

## ARTICLE OPEN

## Daily rhythmicity in coastal microbial mats

Christine Hörnlein<sup>1</sup>, Veronique Confurius-Guns<sup>1</sup>, Lucas J. Stal<sup>1,2</sup> and Henk Bolhuis<sup>1</sup>

Cyanobacteria are major primary producers in coastal microbial mats and provide biochemical energy, organic carbon, and bound nitrogen to the mat community through oxygenic photosynthesis and dinitrogen fixation. In order to anticipate the specific requirements to optimize their metabolism and growth during a day-and-night cycle, Cyanobacteria possess a unique molecular timing mechanism known as the circadian clock that is well-studied under laboratory conditions but little is known about its function in a natural complex community. Here, we investigated daily rhythmicity of gene expression in a coastal microbial mat community sampled at 6 time points during a 24-h period. In order to identify diel expressed genes, meta-transcriptome data was fitted to periodic functions. Out of 24,035 conserved gene transcript clusters, approximately 7% revealed a significant rhythmic expression pattern. These rhythmic genes were assigned to phototrophic micro-eukaryotes, Cyanobacteria but also to Proteobacteria and Bacteroidetes. Analysis of MG-RAST annotated genes and mRNA recruitment analysis of two cyanobacterial and three proteobacterial microbial mat members confirmed that homologs of the cyanobacterial circadian clock genes were also found in other bacterial members of the microbial mat community. These results suggest that various microbial mat members other than Cyanobacteria have their own molecular clock, which can be entrained by a cocktail of Zeitgebers such as light, temperature or metabolites from neighboring species. Hence, microbial mats can be compared to a complex organism consisting of multiple sub-systems that have to be entrained in a cooperative way such that the corpus functions optimally.

npj Biofilms and Microbiomes (2018)4:11 ; doi:10.1038/s41522-018-0054-5

## INTRODUCTION

Phototrophic microbial mats are highly complex and nearly self-sustaining laminated ecosystems that globally develop amongst others in coastal intertidal sandy sediments. The primary producers in microbial mats are oxygenic photoautotrophs, mainly Cyanobacteria and diatoms<sup>1</sup> and they supply the community with organic carbon. Many Cyanobacteria also fix atmospheric nitrogen thereby also providing this important nutrient to the community. Phototrophic microbial mats are driven by the daily cycle of light and dark in which processes such as photosynthesis, respiration, fermentation, migration and nitrogen fixation take place at designated times during a 24-h day.<sup>2–8</sup>

The evolution of a circadian clock enabled Cyanobacteria to anticipate the day and night regime by regulating gene expression and protein synthesis accordingly<sup>9</sup> providing these organisms with an enhanced fitness relative to mutants lacking a circadian clock.<sup>10</sup> The circadian clock is a temperature compensated time-keeping mechanism which shows an approximate 24-h (circadian) cycle at constant environmental conditions.

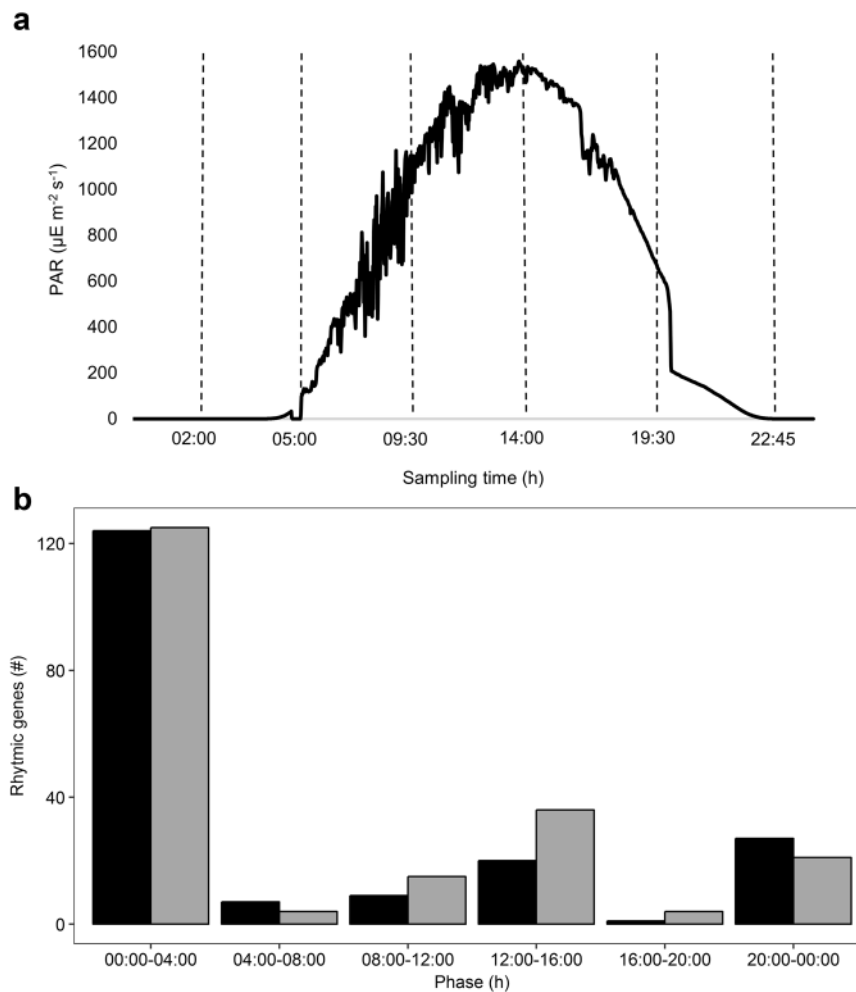
The cyanobacterial circadian clock is encoded by the gene cluster *kaiABC* and the corresponding proteins regulate the transcription of genes involved in various processes such as photosynthesis and nitrogen fixation.<sup>11,12</sup> The phosphorylation state of KaiC acts as a gate keeper to the transcription machinery to which it sends scheduled cues during its 24-h phosphorylation/dephosphorylation cycle. The rhythmic autokinase activity of the ATPase KaiC enhances through binding of KaiA, while KaiC's autophosphatase activity is modulated by KaiB. Entrainment (reset) and, to a certain degree, the output (oscillation period) of

the clock is regulated through a bacteriophytochrome, the circadian input kinase A (CikA).<sup>13,14</sup> The second abundant group of circadian clock controlled photosynthetic primary producers consists mainly of diatoms.<sup>15</sup> However, the underlying mechanism of their circadian clock and its genetic components are not yet fully understood,<sup>15</sup> and will not be further discussed here.

Circadian clocks, which are ubiquitous among Eukarya, were thought to have evolved outside of this domain only in Cyanobacteria.<sup>16–18</sup> However, genome analysis revealed homologs of *kaiB* and *kaiC* but not *kaiA*<sup>17</sup> in several other bacterial phyla such as Proteobacteria, Bacteroidetes, Chloroflexi and also in Archaea.<sup>17,19</sup> Analysis of rhythmicity in gene expression in the Alphaproteobacteria *Rhodospseudomonas palustris*<sup>20</sup> and *Rhodobacter sphaeroides*,<sup>21</sup> revealed similar regulatory patterns as in Cyanobacteria lacking *kaiA* such as in *Prochlorococcus marinus*.<sup>22</sup> The presence of *kaiBC* in both of these Alphaproteobacteria led to higher growth rates and increased fitness under a 24-h light-dark cycle but not under continuous light or dark conditions. This KaiBC based system has been proposed to be part of a 'proto-circadian' oscillator that does not persist under constant conditions<sup>20</sup> and does not contain other defining circadian characteristics, such as temperature compensation and entrainment.<sup>22</sup> This suggests that organisms possessing only *kaiB* and/or *kaiC* are limited to hourglass timekeeping (determination of intervals).<sup>23,24</sup> Potential temperature compensated, endogenous circadian rhythms in gene expression have been observed in the human gut bacterium *Enterobacter aerogenes*<sup>25</sup> and may be entrained by its host's clock-driven signal. Alternatively, the ubiquitous antioxidant peroxiredoxin (PRX) is highly conserved in all domains of life and has been proposed to function as an alternative, universal circadian

<sup>1</sup>Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research, and Utrecht University, Den Hoorn, The Netherlands and <sup>2</sup>Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands  
Correspondence: Henk Bolhuis (henk.bolhuis@nioz.nl)

Received: 9 November 2017 Revised: 26 March 2018 Accepted: 17 April 2018  
Published online: 15 May 2018



**Fig. 1** Photosynthetically active radiation during the sampling period (A) and peak expression phases of rhythmic CGTs in the metatranscriptome (b). **a** Photosynthetically active radiation (PAR) ( $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) measured every minute during the sampling period on the beach of Schiermonnikoog ( $N = 1429$ ). Stippled lines indicate sampling time of metatranscriptomes ( $N = 6$ ). **b** Clustered bar chart plots the peak expression  $F(h)$  against the count of rhythmic protein-coding CGTs. CGT-F (forward reads) = black, CGT-R (reverse reads) = gray. The peak expression phase displays the time frame in which gene expression reaches maximum

oscillator through  $\sim 24$  h cycles of protein oxidation–reduction.<sup>26,27</sup> PRX reacts to a rhythm generation of stress-induced reactive oxygen species (ROS) that are formed in reaction to photosynthesis during the day and to the reducing-, and anoxic conditions during the night. The central timekeeper of this circadian oscillator is, however, hitherto unknown, but it is assumed to be a basic and most ancient circadian timer in all organisms.<sup>27,28</sup> Multiple indications of potential endogenous rhythmicity inducers/ timekeepers in bacterial and archaeal phyla<sup>20,21,27,29</sup> led us to the following fundamental research question: Is rhythmic gene expression in phototrophic microbial mats limited to Cyanobacteria and algae or do other bacteria also exhibit a circadian-like, daily rhythmicity? We used meta-transcriptomics and qPCR to analyze diel rhythmicity in gene expression in six microbial mat samples that were taken at regular time points during a 24-h period.

## RESULTS

Photosynthetically active radiation (PAR) during the full sampling period is presented in Fig. 1a. To cover major changes in light intensity (photon density), two samples were taken at night (02:00 and 22:45 h:  $0 \mu\text{mol m}^{-2} \text{s}^{-1}$ ), two at twilight (05:00 and 19:30 h: 2.97 and  $169.2 \mu\text{mol m}^{-2} \text{s}^{-1}$ ) and two during the day (09:30 and

14:00 h: 1021 and  $1528 \mu\text{mol m}^{-2} \text{s}^{-1}$ ). The microbial mat was not submersed by seawater during the entire sampling period. Metatranscriptomic analysis of the six mat samples generated between 34,016,641 and 47,247,273 paired-end reads per sample with a total of 247,350,938 high quality reads (Phred:  $\sim 38$ ) (Table 1).

### Gene transcript clustering, rhythmicity and annotation

Clustering the complete transcriptome dataset in reads of  $\geq 96\%$  sequence identity yielded 24,035 conserved gene transcripts (CGTs), representing 168,239,136 reads (68% of the total dataset) with  $\sim 10$ – $16$  million reads per sample for the datasets with the forward and reverse reads. Approximately 32% of the reads were either singletons or clustered with less than 120 reads in total and were discarded from further analysis. The remaining datasets were analyzed with the R package metaCycle<sup>30</sup> to identify periodicity within gene expression profiles. The forward and reverse reads were treated as separate datasets. For the dataset with the forward reads (CGT-F), metaCycle analysis revealed 1875 significant ( $p \leq 0.05$ ) rhythmic CGTs representing 5,801,407 reads, 3.4% of the clustered reads and 2.3% of the initial number of reads. Of these rhythmic CGT-Fs, 188 were protein coding. For the reverse dataset (CGT-R), metaCycle analysis revealed 1526 significantly rhythmic CGTs that in total represent 3,467,908 reads,  $\sim 2.0\%$  of the

**Table 1.** General (MG-RAST) statistics of the 6 metatranscriptomes

MG-RAST ID	Time of sampling (h)	Uploaded sequences (#)	Validated reads (%)	rRNA features (%)	Annotated protein features (%)
4611784.3	02:00	40,792,897	39	78	9
4611785.3	05:00	34,016,641	41	85	6
4604210.3	09:30	41,937,839	33	79	9
4604211.3	14:00	44,490,594	30	82	8
4604212.3	19:30	38,865,694	34	81	8
4604213.3	22:45	47,247,273	35	76	10

**Table 2.** Phylogenetic distribution of rhythmic, protein-coding CGTs (#)

Domain	Phyla	Rhythmic CGT (#)
Bacteria	Cyanobacteria	42
	Proteobacteria	8
	Firmicutes	4
	Fusobacteria	2
	Planctomycetes	0
	Verrucomicrobia	0
	Bacteroidetes	0
	Eukarya	Stramenopiles
Opisthokonta	28	
Viridiplantae	25	
Rhodophyta	18	
Alveolata	5	
Euglenozoa	2	
Haptophyceae	2	

**Table 3.** Functional annotation of rhythmic, protein-coding CGTs (F + R)

Function	Rhythmic genes (#)
RuBisCO	113
Hypothetical protein	90
Photosystem II	88
Photosystem I	23
tmRNA	17
Chloroplast	14
Transposase	6
ATP synthase	4
COI	4
ITS	4
LSU (rRNA)	4
arfA/Rf2	3
Mitochondrial protein	4
Nitrogen fixation related gene cluster	3
Cytochrome b	2
DNA starvation protein	2
Rnase P subunit RnpB	2
Cold shock protein	1
cox2	1
CRISPR	1
Ferritin	1
High light inducible protein	1
Lipoprotein	1
ncRNA	1
Phycocyanin	1
Plasmid	1
tRNA	1

clustered reads and ~1.4% of the initial number of reads. 205 rhythmic CGT-Rs were protein coding. For each rhythmic gene predicted by metaCycle also the phase was calculated, the time point at which the fitted curve indicates maximal expression. These maxima are not discrete and not necessarily identical to the sampling time points and were therefore grouped around the sampling time points in order to distinguish day (8:00–12:00 and 12:00–16:00) from night (20:00–00:00 and 00:00–4:00) and dusk (04:00–8:00) from dawn (16:00–20:00). The highest number of protein-coding CGTs had their maximum expression between 00:00 h midnight and 04:00 h in the morning (Fig. 1b) followed by a smaller group of CGTs that had their maximal expression between 12:00 and 16:00 h.

Phylogenetic annotation of rhythmic CGTs revealed approximately one third without a blast hit against the NCBI NR (protein) or NT (nucleotide) database. Reads with a blast hit were divided into ribosomal RNA (rRNA)-coding hits (mostly of micro-eukaryote origin), hits to whole genomes without specific annotation, and hits to mainly protein-coding genes. The number of protein-coding CGTs for the forward and the reverse datasets varied respectively between 66 and 138 (188 unique CGTs in total) and between 142 and 116 (205 unique CGTs in total) when blasted against the protein or nucleotide dataset, respectively. The majority of protein-coding genes in each dataset was derived from algae belonging to the phyla of Stramenopiles, Opisthokonta and Viridiplantae, while approximately one fourth was attributed to Bacteria, mainly Cyanobacteria and Proteobacteria (Table 2). Functional annotation revealed several genes encoding the large (*rbcL*) and small (*rbcS*) subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase (the CO<sub>2</sub>-fixing enzyme RuBisCO) (Table

3) that were maximally expressed between 00:00 and 04:00 h (Table S1). Photosystem I (e.g., *psaA*) and II (e.g., *psbA*) genes were abundant amongst the protein annotated rhythmic reads (Table 3 and S1) and were maximally expressed between 00:00 and 04:00 h. Several hypothetical protein coding genes had their maximal expression either at night or in the afternoon.

#### MG-RAST annotation and rhythmicity

One third of the raw reads submitted to the MG-RAST metagenome server matched to their pair and passed the quality control which includes having a minimum size requirement of 75 nt for analysis. On average 80% (4,011,978) of these reads were annotated as rRNA, ~8% (428,087) were identified as protein-coding with known functions while for ~11% (561,507) no

**Table 4.** (A) Number of total rhythmic genes of forward (F) and reverse (R) reads of protein-coding CGTs and (B) of dataset B<sup>C</sup>, P, Bs, C and Eu. (C) Number of rhythmic genes obtained by recruitment analysis of *L. aestuarii* PCC8106, *C. chthonoplastes* CCY9604, *C. litoralis* KT71, *R. denitrificans* Och114 and *A. vinosum* DSM180. Displayed numbers were filtered for 20–24 h cycle periods and  $p \leq 0.05/0.01$ .  $P = B^C$  and  $Bs = B^C$  displays amount of shared rhythmic genes between the P/Bs and B<sup>C</sup> dataset. The fractions of the total CGTs/genes are given

Database/ reference genome	Rhythmic genes (#) $p \leq 0.05$	Rhythmic genes (#) $p \leq 0.01$	% of initial CGTs/genes
A CGT-F <sup>a</sup>	188	79	0.78
CGT-R <sup>a</sup>	205	8	0.85
B B <sup>C</sup>	265	115	8.5
P	220	109	7.1
Bs	119	47	8.2
C	91	32	5.9
Eu	36	13	5.4
$P = B^C$	137		
$Bs = B^C$	48		
C <i>L. aestuarii</i>	613	203	10
<i>C. chthonoplastes</i>	105	23	1.6
<i>R. denitrificans</i>	81	5	9
<i>A. vinosum</i>	44	2	2
<i>C. litoralis</i>	35	4	5.3

<sup>a</sup>Protein-coding

function could be assigned. The majority of rRNA reads was of eukaryal origin (96%) leaving only 3% of bacterial rRNA reads confirming the efficient bacterial rRNA removal. Of the protein-coding genes, 69% was of bacterial and 31% of eukaryal origin. The bacterial protein-coding reads were assigned to Cyanobacteria, Proteobacteria, and Bacteroidetes while the eukaryal part belonged mainly to Bacillariophyceae.

To get a better insight in the taxonomic distribution of rhythmically expressed genes, the genes and their normalized abundance were assigned by the MG-RAST analysis pipeline and divided in the following taxonomic subsets: B<sup>C</sup> (bacterial reads excluding cyanobacterial reads), P (proteobacterial reads), Bs (bacteroidetal reads), C (cyanobacterial reads) and Eu (eukaryal reads). The highest number of rhythmic genes (RGs) was found in the bacterial dataset B<sup>C</sup> (265 RGs), followed by datasets P (220 RGs), Bs (119 RGs), C (91 RGs) and Eu (36 RGs) (Table 4, Figure S1). About 39% of rhythmic genes found in dataset B<sup>C</sup> also occurred in P whereas 12% matched RGs found in Bs. The combined datasets, revealed between 60 and 80% of the rhythmic genes maximally expressed during the early morning between 00:00 and 04:00 h (210 (B<sup>C</sup>); 166 (P); 94 (Bs); 27 (C); 23 (Eu) RGs) (Fig. 2a, Table S3). The rhythmic genes of dataset B<sup>C</sup> and P that contributed to the highest peak (Fig. 2a) encoded mainly proteins involved in vitamin and pigment biosynthesis (e.g., tetrapyrrole), clustering-based subsystem proteins (i.e., proteins with confirmed functional coupling but of unknown function) and carbohydrate turn-over (e.g., fermentation (B<sup>C</sup>), CO<sub>2</sub> fixation (B<sup>C</sup>, P) and (poly-) saccharide utilization (B<sup>C</sup>, P) (Table S3). A major part of Bs-derived rhythmic genes, expressed during the early morning, encoded proteins involved in clustering-based subsystems and protein metabolism (e.g., ribosomal large subunit (LSU) and small subunit (SSU)).

Thirty percent of the cyanobacterial RGs (29 RGs) was maximally expressed between 00:00 and 04:00 h of which almost half (12 RGs) of the encoded proteins annotated as clustering-based

subsystem proteins or as protein- and DNA metabolism proteins (Table S3). 25% (23 RGs) were maximally expressed between 12:00 and 16:00 h and were mostly annotated as clustering-based subsystem proteins and vitamin and pigment biosynthesis proteins (e.g., tetrapyrrole). The eukaryal rhythmic genes, which were most highly expressed in the early morning, mainly encoded proteins involved in protein metabolism, photosynthesis and respiration (Table S3).

#### Rhythmicity in recruited genomes

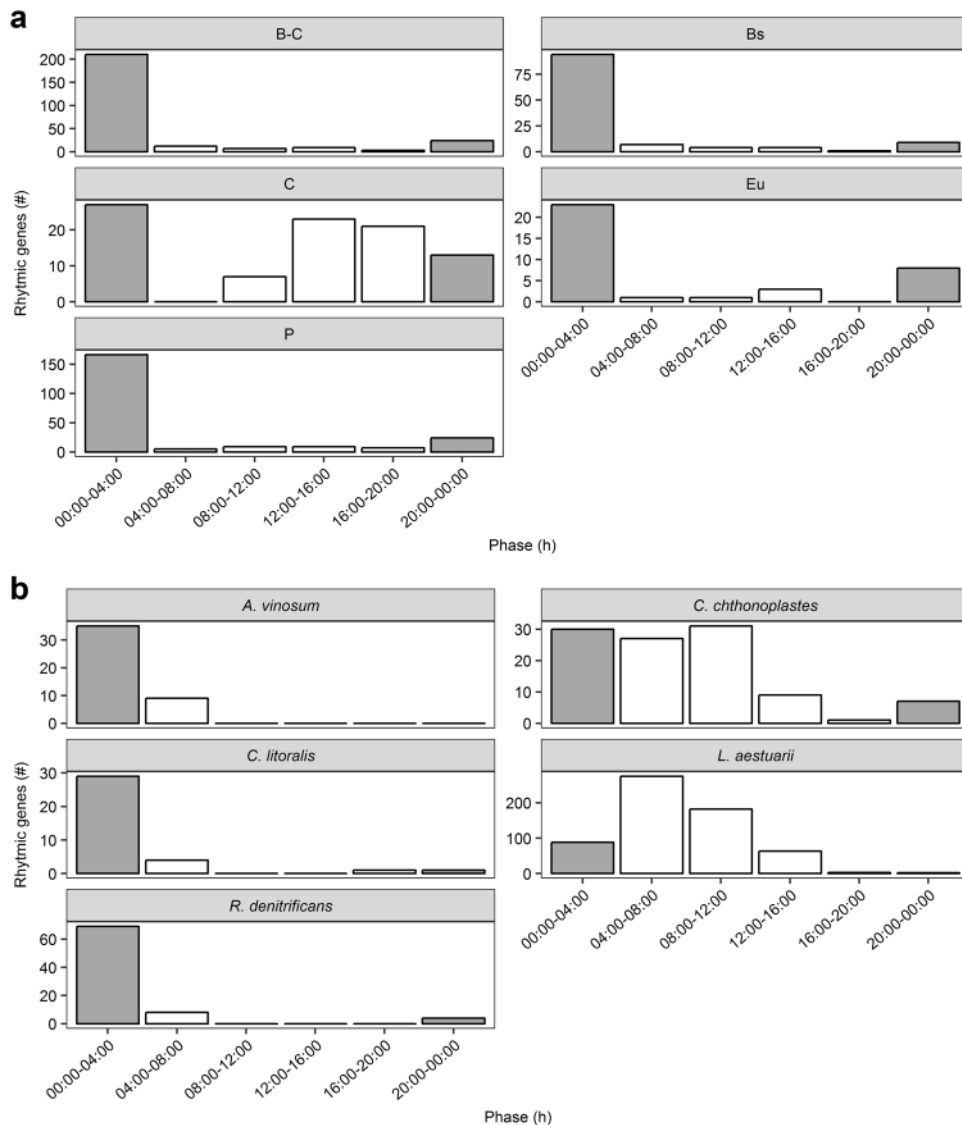
Figure 2b depicts the distribution of the peak expression of rhythmic genes, recruited using five different reference genomes. Recruitment against the genome of the cyanobacterium *Lyngbya aestuarii* PCC8106 (synonym = *Lyngbya* sp. CCY9616) revealed 613 rhythmic genes out of a total of 5811 coding sequences (CDS) (Table 4). The vast majority of rhythmic genes, 275 and 182, revealed peak expressions between 04:00 and 08:00 h, and between 08:00 and 12:00 h respectively (Fig. 2b). The genes encoded mostly for proteins involved in carbohydrate turn-over, amino acid metabolism, protein biosynthesis and biodegradation, photosynthesis (i.e., photosystem II proteins *psbCDJX*) and respiration (between 04:00 and 08:00 h only) (Table S3). Recruitment to the genome of the cyanobacterium *Coleofasciculus (Microcoleus) chthonoplastes* PCC7420 (CCY9604) revealed 105 rhythmic genes out of 8109 CDS (Table 4), the majority of which revealed highest expression between 00:00 and 04:00 h (30 RGs) and between 08:00 and 12:00 h (31 RGs). More than 50% of the rhythmic genes encoded hypothetical proteins while others were involved in nitrogen fixation (00:00h–04:00 h) or stress response (00:00h–04:00, 08:00h–12:00 h).

Recruitment against the genomes of three proteobacterial species, the alphaproteobacterium *Roseobacter denitrificans* Och114 and the gammaproteobacteria *Allochrochromatium vinosum* DSM180 and *Congregibacter litoralis* KT71, revealed the highest number of rhythmic genes transcribed during the night between 00:00 and 04:00 h (Fig. 2b), while no rhythmic genes were detected with peak expression between 12:00 and 20:00 h (*R. denitrificans*, *A. vinosum*) or 12:00 and 16:00 h (*C. litoralis*). As presented in Table 4 and Table S3, *R. denitrificans* revealed 81 rhythmic genes out of 4007 CDS, encoding proteins involved in protein metabolism, carbohydrate turn-over and photosynthesis. *A. vinosum* revealed 44 rhythmic genes out of 3883 CDS, encoding proteins involved in amino acid biosynthesis, carbohydrate turn-over and respiration. *C. litoralis* revealed 35 rhythmic genes out of 3220 CDS, encoding predominantly proteins involved in protein metabolism, photosynthesis and stress response.

An attempt to recruit specimen of Bacteroidetes failed, due to an insufficient amount of sequences.

#### Nitrogenase activity, *nifH*, and *psbA* expression

Circadian control of a number of cyanobacterial genes is well established and includes genes involved in nitrogen fixation and photosynthesis. Total nitrogenase activity (acetylene reduction) in natural microbial mat samples, derived from the same sampling site as the metatranscriptomes, is depicted in Fig. 3a. The overall activity was low during the night (21:30–3:30 h) reached maximum in the morning (9:30 h) becoming minimal at dusk (21:30 h). The amount of chlorophyll *a*-normalized ethylene production varied between 1.1 and 5.9  $\mu\text{mol mg}^{-1} \text{h}^{-1}$ . In order to quantify *nifH* and *psbA* transcripts from *L. aestuarii* PCC8106 in the microbial mat samples we estimated their relative expression obtained by normalization against two housekeeping genes (HKG), by means of reverse transcriptase quantitative PCR (RT-qPCR). Validation of *rnpA* and *ppc*, using BestKeeper, ranked *ppc* ( $r = 0.914$ ,  $p = 0.001$ ) before *rnpA* ( $r = 0.843$ ,  $p = 0.001$ ) as the most stable and evaluated both as suitable HKGs (Table S2). Relative expression of *nifH* from *L. aestuarii* was highest during the night



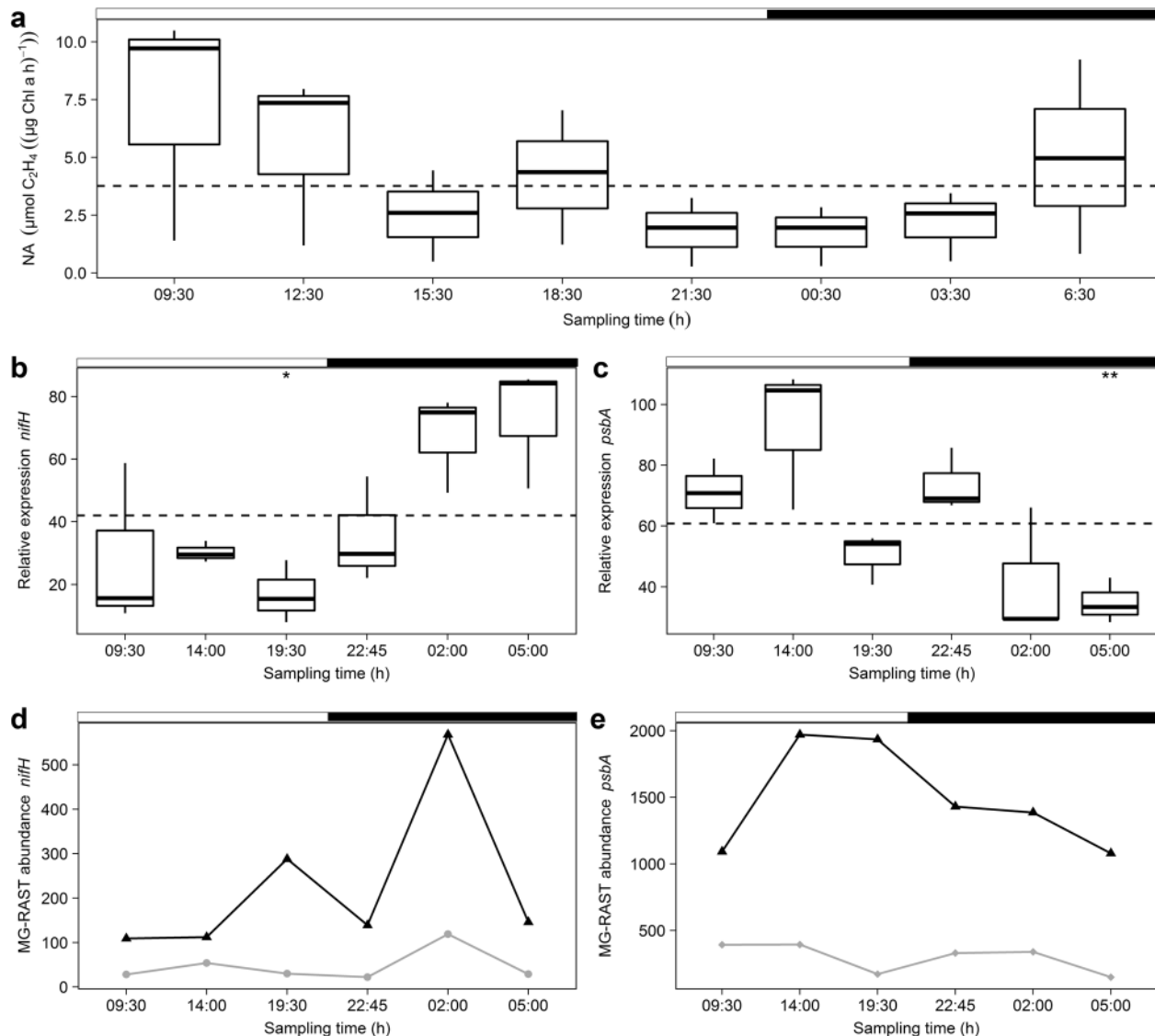
**Fig. 2** **a, b** Peak expression phase distribution (metaCycle) of the rhythmic genes of the MG-RAST datasets (**a**) and of the recruited genomes (**b**). Clustered bar charts in **a** and **b** plot the peak expression phase (h) distribution of rhythmic genes against their count (#). Bar coloration indicates light (white) and dark (gray) sampling times

and early morning (02:00, 05:00 h) (Fig. 3b). At 19:30 h *nifH* was significantly lowest expressed ( $p = 0.0076$ ) (Fig. 3b). The program metaCycle estimated the relative expression of *nifH* of *L. aestuarii* as significantly rhythmic ( $p = 0.037$ ) and predicted its peak expression at ~03:00 h (Table S4). MG-RAST annotated cyanobacterial *nifH* transcripts were derived from unicellular Cyanobacteria of the order Chroococcales (mainly *Cyanothece*) (~48.5%) and filamentous Cyanobacteria of the order Oscillatoriales (*Trichodesmium*) (~27%) (Table S5). The majority of proteobacterial annotated *nifH* transcripts was derived from Deltaproteobacteria (i.e., *Geobacter*, *Pelobacter*, *Rhodospirillum*) (~10.5%) (Table S5). The pooled MG-RAST *nifH* expression was determined as highly significantly rhythmic ( $p = 0.003$ ) with an estimated maximum expression around 12:00 h (Table S4). The MG-RAST derived based dataset displayed a maximum in *nifH* expression dissimilar to the observed maximum in nitrogenase activity with a major peak in the night (02:00 h) and a second minor peak in the evening (19:30 h) (Fig. 3d).

The majority of the MG-RAST annotated *psbA* transcripts was derived from the cyanobacterial order Chroococcales (*Cyanothece*, *Synechococcus*) (~59%) while ~21% originated from the red algae

order Bangiales and specifically the genus *Porphyra* (~14.6%) (Table S5). Both *L. aestuarii* recruited *psbA* transcripts and the MG-RAST annotated *psbA* transcripts reached a maximum expression at 14:00 h (Fig. 3c,e), while being significantly lowest expressed ( $p = 0.0038$ ) in *L. aestuarii* at 05:00 h (Fig. 3c).

Expression of circadian clock (related) genes and peroxiredoxin  
In order to focus on the circadian clock genes, the relative expression of the cyanobacterial circadian clock core genes *kaiA*, *kaiB*, *kaiC* and of the circadian input kinase *cikA* were followed by qPCR and compared to the MG-RAST annotated homologs. In addition, the expression profiles from the conserved circadian marker peroxiredoxin (*prx*) were also investigated. Relative expression was obtained by normalizing target gene transcript concentrations (copies/ $\mu$ l) to the geometric mean of the house-keeping gene transcript concentrations. Figure 4 shows the relative expression of *L. aestuarii* PCC8106 *kaiA*, *kaiB*, *kaiC*, *cikA* and *prx* (Fig. 4a) and their MG-RAST abundance in the metatranscriptomes (Fig. 4b). Relative expression of *L. aestuarii* *kaiA*, *kaiB* and *kaiC* peaked in the morning (9:30 h) and declined in



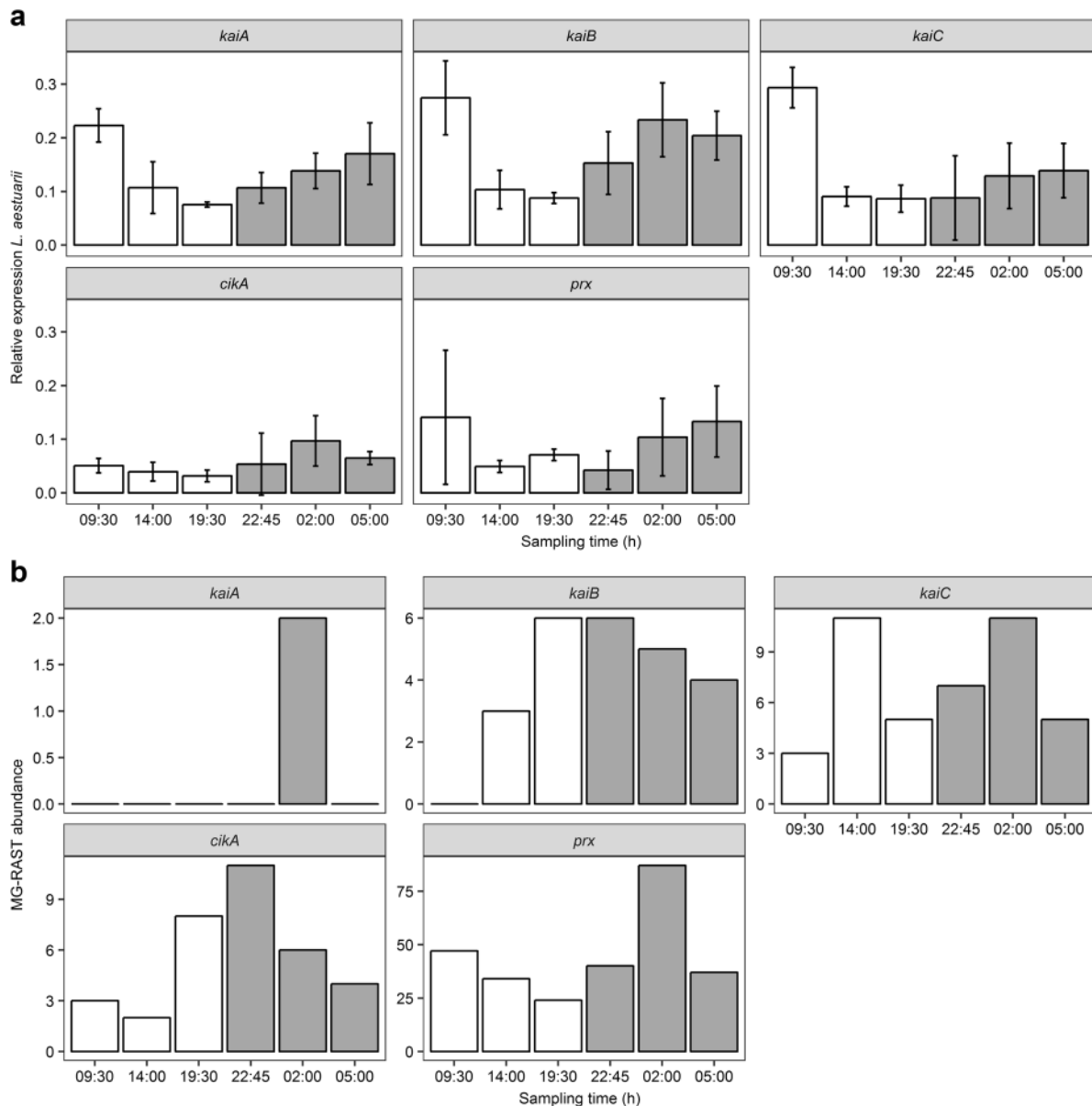
**Fig. 3** **a–e** Nitrogenase activity (**a**), *nifH* expression profiles (**b**, **d**) and *psbA* expression profiles (**c**, **e**) of MG-RAST and qPCR. Boxplots display (**a**) the nitrogenase activity (NA) ( $\mu\text{mol C}_2\text{H}_4 (\mu\text{g Chl a h}^{-1})$ ) and relative expression (qPCR) of (**b**) *nifH* and (**c**) *psbA* in *L. aestuarii* of which the biological triplicates were taken at 8 (NA) and 6 (relative expression) time points within 24 h. ANOVA results are indicated by asterisks displaying significant differences ( $p < 0.05$ ) in gene expression between time points based on their deviation from the mean (dotted line). The line charts show the MG-RAST abundance of (**d**) *nifH* in Cyanobacteria (black triangle) and Proteobacteria (gray circle), and (**e**) *psbA* in Cyanobacteria (black triangle) and Bangiophyceae (gray diamond). White and black bars on top of charts indicate light and dark periods. Error bars display standard deviations (SD) of biological triplicates

the evening (19:30 or 22:45 h (*kaiC*)) (Fig. 4a). The highest relative expression of *L. aestuarii cika* was at night (2:00 h) and the lowest in the evening (19:30 h) (Fig. 4a). *L. aestuarii kaiA* and *cika* expression profiles were significantly rhythmic ( $p = 0.03/0.01$ ) and had their peak expression between 00:00 and 04:00 h and 04:00 and 08:00 h respectively (Table S4). The small pool of MG-RAST annotated *kaiA* transcripts (MG-RAST abundance = 2) of the metatranscriptomes was only detected in the night (2:00 h), while *kaiB* transcripts (MG-RAST abundance = 24) were most abundant in the early evening (19:30 h) and at dusk (22:45 h) (Fig. 4a). Transcripts of *kaiC* (MG-RAST abundance = 42) peaked at noon (14:00 h) and at night (2:00 h) (Fig. 4b) while *cika* transcripts (MG-RAST abundance = 34) displayed the highest abundance in the beginning of the night (22:45 h).

The majority of *kaiA* sequences in the MG-RAST dataset was derived from *Cyanothece*, while *kaiB* transcripts were derived from the cyanobacterial orders Chroococcales (*Cyanothece*,

*Synechococcus*) (~42%), Nostocales (*Nostoc*) (~27%) and Oscillatoriales (*Trichodesmium*) (~23%) (Table S4). The majority of the *kaiC* transcripts also originated from Chroococcales (*Chroocosphaera*, *Cyanothece*, *Synechococcus*) (~56%) and Oscillatoriales (*Trichodesmium*) (~22%) (Table S4). Although the majority of clock genes was derived from Cyanobacteria, about 2–3% of the *kaiB* and *kaiC* transcripts was derived from archaeal Methanomicrobiales (*kaiB* and *kaiC*) and from the bacterial orders Flavobacteriales (*kaiB*) and Myxococcales (*kaiC*) (Table S4). The majority of *cika* transcripts was also of cyanobacterial origin, namely of the orders Chroococcales (*Cyanothece*) (~64%) and Nostocales (*Nostoc*) (~28%). However, ~6% of the *cika* annotated genes was derived from Rhizobiales (*Rhodopseudomonas*, *Methylobacterium*) (Table S5).

The expression of *L. aestuarii prx* gene peaked in the morning (5:00, 9:30 h) and dropped at dusk (22:45 h) (Fig. 4a). MG-RAST annotated *prx* transcripts were traced back to mainly Bacteroidetes of the order Flavobacteriales (*Zunongwangia*) (~87% of



**Fig. 4** Expression profiles of *cika*, *kaiABC* and *prx* of *L. aestuarii* PCC8106 (CCY9616) (**a**) and of the MG-RAST metatranscriptomes (**b**). **a** Bars show the relative expression (normalized to the housekeeping genes *rnpA* and *ppc*) of the *L. aestuarii* circadian clock core genes *kaiABC*, the circadian input kinase *cikA* and peroxiredoxin (*prx*) and (**b**) the MG-RAST abundance of the summed expression of the aforementioned genes. X-axes show the 6 sampling times within the 24 h period while the white and gray coloration of bars symbolize light and dark periods. Error bars display standard deviations (SD) of biological triplicates

transcripts), of which peak expressions preceded those of *L. aestuarii* (Fig. 4a) by about 3–4 h. The *prx* expression profiles of the *L. aestuarii* qPCR, recruitment and MG-RAST data were not significantly rhythmic.

## DISCUSSION

Many organisms synchronize their metabolism with the Earth's rotation cycle. However, apart from the domain Eukarya, it was thought that only Cyanobacteria possess a circadian clock. However, it is difficult to understand why also not other Bacteria or Archaea developed a time-keeping mechanism, while this is obviously an important property. The occurrence of key cyanobacterial circadian clock genes in other Bacteria<sup>17,19</sup> and the occurrence of the circadian peroxiredoxin redox-cycles,<sup>27</sup> may hint to a circadian-like rhythmicity in Bacteria other than Cyanobacteria

as well as in Archaea. To resolve this question, this study used a combined metatranscriptomics- and qPCR based gene quantitation approach to study the daily rhythm in gene expression in a microbial mat community.

Previous metagenomic analysis of coastal microbial mats revealed a low percentage (1–5%) of reads originating from Eukarya.<sup>31–33</sup> As expected, we found rhythmic genes belonging to Cyanobacteria as well as those belonging to diatoms (Bacillariophyceae). Similar to Cyanobacteria and multicellular Eukarya, circadian control of metabolic processes in oxygenic photosynthetic micro-eukaryotes may provide a fitness advantage when anticipating the day/night regime.<sup>34–36</sup> The mechanism of the circadian clock of diatoms is poorly understood and, hence, the identification of potential clock genes has been proven difficult.<sup>37,38</sup> Rhythmic gene expression in diatoms may also be influenced by the tides,<sup>7,39</sup> which would influence endogenously

induced cycles. Nevertheless, the substantial contribution of diatoms to the number of rhythmic CGTs, which was also reflected in the MG-RAST analyzed rRNA pool, showed that these micro-eukaryotes played an important role in the observed daily rhythmicity of gene expression in the microbial mat.

The majority of rhythmic genes in algae and Cyanobacteria encoded proteins involved in CO<sub>2</sub> fixation and oxygenic photosynthesis, two processes that are well-known to be under circadian clock control.<sup>40–42</sup> The RuBisCO genes *rbcl* and *rbcS* and the photosystem II-coding gene *psbA* dominated the rhythmic CGT dataset, of which the majority was derived from oxygenic photoautotrophs. Expression of the RuBisCO genes and *psbA* was maximal at low light, between midnight and early morning in accordance with laboratory studies of cultures of *Synechococcus elongatus*.<sup>43</sup> Nitrogen fixation in non-heterocystous Cyanobacteria is often confined to the dark or to low light conditions<sup>44,45</sup> and this was also found in our study, concomitantly with *nifH* expression in *L. aestuarii*. *nifH* expression profiles of MG-RAST and of *L. aestuarii* were both significantly rhythmic. Unexpectedly, expression of *psbA* in the MG-RAST dataset and recruitment- and qPCR data was not significantly rhythmic. However, arrhythmically expressed genes can still lead to rhythmic protein expression (e.g., *rpoD4*)<sup>43</sup> and vice versa (e.g., *kaiA*) due to post-transcriptional control.<sup>46</sup>

Rhythmic analysis uncovered 2–10% rhythmic genes in *C. chthonoplastes* and *L. aestuarii*, a number far lower than found in the promoter trap studies in *Synechococcus elongatus* that predicted global circadian control of nearly all genes investigated.<sup>47</sup> However, it must be taken into account that results of rhythmicity obtained from homogeneously grown monocultures in laboratory experiments may not necessarily reflect what happens in a multispecies heterogeneous environment such as the microbial mat. However, a microarray study of *Synechocystis* sp. PCC 6803 also showed only up to 9% of the genes under circadian clock control,<sup>48</sup> which is in line with our study. A discrepancy between the amount of rhythmic genes based on mRNA vs. promoter studies was also seen in the eukaryote *Arabidopsis*,<sup>49,50</sup> and was attributed to post-transcriptional control mechanisms that keep the abundance of some mRNAs constant thereby counteracting their rhythmic transcriptional activity.<sup>51</sup> Even at constant transcription and protein expression rates, the actual activity can be rhythmically regulated as was shown for RuBisCO expression in the marine alga *Gonyaulax* sp.<sup>52</sup> Here RuBisCO activity depended on the actual distribution of the enzyme within the chloroplast and was causally related to the rhythmic fixation of carbon dioxide.

Another essential discovery made in our study is that rhythmic gene expression was not restricted to Cyanobacteria and micro-eukaryotes. Notably, more than 50% of the rhythmic genes were derived from Proteobacteria putting this phylum in the spotlight for future studies of daily rhythms. Key to the (genome-wide) control of metabolic processes by a daily rhythm is the presence of potential Zeitgeber.<sup>20,53</sup> In microbial mats, rhythmic control mechanisms such as the circadian clock genes *kaiABC* and the universal oscillator *prx* were not confined to Cyanobacteria. Approximately 2 to 4% of *kaiB* and *kaiC* transcripts in the MG-RAST dataset were assigned to Bacteroidetes, Proteobacteria and to Euryarchaeota. Furthermore, the majority of *prx* transcripts was derived from Bacteroidetes, Proteobacteria, and to a lesser extent Bacteroidetes, expressed several rhythmic genes with a periodicity of approximately 24 h. A similar periodicity was found in metatranscriptome studies of marine microbial communities where functional groups of microorganisms other than Cyanobacteria revealed significant differences in gene expression between day and night<sup>54</sup> and in daily transcript cycles.<sup>22</sup>

In these marine microbial community studies, a large fraction of the cyclic transcripts was derived from *Roseobacter*, a genus belonging to the Alphaproteobacteria. Our recruitment analysis

also yielded twice as many rhythmic genes for *Roseobacter denitrificans* when compared to *C. litoralis* or to *A. vinosum*. Furthermore, *R. denitrificans* displayed a higher fraction of rhythmic genes in its genome than the cyanobacterium *C. chthonoplastes*. The rhythmic transcription of genes involved in carbohydrate turnover, fermentation and polysaccharide utilization during the night in organisms other than Cyanobacteria may hint to the tight coupling of gene transcription to the availability of substrates produced by the phototrophic organisms during the day. A similar tight coupling of gene expression was observed for animal gut microbiomes that displayed daily rhythmicity of gene expression in response to the host feeding schedule.<sup>25,55</sup> Gasol et al.<sup>56</sup> suggested that a day–night cycle in Bacteria implies a tight coupling of the photosynthetically produced dissolved organic carbon and its subsequent consumption. This is in agreement with our results and also confirms the results of Poretsky et al.<sup>54</sup> and Ottesen et al.<sup>22</sup>

Several phototrophic microorganisms such as diatoms and Cyanobacteria show diel vertical migration patterns in microbial mats. Migration to deeper, darker and low oxygen parts of a microbial mat protect cells against high UV radiation and inhibiting oxygen concentrations during the day and allow Cyanobacteria to migrate up to 200 µm in a mat<sup>57,58</sup> while at night sulfide may accumulate to toxic concentrations result in a migration back to the surface.<sup>3</sup> Migration patterns in diatoms in marine sediment are linked to both the diel light cycle and to the tidal cycle with an initial endogenous controlled mechanism followed by a physiological response.<sup>59</sup> In Cyanobacteria, it has not yet been established whether tidal and circadian control contribute to the cyanobacterial migration,<sup>60</sup> or alternatively whether migration through the different micro-gradients in the mats itself affect gene expression. Our analysis revealed rhythmic expressed flagellar protein coding genes in Proteobacteria (*flgG*, *flaA*), this may suggest rhythmic motility but not necessarily vertical migration while benthic Cyanobacteria and diatoms lack flagella and migrate through gliding motility, which speed and direction is mainly controlled by light.<sup>61,62</sup> The sampling depth of ~5 mm as was applied in this study covers the effect of migration.

Manipulation of the natural microbial mat ecosystem would have led to complex unknown environmental effects, which would have made the interpretation difficult or even impossible. Therefore, we decided not to study gene expression while putting the microbial mat under constant dark or light conditions (the latter would have raised questions about the light source and -intensity and would have been impractical in the field). Hence, we were unable to determine with certainty that the observed rhythmicity in the microbial mat was caused by an endogenous Zeitgeber rather than by environmental cues. However, our experiments confirm the expression of cyanobacterial *kai*-genes, bacterial *kai*-homologs (e.g., *A. vinosum*) and *prx* genes (e.g., *R. denitrificans* and *C. litoralis*) and indicate the presence of potentially circadian clock controlled organisms that may have contributed to the observed rhythmicity in gene expression.

The observation of community-wide rhythmic gene expression patterns in a coastal microbial mat led us to propose the Choirmaster-Choir theory. This supposes that rhythmic gene expression in microorganisms possessing a circadian clock induces rhythmic gene expression in other microbial mat members. The Choirmaster-Choir theory is derived from three sub-hypotheses: (i) only Cyanobacteria and algae have a fully functional molecular clock and direct their gene expression patterns to other mat members through the rhythmic release of photosynthate and other metabolites; (ii) other microbial mat members have clocks but are only entrained by the rhythmic release of metabolites from phototrophic microorganisms; (iii) other microbial mat members have their own clock that is entrained by a cocktail of Zeitgebers such as light, temperature and the rhythmic release of photosynthate and metabolites by their neighbors. Cyanobacteria



and algae are not the only organisms in a microbial mat that may possess a circadian clock. The wide spread occurrence of *kaiB* and *kaiC* homologs, the proposed universal circadian oscillator gene *prx*, and the experimental evidence for light-entrained rhythmic expression among several Proteobacteria is strong evidence for the presence of a circadian clock in other microorganisms.<sup>20,21</sup> Nevertheless, it is too early to conclude that all microbial mat members possess a clock. Many members of the microbial mat community are still unknown but the *prx* gene might be a good candidate universal oscillator. The nature of our experimental setup using a natural system under natural conditions, does not allow conclusive proof that light and temperature were the Zeitgebers. However, it is nevertheless likely that these factors entrained the daily gene expression in several mat members as the Cyanobacteria, eukaryotic algae and the Alphaproteobacteria.<sup>20,21</sup> Rhythmic release of organic molecules such as photosynthate (mainly polysaccharides) during the day and fermentation products such as low molecular weight organic acids (lactate, acetate and ethanol)<sup>63</sup> and nitrogen-containing organic compounds during the night<sup>64</sup> may serve as additional Zeitgebers. Regulation of gene expression by external stimuli and fluctuating concentrations of substrate is well documented in Bacteria,<sup>65,66</sup> while rhythmic synchronization through quorum-sensing, as was shown for *Vibrio fischeri*,<sup>67</sup> could also contribute to gene expression in this complex high diverse and biomass community.<sup>68</sup>

In conclusion, it was demonstrated that daily rhythmicity of gene expression in a coastal microbial mat surpasses the well-documented circadian clock controlled system of oxygenic phototrophs such as Cyanobacteria and algae. The microbial mat behaves as a well-coordinated consortium of microorganisms with tightly controlled metabolic networks that exhibit the characteristics of a single living entity. The Choirmaster-Choir theory, proposes that light and temperature induced daily rhythmicity in clock possessing phototrophic microorganisms (the choirmaster) is imposed on the microbial mat community as a whole (the choir) through the rhythmic release of metabolites.

## MATERIAL AND METHODS

### Sampling

Microbial mats from the mid-littoral zone (N 53.484240, E 6.134550) of the sandy beaches of the North Sea coast of the Dutch barrier island Schiermonnikoog were sampled on 6th and 7th of June 2013. Triplicate samples were taken at close proximity during a 24-h period at intervals of 4-h covering the natural light and dark period (Fig. 1a). A sterile plastic 10-ml syringe from which the top was removed was used as corer to collect the top 5 mm of the mat (~2 g wet weight). Samples were immediately transferred to 15-ml bead tubes (RNA PowerSoil Total RNA Isolation kit (MoBio, USA)) containing 4 ml RNA LifeGuard (MoBio, USA) to preserve RNA, thoroughly mixed, and instantly stored in a liquid nitrogen saturated dry shipper type HCl Cryogenics 34HC (Taylor-Wharton, USA). Upon return to the lab the samples were transferred to a -80 °C freezer and analyzed within 3 months. Nitrogenase activity (NA) of the natural mats was measured using the acetylene reduction assay (ARA).<sup>69,70</sup> Triplicate samples of the mat were taken using PVC cores (Ø 50 mm) and incubated with acetylene (10%) under ambient light and temperature, and ethylene production in the head space was measured every 4 h in triplicate according to Severin and Stal (2008). Throughout the sampling campaign, the local light intensity was measured using a LI-190 quantum PAR-sensor and a LI-1000 data logger (LI-COR, USA) (Fig. 1a).

### RNA extraction, sequencing and RT-qPCR

RNA was isolated from each sample using the RNA PowerSoil Total RNA Isolation kit (MoBio, USA) according to the manufacturer's protocol. Residual DNA was removed with TURBO™ DNase (Life Technologies, USA) following the manufacturers recommendations. DNase treated RNA was analyzed using a Bioanalyzer 2100 (Agilent Technologies, USA) and samples with RIN values  $\geq 6$  and total RNA concentration  $\geq 100$  ng/ $\mu$ l were shipped to BaseClear (Leiden, The Netherlands) for Illumina sequencing (TrueSEQ). BaseClear performed rRNA depletion (RIBO-Zero rRNA Removal Kit (bacteria)) (Epicenter, USA), barcoded cDNA library synthesis ('dUTP method'<sup>71,72</sup>) and paired end-sequencing (PE50) and delivered quality trimmed sequence reads. The remaining part of the DNase treated RNA samples was used to quantify the genes of interest (GOI) *kaiA*, *kaiB*, *kaiC*, *cikA*, *nifH*, *psbA* and *prx* of *Lyngbya aestuarii* PCC8106 (CCY9616) (ccy.nioz.nl) by RT-qPCR with species-specific primers and TaqMan probes (Table S2, M&M S1). RT-qPCRs were run with technical and biological triplicates. RT-qPCR protocol can be found in the supplementary material (M&M S1). Standard curves were used to estimate GOIs and HKGs concentrations in copies/ $\mu$ l in the original samples. The program BestKeeper<sup>73</sup> was used to validate the stability of HKGs. Relative expression values of genes were obtained by normalization against the housekeeping genes *mnpA* and *ppc* ( $\left(\frac{GOI}{\text{geomean}(HKGs)}\right)$ ). An analysis of variance (ANOVA) was performed to assess the GOIs' regulation.

### Analysis of metatranscriptomes

The metatranscriptomic reads were analyzed using two different approaches which resulted in two differently sized datasets. First, RNA reads were clustered into conserved gene transcripts (CGT-data set) of  $\geq 96\%$  identity allowing 2 nucleotide (nt) differences within the 50 nt reads using Vsearch v1.11.1<sup>74</sup> and transformed to a CGT table depicting the number of reads per CGT per time series. This was done separately for the forward and reverse reads to obtain technical replicates. Both CGT tables served as input for the metaCycle R package that performs statistical tests to identify rhythmic genes with a significant periodicity ( $p \leq 0.05$ ) of 20–24 h, reminiscent of a circadian clock. Subsequently, reads were annotated at the DNA level using the Blastn algorithms to search against the NCBI nt (nucleotide) reference database containing protein coding and structural RNA coding genes and by using BLASTx against nr protein database to retrieve solely protein coding sequences. The R package 'reutils' (<https://cran.r-project.org/web/packages/reutils/index.html>) was used to couple function to taxonomy. Second, reads were uploaded to MG-RAST<sup>75</sup> and joined with their pairs where possible. Due to MG-RAST read length restrictions ( $\geq 75$  bp) only paired reads that merged with a minimum overlap of 8 bp and a maximum difference of 10% were processed (MG-RASTs Fastq-join utility<sup>76</sup>). Taxonomic annotation was done using the implemented reference databases SILVA SSU for rRNA reads and Genbank (gbk)<sup>77</sup> for protein-coding reads with default cut-off values ( $e\text{-value} = 1 - e^{-5}$ , %-identity = 60%). Protein-coding reads were taxonomically filtered and split into five subsets: Eukarya (Eu), Bacteria (all bacterial reads excluding cyanobacterial reads) (B<sup>-C</sup>), Cyanobacteria (C), Proteobacteria (P) and Bacteroidetes (Bs). The last three datasets represent the most abundant bacterial phyla in the metatranscriptomes. Protein annotation of the subsets was performed using the ontology-based SEED subsystem database. Rhythmic genes in this dataset were identified using metaCycle as described above. In addition, the MG-RAST gbk annotation was used to extract reads annotated as *kaiABC*, *cikA*, *nifH*, *psbA* and *prx* for metaCycle analysis.

### Recruitment analysis

The bacterial reference genomes were selected based on their abundance in the MG-RAST dataset, which were found in similar numbers in each sample. Included genomes were *Roseobacter denitrificans* OCh 114 (NC\_008209.1), *Congregibacter litoralis* KT71 (NZ\_CM002299), *Allochroematium vinosum* DSM 180 (NC\_013851), *Lyngbya aestuarii* PCC8106 (CCY9616) (NZ\_AAVU00000000.1) and *Coleofasciculus chthonoplastes* PCC7420 (CCY9604) (synonym: *Microcoleus chthonoplastes*; NZ\_DS989896.1) (<https://ccy.nioz.nl>). The annotated genes of these genomes were linked to their matching SEED subsystems using the RAST-server.<sup>78</sup> Prior to mapping, the metatranscriptomic reads were depleted of rRNA sequences using sortmeRNA v2.0.<sup>79</sup> The rRNA depleted paired-end reads of the six metatranscriptomes were then mapped to the above-mentioned bacterial reference genomes using Bowtie2 v7.1.0<sup>80</sup> with default parameters in Geneious R8.1.7.<sup>81</sup> For downstream analyses, Geneious was used to normalize mapped reads into transcripts per million (TPM) allowing comparison of gene expression between samples.<sup>82</sup> Rhythmic genes in the recruited dataset were identified using metaCycle as described above.

### Data availability

Raw sequencing data is available under the project entitled: 'Rhythm on the beach', submitted to MG-RAST project number: mgp9582. <https://www.mg-rast.org/linkin.cgi?project=mgp9582>

### ACKNOWLEDGEMENTS

We thank Michele Grego and Lillian van der Linde for their contribution in sampling and strain cultivation. This publication was financed under project number: 821.01.013 by the Earth and Life Sciences program (ALW) of the Netherlands Organization of Scientific Research (NWO) and the MaCuMBA Project 311975 of the European Commission FP7. Additional fieldwork support was granted by the Academy Ecology Fund (KNAW).

### AUTHOR CONTRIBUTIONS

C.H. and H.B. conceived and designed the experiments. C.H. and V.C.G. performed the experiments. C.H. and H.B. analyzed the data. C.H., L.J.S. and H.B. wrote the paper.

### ADDITIONAL INFORMATION

**Supplementary information** accompanies the paper on the *npj Biofilms and Microbiomes* website (<https://doi.org/10.1038/s41522-018-0054-5>).

**Competing interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### REFERENCES

1. Severin, I., Acinas, S. G. & Stal, L. J. Diversity of nitrogen-fixing bacteria in cyanobacterial mats. *FEMS Microbiol. Ecol.* **73**, 514–525 (2010).
2. Happey-Wood, C. M. & Jones, P. Rhythms of vertical migration and motility in intertidal benthic diatoms with particular reference to *Pleurosigma angulatum*. *Diatom Res.* **3**, 83–93 (1988).
3. Richardson, L. L. & Castenholz, R. W. Diel vertical movements of the cyanobacterium *Oscillatoria terebriformis* in a sulfide-rich hot spring microbial mat. *Appl. Environ. Microbiol.* **53**, 2142–2150 (1987).
4. Stal, L. J. & Moezelaar, R. Fermentation in cyanobacteria. *Fems. Microbiol. Rev.* **21**, 179–211 (1997).
5. Berman-Frank, I., Lundgren, P. & Falkowski, P. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Res. Microbiol.* **154**, 157–164 (2003).
6. Jonkers, H. M. et al. Structural and functional analysis of a microbial mat ecosystem from a unique permanent hypersaline inland lake: 'La Salada de Chiprana' (NE Spain). *FEMS Microbiol. Ecol.* **44**, 175–189 (2003).
7. Mitbavkar, S. & Anil, A. C. Vertical migratory rhythms of benthic diatoms in a tropical intertidal sand flat: influence of irradiance and tides. *Mar. Biol.* **145**, 9–20 (2004).
8. Steunou, A.-S. et al. Regulation of *nif* gene expression and the energetics of N<sub>2</sub> fixation over the diel cycle in a hot spring microbial mat. *ISME J.* **2**, 364–378 (2008).
9. Johnson, C. H. & Egli, M. Metabolic compensation and circadian resilience in prokaryotic cyanobacteria. *Annu. Rev. Biochem.* **83**, 221–247 (2014).
10. Ma, P., Woelfle, M. A. & Johnson, C. H. An evolutionary fitness enhancement conferred by the circadian system in cyanobacteria. *Chaos, Solitons and Fractals* **50**, 65–74 (2013).
11. Johnson, C. H., Golden, S. S., Ishiura, M. & Kondo, T. Circadian clocks in prokaryotes. *Mol. Microbiol.* **21**, 5–11 (1996).
12. Mori, T. & Johnson, C. H. Independence of circadian timing from cell division in cyanobacteria. *J. Bacteriol.* **183**, 2439–2444 (2001).
13. Schmitz, O. Cika, A bacteriophytochrome that resets the cyanobacterial circadian clock. *Science* **289**, 765–768 (2000).
14. Kim, Y.-I., Vinyard, D. J., Ananyev, G. M., Dismukes, G. C. & Golden, S. S. Oxidized quinones signal onset of darkness directly to the cyanobacterial circadian oscillator. *Proc. Natl. Acad. Sci.* **109**, 17765–17769 (2012).
15. Ashworth, J. et al. Genome-wide diel growth state transitions in the diatom *Thalassiosira pseudonana*. *Proc. Natl. Acad. Sci.* **110**, 7518–7523 (2013).
16. Johnson, C. H., Golden, S. S., Ishiura, M. & Kondo, T. Circadian clocks in prokaryotes. *Mol. Microbiol.* **21**, 5–11 (1996).
17. Dvornyk, V., Vinogradova, O. & Nevo, E. Origin and evolution of circadian clock genes in prokaryotes. *Proc. Natl. Acad. Sci. USA* **100**, 2495–2500 (2003).
18. Axmann, I. M., Hertel, S., Wiegand, A., Dörrich, A. K. & Wilde, A. Diversity of KaiC-based timing systems in marine Cyanobacteria. *Mar. Genom.* **14C**, 3–16 (2014).
19. Loza-Correa, M., Gomez-Valero, L. & Buchrieser, C. Circadian clock proteins in prokaryotes: hidden rhythms? *Front. Microbiol.* **1**, 130 (2010).
20. Ma, P., Mori, T., Zhao, C., Thiel, T. & Johnson, C. H. Evolution of KaiC-dependent timekeepers: a proto-circadian timing mechanism confers adaptive fitness in the purple bacterium *Rhodospseudomonas palustris*. *PLoS Genet.* **12**, e1005922 (2016).
21. Min, H., Guo, H. & Xiong, J. Rhythmic gene expression in a purple photosynthetic bacterium, *Rhodobacter sphaeroides*. *FEBS Lett.* **579**, 808–812 (2005).
22. Ottesen, E. A. et al. Ocean microbes. Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science* **345**, 207–212 (2014).
23. Mullineaux, C. W. & Stanewsky, R. The rolex and the hourglass: a simplified circadian clock in *Prochlorococcus*? *J. Bacteriol.* **191**, 5333–5335 (2009).
24. Axmann, I. M. et al. Biochemical evidence for a timing mechanism in *Prochlorococcus*. *J. Bacteriol.* **191**, 5342–5347 (2009).
25. Paulose, J. K., Wright, J. M., Patel, A. G. & Cassone, V. M. Human gut bacteria are sensitive to melatonin and express endogenous circadian rhythmicity. *PLoS ONE* **11**, e0146643 (2016).
26. O'Neill, J. S. & Reddy, A. B. Circadian clocks in human red blood cells. *Nature* **469**, 498–503 (2011).
27. Edgar, R. S. et al. Peroxiredoxins are conserved markers of circadian rhythms. *Nature* **485**, 459–464 (2012).
28. Hastings, M. H. & O'Neill, J. S. in *Circadian Medicine* (ed Colwell, C. S.) 1–23, chapter 1 (John Wiley & Sons Inc. Hoboken, 2015). <https://doi.org/10.1002/9781118467831>.
29. Whitehead, K., Pan, M., Masumura, K. I., Bonneau, R. & Baliga, N. S. Diurnally entrained anticipatory behavior in archaea. *PLoS ONE* **4**, e5485 (2009).
30. Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. & Hogenesch, J. B. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32**, 3351–3353 (2016).
31. Armitage, D. W., Gallagher, K. L., Youngblut, N. D., Buckley, D. H. & Zinder, S. H. Millimeter-scale patterns of phylogenetic and trait diversity in a salt marsh microbial mat. *Front. Microbiol.* **3**, 293 (2012).
32. Bolhuis, H., Cretou, M. S. & Stal, L. J. Molecular ecology of microbial mats. *FEMS Microbiol. Ecol.* **90**, 335–350 (2014).
33. White, R. A., Power, I. M., Dipple, G. M., Southam, G. & Suttle, C. A. Metagenomic analysis reveals that modern microbialites and polar microbial mats have similar taxonomic and functional potential. *Front. Microbiol.* **6**, 966 (2015).
34. Enright, J. T. Ecological aspects of endogenous rhythmicity. *Ann. Rev. Ecol. Syst.* **1**, 221–238 (1970).
35. Beaver, L. M. et al. Loss of circadian clock function decreases reproductive fitness in males of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **99**, 2134–2139 (2002).
36. Woelfle, M. A., Ouyang, Y., Phanvijitsiri, K. & Johnson, C. H. The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. *Curr. Biol.* **14**, 1481–1486 (2004).
37. Palmer, J. D. The clocks controlling the tide-associated rhythms of intertidal animals. *Bioessays* **22**, 32–37 (2000).
38. Ragni, M., Apos, D. & Alcalá, M. R. Circadian variability in the photobiology of *Phaeodactylum tricoratum*: pigment content. *J. Plankton Res.* **29**, 141–156 (2007).
39. Round, F. & Happey, C. Persistent, vertical-migration rhythms in benthic microflora. IV. A diurnal rhythm of the epipellic diatom association in non tidal flowing water. *Br. Phycol. Bull.* **2**, 463–471 (1965).

40. Mitsui, A. et al. Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. *Nature* **323**, 720–722 (1986).
41. Pichard, S., Campbell, L., Kang, J., Tabita, F. & Paul, J. Regulation of ribulose biphosphate carboxylase gene expression in natural phytoplankton communities. I. Diel rhythms. *Mar. Ecol. Prog. Ser.* **139**, 257–265 (1996).
42. Suzuki, L. & Johnson, C. H. Algae know the time of day: circadian and photo-periodic programs. *J. Phycol.* **37**, 933–942 (2001).
43. Ito, H. et al. Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. USA* **106**, 14168–14173 (2009).
44. Chen, Y. B., Zehr, J. P. & Mellon, M. Growth and nitrogen fixation of the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium* sp. IMS 101 in defined media: evidence for a circadian rhythm. *J. Phycol.* **32**, 916–923 (1996).
45. Stöckel, J. et al. Global transcriptomic analysis of *Cyanothece* 51142 reveals robust diurnal oscillation of central metabolic processes. *Proc. Natl. Acad. Sci. USA* **105**, 6156–6161 (2008).
46. Ishiura, M. et al. Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria. *Science* **281**, 1519–1523 (1998).
47. Liu, Y. et al. Circadian orchestration of gene expression in cyanobacteria. *Genes Dev.* **9**, 1469–1478 (1995).
48. Kucho, K. et al. Global analysis of circadian expression in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **187**, 2190–2199 (2005).
49. Harmer, S. L. et al. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* **290**, 2110–2113 (2000).
50. Michael, T. P. & McClung, C. R. Enhancer trapping reveals widespread circadian clock transcriptional control in *Arabidopsis*. *Plant Physiol.* **132**, 629–639 (2003).
51. Johnson, C. H. Global orchestration of gene expression by the biological clock of cyanobacteria. *Genome Biol.* **5**, 217 (2004).
52. Nassoury, N., Fritz, L. & Morse, D. Circadian changes in ribulose-1,5-bisphosphate carboxylase/oxygenase distribution inside individual chloroplasts can account for the rhythm in dinoflagellate carbon fixation. *Plant. Cell.* **13**, 923–934 (2001).
53. Johnson, C. H., Zhao, C., Xu, Y. & Mori, T. Timing the day: what makes bacterial clocks tick? *Nat. Rev. Microbiol.* **15**, 232–242 (2017).
54. Poretsky, R. S. et al. Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environ. Microbiol.* **11**, 1358–1375 (2009).
55. Trinder, M., Bisanz, J. E., Burton, J. P. & Reid, G. Bacteria need 'sleep' too?: microbiome circadian rhythmicity, metabolic disease, and beyond. *Univ. Tor. Med. J.* **92**, 52–55 (2015).
56. Gasol, J. M. et al. Diel variations in bacterial heterotrophic activity and growth in the northwestern Mediterranean Sea. *Mar. Ecol. Prog. Ser.* **164**, 107–124 (1998).
57. Bebout, B. M. & Garcia-Pichel, F. UV B-induced vertical migrations of cyanobacteria in a microbial mat. *Appl. Environ. Microbiol.* **61**, 4215–4222 (1995).
58. Nadeau, T. L., Howard-Williams, C. & Castenholz, R. W. Effects of solar UV and visible irradiance on photosynthesis and vertical migration of *Oscillatoria* sp. (Cyanobacteria) in an Antarctic microbial mat. *Aquat. Microb. Ecol.* **20**, 231–243 (1999).
59. Coelho, H., Vieira, S. & Serôdio, J. Endogenous versus environmental control of vertical migration by intertidal benthic microalgae. *Eur. J. Phycol.* **46**, 271–281 (2011).
60. Stal, L. J. in *Ecology of cyanobacteria II: their diversity in space and time* (ed Whitton, B. A.) 65–125 (Springer, Dordrecht, 2012) [https://doi.org/10.1007/978-94-007-3855-3\\_4](https://doi.org/10.1007/978-94-007-3855-3_4).
61. Häder, D.-P. & Hoiczyk, E. *Algal Cell Motility* (Springer, US, 1992).
62. Khayatan, B., Meeks, J. C. & Risser, D. D. Evidence that a modified type IV pilus-like system powers gliding motility and polysaccharide secretion in filamentous cyanobacteria. *Mol. Microbiol.* **98**, 1021–1036 (2015).
63. Stal, L. J. Coastal microbial mats: the physiology of a small-scale ecosystem. *South Afr. J. Bot.* **67**, 399–410 (2001).
64. Finzi-Hart, J. A. et al. Fixation and fate of C and N in the cyanobacterium *Trichodesmium* using nanometer-scale secondary ion mass spectrometry. *Proc. Natl. Acad. Sci. USA* **106**, 6345–6350 (2009).
65. Lambert, G. & Kussel, E. Memory and fitness optimization of bacteria under fluctuating environments. *PLoS. Genet.* **10**, e1004556 (2014).
66. Skanata, A. & Kussel, E. Evolutionary phase transitions in random environments. *Phys. Rev. Lett.* **117**, 38104 (2016).
67. Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463**, 326–330 (2010).
68. Decho, A. W., Norman, R. S. & Visscher, P. T. Quorum sensing in natural environments: emerging views from microbial mats. *Trends Microbiol.* **18**, 73–80 (2010).
69. Stal, L. J. in *Methods in Enzymology* **167**, 474–484 (1988).
70. Severin, I. & Stal, L. J. Light dependency of nitrogen fixation in a coastal cyanobacterial mat. *ISME J.* **2**, 1077–1088 (2008).
71. Parkhomchuk, D. et al. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucl. Acids Res.* **37**, e123 (2009).
72. Levin, J. Z. et al. Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat. Methods* **7**, 709–715 (2010).
73. Pfaffl, M. W., Tichopad, A., Prgomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–515 (2004).
74. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ.* **4**, e2584 (2016).
75. Meyer, F. et al. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinforma.* **9**, 386 (2008).
76. Aronesty, E. Comparison of sequencing utility programs. *Open Bioinforma. J.* **7**, 1–8 (2013).
77. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. GenBank. *Nucl. Acids Res.* **33**, D34–D38 (2005).
78. Aziz, R. K. et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom.* **9**, 75 (2008).
79. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
80. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
81. Kearse, M. et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
82. Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285 (2012).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018