The ISME Journal (2010) 4, 621–632
© 2010 International Society for Microbial Ecology All rights reserved 1751-7362/10 $32.00
www.nature.com/ismej

ORIGINAL ARTICLE

Genome-wide analysis of diel gene expression in the unicellular N_2-fixing cyanobacterium *Crocosphaera watsonii* WH 8501

Tuo Shi[1], Irina Ilikchyan, Sophie Rabouille[2] and Jonathan P Zehr
Department of Ocean Sciences, University of California, Santa Cruz, CA, USA

The unicellular cyanobacterium *Crocosphaera watsonii* is an important nitrogen fixer in oligotrophic tropical and subtropical oceans. Metabolic, energy and cellular processes in cyanobacteria are regulated by the circadian mechanism, and/or follow the rhythmicity of light–dark cycles. The temporal separation of metabolic processes is especially essential for nitrogen fixation because of inactivation of the nitrogenase by oxygen. Using a microarray approach, we analyzed gene expression in cultures of *Crocosphaera watsonii* WH 8501 (*C. watsonii*) over a 24-h period and compared the whole-genome transcription with that in *Cyanothece* sp. ATCC 51142 (*Cyanothece*), a unicellular diazotroph that inhabits coastal marine waters. Similar to *Cyanothece*, regulation at the transcriptional level in *C. watsonii* was observed for all major metabolic and energy processes including photosynthesis, carbohydrate and amino acid metabolisms, respiration, and nitrogen fixation. Increased transcript abundance for iron acquisition genes by the end of the day appeared to be a general pattern in the unicellular diazotrophs. In contrast, genes for some ABC transporters (for example, phosphorus acquisition), DNA replication, and some genes encoding hypothetical proteins were differentially expressed in *C. watsonii* only. Overall, *C. watsonii* showed a higher percentage of genes with light–dark cycling patterns than *Cyanothece*, which may reflect the habitats preferences of the two cyanobacteria. This study represents the first whole-genome expression profiling in cultivated *Crocosphaera*, and the results will be useful in determining the basal physiology and ecology of the endemic *Crocosphaera* populations.
The ISME Journal (2010) **4,** 621–632; doi:10.1038/ismej.2009.148; published online 28 January 2010
Subject Category: microbial population and community ecology
Keywords: *Crocosphaera*; gene expression; microarray; unicellular diazotroph

Introduction

Cyanobacteria are prokaryotic phototrophs that catalyze diverse biochemical reactions, including hydrogen production and degradation of complex organics (Herrero and Flores, 2008), and are responsible for a large fraction of global primary productivity. They are important in ecosystem function in freshwater, marine and terrestrial systems (Whitton and Potts, 2000). Some cyanobacteria are capable of biological nitrogen fixation, the reduction of atmospheric dinitrogen (N_2) gas to ammonium (NH_3) (Postgate, 1998). The fixation of N_2 is a particularly important source of nitrogen in marine habitats, yet few N_2-fixing cyanobacteria are found in the oligotrophic open ocean (Howarth *et al.*, 1988; Michaels *et al.*, 1996; Karl *et al.*, 2002; Stal and Zehr, 2008). The free-living bloom-forming filamentous cyanobacterium *Trichodesmium* (Capone *et al.*, 1997; LaRoche and Breitbarth, 2005) and the heterocystous *Richelia* symbionts of marine pelagic diatoms (Villareal, 1992; Foster and Zehr, 2006) were believed to be the primary N_2-fixers (diazotrophs) until the discovery that the unicellular N_2-fixing cyanobacteria were also generally widespread and important in N_2-fixation in the open ocean (Zehr *et al.*, 2001; Falcón *et al.*, 2002; Mazard *et al.*, 2004; Montoya *et al.*, 2004; Langlois *et al.*, 2005). One of these unicellular cyanobacterial species, *Crocosphaera watsonii* strain WH 8501 (*C. watsonii*), had previously been isolated from tropical waters (Waterbury and Rippka, 1989).

As nitrogenase is extremely sensitive to O_2, diazotrophic cyanobacteria must avoid inactivation of nitrogen fixation by the O_2 produced in photosynthesis (Fay, 1992; Gallon, 1992; Stal and Zehr, 2008). Many filamentous diazotrophic cyanobacteria

Correspondence: J Zehr, Department of Ocean Sciences, University of California, Santa Cruz, CA 95064, USA.
E-mail: zehrj@ucsc.edu
[1]Current address: Algenol Biofuels, Inc., 1700 Union Ave, Baltimore, MD 21211, USA.
[2]Current address: CNRS, UMR 7093, LOV, 06230 Villefranche-sur-Mer, France.
Received 11 August 2009; revised 30 November 2009; accepted 30 November 2009; published online 28 January 2010

develop specialized N_2-fixing cells, called heterocysts (Fay, 1992) that do not evolve O_2. A majority of the unicellular cyanobacteria fix N_2 at night, when O_2 is not being produced by photosynthesis. Cyanobacteria control metabolism during light–dark cycles through complex gene expression patterns that include circadian rhythms (Chen *et al.*, 1996; Kucho *et al.*, 2005; Toepel *et al.*, 2009; Zinser *et al.*, 2009). The expression of nitrogenase genes in *Trichodesmium* and in the intertidal or coastal marine unicellular cyanobacterium *Cyanothece* sp. strain 51142 has been shown to be under circadian control (Schneegurt *et al.*, 1994; Chen *et al.*, 1998; Berman-Frank *et al.*, 2001; Toepel *et al.*, 2008). Genome-wide diel expression analysis may further elucidate the mechanism mediating the spatial/temporal segregation of oxygenic photosynthesis and N_2 fixation in diazotrophic cyanobacteria.

Crocosphaera is unicellular, planktonic and fixes N_2 in the aerobic surface waters of the ocean. It is widely distributed in tropical and subtropical waters (reported from the North and South Atlantic Ocean, North and South Pacific Ocean and the Arabian Sea), but reports on its physiology or ecology are still very recent (Falcón *et al.*, 2004, 2005; Tuit *et al.*, 2004; Dyhrman and Haley, 2006; Zehr *et al.*, 2007a; Fu *et al.*, 2008; Compaoré and Stal, 2009; Hewson *et al.*, 2009; Webb *et al.*, 2009). The draft genome of *C. watsonii* WH 8501 is 6.24 Mb, with 323 contigs encoding a total of 5958 candidate protein-encoding genes (http://microbes.ucsc.edu/cgi-bin/hgGateway?db=crocWats_WH-8501_0). On the basis of the first draft of this genome, a microarray platform was designed (through Roche NimbleGen Inc., Madison, WI, USA) that included 4407 open reading frames (ORFs) identical to those of the latest released *C. watsonii* draft genome (excluding most of the transposases and ambiguous ORFs, see Supplementary Figure S1 for details). Recently, whole-genome expression in the closely related marine, but intertidal, unicellular cyanobacterium *Cyanothece* sp. ATCC 51142 (*Cyanothece*) was reported (Stöckel *et al.*, 2008; Toepel *et al.*, 2008), making it possible to determine common features of genome expression in the unicellular N_2-fixing cyanobacteria and to determine unique gene expression patterns of the open ocean species *C. watsonii*. In this paper, we report whole-genome expression patterns in the cultures of *C. watsonii* that were grown under alternating light–dark cycles and N_2-fixing conditions. Our objectives were to (1) compare transcriptional regulation of major metabolic and cell processes during diel cycle in *C. watsonii* to those in *Cyanothece*, (2) identify genes in *C. watsonii* with a diel expression pattern (DEP) that could be targeted in future research to determine whether expression of these genes is circadian or light-regulated and (3) identify potential target genes to assess the physiological status of unicellular diazotrophic cyanobacteria in the environment.

Materials and methods

Culture and growth conditions

C. watsonii was cultivated in polycarbonate tissue culture flasks with a 0.2 μm pore-size vent cap (Corning Inc., Corning, NY, USA). For the diel experiment, duplicate cultures were maintained under N_2-fixing conditions in a modified YBCII medium (Chen *et al.*, 1996) at 26 °C on a 12:12 h light–dark cycle at 50 μmol photons m^{-2} s^{-1}. Concentrations of phosphate (KH_2PO_4) and iron ($FeCl_3 \cdot 6H_2O$ and ferric ammonium citrate) in the medium were 50 and 5.8 μM, respectively. The cultures were at least 20-fold diluted from the inoculum and were verified axenic by staining with SYBR Gold (Invitrogen, Carlsbad, CA, USA) and visualizing under an epifluorescent microscope (Carl Zeiss, Thornwood, NY, USA). Growth and cell density were monitored with flow cytometry (Cytopia, Seattle, WA, USA) until the cultures reached exponential phase (~10–14 days after inoculation), and then the cells were harvested for the experiments. A total of 16 samples (two replicates taken from different flasks at each of eight time points) were collected, taken every 2–5 h starting with 1 h before the light until 1 h after the light period the following day: D11, L1, L6, L11, D1, D6, 2D11 and 2L1, where L and D stand for light and dark, respectively, followed by the corresponding hour. 2D11 and 2L1 samples were taken from the second light–dark cycle.

Nitrogenase activity

For online monitoring of nitrogenase activity, a sample of 125 ml of *C. watsonii* cultures reaching exponential growth was collected right before the onset of the dark period (L12) together with two 1.5-ml samples for cell counts. The two 1.5-ml replicate samples were fixed with formaldehyde and were stored at 4 °C before cell counts by flow cytometry. The 125-ml sample was immediately placed in a tightly closed, flow-through polycarbonate bottle in a water bath maintained at 26 °C. The bottle was equipped with a magnetic stirrer whose rotation was set at 150 revolutions per minute to ensure homogeneity of the culture. An inlet tube, connected to a gas-mixture system monitored by computer, was inserted into the culture for gas bubbling. The gas mixture, regulated by Brooks 5850S mass flow controllers, flowed at a constant rate of 1.02 l h^{-1} and was composed of 20% oxygen, 10% acetylene and 70% nitrogen. A second tube, connected to an Agilent 6850 gas chromatograph was inserted to the headspace for gas sampling. Apart from the online incubator, nitrogenase activity was recorded according to the acetylene reduction assay as described by Staal *et al.* (2001). The flow-through system was covered with dense, black plastic sheets to ensure total darkness. In the end of the dark period, the sheet was removed and the

system received an ambient, natural light irradiance whose intensity was equivalent to that in the culture incubator.

Sample preparation
At designated times, 50 ml of cultures ($\sim 10^6$–10^7 cells ml^{-1}) were removed and filtered through a 0.2 μm pore-size polycarbonate membrane filter (GE Osmonics Lab Store, Minnetonka, MN, USA) at approximately 20 KPa vacuum pressure. The filters were immediately frozen in liquid nitrogen and stored at −80 °C until they were processed. Total RNA was isolated using the RiboPure-Bacteria RNA isolation Kit (Ambion, Austin, TX, USA), followed by DNA removal with DNase treatment in solution and cleanup with RNeasy spin columns (Qiagen, Valencia, CA, USA), according to the manufacturer's instructions. The quantity and quality of the RNA was analyzed using a NanoDrop ND-1000 UV-VIS spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) respectively. The RNA from two replicates was combined, and 20 μg of the mixed RNA for each time point was reverse transcribed into complementary DNA (cDNA) using the SuperScript double-stranded complementary DNA synthesis kit (Invitrogen Corp., Carlsbad, CA, USA) resulting in eight cDNA samples. The cDNA was labeled with Cy3, and the microarrays were hybridized at the NimbleGen laboratory (NimbleGen).

Microarray
The oligonucleotide expression array of *C. watsonii* (NimbleGen design ID 2007-03-14_EW_C_watsonii) was designed using draft assemblies of the *C. watsonii* genome sequence as of 17 March 2004, but without probes for most of the transposases. Up to 19 60-mer oligonucleotide probes were designed for each of the 4991 ORFs and were quadruplicated on the microarray, which allowed chip-internal evaluation of signals. For this report, we mapped 4407 ORFs to the current draft genome *NCBI* gene models (GenBank accession number AADV00000000 as of 26 August 2009; http://www.ncbi.nlm.nih.gov) (Supplementary Figure S1). To further identify these genes and predict their functions, all the 5958 protein-coding genes in *C. watsonii* draft assembly were compared on the amino acid level to those in the genome of *Cyanothece* (GenBank accession numbers NC_010546 and NC_010547) using the program Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1997) with an E-value cutoff of 10^{-10}. For each of the 4407 *C. watsonii* ORFs, the best BLAST hit to *Cyanothece* was reported (Supplementary Tables S1 and S2). Full microarray platform descriptions and data are available at the ArrayExpress database (http://www.ebi.ac.uk/microarray-as/ae/) under accession number E-TABM-737. Microarray hybridization signals were quantified using a GenePix 4000B Scanner (Molecular Devices, Sunnyvale, CA, USA). Images were analyzed with NimbleScan software v2.4 (NimbleGen). Expression data were generated using the robust multichip average algorithm (Irizarry *et al.*, 2003b).

Data analysis
Microarray data were analyzed with Bioconductor software available at http://www.bioconductor.org/ (Gentleman *et al.*, 2004). The quality of arrays was assessed by examining the distribution of log scale probe intensities and MA-plots constructed by comparing each array to probe-wise medians (Gentleman *et al.*, 2005), where M_i is the log intensity for an *i* probe in an array, and A is the median log intensity for *i* probe in all arrays (Supplementary Figure S2). Differentially expressed genes were selected by comparing gene expression between all possible pairs of time points using the Limma package (Smyth, 2005). The threshold of gene expression change was set to twofold ($P < 0.01$, false discovery rate adjusted) (Benjamini and Hochberg, 1995), and genes that were differentially expressed in pairs with D11 and 2D11 or L1 and 2L1 were omitted. For example, comparing the expression in pair L6/D11 yielded 832 differentially expressed genes, and in pair L6/2D11 yielded 877 differentially expressed genes, but both comparisons shared only 730 genes that were selected for subsequent analysis. This resulted in a total of 2823 potential genes that had a DEP.

To visualize the major expression patterns, the 2823 genes were clustered with hierarchical cluster analysis using Pearson's correlation as distance method and complete linkage agglomeration. The number of clusters was set to 10 based on the condition that all clusters are enriched in annotated genes. Genes in each cluster were assigned to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, if known (Supplementary Table S1). Principal component analysis was performed to assess switching of the key pathways in relation to the light–dark cycles (Supplementary Figure S3). The microarray results were validated with reverse transcription quantitative PCR (Supplementary Information).

Results and discussion

As described in Materials and methods, 2823 genes (64.1% of genes investigated in this study or 47.4% of the whole transcriptome) were identified as genes that had a DEP pool with a more than twofold change ($P < 0.01$, false discovery rate adjusted) in the expression between any two time points (Supplementary Table S1). The percentage of cycling genes detected in *C. watsonii* is less than that in the sympatric oceanic non-N_2-fixing cyanobacterium

Prochlorococcus sp. MED4 (*Prochlorococcus*) (approximately 81% of all protein-coding genes, Zinser *et al*., 2009), but is 17% more than what was detected in *Cyanothece* (approximately 30%, Stöckel *et al*., 2008). The oscilloid hypothesis suggests that a circadian clock mediated by the KaiABC protein complex in cyanobacteria modulates DNA topology, which in turn affects gene expression (Mori and Johnson, 2001); therefore, all actively expressed genes would be expected to show cyclic expression patterns. The difference in number of cycling genes in *Cyanothece*, *Prochlorococcus*, and *C. watsonii* could be due to different approaches to microarray data analysis in these studies, specific limits of different microarray technologies (Miller and Tang, 2009) or to different culture conditions. Alternatively, it is possible that the number of cycling genes reflects the unique habitats and different ecology of these organisms for *C. watsonii* and *Prochlorococcus* both live in the oligotrophic open ocean, whereas *Cyanothece* is found in benthic or intertidal marine environments.

The DEP pool in *C. watsonii* contained ca. 23% of genes assigned to KEGG pathways and 29% of hypothetical proteins (including 168 conserved hypothetical). In contrast, 1584 genes with a non-diel expression pattern (non-DEP pool) (Supplementary Table S2) contained ca. 12% of genes assigned to KEGG pathways and 42% of the hypothetical proteins (including 86 conserved hypothetical). The gene expression values (log 2-based) ranged from 4 to 16 with the first and third quartiles of 5.5 and 10.5, respectively, and a median of 8 (Supplementary Figure S2). The majority of genes encoding hypothetical proteins in the non-DEP pool had low expression values (<5.5) throughout the diel cycle, and these genes may be upregulated during the stationary phase or involved in responses to different stresses. Some genes in the non-DEP pool, for example, photosynthesis genes *psaD*, *psbN* and *petF3*, may have a diel pattern, but were not detected because of the twofold expression threshold chosen to identify the DEP genes.

Cyanothece and *C. watsonii* genomes share 909 genes that have at least 90% BLAST identity in protein sequences and 3364 genes with at least 60% identity. In spite of the genome similarity, the numbers of genes in the cycling and non-cycling pools were significantly different between these two cyanobacteria (Figure 1). The higher rate of DEP in *C. watsonii* could be in part because of multiple gene copies or ORF breakages resulted from the draft genome assembly, which led to several *C. watsonii* genes with a single best BLAST hit in *Cyanothece* (for example, 31 ORFs encoding transposases IS4 in *C. watsonii* match a single gene in *Cyanothece*, *cce_0105*). However, the higher number of regulatory systems in the cycling pool in *Cyanothece* and the higher number of ABC transporters in the DEP pool in *C. watsonii* (Figure 1) could be due to the need for regulation in the fluctuating conditions of the coastal or benthic environment compared with the more stable nutrient-depleted water column of the open ocean.

The transcriptional regulation of cyanobacterial metabolism and cellular processes over light–dark cycles has been described in details previously (Kucho *et al*., 2005; Labiosa *et al*., 2006; Stöckel *et al*., 2008; Zinser *et al*., 2009), and similar co-regulation of major processes in *C. watsonii* was not unexpected. The genes from the DEP pool can be divided into four categories based on the timing of maximal expression (Figure 2). The first category is comprised of clusters 1 and 7 that had maximum expression during the light period. The corresponding genes encode some ABC transporters, ribosomes and

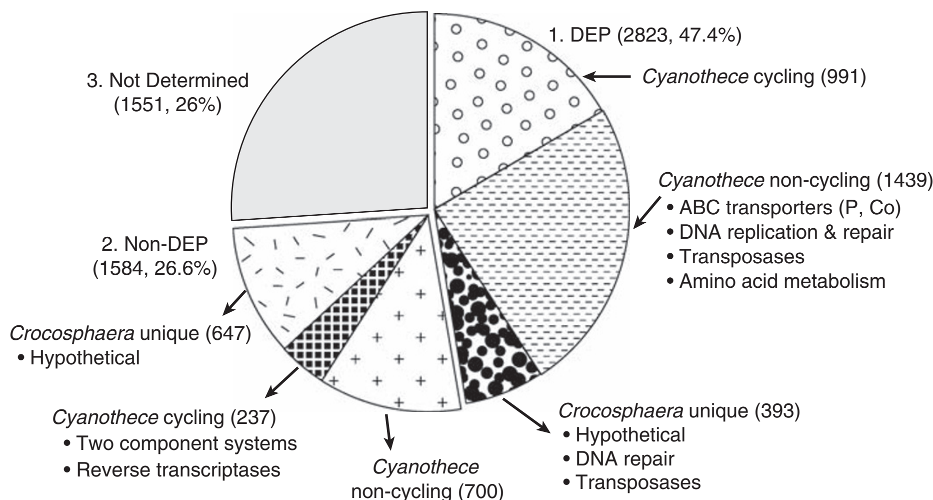

Figure 1 Comparison of the transcriptome profiles of *C. watsonii* (this study) and *Cyanothece* (Stöckel *et al*., 2008). The *C. watsonii* transcriptome is divided into three parts: (1) the DEP, (2) the non-DEP and (3) genes that were not determined in this study. In each category, the number of genes, followed by the corresponding transcriptome percentage, is indicated in parenthesis. The DEP and non-DEP pools in *C. watsonii* were compared with those in *Cyanothece*, and were further decomposed based on the number of genes that were similarly or differentially transcribed in these two organisms. Representative functions of uniquely transcribed genes are also shown.

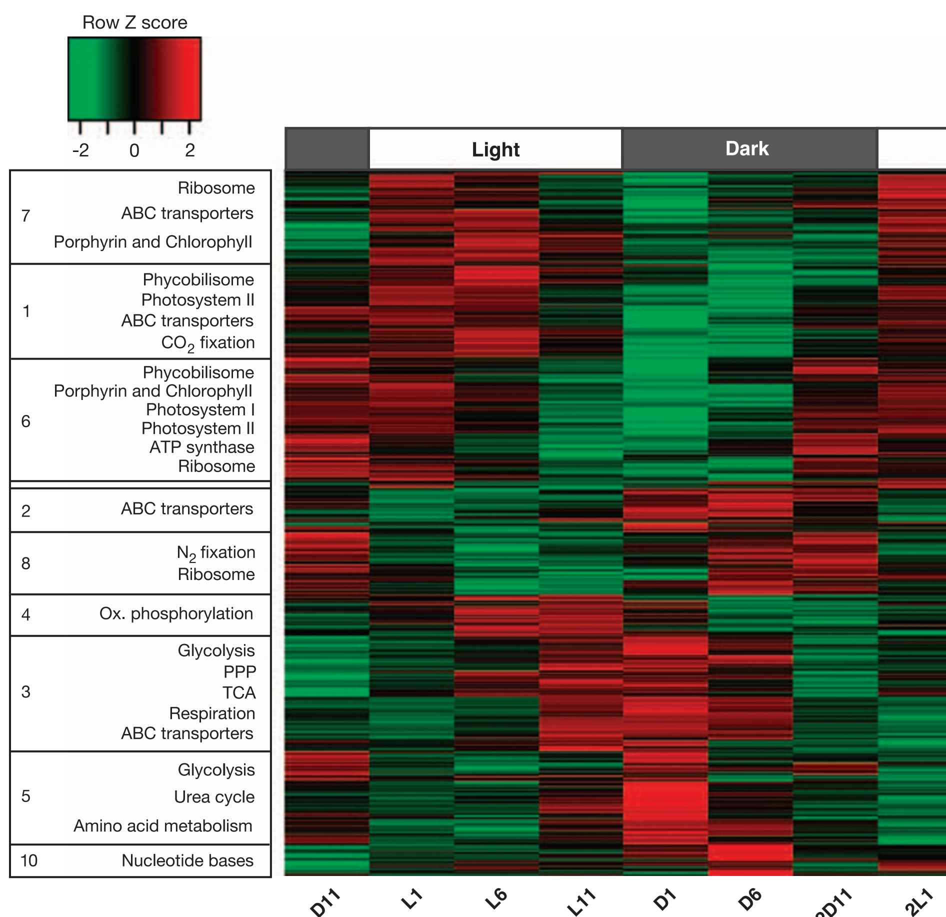

Figure 2 Diel expression patterns for 2823 genes over light–dark cycles. The expression values of genes at each time point were standardized, and the green–red scale shows how many s.d. (Z score) the expression value is lower or higher, respectively, than the mean expression values over the diel cycle. For each of 10 clusters, the dominant KEGG pathways are shown. Cluster 9 is not designated because of its small size and is located between clusters 6 and 2. L and D stand for light and dark, respectively, followed by the corresponding hour. The shaded area represents the dark period.

photosynthesis proteins. The transcript abundance of genes from clusters 3, 4 and 5 increased during the late light—early dark period. This category includes the highest number of genes, 898 in total (Table 1), and is responsible for carbohydrate catabolism, respiration and amino acid metabolism. The third category of genes had highest expression during the dark period and consists of clusters 2, 8 and 10. The dominant pathway in this category is nitrogen fixation. The temporal separation of gene expression for nitrogen fixation and photosynthesis was also supported by principal component analysis (Supplementary Figure S3). Clusters 6 and 9 constitute the fourth category with the highest gene transcription during the transition from dark to light period. A majority of genes involved in photosynthesis, light harvesting and carbon fixation are in the fourth category.

The transcription patterns of many genes from the DEP pool were similar to those in *Cyanothece* (Stöckel *et al*., 2008) including genes for photosynthesis, carbohydrate metabolism, respiration, nitrogen fixation and amino acid biosynthesis. Therefore, expression of these genes may be a common pattern in the unicellular N_2-fixing cyanobacteria, and they may provide good targets for uncultivated populations of this group of cyanobacteria. Below, we focus on genes whose expression differs from that in other cyanobacteria, particularly in *Cyanothece*, and expression patterns that may reflect the adaptation of *Crocosphaera* strains to oligotrophic tropical and subtropical waters.

Table 1 Characterization of clusters of genes with diel expression patterns (2823 genes)

Cluster	*Number of genes*	*Genes with KEGG assignment, %*	*Hypothetical proteins, %*
1	378	26.46	22.49
2	177	14.69	36.72
3	460	18.70	32.83
4	166	21.69	24.10
5	372	19.89	34.68
6	488	37.09	20.49
7	374	16.31	33.42
8	249	28.23	31.05
9	29	10.34	51.72
10	130	13.85	26.15

Nitrogen fixation and hydrogen uptake

In *C. watsonii*, like in other N_2-fixing non-heterocystous cyanobacteria (Welsh *et al*., 2008), nitrogen fixation-related proteins are encoded by a 26-kb cluster containing 29 genes (Postgate, 1998). The expression of all 29 genes was strongly synchronous (Figure 3a), and the pattern is consistent with that previously observed in field samples through *nifH* reverse transcription PCR (Falcón *et al*., 2004; Church *et al*., 2005; Zehr *et al*., 2007b). In addition, nitrogenase activity followed the *nif* expression pattern with a delay in several hours (Figure 3b) required for synthesis of a functional enzyme. Such delay was also observed for *Trichodesmium* IMS101 (Chen *et al*., 1998). On the onset of the light, N_2-fixation transiently showed a peak that denoted the use of new energy, coming from photosynthesis, by the remaining nitrogenase enzyme. This conclusion is supported by data from a similar experiment in which there was no increase in N_2-fixation when the online system stayed in the dark after D12 (data not shown).

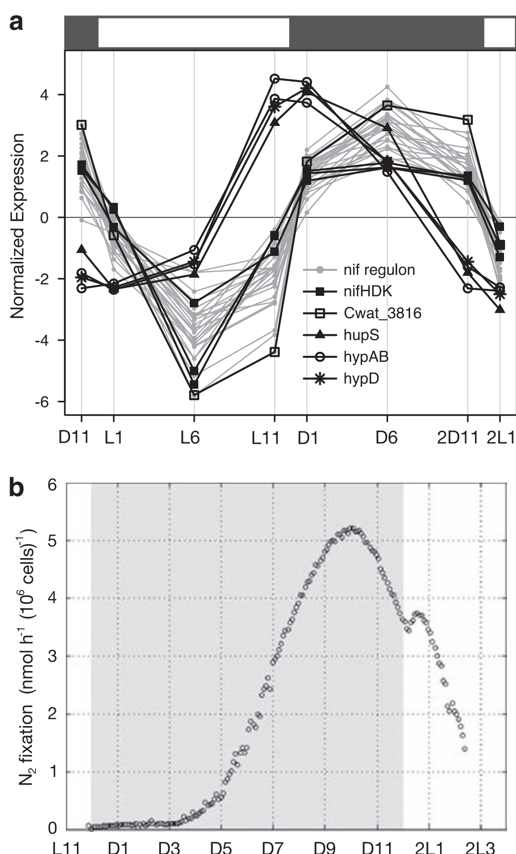

Figure 3 Profiles of expression of nitrogen fixation and hydrogenase uptake genes (**a**) and diel variations in the N_2-fixation rate (**b**). The expression value of each gene at each time point was normalized to the mean at all time points (*Y* axis, log scale). The *X* axis represents time points where D and L stand for dark and light, respectively, followed by the corresponding hour. The shaded area represents the dark period. Nitrogen fixation activity was estimated from acetylene reduction using a conversion factor of 4, and corrected for biomass using cell counts determined in the beginning of the experiment.

The genes from the *nif* cluster had the greatest difference in expression over the 24-h period (Table 2, Supplementary Table S1). The most dramatic difference in expression (more than 600-fold, confirmed by reverse transcription quantitative PCR, see Supplementary Table S3) was observed for the genes *nifU*, CwatDRAFT_3816 and *nifB* that likely constitute a single operon (Table 2). In contrast, these genes had approximately the same expression fold changes as the rest of *nif* genes in *Cyanothece*, in which the maximum difference in the expression of *nif* genes was $<$150-fold (Stöckel *et al*., 2008). The transcripts for the *nifB* gene, encoding nitrogenase cofactor biosynthesis, were also abundant in the metatranscriptomic library of *C. watsonii* from the South Pacific (Hewson *et al*., 2009). Thus, the genes *nifB*, CwatDRAFT_3816 and *nifU* may serve as more sensitive targets for assessing regulation of nitrogen fixation in natural *Crocosphaera* populations in the open ocean.

As a result of the reduction of N_2 to NH_3, molecular hydrogen is released and captured by a hydrogenase, resupplying nitrogenase with electrons, protons, ATP, as well as reducing the O_2 concentration (Tamagnini *et al*., 2007). In *C. watsonii*, the expression of *hup* genes for uptake hydrogenase and *hypAB* and *hypD* genes for hydrogenase maturation proteins (Tamagnini *et al*., 2002) resembled the expression of the *nif* genes (Figure 3a), but the maximum *hup* and *hyp* transcript abundance was 5 h earlier than maximum transcript abundance of *nif* genes. This pattern is similar to the one observed for *hup* genes in other cyanobacteria (Leitão *et al*., 2005). *Cyanothece*, in contrast to *C. watsonii*, has both *hup* and *hox* genes encoding uptake and bidirectional hydrogenases, respectively. The *hup* genes in *C. watsonii* had transcription pattern more similar to the *hox* than to the *hup* genes in *Cyanothece* (Stöckel *et al*., 2008).

Cell cycle

DNA replication was likely initiated in the afternoon or early night as indicated by the increased transcript abundance for the DNA replication genes including genes for the replication initiation protein DnaA, DNA polymerase III DnaE, the adenine-specific DNA methylase gene *Dam*, and DNA topoisomerase I TopA (Figure 4a, Supplementary Table S1). The termination of DNA replication by onset of the day was indicated by the maximum transcript abundance for genes involved in the cellular division ring formation and localization, *ftsZ* and *minE* (Vicente *et al*., 2006). In the early light period, expression of the next set of genes for division ring proteins, *FtsQ* and *FtsY* (Vicente *et al*., 2006), was upregulated suggesting initiation of cell division (Figure 4a, Supplementary Table S1). At the same time, transcripts for two *parB* genes that encode a chromosome segregation protein were also most abundant (Supplementary Table S1). In *Cyanothece*, cell division genes showed similar

Table 2 Log (base-2) fold change (LFC) of the *nif* gene expression

Gene ID	*Gene*	*Annotation*	*LFC*
CwatDRAFT_3849[a]		Nitrogenase-associated protein	5.04
CwatDRAFT_3848[a]		Hypothetical	7.69
CwatDRAFT_3847[a]	*nifT*	NifT/FixU	7.88
CwatDRAFT_3846[a]	*nifZ*	NifZ	5.12
CwatDRAFT_3845[a]	*nifV*	HMG-CoA lyase-like	5.99
CwatDRAFT_3844[a]		Hypothetical	7.96
CwatDRAFT_3843[a]		Hypothetical	7.30
CwatDRAFT_3842[a]	*cysE*	Serine O-acetyltransferase	8.10
CwatDRAFT_3815	*nifB*	Nitrogenase cofactor biosynthesis protein NifB	8.90
CwatDRAFT_3816		Hypothetical	9.44
CwatDRAFT_3817	*nifU*	Aminotransferase, NifU,	8.97
CwatDRAFT_3818	*nifH*	Nitrogenase iron protein	7.08
CwatDRAFT_3819	*nifD*	Nitrogenase Mo-Fe protein alpha chain	6.78
CwatDRAFT_3820	*nifK*	Nitrogenase Mo-Fe protein beta chain	4.41
CwatDRAFT_3821		Mo-dependent nitrogenase, C-terminal	7.42
CwatDRAFT_3822	*nifE*	Nitrogenase MoFe cofactor biosynthesis protein NifE	6.67
CwatDRAFT_3823	*nifN*	Nitrogenase Mo-Fe cofactor biosynthesis protein	6.96
CwatDRAFT_3824	*nifX*	Nitrogen fixation-related protein	8.47
CwatDRAFT_3825		Protein of unknown function DUF269	6.73
CwatDRAFT_3826		Protein of unknown function DUF683	5.57
CwatDRAFT_3827	*nifW*	Nitrogen fixation protein NifW	6.35
CwatDRAFT_3828	*hesA*	UBA/THIF-type NAD/FAD binding fold:MoeZ/MoeB	5.77
CwatDRAFT_3829	*hesB*	HesB/YadR/YfhF	7.80
CwatDRAFT_3830	*fdx*	Ferredoxin	6.31
CwatDRAFT_3831		Hypothetical	5.03
CwatDRAFT_3832	*feoB1*	Small GTP-binding protein domain	5.62
CwatDRAFT_3833	*feoB2*	Similar to Fe^{2+} transport system protein B	5.20
CwatDRAFT_3834	*feoA*	FeoA	5.66
CwatDRAFT_3835		ABC transporter: permease	4.37

[a]Genes are located on the complementary strand.
The genes are arranged in the order as they appear in the genome.

expression profile and found to be under circadian control (Toepel *et al.*, 2008). However, genes for DNA replication were not among the cycling genes in *Cyanothece* (Stöckel *et al.*, 2008). In another cyanobacterium, *Prochlorococcus*, the timing and cycling feature of DNA replication genes were somewhat similar (Zinser *et al.*, 2009) to that observed in *C. watsonii*.

Sigma factors

Almost all identified sigma factors were differentially expressed during the diel cycle in *C. watsonii* (Figures 4b and c). SigA is a primary sigma factor, and its increased expression coincided with DNA replication (Figure 4b). The group II sigma factor SigE is involved in the activation of genes for glycogen breakdown, glycolysis and oxidative pentose phosphate pathway in *Synechocystis* sp. PCC 6803 (Osanai *et al.*, 2005). In this study, transcription of *sigE* gene was similar to the expression of genes for carbohydrate catabolism (Figures 4c and 2). Expression profiles for group II sigma factors SigB and SigD coincided with their suggested role in regulation of global gene expression in the dark and light, respectively (Summerfield and Sherman, 2007). In *Cyanothece*, most of the sigma factors had expression patterns closely resembling those observed in *C. watsonii* (Stöckel *et al.*, 2008). Moreover, differential diel expression was reported for one primary and four alternative sigma factors in *Prochlorococcus* (Zinser *et al.*, 2009) and two sigma factors in *Synechocystis* sp. PCC 6803 (Kucho *et al.*, 2005). Alternative sigma factors activate transcription of many genes in response to various stresses (Los *et al.*, 2008) and are also involved in maintaining the circadian rhythm (Nair *et al.*, 2002). Genes controlled by these sigma factors are not yet well defined in cyanobacteria, but it appears that the sigma factors orchestrate global gene expression over the diel cycle (Summerfield and Sherman, 2007).

Iron uptake

The expression profile of several genes suggested an increase in iron demand during the late day—early night period. These genes include two *feoAB* operons encoding a putative ferrous transporter, *isiA* gene for the iron stress inducible chlorophyll *a*-binding protein, *isiB* encoding flavodoxin, and two *fur* genes (from four genes total) encoding the ferric uptake regulator (Figure 5a, Supplementary Table S1) (Pakrasi *et al.*, 1985; Laudenbach and Straus, 1988; Laudenbach *et al.*, 1988; Katoh *et al.*, 2001; Singh *et al.*, 2003; Chappell and Webb, 2009). The third *feoAB* operon is located in the *nif* cluster and had an expression pattern identical to *nif* genes.

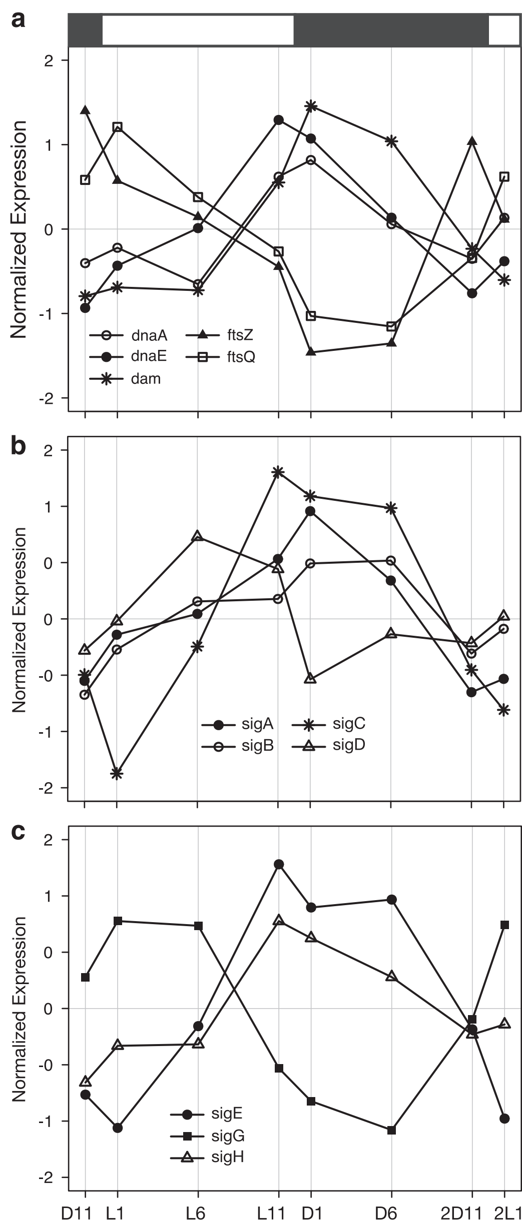

Figure 4 Expression profiles for genes involved in DNA replication and cell division (**a**), and for sigma factor genes (**b, c**). The *X* and *Y* axes follow the legend for Figure 3a.

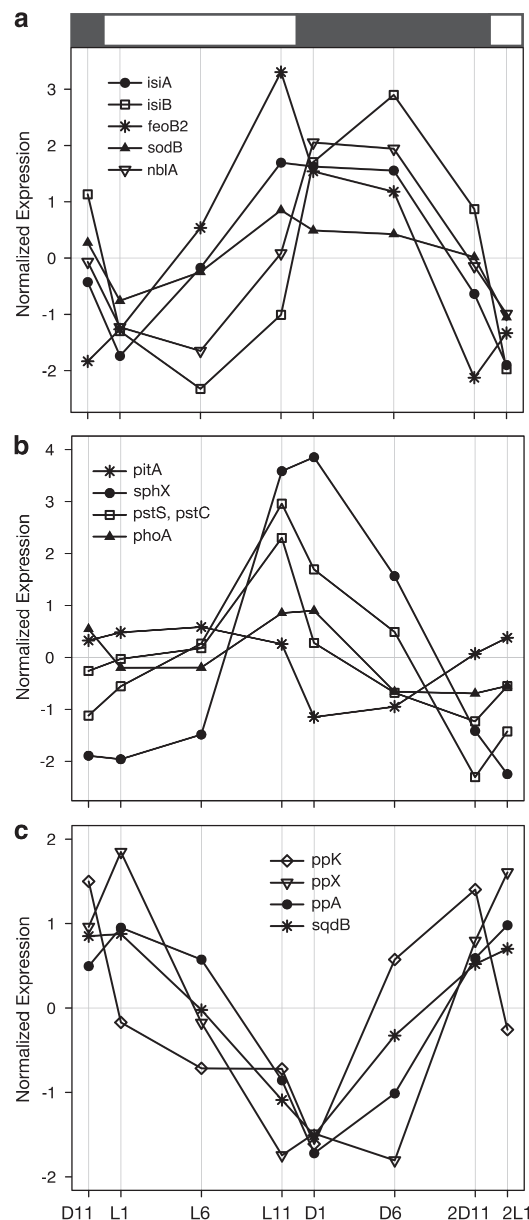

Figure 5 Expression profiles for iron (**a**) and phosphorus (**b, c**) genes. The expression pattern for only one *feoB* gene (CwatDRAFT_3154) is shown. The expression for *pstSC* represents a pattern for whole operon (*pstSCA1A2B1B2*). The gene *sphX* is a *pstS* homolog, and is located ca. 200-bp upstream of *pstSCAB* operon on the complementary strand. Genes *sphX*, *sqdB* and *ppA* were also among the cyclic genes in *Cyanothece* (Stöckel *et al*., 2008). The *X* and *Y* axes follow the legend for Figure 3a.

The *isiA* gene transcripts were also abundant at several locations in natural *C. watsonii* populations in the Pacific and equatorial Atlantic Oceans, as determined by metatranscriptomic and reverse transcription quantitative PCR analysis (Hewson *et al*., 2009).

Simultaneously with the iron genes, transcript abundance of oxidative stress genes significantly increased, including the genes for the phycobilisome degradation protein (*nblA*), peroxidase (*prx*, CwatDRAFT_4095) and superoxide dismutase (*sodB*) (Figure 5a, Supplementary Table S1). The *isiA* and *isiB* genes are also inducible by oxidative stress (Yousef *et al*., 2003; Kanesaki *et al*., 2007). It is noteworthy that in *Cyanothece*, several genes for iron and oxidative stress, three *feoAB* operons, *fur1, isiB* and *nblA,* had expression profiles similar to *C. watsonii* (Stöckel *et al*., 2008). In *C. watsonii*, intracellular iron concentration increases at the onset of darkness because of *de novo* synthesis of nitrogenase (Tuit *et al*., 2004), which requires 38 atoms of iron per holoenzyme (Shi *et al*., 2007). The increased expression of oxidative stress genes may be a response to increased iron uptake, or to the accumulation of reactive oxygen species by the end of the light period, or both.

Phosphate acquisition
Among cyanobacteria, *C. watsonii* show unique adaptations to low P environment with the presence of both low- and high-affinity transport systems and at least four genes for phosphomonoesterases (Dyhrman and Haley, 2006). Nearly all phosphate acquisition and metabolism genes were differentially expressed in *C. watsonii* during the diel cycle (Figures 5b and c); the exception was *phoH* gene encoding phosphate starvation inducible protein (Supplementary Table S2). DNA replication may cause the increased demand for inorganic phosphate in a cell (Klausmeier *et al.*, 2004; Davey *et al.*, 2008) explaining maximum transcript abundance for phosphate acquisition and phosphatase genes during late light—early dark period (Figure 5b), similar to the expression of *pstCAB* and *phoA* genes in *Prochlorococcus* (Zinser *et al.*, 2009). In contrast, only one phosphate transport gene, *pstS,* was cycling in *Cyanothece*, although it had similar to *C. watsonii* maximum expression at L11 (Stöckel *et al.*, 2008). Other phosphorus metabolism genes, including the gene for sulfolipid biosynthesis *sqdB*, were transcribed at a highest level during late dark—early light period (Figure 5c) in parallel with the increased biosynthesis in a cell and beginning of cell division.

Hypothetical proteins
The diel cycling gene pool contained 845 genes annotated as hypothetical or conserved hypothetical proteins. Some of these genes showed significantly increased expression at a specific time point (Table 3). It is noteworthy that expression of some of these genes was also reported in a metatranscriptome study of natural populations in the South Pacific (Hewson *et al.*, 2009) indicating that hypothetical proteins have important cellular functions in *Crocosphaera*. Metatranscriptome analysis

Table 3 Genes encoding hypothetical proteins with peak expression at specific times

L1		*L6*		*L11*	
Gene ID	*LFC*	*Gene ID*	*LFC*	*Gene ID*	*LFC*
CwatDRAFT_0500[a]	6.11	CwatDRAFT_2504[b]	5.75	CwatDRAFT_4081[b]	6.96
CwatDRAFT_1830[b]	6.03	CwatDRAFT_2346[b]	5.20	CwatDRAFT_5119[b]	6.96
CwatDRAFT_5913[b]	5.70	CwatDRAFT_1830[b]	4.90	CwatDRAFT_1413	6.11
CwatDRAFT_0497[a]	5.44	CwatDRAFT_3155	4.51	CwatDRAFT_3790[b]	6.06
CwatDRAFT_2127[b]	5.32	CwatDRAFT_5913[b]	4.50	CwatDRAFT_3254[a]	6.02
CwatDRAFT_0231[a]	5.05	CwatDRAFT_2345[b]	4.34	CwatDRAFT_4068[b]	5.87
CwatDRAFT_0230[a]	4.86	CwatDRAFT_3265[b]	4.21	CwatDRAFT_5461[b]	5.66
CwatDRAFT_5915[b]	4.68	CwatDRAFT_0609[a]	4.15	CwatDRAFT_5544	5.15
CwatDRAFT_0227[b]	4.21	CwatDRAFT_2855[b]	4.15	CwatDRAFT_4592	5.04
CwatDRAFT_0609[b]	4.20	CwatDRAFT_3442[b]	3.89	CwatDRAFT_1822	4.84
CwatDRAFT_3265	4.11	CwatDRAFT_2127[b]	3.80	CwatDRAFT_5545	4.77
CwatDRAFT_2855[b]	4.08	CwatDRAFT_1831	3.75	CwatDRAFT_1972[a]	4.75
CwatDRAFT_6283[b]	4.03	CwatDRAFT_0231[a]	3.74	CwatDRAFT_3978[b]	4.57
CwatDRAFT_1305[b]	4.01	CwatDRAFT_0500[a]	3.64	CwatDRAFT_3424[a]	4.55
CwatDRAFT_6461	3.99	CwatDRAFT_0227[a]	3.59	CwatDRAFT_0845[b]	4.53
D1		*D6*		*D11*	
Gene ID	*LFC*	*Gene ID*	*LFC*	*Gene ID*	*LFC*
CwatDRAFT_1413	9.21	CwatDRAFT_3816[b]	9.44	CwatDRAFT_3816[b]	7.40
CwatDRAFT_3790[b]	8.97	CwatDRAFT_3844[b]	7.96	CwatDRAFT_1830[b]	5.74
CwatDRAFT_4081[b]	8.68	CwatDRAFT_3848[b]	7.69	CwatDRAFT_0382[b]	4.40
CwatDRAFT_5119[b]	7.73	CwatDRAFT_3843[b]	7.30	CwatDRAFT_3843[b]	4.13
CwatDRAFT_5461[b]	7.57	CwatDRAFT_4081[b]	6.67	CwatDRAFT_0830[b]	3.92
CwatDRAFT_5675[b]	7.36	CwatDRAFT_1413	6.20	CwatDRAFT_5642	3.91
CwatDRAFT_0845[b]	6.77	CwatDRAFT_3790[b]	5.89	CwatDRAFT_2504[b]	3.62
CwatDRAFT_1596[a]	6.10	CwatDRAFT_3260	5.56	CwatDRAFT_0500[a]	3.60
CwatDRAFT_6496	5.93	CwatDRAFT_5119[b]	5.29	CwatDRAFT_6283[b]	3.58
CwatDRAFT_4068[b]	5.63	CwatDRAFT_5675[b]	5.18	CwatDRAFT_0497[a]	3.42
CwatDRAFT_4237[b]	5.59	CwatDRAFT_5461[b]	5.18	CwatDRAFT_3844[b]	3.26
CwatDRAFT_1601[b]	5.53	CwatDRAFT_0845[b]	5.12	CwatDRAFT_3586[b]	3.17
CwatDRAFT_5544	5.52	CwatDRAFT_3831	5.03	CwatDRAFT_6306[b]	2.87
CwatDRAFT_4537[b]	5.46	CwatDRAFT_0061[b]	5.03	CwatDRAFT_6677[a]	2.86
CwatDRAFT_4544	5.46	CwatDRAFT_1596[a]	4.63	CwatDRAFT_6437	2.85

[a]Genes are unique to *C. watsonii.*
[b]Homologous genes are cyclic in *Cyanothece*.
LFC—log (base-2) fold change between maximum and minimum expressions. D and L stand for dark and light, respectively, and the corresponding hour in light or dark cycles is indicated.

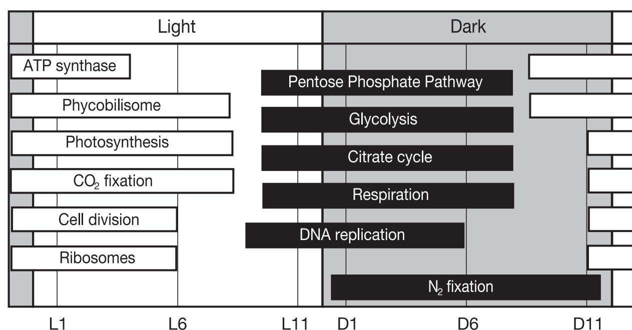

Figure 6 Temporal separation of cellular processes and metabolisms in *C. watsonii* during the dark–light cycles. The timing was set based on the time course when expression of the corresponding genes was higher than the mean expression at all time points. The shaded area represents the dark period.

and reverse transcription quantitative PCR revealed high expression of the CwatDRAFT_0061 gene in samples from the South Pacific during a *Crocosphaera* bloom (Hewson *et al.*, 2009). In this study, the expression of CwatDRAFT_0061 gene was enhanced 32-fold in the middle of the dark period (Table 3). Another hypothetical protein encoded by the CwatDRAFT_3843 gene in the *nif* cluster had the highest transcript abundance at the same time as nitrogen fixation-related genes (D6), and the transcript was abundant in the metatranscriptome from the South Pacific (Hewson *et al.*, 2009). The hypothetical proteins deserve further study, and at least some of them could represent important potential targets for environmental surveys.

Conclusions

Major metabolic pathways and cell processes in *C. watsonii* are co-regulated on transcriptional level similar to what was observed in *Cyanothece* (Stöckel *et al.*, 2008), and the patterns summarized in Figure 6, except for DNA replication, appear to be general for unicellular N_2-fixing cyanobacteria. However, the percentage of genes with diel expression patterns in *C. watsonii* was significantly higher than that in *Cyanothece*. This includes genes for ABC transporters (for example, P and Co transport), DNA replication and repair, and genes encoding hypothetical proteins unique to *C. watsonii*, which all together may reflect the habitats preferences of the two cyanobacteria. The unique genes in both DEP and non-DEP pools of *C. watsonii* would be of great interest for future biochemical and genetic investigations as they may be essential for adaptation of this unicellular N_2-fixing cyanobacterium to oligotrophic environment. Increased expression of iron and oxidative stress genes by the late light period was observed in both *C. watsonii* (this study) and *Cyanothece* (Stöckel *et al.*, 2008), and increased expression of phosphorus genes was observed in both *C. watsonii* (this study) and *Prochlorococcus* (Zinser *et al.*, 2009). Although there is a possibility that in all studies samples were taken in the early stationary phases, the observed differential expression of iron, oxidative and phosphorus stress genes during the diel cycle suggests that sampling time should be carefully chosen when studying stress responses. The diel expression pattern of several alternative sigma factors suggests that these sigma factors may be responsible for the coordinated global gene expression in *C. watsonii* during the light–dark cycles. This pattern is consistent with whole-genome expression analysis in other cyanobacteria (Kucho *et al.*, 2005; Stöckel *et al.*, 2008; Zinser *et al.*, 2009) and with the conclusions from gene mutation studies in *Synechococcus* PCC 7942 (Nair *et al.*, 2002). Results obtained in this study will promote better understanding of physiology and ecology of the unicellular diazotroph, *Crocosphaera*, in oligotrophic ocean.

Acknowledgements

We thank J Waterbury for providing *Crocosphaera watsonii* WH 8501 and E Webb for helping with microarray design. We also thank E Webb, S Dyhrman, I Hewson, K Pennebaker, E Mondragon, P Moisander, B Carter, and UCSC Microbial Environmental Genomics Applications: Modeling, Experimentation and Remote Sensing (MEGAMER)
facility for comments, discussion and technical support. This work was supported by a Gordon and Betty Moore Foundation Marine Microbiology Investigator grant, NSF grant EF0424599 for the Center for Microbial Oceanography: Research and Education (CMORE) and NSF grant OCE-0425363 to J P Z.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* **57**: 289–300.

Berman-Frank I, Lundgren P, Chen YB, Küpper H, Kolber Z, Bergman B *et al.* (2001). Segregation of nitrogen fixation and oxygenic photosynthesis in the marine cyanobacterium *Trichodesmium*. *Science* **294**: 1534–1537.

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ. (1997). *Trichodesmium*: a globally significant marine cyanobacterium. *Science* **276**: 1221–1229.

Chappell PD, Webb EA. (2009). *A molecular assessment of the iron stress response in the two phylogenetic clades of Trichodesmium* Env. Microb. Journal advance online publication 25 August 2009; doi:10.1111/j.1462-2920.2009.02026.x.

Chen Y-B, Dominic B, Mellon M, Zehr JP. (1998). Circadian rhythm of nitrogenase gene expression in the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium* sp. strain IMS 101. *J Bacteriol* **180**: 3598–3605.

Chen Y, Zehr JP, Mellon MT. (1996). 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.

Church MJ, Short CM, Jenkins BD, Karl DM, Zehr JP. (2005). Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. *Appl Environ Microbiol* **71**: 5362–5370.

Compaoré J, Stal LJ. (2009). *Oxygen and the light-dark cycle of nitrogenase activity in two unicellular cyanobacteria*, Environ Microbiol, Journal advance online publication 19 August 2009; doi: 10.1111/j.1462-2920.2009.02034.x.

Davey M, Tarran GA, Mills MM, Ridame C, Geider RJ, LaRoche J. (2008). Nutrient limitation of picophytoplankton photosynthesis and growth in the tropical North Atlantic. *Limnol Oceanogr* **53**: 1722–1733.

Dyhrman ST, Haley ST. (2006). Phosphorus scavenging in the unicellular marine diazotroph *Crocosphaera watsonii*. *Appl Env Microbiol* **72**: 1452–1458.

Falcón LI, Carpenter EJ, Cipriano F, Bergman B, Capone DG. (2004). N_2 fixation by unicellular bacterioplankton from the Atlantic and Pacific Oceans: Phylogeny and *in situ* rates. *Appl Environ Microbiol* **70**: 765–770.

Falcón LI, Cipriano F, Chistoserdov AY, Carpenter EJ. (2002). Diversity of diazotrophic unicellular cyanobacteria in the tropical North Atlantic Ocean. *Appl Environ Microbiol* **68**: 5760–5764.

Falcón LI, Pluvinage S, Carpenter EJ. (2005). Growth kinetics of marine unicellular N_2-fixing cyanobacterial isolates in continuous culture in relation to phosphorus and temperature. *Mar Ecol Prog Ser* **285**: 3–9.

Fay P. (1992). Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* **56**: 340–373.

Foster RA, Zehr JP. (2006). Characterization of diatom-cyanobacteria symbioses on the basis of *nifH*, *hetR*, and 16S rRNA sequences. *Environ Microbiol* **8**: 913–1925.

Fu F, Mulholland MR, Garcia NS, Beck A, Bernhardt PW, Warner ME *et al.* (2008). Interactions between changing pCO_2, N_2 fixation, and Fe limitation in the marine unicellular cyanobacterium *Crocosphaera*. *Limnol Oceanogr* **53**: 2472–2484.

Gallon JR. (1992). Reconciling the incompatible—N_2 fixation and O_2. *New Phytol* **122**: 571–609.

Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S (eds). (2005). *Bioinformatics and Computaional Biology Solutions Using R and Bioconductor*. Springer: New York, NY.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**: R80.1–R80.16.

Herrero A, Flores E (eds). (2008). *The Cyanobacteria: Molecular Biology, Genomics, and Evolution*. Caister Academic Press: Norfolk, UK.

Hewson I, Poretsky RS, Beinart RA, White AE, Shi T, Bench SR *et al.* (2009). *In situ* transcriptomic analysis of the globally important keystone N_2-fixing taxon *Crocosphaera watsonii*. *ISME J* **3**: 618–631.

Howarth RW, Marino R, Lane J, Cole JJ. (1988). Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 1. Rates and importance. *Limnol Oceanogr* **33**: 669–687.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U *et al.* (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat* **4**: 249–264.

Kanesaki Y, Yamamoto H, Paithoonrangsarid K, Shoumskaya M, Suzuki I, Hayashi H *et al.* (2007). Histidine kinases play important roles in the perception and signal transduction of hydrogen peroxide in the cyanobacterium, *Synechocystis* sp. PCC 6803. *Plant J* **49**: 313–324.

Karl DM, Michaels A, Bergman B, Capone DG, Carpenter EJ, Letelier R *et al.* (2002). Dinitrogen fixation in the world's oceans. *Biogeochem* **57/58**: 47–98.

Katoh H, Hagino N, Grossman AR, Ogawa T. (2001). Genes essential to iron transport in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **183**: 2779–2784.

Klausmeier CA, Litchman E, Levin SA. (2004). Phytoplankton growth and stoichiometry under multiple nutrient limitation. *Limnol Oceanogr* **49**: 1463–1470.

Kucho K, Okamoto K, Tsuchiya Y, Nomura S, Nango M, Kanehisa M *et al.* (2005). Global analysis of circadian expression in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **187**: 2190–2199.

Labiosa RG, Arrigo KR, Tu CJ, Bhaya D, Bay S, Grossman AR *et al.* (2006). Examination of diel changes in global transcript accumulation in *Synechocystis* (cyanobacteria). *J Phycol* **42**: 622–636.

Langlois RJ, LaRoche J, Raab PA. (2005). Diazotrophic diversity and distribution in the tropical and subtropical Atlantic Ocean. *Appl Environ Microbiol* **71**: 7910–7919.

LaRoche J, Breitbarth E. (2005). Importance of the diazotrophs as a source of new nitrogen in the ocean. *J Sea Res* **53**: 67–91.

Laudenbach DE, Reith ME, Straus NA. (1988). Isolation, sequence analysis, and transcriptional studies of the flavodoxin gene from *Anacystis nidulans* R2. *J Bacteriol* **170**: 258–265.

Laudenbach DE, Straus NA. (1988). Characterization of a cyanobacterial iron stress-induced gene similar to *psbC*. *J Bacteriol* **170**: 5018–5026.

Leitão E, Oxelfelt F, Oliveira P, Moradas-Ferreira P, Tamagnini P. (2005). Analysis of the *hupSL* operon of the nonheterocystous cyanobacterium *Lyngbya majuscula* CCAP 1446/4: Regulation of transcription and expression under a light-dark regimen. *Appl Environ Microbiol* **71**: 4567–4576.

Los DA, Suzuki I, Zinchenko VV, Murata N. (2008). Stress responses in *Synechocystis*: regulated genes and regulatory systems. In: Herrero A, Flores E (eds). *The Cyanobacteria:Molecular Biology, Genomics, and Evolution*. Caister Academic Press: Wymondham, UK. pp 117–159.

Mazard SL, Fuller NJ, Orcutt KM, Bridle O, Scanlan DJ. (2004). PCR analysis of the distribution of unicellular cyanobacterial diazotrophs in the Arabian Sea. *Appl Environ Microbiol* **70**: 7355–7364.

Michaels AF, Olson D, Sarmiento JL, Ammerman JW, Fanning K, Jahnke R *et al.* (1996). Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean. *Biogeochem* **35**: 181–226.

Miller MB, Tang YW. (2009). Basic concepts of microarrays and potential applications in clinical microbiology. *Clin Microbiol Reviews* **22**: 611–633.

Montoya JP, Holl CM, Zehr JP, Hansen A, Villareal TA, Capone DG. (2004). High rates of N_2 fixation by

unicellular diazotrophs in the oligotrophic Pacific Ocean. *Nature* **430**: 1027–1032.

Mori T, Johnson C. (2001). Circadian programming in cyanobacteria. *Semin Cell Dev Biol* **12**: 271–278.

Nair U, Ditty JL, Min HT, Golden SS. (2002). Roles for sigma factors in global circadian regulation of the cyanobacterial genome. *J Bacteriol* **184**: 3530–3538.

Osanai T, Kanesaki Y, Nakano T, Takahashi H, Asayama M, Shirai M *et al.* (2005). Positive regulation of sugar catabolic pathways in the cyanobacterium *Synechocystis* sp. PCC 6803 by the group 2 sigma factor *sigE*. *J Biol Chem* **280**: 30653–30659.

Pakrasi HB, Riethman HC, Sherman LA. (1985). Organization of pigment proteins in the photosystem II complex of the cyanobacterium *Anacystis nidulans* R2. *Proc Natl Acad Sci USA* **82**: 6903–6907.

Postgate JR. (1998). *Nitrogen Fixation*. 3rd edn. The Press Syndicate of the University of Cambridge, Cambridge, UK.

Schneegurt MA, Sherman DM, Nayar S, Sherman LA. (1994). Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium *Cyanothece* sp. strain ATCC 51142. *J Bacteriol* **176**: 1586–1597.

Shi T, Sun Y, Falkowski PG. (2007). Effects of iron limitation on the expression of metabolic genes in the marine cyanobacterium *Trichodesmium erythraeum* IMS101. *Environ Microbiol* **9**: 2945–2956.

Singh AK, McIntyre LM, Sherman LA. (2003). Microarray analysis of the genome-wide response to iron deficiency and iron reconstitution in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* **132**: 1825–1839.

Smyth GK. (2005). Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (eds). *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer: New York. pp 397–420.

Staal M, Lintel-Hekkert ST, Harren F, Stal L. (2001). Nitrogenase activity in cyanobacteria measured by the acetylene reduction assay: a comparison between batch incubation and on-line monitoring. *Environ Microbiol* **3**: 343–351.

Stal LJ, Zehr JP. (2008). Cyanobacterial nitrogen fixation in the ocean: diversity, regulation and ecology. In: Herrero A, Flores E (eds). *The Cyanobacteria: Molecular Biology, Genomics and Evolution*. Caister Academic Press: Norfolk, UK. pp 423–446.

Stöckel J, Welsh EA, Liberton M, Kunnvakkam R, Aurora R, Pakrasi HB. (2008). Global transcriptomic analysis of *Cyanothece* 51142 reveals robust diurnal oscillation of central metabolic processes. *Proc Natl Acad Sci USA* **105**: 6156–6161.

Summerfield TC, Sherman LA. (2007). Role of sigma factors in controlling global gene expression in light/dark transitions in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **189**: 7829–7840.

Tamagnini P, Axelsson R, Lindberg P, Oxelfelt F, Wunschiers R, Lindblad P. (2002). Hydrogenases and hydrogen metabolism of cyanobacteria. *Microbiol Mol Biol Rev* **66**: 1–20.

Tamagnini P, Leitao E, Oliveira P, Ferreira D, Pinto F, Harris DJ *et al.* (2007). Cyanobacterial hydrogenases: diversity, regulation and applications. *FEMS Microbiol Rev* **31**: 692–720.

Toepel J, McDermott JE, Summerfield TC, Sherman LA. (2009). Transcriptional analysis of the unicellular, diazotrophic cyanobacterium *Cyanothece* sp. ATCC 51142 grown under short day/night cycles. *J Phycol* **45**: 610–620.

Toepel J, Welsh E, Summerfield TC, Pakrasi H, Sherman LA. (2008). Differential transcriptional analysis of the cyanobacterium *Cyanothece* sp. Strain ATCC 51142 during light-dark and continuous-light growth. *J Bacteriol* **190**: 3904–3913.

Tuit C, Waterbury J, Ravizza G. (2004). Diel variation of molybdenum and iron in marine diazotrophic cyanobacteria. *Limnol Oceanogr* **49**: 978–990.

Vicente M, Rico AI, Martinez-Arteaga R, Mingorance J. (2006). Septum enlightenment: assembly of bacterial division proteins. *J Bacteriol* **188**: 19–27.

Villareal TA. (1992). Marine nitrogen-fixing diatom—cyanobacteria symbioses. In: Carpenter EJ, Capone DG, Rueter JG (eds). *Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs*. Kluwer Academic Publishers: The Netherlands. pp 163–175.

Waterbury JB, Rippka R. (1989). Cyanobacteria. Subsection I. Order *Chroococcales*. In: Krieg NR, Holt JB (eds). *Bergey's Manual of Systematic Bacteriology*. Vol.3. Williams & Wilkins: Baltimore. pp 1728–1746.

Webb EA, Ehrenreich IM, Brown SL, Valois FW, Waterbury JB. (2009). Phenotypic and genotypic characterization of multiple strains of the diazotrophic cyanobacterium, *Crocosphaera watsonii*, isolated from the open ocean. *Env Microbiol* **11**: 338–348.

Welsh EA, Liberton M, Stöckel J, Loh T, Elvitigala T, Wang C *et al.* (2008). The genome of *Cyanothece* 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. *Proc Natl Acad Sci USA* **105**: 15094–15099.

Whitton BA, Potts M. (2000). Introduction to the Cyanobacteria. In: Whitton BA, Potts M (eds). *The Ecology of Cyanobacteria*. Kluwer Academic Publishers: The Netherlands. pp 1–11.

Yousef N, Pistorius EK, Michel K-P. (2003). Comparative analysis of *idiA* and *isiA* transcription under iron starvation and oxidative stress in *Synechococcus elongatus* PCC 7942 wildtype and selected mutants. *Arch Microbiol* **180**: 471–483.

Zehr JP, Bench SR, Mondragon EA, McCarren J, DeLong EF. (2007a). Low genomic diversity in tropical oceanic N_2-fixing cyanobacteria. *Proc Natl Acad Sci USA* **104**: 17807–17812.

Zehr JP, Montoya JP, Jenkins BD, Hewson I, Mondragon E, Short CM *et al.* (2007b). Experiments linking nitrogenase gene expression to nitrogen fixation in the North Pacific subtropical gyre. *Limnol Oceanogr* **52**: 169–183.

Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF *et al.* (2001). Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. *Nature* **412**: 635–638.

Zinser ER, Lindell D, Johnson ZI, Futschik ME, Steglich C, Coleman ML *et al.* (2009). Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph, *Prochlorococcus*. *PLoS ONE* **4**: e5135.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)