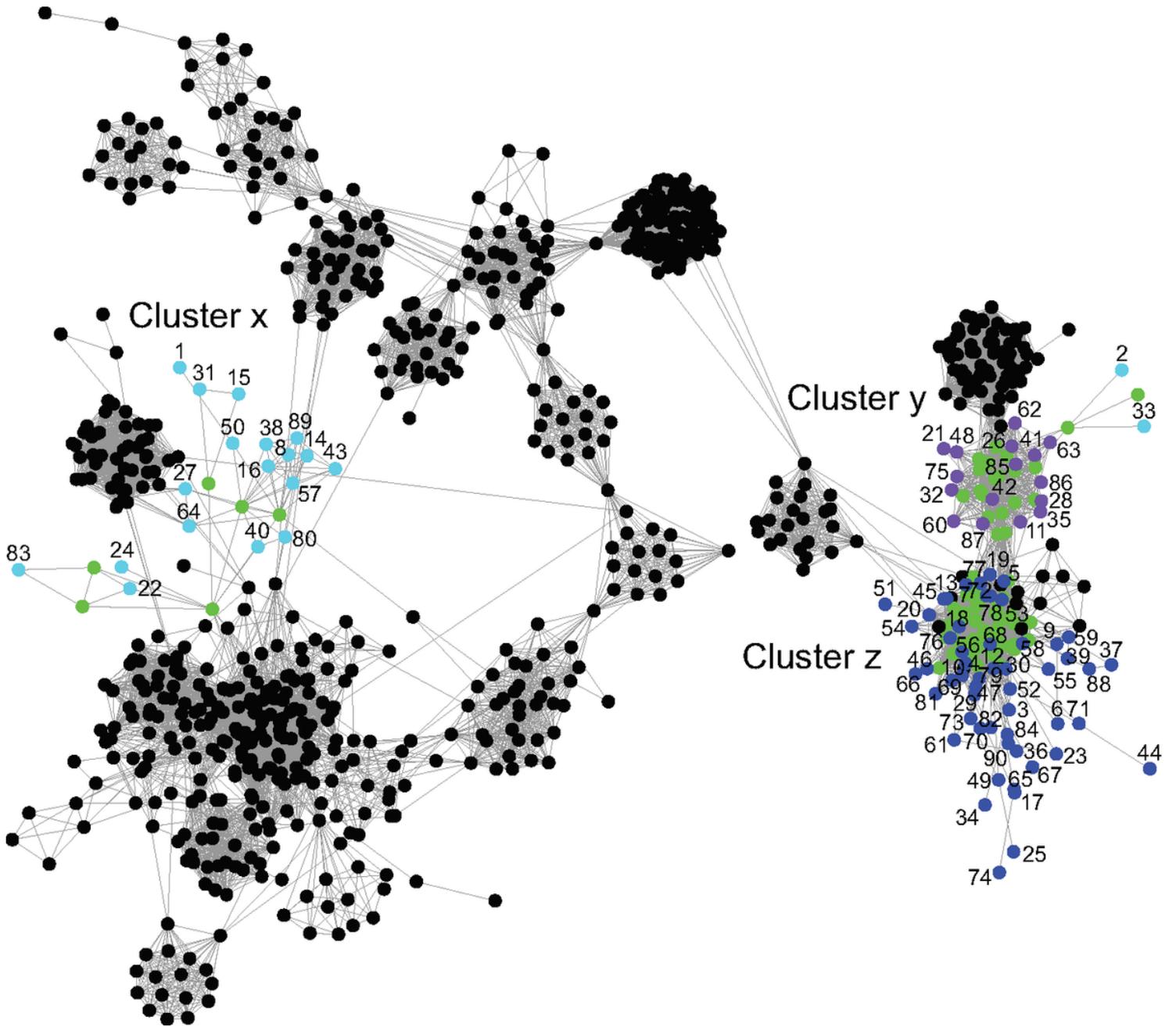
Legend: 0 (yellow), 5 (orange), ≥10 (red)

**B**



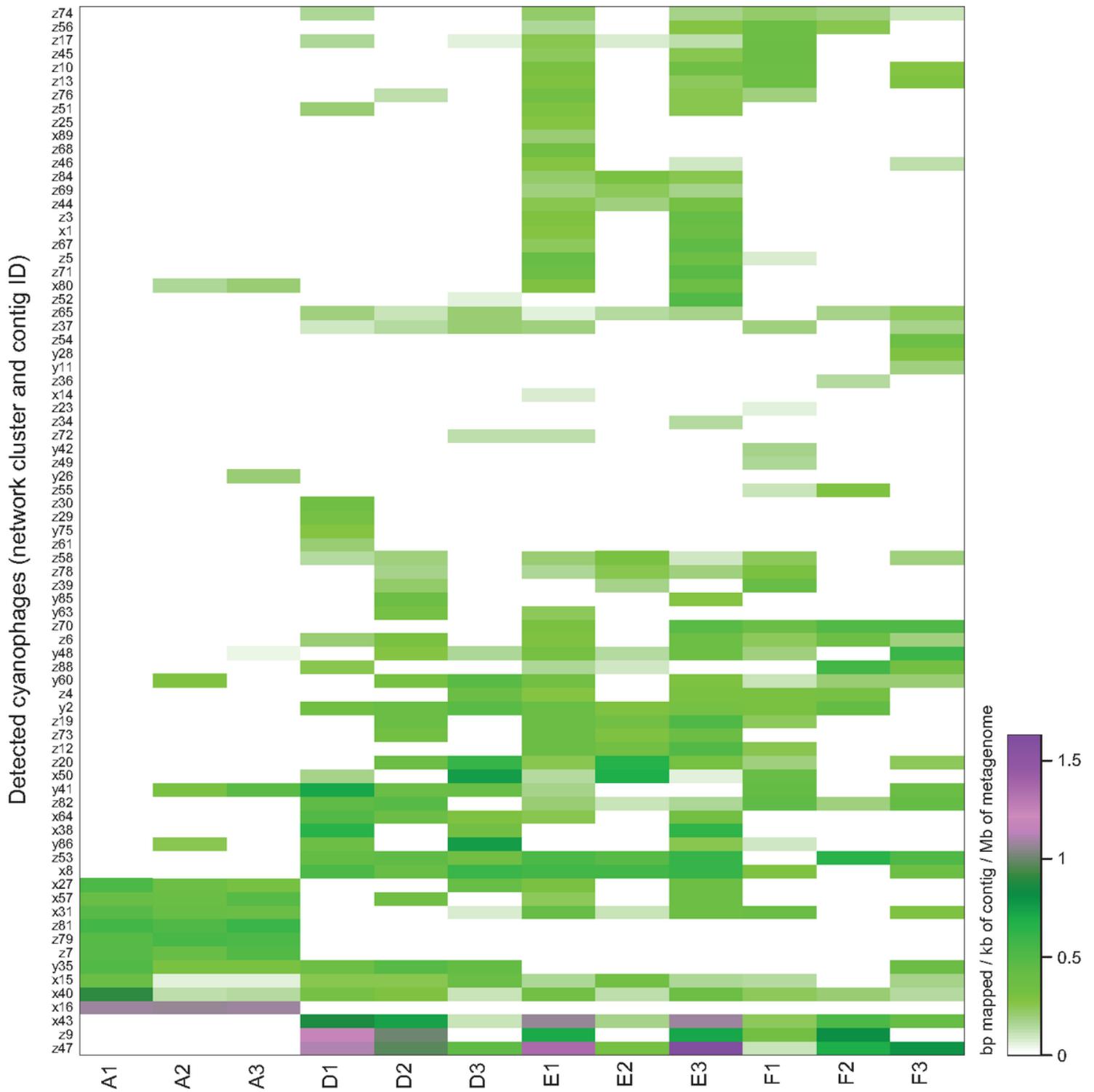
**Figure 4**

A) Cyanobacteria RNA transcripts classified against the UniProtKB/Swiss-Prot linked to KEGG modules. The heatmap shows the top 20 modules and normalized sequence counts (GeTMM). Note that proteins might be associated with more than one module. B) Specific categories among the major metabolic process linked to Cyanobacteria. Data is based on RNA transcripts and their UniProtKB/Swiss-Prot classifications. The photosynthesis subpanel also shows proteins classified as cyanobacterial phytochromes (white circles and right-side y-axis). Note the different y-axes between subpanels.



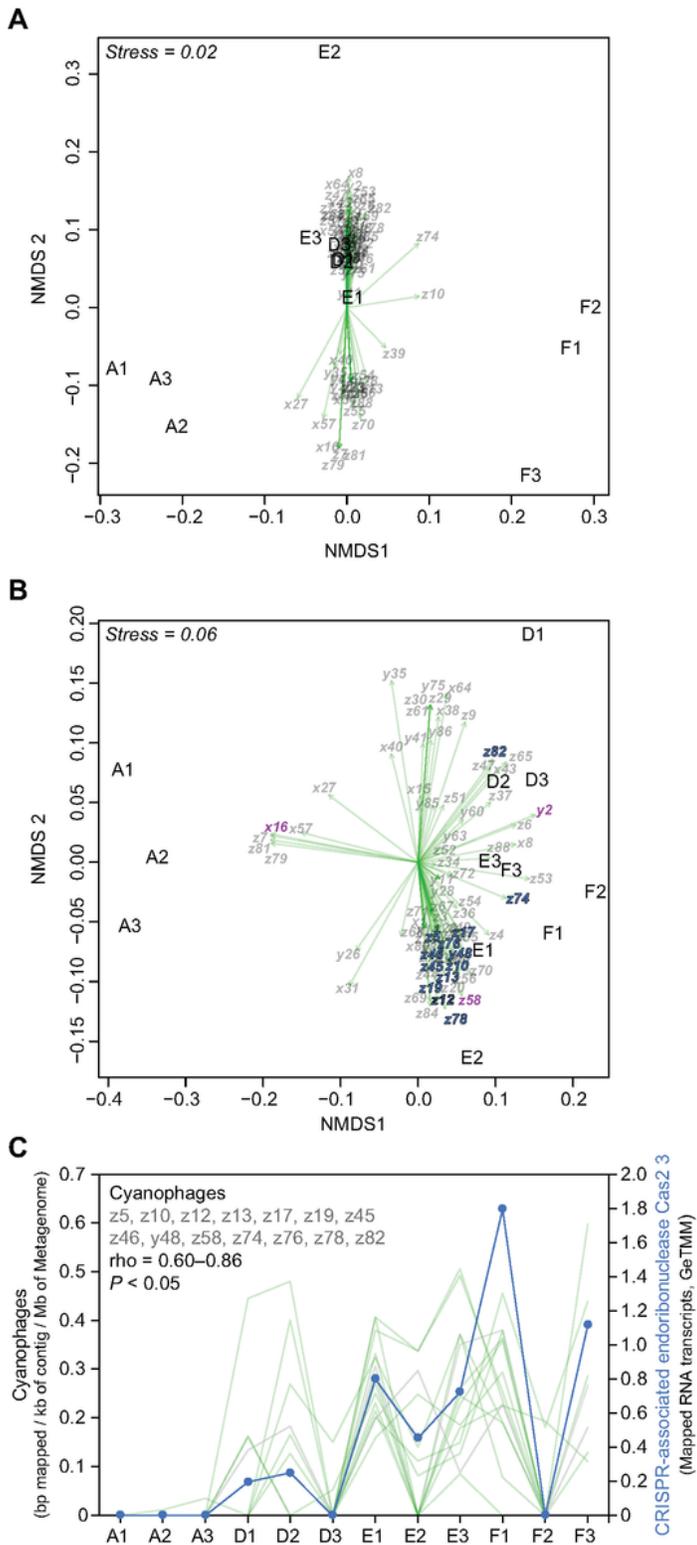
**Figure 5**

Network analysis consisting of reference viruses (NCBI Bacterial and Archaeal Viral RefSeq V88) and the constructed virus genomes. The black nodes show non-cyanophage viruses from the reference database, while green nodes show known cyanophages. The constructed contigs classified as viruses by VIRSorter and associated with cyanophages in the vConTACT network analysis are denoted as turquoise, purple, and blue nodes (followed by the contig ID number) and clustered into three groups in the network (x, y, and z). The colors denote turquoise as cyanophages belonging to the family Siphoviridae; purple Podoviridae; and blue Myoviridae.



**Figure 6**

One-way clustered heatmap by rows showing cyanophage constructed contigs with a minimum of 75% coverage. The green-purple color gradient shows normalized sequence counts (mapped bp per contig / contig length in kb / metagenome in Mb). The row labels show contig IDs and its associated network cluster in Fig. 4 (x, y, or z).



**Figure 7**

A) NMDS showing the Bray curtis beta diversity based on the proportion of reads classified of the cyanobacteria RNA data (lowest taxonomic classification, stations shown on the graph). The overlying triplot shows the distribution of detected cyanophages in relation to the cyanobacteria (light grey labels) beta diversity. The arrow heads denote the data point of the cyanophage. B) NMDS showing the Bray-curtis beta diversity of UniProtKB/Swiss-Prot classified cyanobacteria proteins (mRNA data) identified in

the dataset (stations shown and denoted on the graph). Cyanophages indicated by violet color (y2, x16, and z58) denote the cyanophages that together best explain the beta diversity of the cyanobacteria proteins ( $\rho = 0.80$ ). The blue colored cyanophages (as well as violet z58) denotes to the cyanophages in (C) that correlated with a cyanobacteria-related CRISPR protein. C) Several of the cyanophages in the hypoxic-anoxic sediment correlated positively with the number of mapped RNA transcripts against the metagenome assembly attributed to cyanobacterial protein CRISPR-associated endoribonuclease Cas2 3 (gene cas2-3). The light grey lines denote each cyanophage, while the blue line denotes the GeTMM values for the CRISPR protein.