

## ORIGINAL ARTICLE

# Non-coding RNAs in marine *Synechococcus* and their regulation under environmentally relevant stress conditions

Gregor Gierga, Björn Voss and Wolfgang R Hess  
Faculty of Biology, University of Freiburg, Freiburg, Germany

**Regulatory small RNAs (sRNAs) have crucial roles in the adaptive responses of bacteria to changes in the environment. Thus far, potential regulatory RNAs have been studied mainly in marine picocyanobacteria in genetically intractable *Prochlorococcus*, rendering their molecular analysis difficult. *Synechococcus* sp. WH7803 is a model cyanobacterium, representative of the picocyanobacteria from the mesotrophic areas of the ocean. Similar to the closely related *Prochlorococcus* it possesses a relatively streamlined genome and a small number of genes, but is genetically tractable. Here, a comparative genome analysis was performed for this and four additional marine *Synechococcus* to identify the suite of possible sRNAs and other RNA elements. Based on the prediction and on complementary microarray profiling, we have identified several known as well as 32 novel sRNAs. Some sRNAs overlap adjacent coding regions, for instance for the central photosynthetic gene *psbA*. Several of these novel sRNAs responded specifically to environmentally relevant stress conditions. Among them are six sRNAs changing their accumulation level under cold stress, six responding to high light and two to iron limitation. Target predictions suggested genes encoding components of the light-harvesting apparatus as targets of sRNAs originating from genomic islands and that one of the iron-regulated sRNAs might be a functional homolog of RyhB. These data suggest that marine *Synechococcus* mount adaptive responses to these different stresses involving regulatory sRNAs.**

*The ISME Journal* (2012) 6, 1544–1557; doi:10.1038/ismej.2011.215; published online 19 January 2012

**Subject Category:** integrated genomics and post-genomics approaches in microbial ecology

**Keywords:** cyanobacteria; gene expression regulation; light stress; regulatory RNA

## Introduction

Regulatory small RNA molecules are involved in all major adaptive responses in bacteria. Among the controlled processes is quorum sensing (Lenz *et al.*, 2004; Tu *et al.*, 2008, 2010; Sonnleitner *et al.*, 2011), the response to iron limitation (Wilderman *et al.*, 2004; Masse *et al.*, 2005), low temperature (Sledjeski *et al.*, 1996), photosynthesis under iron limitation (Dühring *et al.*, 2006), pathogenicity and virulence (Lenz *et al.*, 2004; Pfeiffer *et al.*, 2007). Regulatory RNAs in prokaryotes are very diverse and can act through a multitude of different mechanisms (Hess and Marchfelder, 2011). The majority of regulatory RNAs recognize their target mRNAs by base pairing, frequently sequestering the ribosome-binding site or start codon (AUG) sequences, repressing gene expression by modulating the initiation of

translation or mRNA stability (Waters and Storz, 2009). However, several examples have shown that the window for target recognition can also be located upstream or downstream of ribosome-binding site/AUG sequences (Vecerek *et al.*, 2007; Bouvier *et al.*, 2008; Desnoyers *et al.*, 2009; Pfeiffer *et al.*, 2009), and that some small RNA (sRNA)–mRNA interactions lead to an activation of gene expression rather than repressing it. Consequently, hundreds of such potential riboregulators have been predicted in various groups of bacteria, but their actual verification and functional characterization is still mostly limited to human pathogens, mainly enterobacteria, *Staphylococcus aureus* and *Pseudomonas* species, whereas environmentally relevant bacteria are much less investigated in this respect.

Marine picocyanobacteria numerically dominate the euphotic zone of most oceanic waters and are widely recognized as being major contributors to the marine primary production and global photosynthetic biomass. There are only two genera of marine picocyanobacteria, *Prochlorococcus* (Partensky *et al.*, 1999) and *Synechococcus* (Scanlan *et al.*, 2009). Both genera comprise a number of different clades or ecotypes (Coleman and Chisholm, 2007), which exhibit distinct ecophysiologicals,

Correspondence: WR Hess, Faculty of Biology, University of Freiburg, Schaenzlestr. 1, Freiburg D-79104, Germany.  
E-mail: Wolfgang.Hess@biologie.uni-freiburg.de

Microarray data associated with this manuscript can be downloaded from the GEO database with the accession number GSE28263.

Received 26 September 2011; revised 14 December 2011; accepted 14 December 2011; published online 19 January 2012

geographical and spatial distributions (Johnson *et al.*, 2006; Zwirgmaier *et al.*, 2008). The genetic and regulatory basis behind these patterns is only partially understood (Coleman and Chisholm, 2007; Scanlan *et al.*, 2009).

Previously unrecognized putative sRNAs have been detected in microbial metatranscriptomic data sets from the ocean (Shi *et al.*, 2009), including several such sRNAs most likely belonging to *Prochlorococcus*. Indeed, based on the total genome sequences of four marine cyanobacteria and using comparative genomics, systematic searches for sRNAs in three laboratory strains of *Prochlorococcus* and one *Synechococcus* identified seven different sRNAs in these marine cyanobacteria (Axmann *et al.*, 2005), which were called Yfr1–7 for *c*Yano-bacterial Functional RNA. Complementing this approach by using high-density microarrays, additional 14 sRNAs and 24 *cis*-antisense RNAs were found in *Prochlorococcus* sp. MED4 (Steglich *et al.*, 2008). On the basis of RNA profiling and detailed biochemical experiments, the involvement of several of these transcripts was suggested in light stress adaptation or the response to phage infection (Stazic *et al.*, 2011), consistent with the location of most sRNA genes in highly variable genomic islands (Steglich *et al.*, 2008). Whereas these data indicate that a whole class of regulatory genes and their functions are still to be identified in marine cyanobacteria, their further functional characterization is hampered by the fact that most *Prochlorococcus*, including the MED4 strain, are not genetically tractable. Therefore, we here set out to identify the suite of potentially regulatory sRNAs in *Synechococcus* sp. WH7803, a strain which can be manipulated (Brahamsha, 1996). Crucial for the successful computational prediction of sRNA genes is the availability of well-annotated genome sequences. The more than 20 genome sequences available for marine *Synechococcus* and *Prochlorococcus* (Kettler *et al.*, 2007; Scanlan *et al.*, 2009) provide an excellent data set for this purpose. Although phylogenetically closely related, these strains differ by their pigmentation and light-harvesting complexes and span the complete rRNA diversity (96% to 99.9% similarity) among the cultured representatives from the marine picocyanobacterial clade.

## Materials and methods

### *Cultivation of Synechococcus sp. WH7803*

*Synechococcus* sp. WH7803 was grown in ASW medium and total RNA was isolated as previously described (Axmann *et al.*, 2005). The cyanobacteria were exposed to several stresses. For high light stress, cultures were transferred from 30  $\mu\text{mol quanta m}^{-2} \text{s}^{-1}$  to 280  $\mu\text{mol quanta m}^{-2} \text{s}^{-1}$  white light for 1 h. *Synechococcus* was cold stressed by incubating the culture flask in a water tank with

10 °C cold running water for 2 h. To expose the culture to iron or phosphate limitation, the cells were spun twice for 10 min at room temperature at 10 000 g. Each time, the media were replaced and the cells were resuspended in the corresponding depleted media. For iron stress, ASW without  $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$  and in the case of phosphate stress, medium without  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$  was prepared. RNA from iron-limited cells was extracted 12 h after initiating stress conditions. Cyanobacteria growing without added phosphate were collected after 24 h. We began all experiments with an exponentially growing culture diluted to a titer of  $\sim 1 \times 10^7$  cells per ml, before applying the respective stress. Control cultures were treated like the stressed ones in terms of centrifugation and dilution but then incubated under standard conditions.

### *RNA extraction and analysis*

*Synechococcus* were filtered rapidly onto Supor 0.45  $\mu\text{m}$  membrane filters (Pall, New York, NY, USA). The filters were directly placed in RNA resuspension buffer (10 mM sodium acetate, 200 mM sucrose, 5 mM EDTA), flash frozen in liquid nitrogen and stored at  $-80$  °C until extraction. The frozen material was thawed under addition of 3 vol Z6 buffer (8 M guanidinium hydrochloride, 50 mM mercaptoethanol, 20 mM EDTA, 20 mM MES, pH 7 (NaOH)) for 60 min at room temperature and gentle shaking. Acidic phenol was added (0.5 vol) and the cells were incubated for 15 min in a 60 °C water bath with occasional shaking by hand, followed by extraction with 0.5 vol of chloroform:isoamyl alcohol (24:1). After shaking, the phases were separated by centrifugation for 5 min. The aqueous phase was extracted and purified by phase extraction with 1 vol phenol:chloroform:isoamyl alcohol (25:24:1) and finally 1 vol chloroform:isoamyl alcohol (24:1). The RNA in the aqueous phase was precipitated with 1 vol isopropanol at  $-20$  °C overnight. After centrifugation at 13 000 g for 30 min at 4 °C the pellet was washed with 70% ethanol and subsequently centrifuged at 13 000 g for 5 min at 4 °C. All liquid was removed and the pellet dried for 10 min, dissolved in autoclaved water and stored at  $-80$  °C. Quality of RNA was checked visually on 1.5% denaturing agarose gels stained with ethidium bromide.

High resolution northern hybridization and 5'RACE experiments were performed from the separation of 5 to 50  $\mu\text{g}$  of total RNA as described by Georg *et al.* (2009). Single-stranded RNA probes were constructed by *in vitro* transcription from the T7 promoter as described (Georg *et al.*, 2009). DNA probes for the detection of transcript levels of control genes for cold stress and high light conditions were labelled by random priming. 12–25 ng of PCR fragment in 5  $\mu\text{l}$  of aqueous solution were mixed with 7  $\mu\text{l}$  of 1 M Hepes-KOH (pH = 6.7, final concentration 20  $\mu\text{M}$ ; final volume 35  $\mu\text{l}$ ). A volume

of 2  $\mu$ l of random hexamer primer (100  $\mu$ M, final concentration 5.7  $\mu$ M) (Fermentas, St. Leon-Rot, Germany), 4.5  $\mu$ l of TE buffer and 8  $\mu$ l H<sub>2</sub>O were added. The solution was denatured for 10 min at 95 °C and cooled for 15 min at room temperature. The preparation was continued by the addition of 3.5  $\mu$ l mercaptoethanol (100 mM, final concentration 10 mM), 2.5  $\mu$ l dNTPs (200  $\mu$ M of each dATP, dCTP and dGTP, 500 mM Tris-HCl pH 7.8, 50 mM MgCl<sub>2</sub>), 1  $\mu$ l of Klenow exo- (2 U per  $\mu$ l) (Fermentas) and 2.5  $\mu$ l of [ $\alpha$ -<sup>32</sup>P]dTTP (10  $\mu$ Ci per  $\mu$ l). The solution was incubated for 10 min at 37 °C followed by the addition of 0.5  $\mu$ l of dNTP mix (1 mM of dATP, dTTP, dCTP, dGTP) and incubation for additional 5 min at 37 °C. The reaction was stopped by the addition of 3.5  $\mu$ l of EDTA (0.2 M). The labelled DNA was denatured for 10 min at 95 °C and stored on ice until use. Sequences of all DNA oligonucleotides are given in Supplementary Table S1.

#### Microarray analysis

To facilitate global experimental analysis, a tiling microarray was designed consisting of 12 000 overlapping oligonucleotides for both DNA strands. We chose the 12K Electrasense CombiMatrix (CombiMatrix Diagnostics, Irvine, CA, USA) microarray system for this experiment. The 12 000 overlapping oligonucleotide probes covered almost all intergenic regions (IGRs) of the *Synechococcus* sp. WH7803 genome on both strands with an average probe length of 31.2 nt, resulting in an average overlap of the probes of 12 nt.

DNase treated RNA (4  $\mu$ g) was directly labelled with the Kreatech Diagnostics (Amsterdam, The Netherlands) ULS labelling kit for CombiMatrix arrays with Biotin-ULS, according to the manufacturer's protocol (CombiMatrix Diagnostics). The RNA fragmentation after labelling was performed as advised for aRNA fragmentation reactions by the manufacturer's instructions (Ambion, Austin, TX, USA). The arrays were hybridized with 3  $\mu$ g of labelled, fragmented RNA, scanned and stripped according to the manufacturer's protocol using chemicals provided by CombiMatrix (CombiMatrix Diagnostics). ElectraSense Application Software (CombiMatrix) was used for feature extraction. Advanced analyses were done with the software Genespring GX 10.0.2 (Agilent Technologies, Santa Clara, CA, USA). The feature intensities of all experiments and its controls were quantile normalised. The microarray data are available in the GEO database with the accession number GSE28263.

#### Sequence data and computational analysis

A comparative prediction of ncRNA elements within IGRs was performed as previously described (Voss *et al.*, 2009). In short, all IGRs longer than 50 nt were compared among the five different

genomes using BlastN. The genome sequence for *Synechococcus* sp. WH7803 (Dufresne *et al.*, 2008) was obtained from the Genoscope (<http://www.genoscope.cns.fr/>). All other sequences were obtained from the finished or unfinished genomes website at Genbank <http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi> with the following accession numbers: *Synechococcus* CC9605, NC\_007516; WH8102, NC\_005070; RS9916, NZ\_AAUA000000; *Synechococcus* CC9311, NC\_008319. Based on the results of BlastN, homologous sequence regions were clustered, aligned using ClustalW, and analyzed for structural significance by RNAz (Washietl *et al.*, 2005). Alignments were postprocessed using the tool rnazSelectSeqs.pl with default parameters from the RNAz package. RNAz was applied in a sliding window approach (a step size of 10 nt and different window sizes were used, namely 80, 100, 120, 140 and 160 nt), of which the window with highest probability was selected. Elements were considered high-scoring when they achieved an RNAz probability of  $\geq 0.5$  or their Z-score was  $\leq -2.0$ . The Z-score is a commonly used measure for structure stability. As we had no prior knowledge about the orientation of an sRNA gene (forward or reverse strand), we computed the measure for both possibilities, given as Z and Zrev, respectively.

In order to assess the coding potential of the candidate sRNAs we used RNAcode (Washietl *et al.*, 2011), which predicts the coding potential for segments of a multiple sequence alignment. We applied RNAcode with default parameters to the same alignments as RNAz and restricted the output to the best hit (option -b). The P-value is given in the comment column, where significant predictions are expected to have a  $P < 0.05$ .

## Results

#### Computational prediction of sRNAs in marine *Synechococcus*

Marine *Synechococcus* have developed specific adaptations to cope with different environmental conditions, especially with regard to the availability of various nutrients, light quantity and quality. Based on 16S rRNA sequences, marine *Synechococcus* have been divided into three subclusters, with the major one (subcluster 5.1) consisting of nine genetically distinct clades, I to IX (Scanlan *et al.*, 2009). Here, genome sequences of five marine *Synechococcus* were included (strains CC9311, CC9605, WH 8102, WH 7803 and RS9916), representing clades I, II, III, VI and IX of subcluster 5.1 (Scanlan *et al.*, 2009).

The comparative prediction was performed based on the sequence-structure conservation of possible RNA elements as described before (Voss *et al.*, 2009) and yielded information on putative sRNAs, possible riboswitches, as well as highly structured 5' and 3' mRNA leader and trailer sequences.

This analysis resulted in 156 sequence clusters above threshold ( $P$  or  $P_{rev} > 0.5$ ; (Washietl *et al.*, 2005)) for marine *Synechococcus*. From these, 70 included a direct hit in *Synechococcus* sp. WH7803 and 24 included sequences from all five compared genomes. A summary of the highest scoring clusters is given in Table 1. Comprehensive information on all predicted RNA elements can be accessed at: <http://www.cyanolab.de/paper/suppl/synechococcus/summary.html>.

Short transcripts from IGRs can also be coding because gene modelling for short protein-coding genes is less accurate than for longer genes. Therefore, the RNaCode algorithm (Washietl *et al.*, 2011) was used to differentiate between non-coding and possibly coding regions based on nucleotide substitution patterns in an alignment of at least three homologous sequences. This approach lead to the identification of three candidates for peptide-coding or dual function sRNAs (i519, i1825 and i1971 (Yfr2a)). The latter sRNA is ubiquitous in cyanobacteria (Gierga *et al.*, 2009), a potential peptide-coding function has not been described thus far. Moreover, there is no suitable AUG in Yfr2a, therefore it still is more likely a non-coding sRNA rather than a short mRNA. However, the other two sRNAs are indeed likely coding for short peptides. Their deduced amino acid sequences are presented in Table 2.

#### Verification experiments

To facilitate global experimental analysis, a tiling microarray was designed. This array consisted of 12 000 probes targeting exclusively the non-protein-coding IGRs of the genome (average overlap of 12 nt). To avoid artefacts of cDNA synthesis, RNA was directly fluorescence-labeled without cDNA synthesis (Georg *et al.*, 2009). RNA was isolated from cyanobacteria cultivated under standard growth conditions or applying environmentally relevant stress conditions, low temperature, high light, iron and phosphate limitation. Samples were taken shortly after the initiation of the respective growth condition in order to detect an early response of the cells and to avoid secondary effects. The successful stress induction was verified by tracking the expression of known marker genes.

For the verification of high light stress, the genes *psbA*, *psaA* and *groEL* were chosen, encoding the photosystem II (PSII) D1 protein, the photosystem I (PSI) reaction center protein PsaA and the GroEL chaperone. These choices were based on previous observations that at least one of the four *Synechococcus* sp. WH7803 *psbA* genes is strongly induced by high light (Garczarek *et al.*, 2008), that *psaA* expression is downregulated upon transfer from low to high light in many cyanobacteria (Hihara *et al.*, 2001), whereas *groEL* is upregulated under high light in *Synechocystis* PCC 6803 (Hihara *et al.*, 2001). Marker genes for the detection of cold stress

**Table 1** List of selected top-scoring RNA elements

CLID	Sequence no.	Species					A	Z	Zrev	P	Prev	Location	Comment	Reference
		7803	9311	9605	8102	9916								
98	3	0	0	2	1	0	199	-2.71	-4.17	1	1	GHP-x-CHP		
105	2	1	1	1	1	0	144	-4.11	-3.77	1	1	cpcG-x-CHP		
256	2	1	0	0	0	1	52	-3.98	-2.94	1	1	rplL-x-dnaB		
27	5	1	1	1	1	1	242	-3.85	-2.56	1	1	apcE-x-apcA		
197	4	0	1	1	1	1	365	-3.75	-2.72	1	1	GHP-x-psaA	Expression repressed by high light	i1692 in Table 3
0	3	1	1	0	0	1	104	-2.4	-2.34	1	1	dead-x-wbpW	Possible leader of <i>DeadRNA helicase</i> (this gene is found in marine picocyanobacteria and $\gamma$ proteobacteria)	
141	4	1	0	1	0	1	203	-3.13	-2.44	1	1	purK-x-CHP	6S RNA (Yfr7)	Axmann <i>et al.</i> (2007)
14	5	1	1	1	1	1	396	-2.80	-3.01	1	1	rpl3-x-ndhN	rplCD Operon leader, corresponds to E. coli S10 r-operon	Axmann <i>et al.</i> (2005)
3	3	1	0	1	1	1	188	-2.59	-2.65	1	1	GHP-x-irpB	Possible leader of HrpB RNA helicase	
157	5	1	1	1	1	1	205	-2.04	-1.77	1	1	hik-x-CHP		
66	10	2	2	2	2	2	376	-2.43	-2.98	0.99	1	GHP-x-psbA	Overlaps the 5'UTR of <i>psbA</i>	i298 in Table 3
84	5	1	1	1	1	1	246	-1.66	-2.99	0.69	1	groES-x-atpB	Position consistent with CIRCE element	(Nakamoto <i>et al.</i> , 2003; Voss <i>et al.</i> , 2009)
269	3	1	1	0	0	1	147	-2.3	-2.17	1	0.52	GHP-x-recA		
111	5	1	1	1	1	1	129	-2.23	-2.13	0.96	0.82	guaB-x-trxA	Yfr1 (positive control)	Voss <i>et al.</i> (2007)

Abbreviations: CHP, conserved hypothetical protein; CLID, cluster identity number.

The total number of sequences in each cluster (Seq no.) and the distribution within the five compared genomes (shaded boxes) plus the total alignment (A) length in nt is given. The elements are ordered according to the highest RNAz probability in either forward (P) or reverse (P rev) orientation. The closer the probability to 1.0, the more support for structural conservation as an RNA element. The Z-score is a measure for structure stability. For sRNAs one expects low values ( $< -2.0$ ). As there was no prior knowledge about the orientation of an sRNA gene (forward or reverse strand), we computed RNAz probabilities and Z-scores for both possibilities, given as Z or P and Zrev or Prev, respectively. The analyzed genomes are indicated as follows: 7803, *Synechococcus* WH7803; 9311, *Synechococcus* CC9311; 9605, *Synechococcus* CC9605; 8102, *Synechococcus* WH8102; 9916, *Synechococcus* RS9916.

**Table 2** Newly identified protein-coding genes or gene segments

ID	sRNA	Start	End	S	TSS	-10	Sequence	Comments
Norf1	i855	928244	928101	r	928264	TCTATT	MLRFGDRDRLTLEALTMWANHQQLPRKASDRSARRPR FRTSRVR	Conserved in <i>Synechococcus</i> RS9916, RCC307 and CC9311
Norf2	i519	613693	613804	f	613670	GACAGT	MQCQGLSQTGEPMGKPFWFLRVFPDGLHCSDNSL	Conserved in <i>Synechococcus</i> CC9311
Norf3	i1825	2009119	2009220	f	2009096	TAGAAC	MATLRFAPLIVRLAAMGALSLLLLGLVLELLS	Conserved in <i>Synechococcus</i> WH7803, CC9311, RS9916, RS9917, BL107
Syn WH7803_0111	i95	117380	117054	r		(Extension of annotated N terminus by 109 residues)	MDNIQLRFLGRLELPCAESISYSSSTAFVYGGSSLYIVDLS NPFSPSRDITINIGDDIQSVAINANGLLALAVGKEANVGS TQVSENIIGEVQFFEWGSSLSAGE	Close homologs in <i>Pelodictyon phaeoclathratiforme</i> , <i>Chlorobium phaeobacteroides</i> DSM 266, <i>Synechococcus</i> WH 8102 and <i>Thalassiosira pseudonana</i> CCMP1335

Abbreviations: sRNA, small RNAs; TSS, transcriptional start site.

The respective putative protein name (ID), start and end coordinates, TSS, and putative peptide sequence is given, together with strand orientation (S) forward (f) or reverse (r).

were *synWH7803\_1511*, *synWH7803\_0505* and *synWH7803\_2417*, encoding an ATP-dependent RNA helicase, the phycobilisome rod-core linker polypeptide CpcG and one of two *desC*-like desaturase genes. Homologs of these genes were previously found to be upregulated under cold stress (the RNA helicase CrhL (Inaba *et al.*, 2003) and *desC* (Los and Murata, 1999)) or repressed (CpcG). Marker gene for phosphate limitation was *pstSII*, well known to be upregulated under this condition in *Synechococcus* sp. WH7803 (Scanlan *et al.*, 1993, 1997), whereas the iron-stress-induced gene *futA* (Webb *et al.*, 2001; Singh *et al.*, 2003) served as marker for iron limitation. These experiments confirmed the intended effects of gene expression under the respective growth conditions (Supplementary Figure 1).

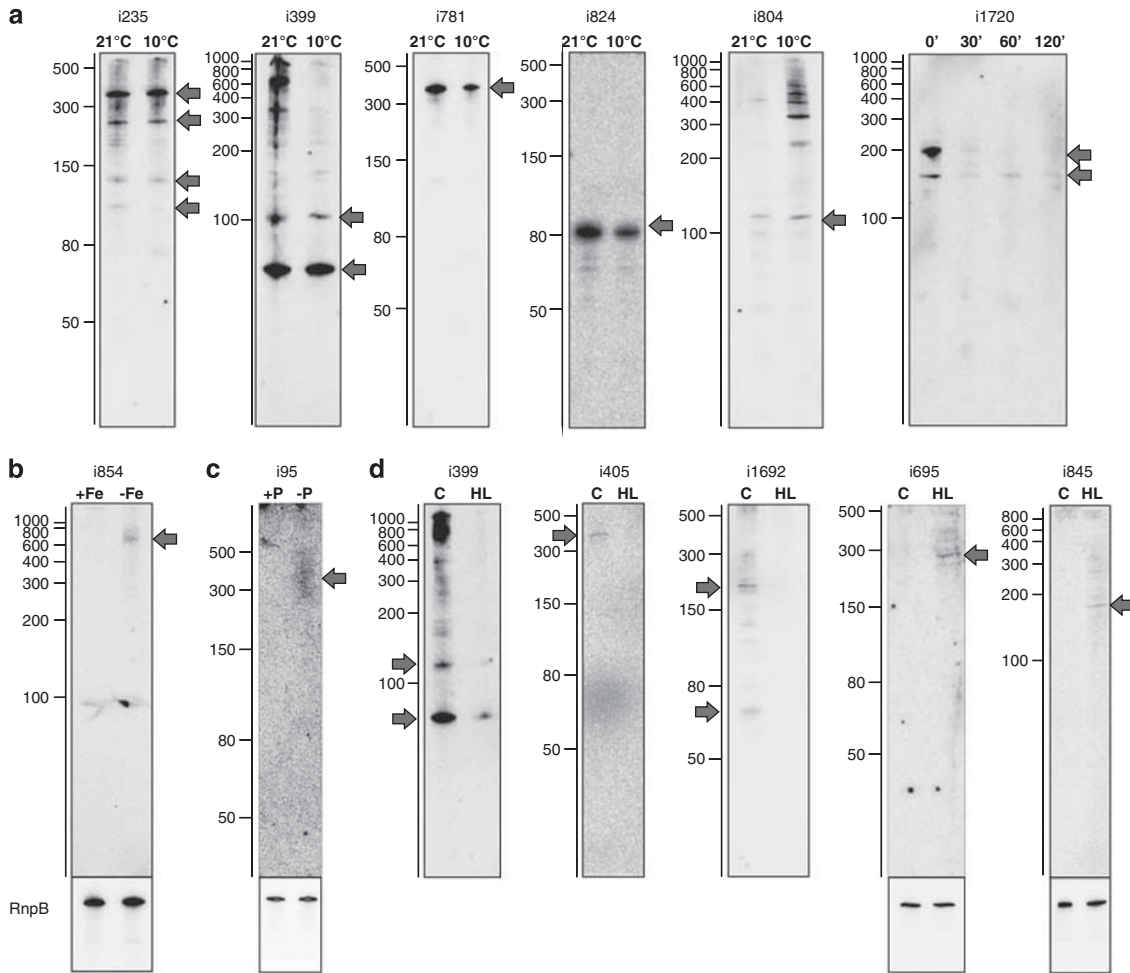
#### Validation of novel transcripts from IGRs and their regulation

RNA from the following conditions was directly labelled and hybridized to the microarrays: light shift from 30  $\mu$ E to 280  $\mu$ E for 1 h, temperature shift from 21  $^{\circ}$ C to 10  $^{\circ}$ C for 2 h, iron and phosphate limitation for 24 h each.

The median intensities of all experiments and controls were quantile normalised, *P*-values calculated by application of an unpaired *t*-test and corrected according to Benjamini–Hochberg adjustment using the Genespring GX 10.0.2 software (Agilent Technologies). The microarray data are available in the GEO database with the accession number GSE28263.

The detected sRNA candidates were validated by Northern blots (Figures 1 and 2) and in many cases by additional 5' RACE experiments. We found six sRNAs differentially regulated under cold stress, six ncRNAs differentially regulated under high light stress, two sRNAs responding to iron limitation and one transcript differentially regulated under phosphate limitation. The homolog of the 6S RNA of *Escherichia coli* (previously called Yfr7 in marine cyanobacteria (Axmann *et al.*, 2005, 2007) was also detected, but revealed no variation in RNA levels. Probes for tmRNA (*ssrA*) and *rnpB* were not included on the microarray because of the good characterization of these RNAs in the literature. Subsequent to the Northern experiments, the analyses of the IGRs were complemented by 5'RACE experiments.

Nineteen additional sRNAs exhibited an overall high expression, verified by northern hybridization (Figure 2), but their accumulation did not change under the tested conditions. There are several bacterial sRNAs known with constitutive high expression that are physiologically very relevant (for example, tmRNA, RNase P RNA or the 4.5S RNA component of the signal recognition particle) but in this work we focused on those which were differentially expressed as these are more likely to have a



**Figure 1** Verification of microarray data by northern hybridization of differentially expressed sRNAs from IGRs. **(a)** Transcript levels for sRNAs affected by cold stress. 21 °C marks the control, 10 °C the cold stress conditions and for i1720, a time series is shown with transcript levels after 0, 30, 60 and 120 min of incubation. **(b)** Levels of sRNA affected by iron limitation. **(c)** Phosphate limitation. **(d)** High light. Major sRNA bands are marked by an arrow. Size markers are given to the left. Control hybridizations against the RNase P RNA (RnpB) are shown at the bottom.

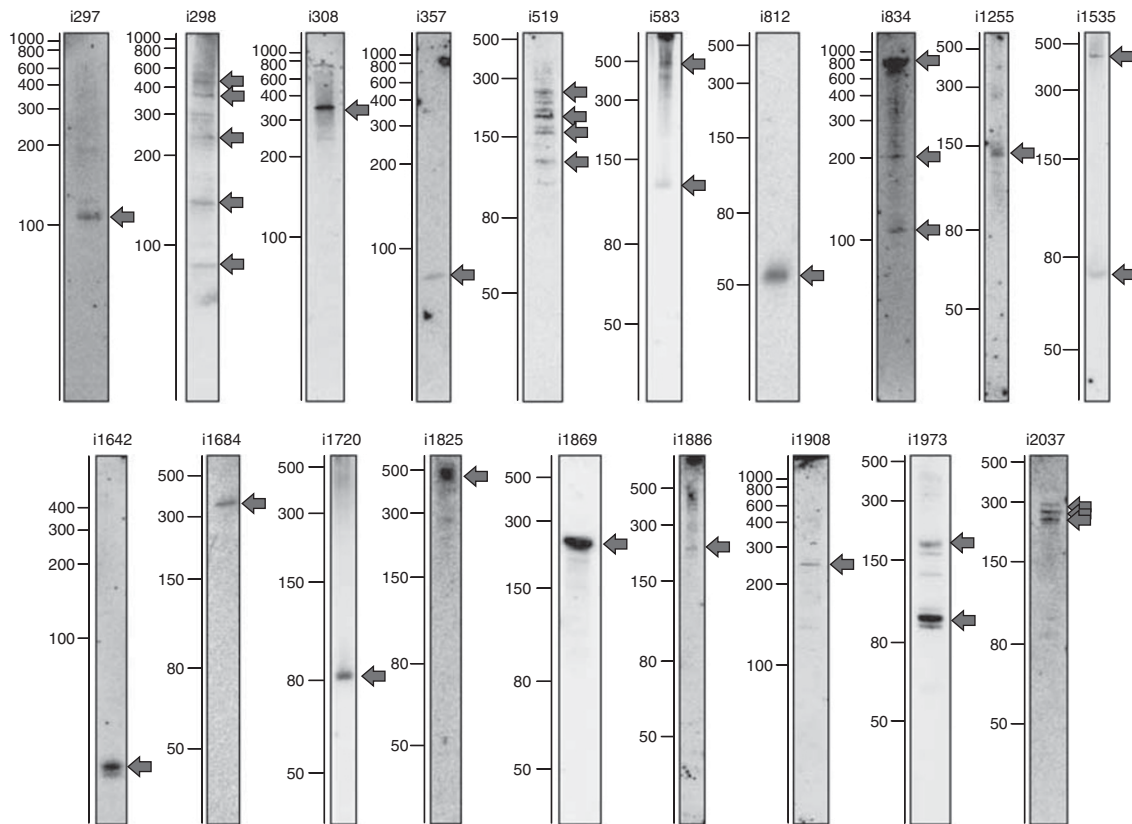
regulatory function. All verified sRNAs are summarized in Table 3 together with additional data, such as genome location, orientation and transcriptional start site (TSS), if determined.

One of the identified sRNAs, i298, is located antisense to the 5' UTR of *psbA*. The hybridization results, with several bands from ~80 to about 500 nt (Figure 2), resemble more the pattern observed for bacterial *cis*-antisense RNAs (Georg *et al.*, 2009) rather than that of a *trans*-acting sRNA originating from an intergenic spacer. Indeed, the TSS of i298 is located in antisense orientation within the coding region of *psbA*, overlapping 109 of its codons. Another, very short sRNA (i1642) was found partly antisense to the 5' end of *kaiA* (Figure 2), encoding one of three core components of the cyanobacterial circadian clock (Ishiura *et al.*, 1998). Like these examples, many identified sRNAs offer an interesting starting point for more detailed analyses.

#### *Differentially expressed sRNAs in Synechococcus sp. WH7803*

Figure 3 illustrates the identification of differentially expressed sRNAs by microarray analysis. The sRNA i403 originates 3' to the gene *mpeU* (Figure 3) coding for a bilin biosynthesis protein, likely involved in binding phycobilins to one or both types of phycoerythrin in this organism (Six *et al.*, 2007). The Northern blot verified a clear increase in transcript levels after exposure to high light for two hours. The sRNA i403 has a size of ~60 nt. This sRNA is located in a genomic island together with several genes encoding phycobilisome rod components (Dufresne *et al.*, 2008). As changes in RNA levels under high light are likely for genes associated with light harvesting, this sRNA might be involved in this regulation (see also sRNA i399 and possible targets below).

The sRNA i855 accumulated more strongly under iron deficient conditions (Figure 3). The size of



**Figure 2** Northern blotting analyses of sRNAs with high microarray probe intensities. Positions are labelled by arrows at which the main sRNA species accumulate.

about 220 nt detected in the Northern blot, together with the microarray data suggested a partial antisense location to gene *synWH7803\_1012*. However, further inspection of this sRNA revealed its likely coding function and the presence of a 48 codons reading frame (Table 2). Indeed, a related protein-coding gene, *RS9916\_26469*, is annotated in *Synechococcus* RS9916 (Supplementary Figure 2a). Consequently, we annotated the deduced gene product of i855 as nORF1 (novel ORF1). Blast searches on the Marine Cyanobacteria Blast Server (<http://www.sb-roscoff.fr/SIG/blast/blastWH7803.html>) detected another homolog in *Synechococcus* sp. WH7805. This sequence contains an ORF of the same length as nORF1 and the gene *RS9916\_26469* and shares 92% identical amino-acid residues with the deduced *Synechococcus* sp. WH7803 nORF1 sequence (Supplementary Figure 2b). In *Synechococcus* sp. WH7805, this gene is flanked by *WH7805\_01707* and *WH7805\_01712*. The latter gene codes for a putative  $Mn^{2+}/Fe^{2+}$  transporter of the NRAMP family and constitutes another possible iron stress-response gene. Further sequence analysis with the upstream promoter sequences revealed a conserved sequence motif also found in front of putative transcription factor genes, belonging to cluster 1390 of the Cyanorak database (<http://www.sb-roscoff.fr/Phyto/cyanorak/>). We detected a strong conservation of this likely promoter element

among the cluster 1390 and nORF1 genes from the mentioned strains, especially in the  $-35$  region (Supplementary Figure 2c). Cluster 1390 proteins constitute with high probability a class of CRP-type transcription factors involved in iron regulatory processes (Scanlan *et al.*, 2009), thus there is a strong link for nORF1 to be involved in the regulation of iron homeostasis.

Expression of the two previously identified highly expressed sRNAs Yfr2a and Yfr2b (Axmann *et al.*, 2005; Gierga *et al.*, 2009) was downregulated under cold stress (Figure 3). Yfr2a and b belong to a conserved sRNA class, found in freshwater as well as marine cyanobacteria (Gierga *et al.*, 2009). Their function is unknown but their differential regulation is here observed for the first time.

Another sRNA (i399) with decreased expression under cold stress and even more reduced levels under high light was detected in the 5'UTR of *cpeB* (Figure 4). The Northern analyses confirmed the regulation inferred from the microarray data and indicated a prominent transcript of about 80 nt, but also several transcripts of greater length. Because of the proximity to *cpeB*, transcribed in the same direction, the TSSs of i399 and *cpeB* were mapped. Interestingly, the *cpeB* TSS was located to genome position 493029, 20 nt upstream of the TSS mapped for i399 (Table 1). We conclude that there are two promoters in front of *cpeB*. The more distal one

**Table 3** Location of Northern-verified sRNAs expressed in *Synechococcus* WH7803

IGR	Transcript start	Flanking genes	Orientation	Genomic island	Expression	Comparative prediction
i235	gtcttaagatcagatGAAGACCAA	<i>syn_0290; syn_0291</i>	← → →	No	Cold down	a,b,c,d
i781	/.	<i>syn_0921; syn_0922</i>	→ ← →	<b>3</b>	Cold down	a,b,c,d
i804	tggcaagatgttgagGTGAAGCAGC	<i>syn_0948; dcd</i>	→ ← →	<b>3</b>	Cold up	/.
i824	cgtgacagtgattGTGTGAGGAA	<i>syn_0975; syn_0976</i>	→ ← →	<b>3</b>	Cold down	a,b,c,d, <i>yfr2b</i>
i1720	gtgtaagatgtttaTTGGCCTGGC	<i>rpoB; tatD</i>	← ← ←	No	Cold down	a,b,c,d
i1971	tcttaagtgattcGTGTGAGGAG	<i>syn_2370; rplL</i>	→ → ←	No	Cold down	a,b,c,d, <i>yfr2a</i>
i854	/.	<i>syn_1011; syn_1012</i>	← ← →	<b>3</b>	— Fe up	b
i855 <sup>a</sup>	tgctctattgagagcCAATCTCAAGT	<i>syn_1012; syn_1013</i>	→ ← →	<b>3</b>	— Fe up	d
i95 <sup>a</sup>	/.	<i>syn_0111; syn_0112</i>	← ← →	No	— P up	/.
i399	tcatgttcggtcgATTCGCCAGCA	<i>syn_0484; cpeB</i>	← → →	<b>PBS</b>	High light down	/.
i403	gctcctaaactcgt(t)GCGTCAATTC	<i>cpeZ; mpeU</i>	→ ← ←	<b>PBS</b>	High light up	/.
i405	/.	<i>syn_0490; cpeY</i>	→ ← →	<b>PBS</b>	High light down	/.
i695	/.	<i>syn_0824; syn_0825</i>	→ ← →	No	High light up	/.
i845	/.	<i>syn_1000; syn_1001</i>	← ← →	<b>3</b>	High light up	/.
i1692	ctgagaatcattggtGATCGGCCTC	<i>apcE; apcA</i>	→ → →	No	High light down	a,b,c,d
i297	/.	<i>nctP; psbA</i>	← → ←	No	—	a,b,c,d
i298	gatagaagtgaagcCGATGGCGTT	<i>psbA; syn_0367</i>	← → →	No	—	a,b,c,d
i308	/.	<i>fusA; rpsG</i>	← → ←	No	—	a,b,c,d
i357	/.	<i>rplC; syn_0435</i>	← → →	No	—	a,b,c,d
i519 <sup>a</sup>	cgagacagtgatcaaGGTCTAAGAG	<i>rpoD; priA</i>	← → →	No	—	a,d
i583	/.	<i>syn_0683; csoS1/ccmK2</i>	→ ← →	No	—	/.
i812	ggttatctgtgatggAGGGAGCGAA	<i>syn_0956; syn_0957</i>	→ → ←	<b>3</b>	—	/.
i834	cgccagggtgtgggATCCGCGCCA	<i>znuC; syn_0990</i>	→ ← →	<b>3</b>	—	/.
i1017	cctaaattgattgtGAACTACCCC	<i>purK; syn_1207</i>	→ → ←	No	—	b,d, 6SRNA
i1255	/.	<i>syn_1498; syn_1499</i>	← ← →	<b>7</b>	—	/.
i1331	/.	<i>guaB; trxA</i>	→ → →	No	—	a,b,c,d, <i>yfr1</i>
i1535	/.	<i>syn_1835; gltX</i>	→ → →	No	—	b,c,d
i1642	gcgatcgtgagtctGGACGAGCCA	<i>kaiA; rplU</i>	← → →	No	—	b,c,d
i1684	/.	<i>atpA; atpD</i>	← → ←	No	—	/.
i1720	/.	<i>rpoB; tatD</i>	← → ←	No	—	a,b,c,d
i1825 <sup>a</sup>	cctagaacagaatcaCTTCGGCCTT	<i>syn_2189; syn_2190</i>	→ → →	No	—	b,c
i1869	cgttatggttcaagtGTTTCAGAGCG	<i>syn_2237; syn_2238</i>	→ ← →	No	—	a,b,c,d
i1886	/.	<i>syn_2257; syn_2258</i>	→ ← ←	No	—	/.
i1908	/.	<i>ndhI/frxB; ndhA</i>	← → ←	No	—	/.
i1973	atgtaggtgagtgCTGGCGAAAG	<i>rplJ; rplA</i>	← ← ←	No	—	a,b,c,d
i2037	ggttaaggttcggtAGGAGTGATT	<i>syn_2450; syn_2451</i>	← ← ←	No	—	/.

Abbreviations: No, location is not in a genomic island; PBS, phycobilisome rod gene cluster.

<sup>a</sup>Putative protein-coding function (see also Table 2).

5' RACE results are shown in 5' to 3' direction, with capital letters indicating the first transcribed nucleotide. The classification into genomic islands is based on Dufresne *et al.* (2008).

The observed regulation of expression is given (—, constitutive under the tested conditions) and the predicted orthologs in closely related marine picocyanobacteria (*a* = *Synechococcus* CC9311, *b* = *Synechococcus* CC9605, *c* = *Synechococcus* WH8102, *d* = *Synechococcus* RS9916, / = no orthologs found).

controls the transcription of the gene, whereas the more proximal located promoter controls the transcription of i399. The expression of *cpeB* appears to be downregulated under cold stress and high light, but the expression of i399 is decreased only under high light.

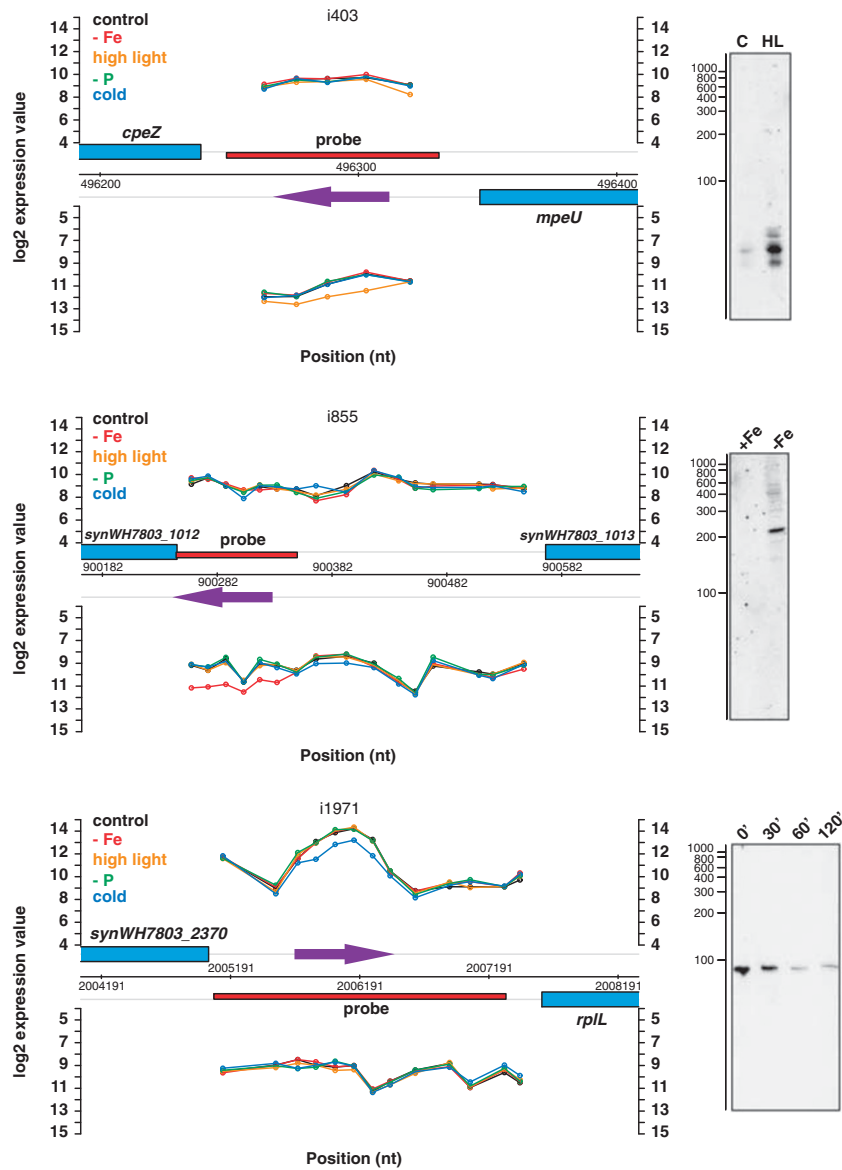
#### Target prediction for selected ncRNAs

We determined possible target genes with the program IntaRNA (Busch *et al.*, 2008), taking the sRNA as query and all annotated protein-coding genes of *Synechococcus* sp. WH7803 as the subject. The complete list of target predictions is presented in Supplementary Table S2. Interestingly, *cpeB* appeared as a possible target gene for two sRNAs, i403 (Figure 3) and the ~60nt long sRNA i812 (Figure 2). The i403 sRNA showed significantly higher transcript levels under high light, possibly

caused by an upregulation of transcription. As a consequence, target expression could be downregulated by base-pairing interactions with this sRNA under these conditions. The two potential target mRNAs *cpeB* and *mpeB* with their possible interaction sites are shown in Figure 5. The predicted interactions consist of a perfect match of 10nt between sRNA and mRNA around the AUG. The biological interpretation of such a suggested downregulation of phycobilisome components could be the prevention of photosystem damage by an excess of energy.

Bioinformatic analysis of the second aforementioned sRNA, i812, also suggested *cpeB* as a possible target (Figure 6a). A similarly extended complementarity was observed between i812 and two other mRNAs involved in photosynthesis, *petA* (Figure 6b) and *cpeE* (Figure 6d). A fourth potentially interesting target is *synWH7803\_1709*, coding





**Figure 3** Graphical output of microarray data and corresponding northern hybridization results for i403, i855 and i1971. Displayed are the normalized log<sub>2</sub> intensity values for the single probes from both strands. The x axis shows a schematic representation of the genome of *Synechococcus* sp. WH7803. The values for the five different growth conditions are presented in different colors. The values for the control growth condition were calculated as the mean intensity from the controls of all experiments. Every dot represents the middle of the respective probe (the strands for probe orientation have been reversed for easier visualization). The gene for the respective sRNA is indicated by a pink arrow and the location of the single-stranded RNA probes for northern hybridization is marked as a red box.

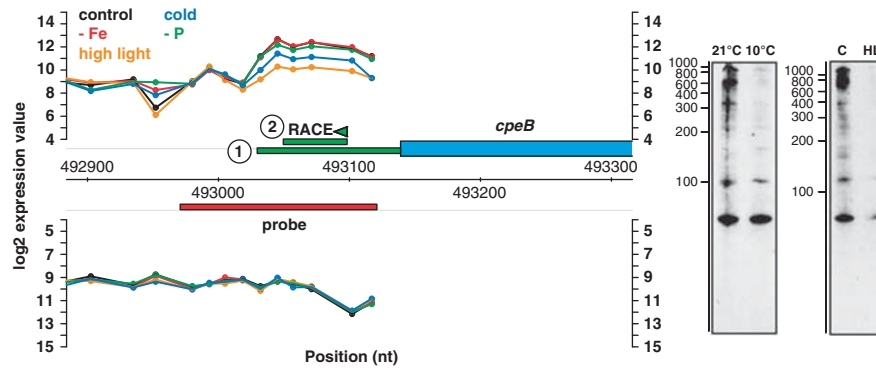
for an OmpR-type two-component-system response regulator (Figure 6c).

## Discussion

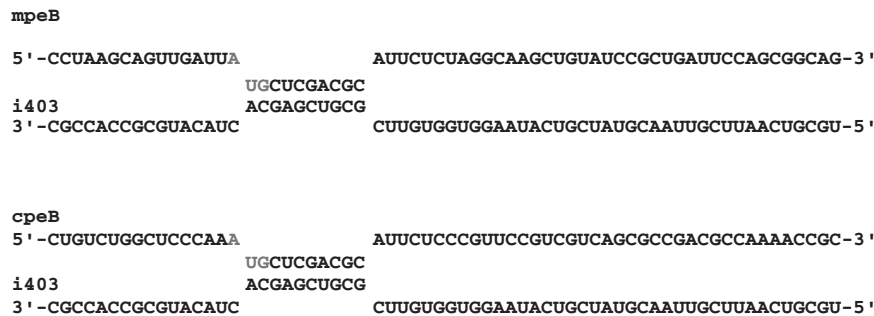
### *New sRNAs and novel protein-coding genes in Synechococcus sp. WH7803*

The necessity for physiological adaptation to fluctuating mesotrophic conditions, in combination with its streamlined genome (Dufresne *et al.*, 2008) and an existing genetic system, make *Synechococcus* sp. WH7803 an attractive model for marine

picocyanobacteria. The microarray analysis did not only confirm many of the predicted RNA elements but in addition revealed the presence of transcripts that had not been computationally predicted. In summary, we identified 15 transcripts differentially regulated under environmentally relevant conditions. Additional 21 transcripts were found with high expression levels. From these 36 transcripts, four are likely protein coding (three novel short peptide-coding genes and one 5'-prolonged long peptide-coding gene), the other 32 are likely non-coding sRNAs. One of the peptide-coding mRNAs and two of the sRNAs are to a substantial



**Figure 4** Microarray data and corresponding northern hybridizations for i399 (color code as in Figure 3). The northern blots show the transcript levels under standard and low temperatures (21 °C and 10 °C) and under standard and high light (C and HL). The 5' RACE results are shown as green bars for (1) *cpeB* and (2) the small RNA in front of *cpeB* with the 3' end of (2) defined by the location of the RACE primers (green triangle) and not reflecting the natural transcript 3' end. The location of the single-stranded RNA probe for northern hybridization is marked as a red box.



**Figure 5** Potential interactions between the sRNA i403 and the mRNAs *cpeB* and *mpeB* involving 2 nt of the AUG (highlighted in red) and eight following positions. The interaction was predicted using the program IntaRNA (Busch *et al.*, 2008). The color reproduction of this figure is available at *ISME* journal online.

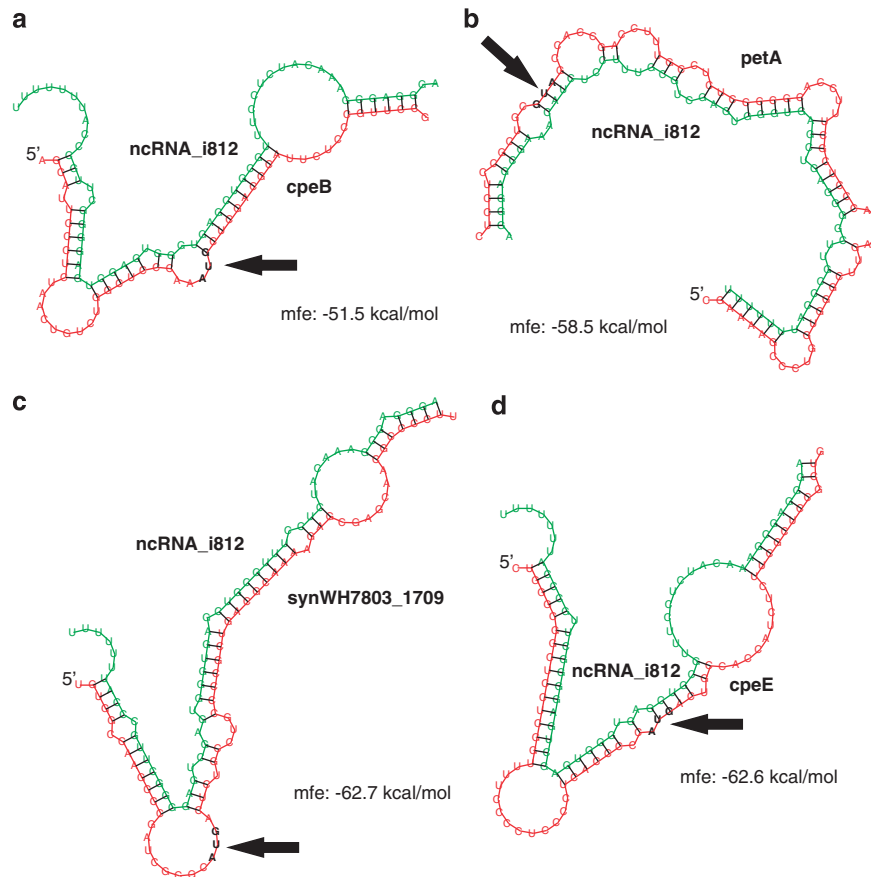
degree *cis*-antisense RNAs, whereas all other appear to be candidates for *bona fide* *trans*-acting, potentially regulatory sRNAs. Their functions are entirely unknown at present but some of them may be involved in the response to distinct types of nutrient limitation, redox stress, or phage infection.

Very short protein-coding genes are notoriously difficult to predict during standard genome annotation, therefore it is not surprising to identify some transcripts, which are, in fact, rather short mRNAs than regulatory non-coding sRNAs. From the validated sRNAs in this work, only three possessed the potential to code for a protein. The sRNA i855 coding for nORF1 was not classified as potentially protein coding, because it substantially overlaps *synWH7803\_1012*, as described in the results section. The deduced amino-acid sequences of these three peptide candidates are presented in Table 3.

#### A *cis*-antisense RNA to *psbA*

The combination of computational results with experimental data revealed interesting features of some of the newly identified sRNAs. We did not only define a set of sRNAs for *Synechococcus* sp. WH7803, but provide information about the distribution of these sRNAs within a subset of

marine picocyanobacteria. With i298 we identified an sRNA *cis*-antisense to the 5'UTR of *psbA*, for all five tested cyanobacteria. The gene *psbA* encodes the photosystem II reaction center protein D1 and frequently occurs in the form of a small gene family. *Synechococcus* sp. WH7803 possesses four *psbA* copies. One copy (*synWH7803\_0784*) encodes the D1:1 isoform of the protein, mainly expressed under non-stress conditions and providing the majority of the D1 protein (Garczarek *et al.*, 2008). The other three *psbA* copies encode the D1:2 isoform of the protein. Using 5' UTR-specific primers Garczarek *et al.* (2008) identified one of the three copies (*synWH7803\_0790*) to become strongly expressed under high light whereas the other two copies (*synWH7803\_0366* and *synWH7803\_2084*) were only slightly induced. Our microarray data gave evidence for sRNA i298 to be specifically associated with *synWH7803\_0366* and *synWH7803\_2084*. Prediction and microarray data were confirmed further by Northern analysis. Mapping the TSS by 5' RACE revealed i298 to originate within the coding region of *psbA* and the Northern results showed a small cohort of bands (Table 1 and Figure 2). Thus, i298 resembles more a *cis*-antisense RNA to *psbA* than a typical *trans*-acting sRNA. Antisense RNAs were also reported to the *psbA* gene of *Synechococcus* phage S-PM2 (Millard *et al.*, 2010) and the



**Figure 6** Potential targets for the sRNA i812. Targets were identified by IntaRNA (Busch *et al.*, 2008) and interactions between sRNA and the respective mRNAs (orange) for *cpeB* (a), *petA* (b), *synWH7803\_1709* (c) and *cpeE* (d). Especially noticeable is the pronounced complementarity in the region directly 5' of the mRNA AUG (boldface letters and arrow).

*psbA2* and *psbA3* genes of *Synechocystis* PCC6803 (Mitschke *et al.*, 2011). In *Synechococcus* sp. WH7803, this sRNA could be a factor contributing to the observed lesser expression of the two *synWH7803\_0366* and *synWH7803\_2084* *psbA* gene copies.

#### Novel sRNAs and the adaptation to changes in the environment

Regulatory RNAs that control the response to changing physical parameters such as light or temperature are frequently controlled by this environmental factor itself. Therefore, we included experiments to test the sRNA expression under various conditions. Temperature, in particular cold stress, is one of the most important environmental factors driving the distribution and composition of natural populations of marine cyanobacteria at a global scale (Bouman *et al.*, 2006; Johnson *et al.*, 2006; Zwirgmaier *et al.*, 2007). There is solid evidence for the control of adaptation to low temperatures in bacteria through regulatory RNA (Sledjeski *et al.*, 1996). Indeed, here we found six sRNAs specifically responding to cold. From these,

five became repressed and one activated, suggesting sRNA involvement in the response to colder temperatures also in marine picocyanobacteria. Another point for future analysis is how regulation of differential sRNA expression is achieved. There is no doubt that promoters for sRNA genes are fully integrated into the fabric of the intracellular regulatory network. However, cold or heat stress should lead to conformational changes of the regulatory RNA molecules themselves, as well as of other transcripts. Indeed, there is evidence that this is another level of posttranscriptional control of gene expression. Expression of the RNA helicase CrhL becomes induced upon cold stress (Supplementary Figure 1d) and the mutation of its homolog CrhR in *Synechocystis* sp. PCC6803 leads to a severe phenotype at all temperatures but most pronounced in the cold (Rowland *et al.*, 2011).

Besides cold stress, we studied the responses to high light, phosphate and iron limitation during exponential growth. Regulatory sRNAs involved in the regulation of iron homeostasis are known from several different groups of bacteria. The best studied representative is the 90 nt sRNA RyhB that was first identified in a genome-wide screen in *E. coli*

(Wassarman *et al.*, 2001) and regulates probably more than 20 iron-storage, iron-using and other proteins when iron becomes limiting (Masse and Gottesman, 2002; Modi *et al.*, 2011; Salvail and Masse, 2011). Functional homologs of RyhB were also identified in several non-enterobacteria, among them PrrF1 and PrrF2 in *Pseudomonas aeruginosa* (Wilderman *et al.*, 2004). Interestingly, all these sRNAs have in common that their expression becomes activated when iron becomes limiting and share the iron-dependent superoxide dismutase (*sodB* gene) as validated target (Afonyushkin *et al.*, 2005; Masse *et al.*, 2005). From the two sRNAs strongly induced by iron limitation in *Synechococcus* sp. WH7803 i854 and i855, one (i855) is likely coding for a short peptide involved in iron homeostasis. Nevertheless, it qualifies as a candidate for a functional homolog of RyhB and PrrF, even more so as *sodB* ranks relatively high on the list of predicted targets (predicted energy  $-14.35 \text{ kcal mol}^{-1}$ ; Supplementary Table S2).

There is no example for an sRNA involved in the bacterial phosphate stress response in the literature and the i95 transcript identified here as induced upon phosphate limitation is no exception. Closer inspection of our data revealed i95 to belong to the *SynWH7803\_0111* gene, which was incorrectly annotated and codes for a PhoA-type phosphatase. Based on the transcript data we extended the predicted protein sequence by 109 additional amino acids (Table 2), yielding a giant protein of 1855 amino acids. This protein has homologs in chlorobia (*Pelodictyon phaeoclathratiforme* and *Chlorobium phaeobacteroides* DSM 266), *Synechococcus* WH-8102 and the diatom *Thalassiosira pseudonana* CCMP1335. Notable features include the 5' nucleotidase and a phytase-like domain (Scanlan *et al.*, 2009).

Most previous work on bacterial sRNA regulators has been focused on human pathogens. However, for cyanobacteria, as photosynthetic organisms, light is a key environmental factor. Therefore, the finding of six light-regulated sRNAs is truly intriguing as some of them might be involved in the control of photosynthesis. Among the predicted targets of these six light-regulated sRNAs are several components of the light harvesting apparatus, some of which would be consistent with the effects observed upon mutation of the high light-inducible SyR1 sRNA in the cyanobacterium *Synechocystis* PCC 6803 (Mitschke *et al.*, 2011). Future work needs to be done to examine the specific functions of these sRNAs *in vivo*, by manipulating their expression and the physiological characterization of deletion mutants. The identification of a small set of sRNAs here as particularly interesting targets is an essential precondition for such experiments. Hence, this work provides a solid fundament for further experimental analysis of sRNAs, which are likely interacting with proteins involved in essential cellular processes.

## Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft Focus program 'Sensory and regulatory RNAs in Prokaryotes' SPP1258 (project HE 2544/4-2 to WRH) and the BMBF grant 03F0640A.

## References

- Afonyushkin T, Vecerek B, Moll I, Bläsi U, Kaberdin VR. (2005). Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res* **33**: 1678–1689.
- Axmann IM, Holtzendorff J, Voss B, Kensche P, Hess WR. (2007). Two distinct types of 6S RNA in *Prochlorococcus*. *Gene* **406**: 69–78.
- Axmann IM, Kensche P, Vogel J, Kohl S, Herzel H, Hess WR. (2005). Identification of cyanobacterial non-coding RNAs by comparative genome analysis. *Genome Biol* **6**(R73): 1–16.
- Bouman HA, Ulloa O, Scanlan DJ, Zwirgmaier K, Li WK, Platt T *et al.* (2006). Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. *Science* **312**: 918–921.
- Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J. (2008). Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* **32**: 827–837.
- Brahamsa B. (1996). A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl Environ Microbiol* **62**: 1747–1751.
- Busch A, Richter AS, Backofen R. (2008). IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**: 2849–2856.
- Coleman ML, Chisholm SW. (2007). Code and context: *Prochlorococcus* as a model for cross-scale biology. *Trends Microbiol* **15**: 398–407.
- Desnoyers G, Morissette A, Prevost K, Masse E. (2009). Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J* **28**: 1551–1561.
- Dufresne A, Ostrowski M, Scanlan DJ, Garczarek L, Mazard S, Palenik BP *et al.* (2008). Unravelling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol* **9**: R90.
- Dühring U, Axmann IM, Hess WR, Wilde A. (2006). An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc Nat Acad Sci USA* **103**: 7054–7058.
- Garczarek L, Dufresne A, Blot N, Cockshutt AM, Peyrat A, Campbell DA *et al.* (2008). Function and evolution of the *psbA* gene family in marine *Synechococcus*: *Synechococcus* sp. WH7803 as a case study. *ISME J* **2**: 937–953.
- Georg J, Voss B, Scholz I, Mitschke J, Wilde A, Hess WR. (2009). Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol Syst Biol* **5**: 305.
- Gierga G, Voss B, Hess WR. (2009). The Yfr2 ncRNA family, a group of abundant RNA molecules widely conserved in cyanobacteria. *RNA Biol* **6**: 222–227.
- Hess WR, Marchfelder A (eds) (2011). *Regulatory RNAs in Prokaryotes*. Springer-Verlag: Wien New York.
- Hihara Y, Kamei A, Kanehisa M, Kaplan A, Ikeuchi M. (2001). DNA microarray analysis of cyanobacterial

- gene expression during acclimation to high light. *Plant Cell* **13**: 793–806.
- Inaba M, Suzuki I, Szalontai B, Kanasaki Y, Los DA, Hayashi H *et al.* (2003). Gene-engineered rigidification of membrane lipids enhances the cold inducibility of gene expression in *Synechocystis*. *J Biol Chem* **278**: 12191–12198.
- Ishiura M, Kutsuna S, Aoki S, Iwasaki H, Andersson CR, Tanabe A *et al.* (1998). Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria. *Science* **281**: 1519–1523.
- Johnson ZI, Zinser ER, Coe A, McNulty NP, Malcolm E, Woodward S *et al.* (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**: 1737–1740.
- Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S *et al.* (2007). Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet* **3**: e231.
- Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. (2004). The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118**: 69–82.
- Los DA, Murata N. (1999). Responses to cold shock in cyanobacteria. *J Mol Microbiol Biotechnol* **1**: 221–230.
- Masse E, Gottesman S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci USA* **99**: 4620–4625.
- Masse E, Vanderpool CK, Gottesman S. (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* **187**: 6962–6971.
- Millard AD, Gierga G, Clokie MR, Evans DJ, Hess WR, Scanlan DJ. (2010). An antisense RNA in a lytic cyanophage links *psbA* to a gene encoding a homing endonuclease. *ISME J* **4**: 1121–1135.
- Mitschke J, Georg J, Scholz I, Sharma C, Dienst D, Bantscheff J *et al.* (2011). An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC 6803. *Proc Natl Acad Sci USA* **108**: 2124–2129.
- Modi SR, Camacho DM, Kohanski MA, Walker GC, Collins JJ. (2011). Functional characterization of bacterial sRNAs using a network biology approach. *Proc Natl Acad Sci USA* **108**: 15522–15527.
- Nakamoto H, Suzuki M, Kojima K. (2003). Targeted inactivation of the *hrcA* repressor gene in cyanobacteria. *FEBS Lett* **549**: 57–62.
- Partensky F, Hess WR, Vaulot D. (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106–127.
- Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J. (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**: 840–846.
- Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, Vogel J. (2007). A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol Microbiol* **66**: 1174–1191.
- Rowland JG, Simon WJ, Prakash JS, Slabas AR. (2011). Proteomics reveals a role for the RNA helicase *crhR* in the modulation of multiple metabolic pathways during cold acclimation of *Synechocystis* sp. PCC6803. *J Proteome Res* **10**: 3674–3689.
- Salvail H, Masse E. (2011). Regulating iron storage and metabolism with RNA: an overview of posttranscriptional controls of intracellular iron homeostasis. *Wiley Interdiscip Rev RNA* **3**: 26–36. DOI: 10.1002/wrna.102.
- Scanlan DJ, Mann NH, Carr NG. (1993). The response of the picoplanktonic marine cyanobacterium *Synechococcus* species WH7803 to phosphate starvation involves a protein homologous to the periplasmic phosphate-binding protein of *Escherichia coli*. *Mol Microbiol* **10**: 181–191.
- Scanlan DJ, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess WR *et al.* (2009). Ecological genomics of marine picocyanobacteria. *Microbiol Mol Biol Rev* **73**: 249–299.
- Scanlan DJ, Silman NJ, Donald KM, Wilson WH, Carr NG, Joint I *et al.* (1997). An immunological approach to detect phosphate stress in populations and single cells of photosynthetic picoplankton. *Appl Environ Microbiol* **63**: 2411–2420.
- Shi Y, Tyson GW, DeLong EF. (2009). Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* **459**: 266–269.
- 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.
- Six C, Thomas JC, Garczarek L, Ostrowski M, Dufresne A, Blot N *et al.* (2007). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol* **8**: R259.251–R259.222.
- Sledjeski DD, Gupta A, Gottesman S. (1996). The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J* **15**: 3993–4000.
- Sonnleitner E, Gonzalez N, Sorger-Domenigg T, Heeb S, Richter AS, Backofen R *et al.* (2011). The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol Microbiol* **80**: 868–885.
- Stazic D, Lindell D, Steglich C. (2011). Antisense RNA protects from RNase E degradation by RNA-RNA duplex formation during phage infection. *Nucleic Acids Res* **39**: 4890–4899.
- Steglich C, Futschik ME, Lindell D, Voss B, Chisholm SW, Hess WR. (2008). The challenge of regulation in a minimal phototroph: Non-coding RNAs in *Prochlorococcus*. *PLoS Genetics* **4**: e1000173.
- Tu KC, Long T, Svenningsen SL, Wingreen NS, Bassler BL. (2010). Negative feedback loops involving small regulatory RNAs precisely control the *Vibrio harveyi* quorum-sensing response. *Mol Cell* **37**: 567–579.
- Tu KC, Waters CM, Svenningsen SL, Bassler BL. (2008). A small-RNA-mediated negative feedback loop controls quorum-sensing dynamics in *Vibrio harveyi*. *Mol Microbiol* **70**: 896–907.
- Vecerek B, Moll I, Bläsi U. (2007). Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J* **26**: 965–975.
- Voss B, Georg J, Schon V, Ude S, Hess WR. (2009). Biocomputational prediction of non-coding RNAs in model cyanobacteria. *BMC Genomics* **10**: 123.
- Voss B, Gierga G, Axmann IM, Hess WR. (2007). A motif-based search identifies the ortholog of the small RNA Yfr1 in all major lineages of cyanobacteria. *BMC Genomics* **8**: 375.
- Washietl S, Findeiß S, Müller SA, Kalkhof S, von Bergen M, Hofacker IL *et al.* (2011). RNAcode: Robust discrimina-

- tion of coding and noncoding regions in comparative sequence data. *RNA* **17**: 578–594.
- Washietl S, Hofacker IL, Stadler PF. (2005). Fast and reliable prediction of noncoding RNAs. *Proc Natl Acad Sci USA* **102**: 2454–2459.
- Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* **15**: 1637–1651.
- Waters LS, Storz G. (2009). Regulatory RNAs in bacteria. *Cell* **136**: 615–628.
- Webb EA, Moffett JW, Waterbury JB. (2001). Iron stress in open-ocean cyanobacteria (*Synechococcus*, *Trichodesmium*, and *Crocospheara* spp.): identification of the IdiA protein. *Appl Environ Microbiol* **67**: 5444–5452.
- Wilderman PJ, Sowa NA, FitzGerald DJ, FitzGerald PC, Gottesman S, Ochsner UA *et al*. (2004). Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. *Proc Natl Acad Sci USA* **101**: 9792–9797.
- Zwirgmaier K, Heywood JL, Chamberlain K, Woodward EM, Zubkov MV, Scanlan DJ. (2007). Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ Microbiol* **9**: 1278–1290.
- Zwirgmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D *et al*. (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol* **10**: 147–161.

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