Subject category: Integrated genomics and post-genomics approaches in microbial ecology

PtrA is required for coordinate regulation of gene expression during phosphate stress in a marine *Synechococcus*

Martin Ostrowski[1], Sophie Mazard[1], Sasha G. Tetu[2], Katherine Phillippy[3#], Daniel Johnson[3], Brian Palenik[4], Ian T. Paulsen[2,3] & Dave Scanlan[1*]

[1] Department of Biological Sciences, University of Warwick, Gibbet Hill Rd, Coventry, CV4 7AL, United Kingdom.

[2] Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales, Australia.

[3]J Craig Venter Institute, Rockville, MD, USA.

[4] Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA

Current address: National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Bethesda, MD, USA

*Corresponding author
Dave Scanlan
Department of Biological Sciences,
University of Warwick,
Gibbet Hill Rd,
Coventry,
CV4 7AL,
United Kingdom.

Tel. +44 (0) 2476 528363
Fax. + 44 (0) 2476 523701

Running title: Phosphate regulation in *Synechococcus* sp. WH8102

Key words: Cyanobacteria/*Synechococcus*/Phosphorus metabolism/Transcriptional activators/PtrA

SUMMARY

Previous microarray analyses have demonstrated a key role for the two-component system PhoBR (SYNW0947, SYNW0948) in the regulation of P transport and metabolism in the marine cyanobacterium *Synechococcus* sp. WH8102. However, there is some evidence that another regulator, SYNW1019 (PtrA), probably under the control of PhoBR, is involved in the response to P-depletion. PtrA is a member of the CRP transcriptional regulator family that shows homology to NtcA, the global nitrogen regulator in cyanobacteria. To define the role of this regulator we constructed a mutant by insertional inactivation and compared the physiology of wild-type *Synechococcus* sp. WH8102 with the *ptrA* mutant under P-replete and P-stress conditions. In response to P stress the *ptrA* mutant failed to up-regulate phosphatase activity. Microarrays and quantitative RT-PCR indicate that a subset of the Pho regulon is controlled by PtrA, including two phosphatases, a predicted phytase and a gene of unknown function *psiP1*, (SYNW0165), all of which are highly up regulated during P-limitation. Electrophoretic mobility shift assays indicate binding of over-expressed PtrA to sequences upstream of the induced genes. This work suggests a two-tiered response to P-depletion in this strain, the first being PhoB-dependent induction of high affinity PO_4 transporters, and the second the PtrA-dependent induction of phosphatases for scavenging organic P. The levels of numerous other transcripts are also directly or indirectly influenced by PtrA, including those involved in cell surface modification, metal uptake, photosynthesis, stress responses and other metabolic processes, which may indicate a wider role for PtrA in cellular regulation in marine picocyanobacteria.

INTRODUCTION

Marine cyanobacteria of the genera *Prochlorococcus* and *Synechococcus* are globally distributed and ecologically significant (Partensky *et al*., 1999; Scanlan *et al*., 2009). These two genera are genetically diverse and numerically abundant across global scales, from the oligotrophic ocean gyres where *Prochlorococcus* predominate, to subtropical, temperate and coastal systems dominated by *Synechococcus* (Zwirglmaier *et al*., 2008). As the base of the food web and as a source of photosynthetically fixed carbon, cyanobacterial limitation by macronutrients such as nitrogen and phosphorus would therefore impact all trophic levels.

Reports of P-limitation in some oceanic regimes are accumulating e.g. in the north-western Atlantic (Sargasso Sea) (Cotner *et al*., 1997; Ammerman *et al*., 2003), North Pacific Subtropical Gyre (Karl and Tien 1997) whilst the Mediterranean Sea displays P limitation during summer stratification (Thingstad *et al*., 2005). Since dissolved mineral phosphate concentrations fall into the low nanomolar range in oligotrophic regions and the nutrient is rapidly turned over (Karl and Tien 1997; Thingstad *et al*., 2005; Zubkov *et al*., 2007) there is likely an intense competition for bioavailable P.

Marine cyanobacteria have adopted distinct strategies to cope with low and fluctuating levels of P in the environment. Comparative genomics (Palenik *et al*., 2003; Palenik *et al*., 2006; Dufresne *et al*., 2008; Scanlan *et al.,* 2009) and metagenomic surveys (Venter *et al*., 2004) highlight that there are multiple genome-encoded copies of the gene for the high-affinity, periplasmic P-binding protein, PstS, indicating the importance of the affinity capture of inorganic P, even for strains that have undergone significant genome streamlining (Dufresne *et al*., 2003). It is also evident that distinct isolates, or ecotypes, exhibit different physiological and genetic capabilities to utilise organic P (Moore *et al*., 2005; Martiny *et al*., 2006; Scanlan *et al*., 2009). Picocyanobacteria that inhabit low P environments display an overall low requirement for P (Bertilsson *et al*., 2003; Heldal *et al*., 2003) and are also capable of economising their use of P in cellular constituents, e.g. by substituting phospholipids for sulfolipids (Van Mooy *et al*., 2006). In addition, in contrast to many other bacteria, marine cyanobacteria have evolved forms of lipopolysaccharide, a major cellular constituent, that lack phosphate (Snyder *et al*., 2009).

In model freshwater cyanobacteria a P-regulatory system composed of a two-component response regulator and a sensory kinase, which is homologous to the *E.coli* PhoBR, has been well characterised (Aiba *et al.,* 1993; Aiba and Mizuno 1994; Hirani *et al*., 2001). This system activates the transcription of genes of the P-regulon, through activation of the response regulator, SphS (PhoB), by its cognate sensor histidine kinase, SphR (PhoR), presumably in response to a low external concentration of P. In *Synechocystis* sp. PCC6803, the Pho regulon is composed of at least 12 genes, including two clusters of high-affinity ABC transport systems as well as two co-localised genes encoding an alkaline phosphatase and an extracellular nuclease (Suzuki *et al*., 2004).

The composition and arrangement of genes in the P-regulons of marine cyanobacteria appear to be highly variable (Martiny *et al*., 2006; Scanlan *et al*., 2009; Tetu *et al*., 2009). Amongst available genome sequences PhoBR homologues are conserved in some, but not all, indicating that P-sensing and regulation may have been lost in some lineages. The loss of *phoBR* in some strains may reflect adaptation to environments with a relatively stable and adequate supply of P such as coastal environments (Palenik *et al*.,

2006). In the open ocean *Synechococcus* sp. WH8102 the P-regulon has been characterised by computational predictions (Su *et al*., 2003; Su *et al*., 2007), as well as physiological experiments (Moore *et al*., 2005) and microarray analyses (Tetu *et al*., 2009). Comparisons of gene expression in knockout mutants constructed in *phoB* and *phoR* confirm a role for P sensing and regulation for these genes in this strain (Tetu *et al*., 2009). Central elements of the *Synechococcus* sp. WH8102 P-regulon were identified, including four genes encoding paralogues of the periplasmic P binding proteins, PstS (SYNW1018, SYNW1815, SYNW2507) and SphX (SYNW1286), genes of the ABC transport system for P (SYNW1270, SYNW1271), porins (SYNW2224, SYNW2223), at least three diverse predicted phosphatases (SYNW0196, SYNW2390 and SYNW2391) and a number of genes with weakly associated or no functional predictions (SYNW0165, SYNW0762 and SYNW1333). These experiments also highlight the involvement of another regulator, PtrA, in the P stress response that was hypothesised to be a potential P regulator in *Synechococcus* sp. WH7803 (Scanlan *et al*., 1997). In *Synechococcus* sp. WH7803 and WH8102 *ptrA* is located downstream of the gene for the P stress induced periplasmic phosphate binding protein, PstS. A putative *pho* box upstream of *ptrA* suggests that this gene is regulated by PhoB (Su *et al*., 2007), while, microarray analysis of *Synechococcus* sp. WH8102 in response to P depletion show that *ptrA* is highly expressed in response to P-stress, and confirms that it is most-likely under the influence of PhoB (Tetu *et al*., 2009).

CRP-family regulators consist of a C-terminal DNA binding domain that forms a helix-turn-helix (HTH) motif which slots into the major groove of the DNA and binds to the promoter region of target genes to either act as an activator or repressor (Korner *et al.,* 2003). The N-terminus is composed of a nucleotide binding domain, which acts as a sensor module that interacts with a signal molecule and is responsible for dimerisation and activation. Comparison of amino acid sequences of putative CRP-family regulators from marine cyanobacterial genomes highlights at least 4 distinct clusters of orthologues (Scanlan *et al*., 2009). In *Synechococcus* sp. WH8102 PtrA is one of only two genome-encoded regulators of the CRP-family and a homologue of NtcA, the global nitrogen regulator that is conserved in all cyanobacteria (Luque *et al*., 1994; Lindell *et al*., 1998; Herrero *et al*., 2001). The *E.coli* CRP activates transcription in the presence of cAMP at more than 100 different promoters (Salgado *et al*., 2004) while there are a predicted 17-54 NtcA targets in *Synechococcus* sp. WH8102 (Su *et al*., 2005).

Given the sequence similarity, PtrA could be the P 'equivalent' of NtcA, and therefore a general phosphate regulator in marine cyanobacteria that modulates the expression of part of the P regulon with potentially wider influence on cellular processes. To define the role of this regulator we constructed a mutant in *Synechococcus* sp. WH8102 by insertional inactivation and compared the physiology of the *ptrA* mutant with the isogenic parent strain in response to P stress. Microarrays and quantitative RT PCR were used to compare patterns of global gene expression in response to P stress. Comparison of the *ptrA* mutant transcriptome with data from parallel studies (Tai *et al*., 2009; Tetu *et al*., 2009) indicate that *ptrA* is an important component of the WH8102 P regulon directly controlling a subset of P stress genes but also may have a wider influence on the expression of a variety of gene clusters throughout the genome, implicating it in co-ordinately regulating cellular metabolism during the P stress response.

MATERIALS and METHODS

***Bacterial strains, growth conditions and P stress experiments*.** For genetic transformations *Synechococcus* sp. WH8102 was grown in SN medium (Waterbury & Willey 1988) prepared with Sargasso Sea water (Sigma Chemical Co.) with constant illumination at 25 µmol photons m^{-2} s^{-1} at 25˚C. For growth experiments *Synechococcus* sp. WH8102 was grown in synthetic seawater medium with salts based on Aquil (Morel *et al*., 1979). The final concentration of nutrients in Aquil were as follows, 4.5 mM $NaNO_3$, 90 µM Na_2PO_4, 100 µM NH_4Cl, 13.4 µM Na_2EDTA, 10 µM Na_2CO_3 (N:P = 50:1). For P stress experiments the PO_4 concentration was amended to 10 µM, yielding an N:P ratio of 450:1. Cultures were acclimated to 40 µmol photons m^{-2} s^{-1} and 10 µM PO_4 for a minimum of three serial transfers prior to the start of each experiment. Growth was monitored on a daily basis by spectrophotometry and flow cytometry, the latter using a FACScan flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA). Culture axenicity was monitored regularly by plating an aliquot (200 µl) onto solid Aquil containing 500 mg l^{-1} yeast extract.

Insertional inactivation of ptrA in Synechococcus sp. WH8102. Directed inactivation of *ptrA* (SYNW1019) was accomplished according to methods previously described (Brahamsha 1996). Briefly, a 297 bp internal fragment, corresponding to nucleotides 213 to 488 of the 684 bp gene, was cloned into the suicide vector pMUT100. The *ptrA* fragment was amplified with primers PtrAF9 (5'-TGTGCGCGGCATGGTCAAGCTTG-3') and PtrAR9 (5'-CATTTCCAGAAAACCTCTCACCCG-3'), TA cloned into pCR2.1TOPO (Invitrogen, Carlsbad, CA) and then sub cloned into the *Eco*RI site of pMUT100. Biparental conjugations were conducted exactly as described by Brahamsha (1996) with the exception that SN medium was prepared with Sargasso seawater. Exconjugants were selected on plates and clonal isolates were propagated in liquid SN with kanamycin (25 µgml^{-1} and 15 µg ml^{-1} respectively). Segregation of mutant chromosomes, arising from a single cross-over insertion of the plasmid construct, was confirmed by PCR utilising primer sets targeting sites on the *Synechococcus* sp. WH8102 genome flanking *ptrA* and sites internal to the pMUT100 vector (as outlined in Supplementary Figure 1).

Phosphatase and soluble reactive phosphate assays. Phosphatase activity was measured throughout the growth of mutant and wild type cultures using the para-nitrophenyl phosphate (*p*-NPP) assay (Bessey *et al*., 1946) adapted for use in a microplate reader as described in Moore *et al*., (2005). The concentration of extracellular soluble reactive phosphate was determined spectrophotometrically by the ammonium molybdate assay (Itaya and Ui, 1966) in a quartz cuvette with an optical path length of 5 cm. The limit of detection with this method was 50 nM.

SDS-PAGE analysis of protein expression in response to P stress. Protein expression in low P cultures was monitored by pulse-labelling of a culture aliquot (5.0 ml) with ^{35}S-methionine (1.0 µCi ml^{-1}) for 4 h under identical incubation conditions as the source culture as described previously (Scanlan *et al*., 1993). At the end of the incubation period the cells were harvested by centrifugation and resuspended in SDS-PAGE loading buffer (90 mM Tris-HCl pH 6.8, 20% (v/v) glycerol, 2% (w/v) SDS, 0.02% (w/v) bromophenol blue, 100 mM DTT). Samples were boiled for 10 min and centrifuged (10,000 x *g*, 5 min) before loading onto a 12% SDS-PAGE gel. After electrophoresis the gel was stained with coomassie blue and the labelled proteins were visualised by exposing the dried gel to a phosphor screen (Fujifilm LifeScience) analysed using a phosphorimager (Fuji FLA5000). Protein bands of interest were excised from a replicate gel prepared with

unlabelled cells extracted from the source culture, subjected to tryptic digest and identified by MALDI-TOF Mass spectrometry using a Micromass MALDI-LR (Waters Corporation, Milford, MA, USA).

RNA isolation and microarray analysis. Total RNA from 1 litre of exponential phase cells was extracted using a Trizol-based method and purified using the Qiagen RNeasy kit following manufacturer's instructions as previously described (Tetu *et al*., 2009). cDNA was labeled, hybridised to microarrays and the results captured as previously described (Tetu *et al*., 2009). RNAs from two wild type and two *ptrA* mutant cultures, harvested at ~125-150 h after inoculation were used. Six different hybridisations were carried out, of which two used different RNA pools (biological replicates) and four were replicates of these two experiments, either dye swapping experiments or direct replicates. Statistical analyses were carried out on the mean of log2-transformed signal ratios of all replicates using the Significance Analysis of Microarrays (SAM) algorithms (Tusher *et al*., 2001) with a false discovery rate of less than 1%.

The microarray data presented here is in accordance with the Microarray Gene Expression Data (MGED) Society's minimum information about a microarray experiment (MIAME) recommendations (Brazma *et al*., 2001). A description of the experiments, quantitation data, and array design has been deposited into the gene expression omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) with the assigned accession number GSE18511.

Quantitative RT-PCR. DNA was removed from total RNA using the Turbo-DNase Digestion kit (Ambion). The absence of DNA was confirmed by PCR and then RNA (200-500 ng) was reverse transcribed with SuperScript II (Invitrogen, Carlsbad, CA) in the presence of 200 units of SuperaseIN (Ambion) and random hexamers (25 pM). The resulting cDNA was diluted more than 20-fold and used as a template (5-10 ng equivalent of starting RNA). qRT-PCR reactions were carried out in triplicate for biological replicates with *Power* SYBR green PCR master mix on an ABI Prism® 7000 sequence detection system (Applied Biosystems) according to the manufacturers recommendations. Primers were designed with ABI-Prism Primer Express® version2.0 software for products of 50-65 bp length (Supplementary Table 1). The gene for the class B RNaseP (*rnpB*) was used as an internal reference. Analysis of RNA transcript abundance was carried out using the $\Delta\Delta C_T$ method using the ABI-Prism SDS 2.1 Software (Applied Biosystems).

Over-expression of PtrA in E. coli. The *ptrA* gene was PCR amplified with primers 1019pet15bf (5′-AGCCATATGCATGTTGCGCTCCATAC-3′) and 1019pet15br (5′-GGATCCTAGCGACGTGGCAGGTGGGCG-3′) with engineered restriction sites compatible with the in-frame insertion into the *Nde*I and *Bam*HI sites of the 6xHis tag expression vector pET15b (Novagen). After the sequence of the construct was confirmed, the plasmid was transformed into the *E. coli* expression host Rosetta (DE3). Expression cultures were inoculated (1%, 250 ml) from overnight pre-cultures and grown in LB medium supplemented with chloramphenicol (30 μg ml^{-1}) and ampicillin (100 μg ml^{-1}) at 37˚C in an orbital shaker at 200 rpm. Upon reaching an OD_{600} of 0.4 the expression of recombinant protein was induced with IPTG (20 μM) and the culture transferred to 25˚C with shaking (200 rpm) before the cells were harvested by centrifugation 4-5 h later. These conditions favoured the production of soluble recombinant protein despite the decrease in overall yield as a result of using sub-saturating concentrations of IPTG. Harvested cell pellets were washed once and then resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% Triton X100), frozen for a minimum of 2 h, and then lysed with three passages through a French pressure cell followed by sonication (3x 30 s). The

lysate was fractionated by centrifugation (14,000 x *g*, 30 min, 4˚C). Imidazole was added to the soluble fraction to a final concentration of 50 mM before being applied to a pre-equilibrated Ni(II) immobilised metal affinity column (5 ml, HiTrap Chelating HP, GE Health Sciences). The column was washed with 1 volume each of 20 mM Tris-HCl pH 8.0, 500 mM NaCl (Buffer A) containing 50 and 100 mM imidazole. Bound PtrA was eluted with Buffer A containing 300 mM imidazole and the purity of the fraction was assessed by SDS-PAGE. The purified protein was concentrated and the buffer exchanged for DNA Binding Buffer (20 mM Tris-HCl pH 7.4, 120 mM KCl, 10 mM $MgCl_2$, 5 mM $ZnCl_2$, 100 µM EDTA) supplemented with 20% v/v glycerol by successive concentration and dilution with a Centricon YM-10 centrifugal filter device (Millipore). The concentration of recombinant protein was determined with the BCA kit against known quantities of BSA (Sigma Chemical Co.).

Electrophoretic mobility shift assays (EMSA). DNA fragments for EMSA, encompassing approximately 350 bp upstream and 50 bp downstream of the start codon of target genes, were PCR amplified from *Synechococcus* sp. WH8102 using primers listed in Supplementary Table 2 and cloned into pCR2.1TOPO. Each fragment was excised from the plasmid by restriction digest with *Eco*RI, agarose gel purified and end-labelled with [γ-^{32}P]dATP using polynucleotide kinase. DNA fragments were then incubated with purified proteins at 25˚C for 20 min in DNA Binding Buffer supplemented with non-specific competitor DNA, either 25-50 µg.ml^{-1} polydI/dC or 12.5 µg.ml^{-1} herring sperm DNA. Protein stocks of PtrA were diluted in buffer containing competitor DNA prior to addition to the binding reaction. Reactions were loaded onto a 5.0% polyacrylamide gel run in 0.5 x Tris-Borate-EDTA buffer for 4-5 h at 180 V. The results were visualised by exposing the dried gel to a phosphor screen (Fujifilm LifeScience) and analysed using a phosphorimager (Fuji FLA5000).

RESULTS AND DISCUSSION

Phenotype of the Synechococcus sp. WH8102 ptrA mutant. To define the precise role of *ptrA* we constructed a mutant by insertional inactivation in *Synechococcus* sp. WH8102 and examined the physiology of the mutant relative to the wild type strain during growth in low P medium (10 µM PO_4). P stress growth experiments were conducted exactly as described in previous work (Tetu *et al*., 2009) to enable comparison of gene expression with independent microarrays of *phoB* and *phoR* mutants. The maximum rates of growth and yield were similar for wild type and mutant under P replete (90 µM PO_4) and low P growth conditions (0.41 and 0.39 d^{-1} respectively, Figure 1). Extracellular P was depleted to undetectable levels (50 nM) within the first 120 h of growth, i.e. approximately 100-150 h before the onset of stationary phase (Figure 1) which is similar to previous observations for this strain (Moore *et al*., 2005; Tetu *et al*., 2009). The disappearance of P was identical in wild type and mutant, indicating that P-uptake was not significantly impaired in the mutant.

While P-uptake was apparently unaffected, the production of phosphatase activity was markedly different in response to P stress (Figure 1B). Wild type *Synechococcus* sp. WH8102 displays three distinct levels of phosphatase activity in batch culture, a low constitutive level, followed by an intermediate level approximately 50 h after the disappearance of extracellular P, and a maximum level at the point where cell division ceases. The induction of phosphatase activity was significantly affected in the *ptrA* mutant, where

the absolute amount of activity did not increase above a constitutive level. This result suggests that at least one, or more, of the four identified phosphatase genes in *Synechococcus* sp. WH8102 (SYNW0196, SYNW1799, SYNW2390 and SYNW2391) are directly or indirectly under the control of PtrA, although the *p*-NPP assay used here may not account for any highly-specific nucleotidase activity of phosphatases that possess nucleotidase domains. Insertional inactivation of the *ptrA* orthologue in *Synechococcus* sp. WH7803 also resulted in a similar phenotype (Ostrowski and Scanlan, unpublished data) where induction of phosphatase activity was impaired in response to P stress.

Changes in protein expression in low P cultures were monitored using ^{35}S–methionine labelling and SDS polyacrylamide gel electrophoresis. It is particularly interesting to note the lag between the induction of PstS (SYNW1018) and porin (SYNW2224) expression after 96 h (Figure 1C), and the time when intermediate and maximum levels of phosphatase activity are observed in the wild type *Synechococcus* sp. WH8102 (at ~ 172 and 240 h, respectively). There was no significant time difference in the induction of PstS expression between the mutant (not shown) and wild type, which produced identical expression profiles, indicating that the initial response to P stress is apparently unaffected by the inactivation of *ptrA*.

Comparison of expression of P stress genes in the Synechococcus sp. WH8102 ptrA mutant. To characterise the PtrA regulon in more detail we used a whole genome microarray to compare patterns of global gene expression in the mutant and isogenic parent strain. RNA was harvested from mutant and wild type strains under low P conditions just after the onset of induction of phosphatase activity in wild type (~125-150 h after inoculation). The expression of 516 genes were negatively affected in the mutant. Genes whose expression was lower by 2-fold or more are listed in Table 1. Six hundred and forty five genes were up-regulated in the *ptrA* mutant and those up-regulated more than 2-fold are shown in Table 2.

The amount of overlap between sets of genes negatively affected by *ptrA* and the genes previously reported as up-regulated by early P stress or affected by *phoB* inactivation (Tetu *et al*., 2009) was surprisingly small (Figure 2). Of the 97 genes that were down-regulated by more than 2-fold only seven were shown to be correspondingly up-regulated in response to early P stress in wild type (highlighted in colour in Figure 2). Tetu *et al*., (2009) described nine genes of the PhoB regulon that were highly expressed in response to early P stress and were not up-regulated in a *phoB* knockout. These genes include a possible porin (SYNW2224), three phosphatases (SYNW0196, SYNW2390, SYNW2391), one PstS gene (SYNW1018) as well as *ptrA*. Five of those nine genes were also significantly down regulated in the *ptrA* mutant (Figure 2), encompassing two phosphatase genes (SYNW0196 and SYNW2390) a possible phytase (SYNW0762), a conserved hypothetical gene (SYNW1333) and the gene for a P starvation inducible polypeptide (*psiP1*, SYNW0165) which is localised to the cell wall (West and Scanlan unpublished data).

The direct comparison between the expression profiles of *ptrA* and *phoB* knockouts in Figure 2 provides good support for the hypothesis that PtrA may directly regulate a subset of the P-responsive genes in *Synechococcus* sp. WH8102 (i.e. SYNW0165, SYNW0196, SYNW0762, SYNW1333 and SYNW2390) but not others, including the *pstS* gene directly upstream of *ptrA* (SYNW1018), a phosphatase/nucleotidase gene (SYWN2391) upstream of SYNW2390, and a porin (SYNW2224). It is highly likely that PtrA is directly regulated by PhoB, either from a predicted *pho* box promoter in the intergenic region between *pstS* (SYNW1018) and *ptrA* (Su *et al*., 2003; Su *et al*., 2007) or by co-transcription directed from the *pho* box of

pstS. Using primers spanning the 3′ end of *pstS* and the 5′ end of *ptrA* we were able to detect RT-PCR products of expected size indicating that these two genes are, at least partially, co-transcribed (data not shown).

The remaining 92 genes with significantly down-regulated expression levels in the mutant are involved in a range of cellular processes. It is interesting to note that many of these potential PtrA regulated genes are gathered into five distinct genomic clusters (Table 1, Figure 3). One cluster (Cluster 9, SYNW2477-SYNW2485) includes genes for a predicted Zn^{2+} ABC transport system, a ferredoxin-nitrite reductase, a putative cyanate transporter and four conserved hypothetical genes of unknown function. Zn is found at sub-nanomolar concentrations in surface waters of the Pacific and Atlantic Oceans yet it is unclear whether this element limits phytoplankton growth (Lohan *et al*., 2002). In cyanobacteria, Zn is required for the activity of carbonic anhydrase as well as a range of metalloproteins involved in many aspects of metabolism (Blindauer 2008). Since phosphatases also have a known Zn requirement for activity (Coleman 2003), the elevated levels of the Zn^{2+} transport system in wild type may simply reflect higher Zn requirements to match the increased phosphatase production, rather than direct regulation by PtrA. *Synechococcus* sp. WH8102 possesses Zur (SYNW2401) and SmtA (SYNW0359) orthologues which have the potential to act as Zn sensors and/or regulators (Blindauer 2008). Although a direct involvement in Zn regulation has not been experimentally determined, either of these proteins appears to be better candidates to regulate the Zn transport system in response to Zn demand.

Cluster 6 (Table 1, Figure 3) is composed of 11 consecutive genes (SYNW0952-SYNW0962) mostly of unknown function. These genes include three that were shown by transposon mutagenesis to have a role in swimming motility in *Synechococcus* sp. WH8102, including the 34 kb gene for SwmB (SYNW0953) and two associated orfs (SYNW0958 and SYNW0960), that may encode a multicomponent transport apparatus (McCarren and Brahamsha 2005, 2007). While SwmB, which is known to be associated with the outer membrane, and surrounding genes in this cluster are required for motility, the genes for other integral components of the swimming apparatus, such as SwmA, do not show a similar pattern of expression in the mutant. It is possible that the genes in this cluster serve an additional role unrelated to motility that involves modifying the cell surface. Indeed, the expression of 27 genes (cluster 3, SYNW0424-SYNW0458), which are all related to cell surface modification and polysaccharide production, are all down regulated in the mutant. As stated previously (Tetu *et al*., 2009), reorganisation and/or strengthening of the cell envelope may be required to accommodate additional P stress induced porins (SYNW2224 and SYNW2223). An alternative explanation could be that, as a result of a reduction in cellular growth rate due to limiting P, the cell is capable of diverting non-limiting resources, e.g. C and N, to LPS and carbohydrate modification of the cell wall as a means of avoiding grazers and phage.

A number of genes involved in phycobilisome biosynthesis are also down-regulated in the mutant (cluster 7). Other genes of interest with predicted functions that display lower levels of expression in the mutant include genes for cytochrome *c* oxidase subunits I and II (SYNW1529, SYNW1861-1862), a PhoH family protein (SYNW1946), a putative transcriptional regulator (SYNW2105) and the gene for the plastoquinol terminal oxidase (SYNW0887) that may play a role in diverting electrons from PSII out of the photosynthetic electron transport chain under certain conditions (Mackey *et al*., 2008).

As an independent means of observing trends in gene expression qRT-PCR was performed on a

selection of genes. In each case qRT-PCR resulted in the same directional trend as for the microarray analysis in the mutant relative to wild type, while the magnitude of up-regulation during P stress in wild type was equivalent to the values reported in previous work (Figure 2, and Tetu *et al*., 2009).

Genes with relatively higher levels of expression in the PtrA mutant. The transcript levels of 63 genes recorded higher levels in the *ptrA* mutant (Table 2). Among the genes with known functions are the RNA polymerase β' subunit (SYNW0615) and 16 ribosomal proteins (SYNW2067-SYNW2091, SYNW2135-6), many of which were significantly repressed in *Synechococcus* sp. WH8102 (Tetu *et al*., 2009) and in other organisms in response to P-limitation (Martiny *et al*., 2006) and N-limitation (Silberbach *et al*., 2005). Since the ribosomes themselves are a significant sink for P-rich ribonucleotides it would be advantageous for a cell to economise the biosynthesis of ribosomes, and associated components, in response to P stress. Given that the mutant and parent strain were grown side-by-side under identical conditions the higher levels of expression of ribosomal proteins suggest the mutant fails to down-regulate its translation machinery. It is possible that the expression of central metabolism genes, such as ribosomal proteins, is influenced by PtrA. However, a perhaps more likely explanation is that the growth of the parent and mutant was subtly different and that they were simply in different growth phases, possibly because the mutant utilised intracellular P at a slower rate.

Ribonucleotide reductase (ClassII), (SYNW1147) also displayed a higher level of expression. This enzyme catalyses the conversion of ribo- to deoxy-ribonucleotide-diphosphates (Elledge *et al*., 1992) and thus occupies a central role channelling nucleotides to DNA synthesis. The expression of several stress response proteins was also higher, including GroES (SYNW0513), two copies of GroEL (SYNW0514, SYNW1854), DnaK2 (SYNW2508) and the heat shock protein HtpG (SYNW1278). These stress response genes suggest the *ptrA* mutant may have been more subtly stressed than wild type possibly due to the inability to mount an effective response to early phosphate depletion, and this might explain some gene expression differences not directly related to phosphate. While a number of genes involved in phycobilisome biosynthesis and cytochrome *c* oxidase subunits were repressed in the mutant, genes for photosystem II reaction centre proteins L, T (SYNW0202 and SYNW1983) and all four D1 forms I and II (SYNW0983, SYNW1470, SYNW1919 and SYNW2151) displayed significantly higher levels. This was also the case for NADH Dehydrogenase I chain 2 (SYNW1873), ATP synthase subunit c, cytochrome b_6f complex subunit 4, apocytochrome b_6 and the cytochrome b_{559} beta chain. However, the reason for these differences is unclear.

PtrA binding to promoters. In light of the involvement of additional regulators in the P stress response (Tetu *et al*., 2009) we employed electrophoretic mobility shift assays (EMSA) as an alternative means of confirming that PtrA is a functional regulator that binds to the upstream regions of regulated genes. Despite attempts to optimise expression of recombinant PtrA in *Escherichia coli* a soluble product was produced in low quantities. Despite low yield, soluble PtrA was purified via metal affinity chromatography, concentrated and re-suspended in DNA binding buffer supplemented with potential cofactors, 5 mM $ZnCl_2$ and 10 µM cAMP. Figure 4 shows an EMSA experiment to investigate the binding of increasing amounts of purified PtrA to ~400 bp DNA fragments comprising the upstream regions of SYNW0165, SYNW0196, SYNW1018 and SYNW2390. Up to three distinct PtrA-DNA fragments can be observed for the upstream region of

SYNW0165, corresponding to 1, 2 and 3 PtrA target sites in this DNA fragment. Pre-incubation with non-specific competitor or 125x excess of unlabelled competitor DNA demonstrates that PtrA-binding to this DNA fragment is sequence-specific. While we were able to detect PtrA binding to specific promoter elements in SYNW0165, SYNW0196 and SYNW2390, no mobility shift was observed for the DNA fragment corresponding to the upstream region of *pstS* (SYNW1018) (Figure 4) which lends support to the hypothesis that PtrA is a transcription factor that regulates *psiP1* and these two phosphatases, but not *pstS*. In general the amount of DNA specifically bound to recombinant PtrA in each assay was low which might reflect a low proportion of correctly folded, active protein in the preparation. The relatively weak binding signals observed *in vitro* also suggest that additional co-factors or cooperative interactions with other transcription factors and possibly RNA polymerase are required for binding *in vivo.* Indeed, this could even involve cooperative interactions with PhoB at some target sites since a number of PtrA regulated genes display predicted pho boxes, e.g. three operons spanning cluster 3 (SYNW0440-0458) and a weak prediction for *psiP1* (Su *et al*., 2007). However, the majority of genes potentially regulated by PtrA, including the most highly expressed genes, do not possess pho boxes but have regions bound by PtrA, although with motifs we have not been able to determine. Thus despite our hypothesis that PtrA would be a phosphate regulatory factor with a mechanism similar to NtcA for nitrogen it seems to have distinct differences. Further bioinformatic and experimental work is required to identify PtrA-binding motifs in *Synechococcus* sp. WH8102.

The role of PtrA in response to P limitation in WH8102. Taken together, the respective organisation of the PhoB and PtrA regulons suggests a two-tiered response to the level of P-limitation in *Synechococcus* sp. WH8102. The first level of response is manifest by the affinity scavenging of inorganic P during early P stress involving PhoB-induced expression of *pstS* genes (SYNW1018, SYNW1286 and SYNW1815) as well as an elevation in the level of PtrA, either from co-transcription of *ptrA* with SYNW1018 or from a dedicated pho box immediately upstream of *ptrA*. Elevated levels of PtrA lead to the second level of response characterised by the scavenging of organic P involving PtrA-induced expression of two phosphatases (SYNW0196 and SYNW2390), and a predicted phytase (SYNW0762). Although SYNW2391-SYNW2390 is a predicted transcriptional unit with two tandem pho boxes found upstream of SYNW2391 (Su *et al*., 2007) the expression of these genes appears to be influenced separately by PtrA and PhoB. The EMSA presented here shows that PtrA binds to the promoter region of SYNW2390. This is further supported by the comparison of *ptrA* and *phoB* array data which confirms that SYNW2390 is predominantly regulated by PtrA and SYNW2391 regulated by PhoB. More complicated regulation could also be occurring for some genes with potentially both PtrA and PhoB binding sites. Thus, PtrA may be involved in a signal cascade that grades the cellular response initially to P stress, through P limitation and finally to chronic P starvation.

The role of PtrA in adaptation to P limitation in marine Synechococcus and Prochlorococcus. The gene for PtrA was first discovered in *Synechococcus* sp. WH7803 downstream of the gene for the periplasmic P-binding protein (PstS) which is highly induced in response to P stress in WH7803 (Scanlan *et al*., 1993; Scanlan *et al*., 1997). Since then *ptrA* homologues have been identified in the genomes of 10 out of 12 oceanic *Synechococcus* and 5 out of 12 *Prochlorococcus.* The phylogenetic profile of *ptrA* and *phoBR*

provides some interesting insights. Out of 12 complete, or nearly complete *Synechococcus* genomes, *ptrA* is absent from only two coastal strains isolated from the California Current, CC9311 (Clade I) and CC9902 (Clade IV) (Palenik *et al*., 2006). Both strains also lack *phoBR* and encode a single copy of *pstS* suggesting that they do not encounter P limitation in their environments. The *ptrA* gene is represented in fewer *Prochlorococcus* genomes (5/12) where it is predominantly found in low light-adapted ecotypes. It is also interesting to note the presence of *ptrA* in at least three strains where *phoBR* is absent or incomplete, BL107 (Clade IV) and *Prochlorococcus* strains SS120 and MIT9303. This may indicate that PtrA can function in the absence of PhoBR and possibly substitute for some form of P sensing and regulation in these strains.

When present, *ptrA* is found in a relatively conserved genome context in *Synechococcus* (Figure 5) where it is almost exclusively found downstream of *pstS*. *Synechococcus* sp. RCC307 is the only exception, where the *ptrA* gene is located in a unique context with no apparent linkage with co-located genes in other strains. Apart from oceanic cyanobacteria, PtrA homologues have not been identified in any other organism, although a potential, yet distant, candidate exists in *Cyanobium* sp. PCC 7001. PtrA is one of four distinct clusters of CRP-regulators found in marine *Prochlorococcus* and *Synechococcus*. The similarities in gene context suggest that *ptrA* has been vertically inherited in *Synechococcus*, where the surrounding genes, especially the upstream *pstS*, have been largely conserved. Moreover, *ptrA* is not located in any predicted island in the *Synechococcus* genomes sequenced so far (Dufresne *et al*., 2008). In comparison, in *Prochlorococcus* the gene order in the vicinity of *ptrA* is not well conserved. However, the gene is often found in a genomic region that contains all of the P uptake machinery (e.g. as found in *Prochlorococcus* MED4, NATL1A, NATL2A and MIT9312, Martiny *et al*., 2006). The implication here is that P uptake machinery, including the regulatory components *phoBR,* and *ptrA*, resides on a genomic region that may have been horizontally acquired by *Prochlorococcus* ecotypes that inhabit P deplete regions of some oceans (Martiny *et al*., 2006).

As noted previously (Tetu *et al*., 2009) those few genes that are significantly influenced by PtrA during P stress in *Synechococcus* sp. WH8102 (SYNW0165, SYNW0196, SYNW0762 and SYNW2390) are represented in a minority of other marine cyanobacterial genomes, suggesting that they have been laterally acquired and their inclusion in the P-regulon has been selected by environmental conditions (Palenik *et al*., 2003; Moore *et al*., 2005; Dufresne *et al*., 2008; Scanlan *et al*., 2009). The clustering of PtrA regulated genes throughout the genome also has implications for the mechanisms of gene gain (Figure 3). Firstly, these key genes are spread throughout the genome in WH8102 and do not occur in any recognisable island (Palenik *et al*., 2003; Dufresne *et al*., 2008), nor are they linked on the chromosome with their cognate regulator. This implies that each gene was independently acquired along with its respective PtrA binding sequence. On the other hand there are several clusters of genes that appear to be co-regulated, including the *swmB* cluster (SYNW0952-SYNW0960) and a large region of glycosyltransferases, LPS and cell wall biogenesis related genes (SYNW0424-SYNW0458) that are clearly genomic islands (Palenik *et al*., 2003; Dufresne *et al*., 2008). Whether they are directly regulated by PtrA is not clear but it is clear that the genes within these regions are co-regulated, in this case exhibiting lower levels of expression in the *ptrA* knockout in comparison to wild type.

Given that the genomic complement of P acquisition and regulation genes (Scanlan *et al*., 2009) and the responses to P stress (Martiny *et al*., 2006; Tetu *et al*., 2009) varies quite considerably between marine

picocyanobacteria it will be interesting to determine the role of PtrA in other marine strains. Overall, it is intriguing that PtrA appears to regulate accessory genes in WH8102, providing a potentially novel example of how laterally acquired genes have been recruited into an existing regulon in an environmental isolate. This example illustrates the importance of regulatory networks in coordinating the expression of the P stress response, despite relatively low regulatory capacity in this model oligotrophic cyanobacterium.

Acknowledgements

We would like to thank Bianca Brahamsha for assistance and guidance in preparing the *ptrA* gene knockout and for supplying axenic cultures of *Synechococcus* sp. WH8102. MO and SM were supported by the EU FP5 Program Margenes (QLRT-2001-0226), the FP6 EU Marine Genomics Network and NERC grants NE/C0005361/1, NE/F004249/1 and NE/D003385/1 to DJS.

References

Aiba H, Mizuno T (1994). A novel gene whose expression is regulated by the response-regulator, SphR, in response to phosphate limitation in *Synechococcus* species PCC7942. *Mol Microbiol* **13:** 25-34.

Aiba H, Nagaya M, Mizuno T (1993). Sensor and regulator proteins from the cyanobacterium *Synechococcus* species PCC7942 that belong to the bacterial signal-transduction protein families - implication in the adaptive response to phosphate limitation. *Mol Microbiol* **8:** 81-91.

Ammerman JW, Hood RR, Case DA, Cotner JB (2003). Phosphorus deficiency in the Atlantic: an emerging paradigm in oceanography. *EOS* **84**: 165-170.

Bertilsson S, Berglund O, Karl DM, Chisholm SW (2003). Elemental composition of marine *Prochlorococcus* and *Synechococcus*: Implications for the ecological stoichiometry of the sea. *Limnol Oceanogr*. **48:** 1721-1731.

Bessey OA, Lowry OH, Brock MJ (1946). A method for the rapid determination of alkaline phosphatase with five cubic millimeters of serum. *J Biol Chem* **164:** 321–329.

Blindauer CA (2008). Zinc-handling in cyanobacteria: an update. *Chem Biodiv* **5:** 1990-2013.

Brahamsha B (1996). A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl Environ Microbiol* **62:** 1747-1751.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P *et al*. (2001). Minimum information about a microarray experiment (MIAME) toward standards for microarray data. *Nat Genet* **29:** 365-371.

Coleman JE (2003). Structure and mechanism of alkaline phosphatase. *Ann Rev Biophys Biomol Struct* **21:** 441-483.

Cotner JB, Ammerman JW, Peele, ER, Bentzen E (1997). Phosphorus-limited bacterioplankton growth in the Sargasso Sea. *Aquat Microb Ecol* **13**: 141-149.

Dufresne A, Ostrowski M, Scanlan DJ, Garczarek L, Mazard S, Palenik BP *et al.* (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol* **9**: R90.

Dufresne A, Salanoubat M, Partensky F, Artiguenave F, Axmann IM, Barbe V *et al*. (2003). Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal

oxyphototrophic genome. *Proc Natl Acad Sci USA* **100:** 10020-10025.

Elledge SJ, Zhou Z, Allen JB (1992). Ribonucleotide reductase: regulation, regulation, regulation. *Trends Biochem Sci* **17:** 119-123.

Heldal M, Scanlan DJ, Norland S, Thingstad F, Mann NH (2003). Elemental composition of single cells of various strains of marine *Prochlorococcus* and *Synechococcus* using X-ray microanalysis. *Limnol Oceanogr* **48:** 1732-1743.

Herrero A, Muro-Pastor AM, Flores E (2001). Nitrogen control in cyanobacteria. *J Bacteriol* **183:** 411-425.

Hirani TA, Suzuki I, Murata N, Hayashi H, Eaton-Rye JJ (2001). Characterization of a two-component signal transduction system involved in the induction of alkaline phosphatase under phosphate-limiting conditions in *Synechocystis* sp. PCC 6803. *Plant Mol Biol* **45**: 133-44.

Karl DM, Tien G (1997). Temporal variability in dissolved phosphorus concentrations in the subtropical North Pacific Ocean. *Mar Chem* **56**: 77-96.

Korner H, Sofia HJ, Zumft WG (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* **27:** 559-592.

Lindell D, Padan E, Post AF (1998). Regulation of *ntcA* expression and nitrite uptake in the marine *Synechococcus* sp. strain WH7803. *J Bacteriol* **180:** 1878-86.

Lohan MC, Statham PJ, Crawford, DW (2002). Total dissolved zinc in the upper water column of the subarctic North East Pacific. *Deep-Sea Res Part II* **49:** 5793-5808.

Luque I, Flores E, Herrero A (1994). Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* **13:** 2862-2869.

Mackey KRM, Paytan A, Grossman AR, Bailey S (2008). A photosynthetic strategy for coping in a high-light, low-nutrient environment. *Limnol Oceanogr* **53:** 900-913.

Martiny AC, Coleman ML, Chisholm SW (2006). Phosphate acquisition genes in *Prochlorococcus* ecotypes: Evidence for genome-wide adaptation. *Proc Natl Acad Sci USA* **103**: 12552-12557.

McCarren J, Brahamsha B (2005). Transposon mutagenesis in a marine *Synechococcus* strain: Isolation of swimming motility mutants. *J Bacteriol* **187:** 4457-4462.

McCarren J, Brahamsha B (2007). SwmB, a 1.12-megadalton protein that is required for nonflagellar swimming motility in *Synechococcus*. *J Bacteriol* **189:** 1158-1162.

Moore LR, Ostrowski M, Scanlan DJ, Feren K, Sweetsir T (2005). Ecotypic variation in phosphorus acquisition mechanisms within marine picocyanobacteria. *Aquat Microb Ecol* **39:** 257-269.

Morel FMM, Rueter JG, Anderson DM, Guillard RRL (1979). Aquil - Chemically defined phytoplankton culture-medium for trace-metal studies. *J Phycol* **15:** 135-141.

Palenik B, Brahamsha B, Larimer FW, Land M, Hauser L, Chain *et al*. (2003). The genome of a motile marine *Synechococcus*. *Nature* **424:** 1037-1042.

Palenik B, Ren QH, Dupont CL, Myers GS, Heidelberg JF, Badger JH *et al*. (2006). Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. *Proc Natl Acad Sci USA* **103:** 13555-13559.

Partensky F, Hess WR, Vaulot D (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106-127.

Salgado H, Gama-Castro S, Martinez-Antonio A, Diaz-Peredo E, Sanchez-Solano F, Peralta-Gil M *et al*. (2004). RegulonDB (version 4.0): transcriptional regulation, operon organization and growth conditions in *Escherichia coli* K-12. *Nucleic Acids Res* **32:** D303-306.

Scanlan D, Bourne J, Mann N (1997). A putative transcriptional activator of the Crp/Fnr family from the marine cyanobacterium *Synechococcus* sp. WH7803. *J Appl Phycol* **8:** 565-567.

Scanlan D, Mann N, Carr N (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 D, Silman N, Donald K, Wilson W, Carr N, Joint I, Mann N (1997). An immunological approach to detect phosphate stress in populations and single cells of photosynthetic picoplankton. *Appl Environ Microbiol* **63:** 2411-2420.

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:** 249-99.

Snyder DS, Brahamsha B, Azadi P, Palenik B (2009). Structure of compositionally simple lipopolysaccharide from marine *Synechococcus*. *J Bacteriol* **191:** 5499-5509.

Silberbach M, Hüser A, Kalinowski J, Pühler, A, Walter B, Kramer R, Burkovski A (2005). DNA microarray analysis of the nitrogen starvation response of *Corynebacterium glutamicum*. *J Biotech* **119:** 357-367.

Su ZC, Dam P, Chen X, Olman V, Jiang T, Palenik B, Xu Y (2003). Computational inference of regulatory pathways in microbes An application to phosphorus assimilation pathways in *Synechococcus* WH8102. *Genome Informatics* **14:** 3-13.

Su, ZC, Olman V, Mao FL, Xu Y (2005). Comparative genomics analysis of NtcA regulons in cyanobacteria: regulation of nitrogen assimilation and its coupling to photosynthesis. *Nucleic Acids Res* **33:** 5156-5171.

Su ZC, Olman V, Xu Y (2007). Computational prediction of Pho regulons in cyanobacteria. *BMC Genomics* **8**: 156.

Suzuki S, Ferjani A, Suzuki I, Murata N (2004). The SphS-SphR two component system is the exclusive sensor for the induction of gene expression in response to phosphate limitation in *Synechocystis*. *J Biol Chem* **279:** 13234-13240.

Tai V, Paulsen IT, Phillippy K, Johnson DA, Palenik B (2009). Whole-genome microarray analyses of *Synechococcus*-*Vibrio* interactions. *Environ Microbiol* **11:** 2698-2709.

Tetu SG, Brahamsha B, Johnson DA, Tai V, Phillippy K, Palenik B, Paulsen IT (2009). Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp WH8102. *ISME J* **3:** 835-849.

Thingstad TF, Krom MD, Mantoura RFC, Flaten GAF, Groom S, Herut B *et al*. (2005). Nature of phosphorus limitation in the ultraoligotrophic eastern Mediterranean. *Science* **309**: 1068-1071.

Tusher VG, Tibshirani R, Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* **98:** 5116-5121.

Van Mooy BAS, Rocap G, Fredricks HF, Evans CT, Devol AH (2006). Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proc Natl Acad Sci*

USA **103:** 8607-8612.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al*. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66-74.

Waterbury JB, Willey JM (1988) Isolation and growth of marine planktonic cyanobacteria. *Methods Enzymol* **167:** 100-105.

Zubkov MV, Mary I, Woodward EMS, Warwick PE, Fuchs BM, Scanlan DJ, Burkill PH (2007). Microbial control of phosphate in the nutrient-depleted North Atlantic subtropical gyre. *Environ Microbiol* **9**: 2079-2089.

Zwirglmaier 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.

Figure 1. A) Comparison of growth and P-uptake of the *ptrA* mutant compared with the wild type strain in low P media in batch cultures and, B) The development of phosphatase activity in mutant and wild type cultures. C) The induction of PstS expression in wild type during early P stress

Figure 2. Comparison of the relative expression of P regulon genes from microarray analysis and qRT-PCR that were differentially expressed in the *ptrA* (SYNW1019) and *phoB* (SYNW0947) mutants in comparison to early P-stress in the wild type strain. Negative values indicate genes that were down-regulated in each mutant. Positive values indicate genes that were up-regulated in the wild-type strain during early P stress. Genes with relative expression greater than 2-fold are highlighted in colour, (red and green for positive and negative values, respectively).

Figure 3. Genomic distribution of genes that were differentially expressed in the *ptrA* mutant relative to wild type in response to P stress in *Synechococcus* sp. WH8102. Negative values indicate genes that were down-regulated in the mutant. Positive values indicate genes that were up-regulated in the mutant. All genes with significant SAM values are plotted. Genes with altered expression levels greater than 2-fold are highlighted with triangles. Single genes and clusters of genes of interest discussed in the text are annotated with cluster numbers.

Figure 4. Electrophoretic mobility shift assay demonstrating specific binding of recombinant PtrA (6His-PtrA) to the upstream regions (-350 to +50 bp relative to the start codon) of SYNW0165, SYNW0196 and SYNW2390 but not SYNW1018. The sloping bar indicates increasing concentration of the recombinant protein.

Figure 5. The conserved genomic context of *ptrA* in representative *Synechococcus* and *Prochlorococcus* genomes. This figure highlights the genes surrounding *ptrA* (labelled B) in marine picocyanobacterial genomes. The position and spacing of *ptrA* downstream of *pstS* (A) is well conserved in all *Synechococcus* genomes except for BL107. When *ptrA* is present in *Prochlorococcus* the gene order is less conserved, however, the gene is found in a genomic region which contains many genes required for P-uptake (e.g. *pstS, pstCAB, som* and a secreted phosphatase) and regulation (*phoBR*). Gene names and descriptions are given in the legend.

Supplementary Figure 1. Insertional inactivation of *ptrA*. PCR with primers spanning the gene confirm the absence of a wild type copy in the mutant (ptrAF4-ptrAR1) and the integration of the construct into the genome (ptrAF4-ecoF1). Primer combinations PtrAF4 (5′-AGGCTGCTGTGAACAAGATCGGC-3’), ptrAR1 (5′-GGATCAGCTGACCACTCGCATG-3’) and ecoF1 (5′-CCCGAAAAGTGCTCCGAGAACGG-3′) amplified fragments of the expected size from mutant or wild type DNA, and not the other, confirming the integration of the plasmid into the chromosome. Failure to amplify a fragment of the wild-type size from the mutant strain’s DNA confirmed the absence of the intact gene among the clonal population of mutant cells. This same DNA sample was used as a template in another PCR, utilising primers directed to the *petB* gene as a positive control to confirm that the DNA was of sufficient quality for PCR amplification.

625

Table 1. *Synechococcus* sp. WH8102 genes whose expression was down-regulated by more than 2.0-fold in the *ptrA* knockout relative to wild type during P-stress. For comparison, results from an independent experiment comparing the gene expression in wild type WH8102 during early P stress compared with P replete growth conditions are included.

			Log_2 - fold change	
Gene ID	**Gene or predicted function**	**SAM Score[1]**	**PtrA mutant vs wild type**	**wild type early P stress[2]**
SYNW2390	alkaline phosphatase/5' nucleotidase	-8.22	-3.39	3.98
SYNW0196	alkaline phosphatase	-12.3	-3.22	3.07
SYNW0165	conserved hypothetical (*psiP1*)	-8.03	-3.17	3.98
SYNW0762	conserved hypothetical (predicted phytase)	-7.6	-1.93	2.47
SYNW1333	conserved hypothetical	-3.03	-1.1	1.99
Cluster of genes centred on Zn ABC transport system (cluster 9 on Figure 3)				
SYNW2477	Ferredoxin--nitrite reductase	-5.14	-1.57	-0.39
SYNW2478	conserved hypothetical protein	-3.77	-1.6	-[3]
SYNW2479	ABC transporter component, possibly Zn transport.	-7.05	-1.57	-
SYNW2480	ABC transporter, ATP binding component, possibly zinc transport	-15.31	-1.82	-
SYNW2482	conserved hypothetical protein	-4.32	-1.48	-
SYNW2483	conserved hypothetical protein	-4.0	-1.31	0.24
SYNW2484	hypothetical	-5.54	-1.59	-
SYNW2485	putative cyanate ABC transporter	-3.17	-1.03	-
Cluster of genes around *swmB*, mostly unknown function (cluster 6)				
SYNW0952	conserved hypothetical protein	-5.57	-1.46	-0.27
SYNW0953	*swmB*	-6.04	-1.97	1.24
SYNW0954	conserved hypothetical protein	-4.41	-1.96	0.33
SYNW0955	hypothetical	-5.9	-1.75	-
SYNW0956	conserved hypothetical protein	-6.15	-1.81	0.45
SYNW0957	conserved hypothetical protein	-11.37	-1.78	-
SYNW0958	similar to leukotoxin secretion protein	-5.9	-2.19	1.07
SYNW0959	putative multidrug efflux ABC transporter	-2.83	-1.75	0.87
SYNW0960	conserved hypothetical protein	-2.73	-1.57	-
SYNW0961	hypothetical protein	-2.78	-1.03	-
SYNW0962	Putative 4-alpha-glucanotransferase	-9.07	-1.3	0.32
cluster of genes involved in cell surface modification (cluster 3)				
SYNW0424	Possible HMGL-like family protein	-3.2	-1.43	-0.59
SYNW0425	Putative CMP-KDO synthetase	-3.3	-1.28	-
SYNW0426	Possible haloacid dehalogenase-like hydrolase family protein	-3.81	-1.36	-0.68
SYNW0427	possible multidrug efflux ABC transporter	-3.99	-1.09	-0.43
SYNW0429	hypothetical	-3.25	-1.16	-
SYNW0431	hypothetical	-3.88	-1.12	-0.43
SYNW0432	Putative short-chain dehydrogenase family protein	-2.25	-1.21	-
SYNW0433	hypothetical	-3.76	-1.17	-0.53
SYNW0434	conserved hypothetical protein	-5.7	-1.69	-0.43
SYNW0435	putative glutamine amidotransferase	-4.78	-1.36	-0.36
SYNW0436	putative cyclase *hisF*	-4.24	-1.55	-0.3
SYNW0438	possible polysaccharide deacetylase (xylanase, chitin deacetylase)	-4.47	-1.3	-0.39
SYNW0440	hypothetical	-4.51	-1.2	-0.15
SYNW0441	conserved hypothetical protein	-3.3	-1.39	-0.25
SYNW0442	conserved hypothetical protein	-3.77	-1.3	-
SYNW0445	putative nucleotide sugar epimerase	-1.86	-1.09	-0.45
SYNW0446	putative aminotransferase (DegT family)	-1.99	-1.14	-0.25
SYNW0447	putative hexapeptide transferase family protein	-2.46	-1.01	-0.22
SYNW0448	putative N-acetylneuraminic acid synthetase	-2.08	-1.28	-0.58
SYNW0450	putative sugar-phosphate nucleotide transferase	-4.9	-1.77	-0.2
SYNW0451	putative O-acetyltransferase	-4.58	-1.26	-0.24
SYNW0452	hypothetical	-4.61	-1.53	-

SYNW0453	possible glycosyltransferase	-2.76	-1.35	-
SYNW0454	possible glycosyltransferase	-2.97	-1.67	-0.15
SYNW0456	possible glycosyltransferase	-4.43	-1.49	0.24
SYNW0457	hypothetical	-2.89	-1.7	0.14
SYNW0458	possible glycosyltransferase group I	-2.5	-1.29	-0.34
phycobilisome biosynthesis genes (cluster 7)				
SYNW2003	CpeT homolog	-3.09	-1.23	-0.12
SYNW2004	CpeR homolog, phycoerythrin linker-proteins region	-8.78	-1.37	-0.78
SYNW2006	hypothetical	-3.65	-1.16	-0.52
SYNW2011	bilin biosynthesis protein MpeU (PBS lyase HEAT-like repeat)	-4.12	-1.39	-0.17
SYNW2013	putative bilin biosynthesis protein (CpeY)	-3.22	-1.41	-
further genes of interest				
SYNW0645	putative glycosyltransferase family 2 protein	-2.68	-1.41	-
SYNW0882	Sodium/glutamate symporter	-3.32	-1.35	0.24
SYNW0887	possible oxidase	-6.07	-1.12	-
SYNW1529	cytochrome *c* oxidase subunit I	-2.63	-1.23	-
SYNW1660	possible transcription regulator	-2.77	-1.02	-
SYNW1662	phage integrase family	-4.27	-1.04	-
SYNW1861	possible cytochrome c oxidase subunit II	-1.96	-1.07	-
SYNW1862	cytochrome c oxidase subunit I	-2.02	-1.14	0.43
SYNW1946	PhoH family protein	-7.4	-1.36	-
SYNW2105	putative transcriptional regulator	-2.96	-1.07	-
SYNW2293	possible hemolysin-type calcium-binding protein	-2.3	-1.06	-
SYNW2409	putative hemolysin-type calcium-binding protein; similar to HlyA	-7.75	-1.53	-
A further 29 genes of unknown function are listed in supplementary table 1				

1. Statistical Analysis of Microarray score.

2. Data from Tetu *et al*., 2009

3. no significant difference in expression levels

Table 2. *Synechococcus* sp. WH8102 genes whose expression was up-regulated by more than 2.0-fold in the *ptrA* knockout relative to wild type during P-stress. For comparison, results from an independent experiment comparing the gene expression in wild type WH8102 during early P stress compared with P replete growth conditions are included.

			Log_2 - fold change	
Gene ID	**Gene Name**	**SAM Score**	**PtrA mutant vs wild type**	**wild type early P stress[1]**
SYNW0189	conserved hypothetical protein	4.29	1.02	-0.11
SYNW0202	photosystem II reaction center L protein (PSII 5 kDa protein)	3.6	1.11	-[2]
SYNW0203	cytochrome b_{559} beta chain	3.27	1.08	-
SYNW0253	ammonium transporter family	2.08	1.13	-
SYNW0331	conserved hypothetical protein	4.82	1.84	0.3
SYNW0490	ATP synthase subunit c	3.83	1.08	-
SYNW0510	conserved hypothetical protein	10.45	1.03	-
SYNW0513	GroES chaperonin	2.67	1.19	0.44
SYNW0514	GroEL chaperonin	2.18	1.07	0.5
SYNW0615	RNA polymerase beta prime subunit	6.38	1.09	0.38
SYNW0670	conserved hypothetical protein	4.38	1.27	0.32
SYNW0673	O-acetylserine (thiol)-lyase A	3.06	1.01	0.45
SYNW0778	conserved hypothetical	7.87	1.14	-0.33
SYNW0810	conserved hypothetical protein	5.65	1.06	-
SYNW0983	photosystem II D1 protein form II	2.13	1.23	-
SYNW1019	possible transcriptional regulator	1.88	1.81	2.71
SYNW1147	ribonucleotide reductase (Class II)	4.83	1.3	-0.57
SYNW1278	heat shock protein HtpG	2.54	1.2	0.6
SYNW1405	conserved hypothetical protein	3.27	1.23	-0.34
SYNW1443	hypothetical	3.15	1.36	-
SYNW1470	photosystem II D1 protein form I	2.74	1.3	-
SYNW1511	conserved hypothetical	7.84	1.74	-
SYNW1512	conserved hypothetical	4.03	1.38	0.27
SYNW1524	putative sulfate transporter	4.64	1.24	-0.14
SYNW1645	putative Ketopantoate hydroxymethyltransferase	6.37	1.21	-
SYNW1687	NifU-like protein	8.83	1.22	-
SYNW1694	30S ribosomal protein S4	4.68	1.17	-
SYNW1778	conserved hypothetical protein	7.29	1.17	0.39
SYNW1790	hypothetical	3.24	2.01	-0.41
SYNW1796	conserved hypothetical protein	2.91	1.15	-0.33
SYNW1854	60 kD chaperonin 2, GroEL homolog 2	6.98	1.2	0.36
SYNW1873	NADH dehydrogenase I chain 2 (or N)	6.8	1.06	-
SYNW1919	photosystem II D1 protein form II	2.49	1.34	-
SYNW1950	hypothetical	2.28	1.61	0.39
SYNW1966	cytochrome b_6f complex subunit 4 (17 kDa polypeptide)	4.4	1.1	0.36
SYNW1967	apocytochrome b_6	7.85	1.34	-
SYNW1983	photosystem II reaction center T protein	2.65	1.28	-0.5
SYNW2016	C-phycoerythrin class I alpha chain	2.32	1.11	-
SYNW2038	conserved hypothetical protein	8.02	1.07	-0.55
SYNW2067	50S ribosomal protein L3	2.26	1.09	-
SYNW2068	50S ribosomal protein L4	4.93	1.29	-0.44
SYNW2069	50S ribosomal protein L23	4.29	1.07	-0.35
SYNW2071	30S ribosomal protein S19	9.03	1.14	-0.48
SYNW2072	50S ribosomal protein L22	7.46	1.81	-0.79
SYNW2073	30S ribosomal protein S3	4.01	1.19	-0.29
SYNW2074	50S ribosomal protein L16	13.74	1.61	-0.62
SYNW2075	50S ribosomal protein L29	7.42	1.04	-0.33
SYNW2076	30S ribosomal protein S17	17.3	1.23	-0.44
SYNW2077	50S ribosomal protein L14	9.43	1.29	-0.42

SYNW2082	50S ribosomal protein L18	5.29	1.81	-1.03
SYNW2083	30S ribosomal protein S5	6.17	2.01	-0.62
SYNW2084	50S ribosomal protein L15	10.43	1.46	-
SYNW2091	50S ribosomal protein L17	5.03	1	-
SYNW2135	30S ribosomal protein S12	7.64	1.56	-
SYNW2136	30S ribosomal protein S7	10.83	1.2	-0.29
SYNW2151	photosystem II D1 protein form II	3.49	2.63	-
SYNW2172	conserved hypothetical protein	6.6	1.94	-0.73
SYNW2173	conserved hypothetical protein	8.73	1.8	-
SYNW2174	conserved hypothetical protein	6.84	1.63	-
SYNW2176	possible serine protease	3.23	1.34	0.53
SYNW2180	possible high light inducible protein	2.18	1.44	0.32
SYNW2238	thymidylate kinase	3.62	1.1	-
SYNW2312	conserved hypothetical protein	3.62	1.09	-
SYNW2508	Molecular chaperone DnaK2, heat shock protein hsp70-2	3.28	1.23	-

1. Data from Tetu *et al*., 2009

2. no significant difference in expression levels

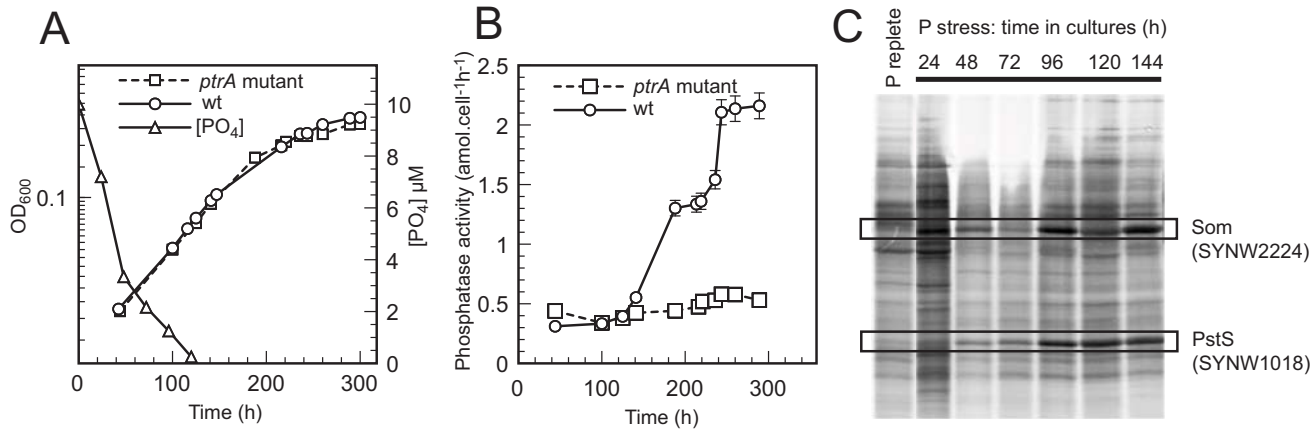
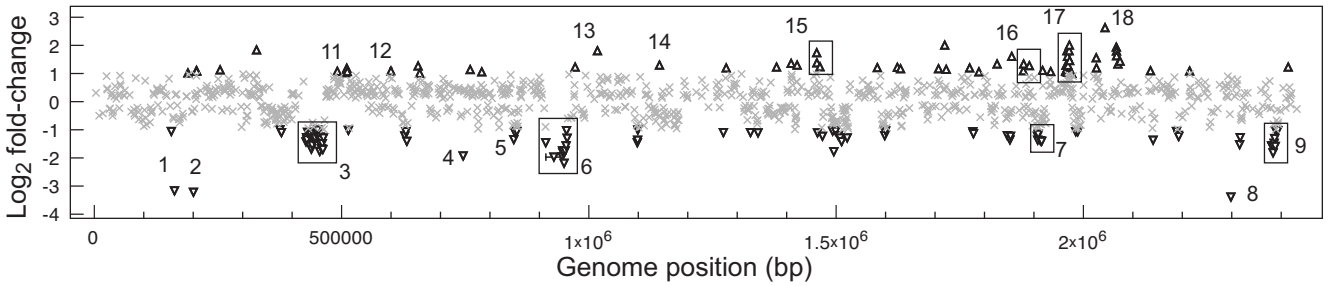

A
ptrA mutant
wt
[PO4]
OD600
0.1
[PO4] µM
Time (h)
B
ptrA mutant
wt
Phosphatase activity (amol.cell-1h-1)
Time (h)
C
P replete
P stress: time in cultures (h)
24 48 72 96 120 144
Som
(SYNW2224)
PstS
(SYNW1018)

figure 1.

Gene ID	Gene Name or Function	qRT-PCR[1]	SAM Score[2]	Log_2 fold change PtrA mutant	PhoB mutant[5]	wild type[5]
SYNW2390	alkaline phosphatase/5 nucleotidase	-2.41	-8.22	**-3.39**	**-2.45**	**3.98**
SYNW0196	alkaline phosphatase	-0.77	-12.30	**-3.22**	**-1.13**	**3.07**
SYNW0165	*psiP1*: P-starvation inducible polypeptide	-5.88	-8.03	**-3.17**	**-1.06**	**3.98**
SYNW0762	possible phytase	nd	-7.60	**-1.93**	**-1.13**	**2.47**
SYNW1333	conserved hypothetical	nd	-3.03	**-1.10**	**-1.78**	**1.99**
SYNW1019	*ptrA*: potential transcriptional regulator		na[3]	na[3]	**-1.22**	**2.71**
SYNW2391	alkaline phosphatase	nd	2.59	0.80	**-2.11**	**3.95**
SYNW1018	*pstS*: ABC transporter, substrate binding protein, phosphate	-0.49	-[4]	-	**-2.63**	**4.08**
SYNW0953	*swmB*: cell surface protein required for swimming motility	nd	-6.04	**-1.97**	-	**1.2**
SYNW0958	similar to leukotoxin secretion protein	nd	-5.90	**-2.19**	-	**1.1**
SYNW1286	ABC transporter, substrate binding protein, phosphate	nd	-	-	-0.31	**1.02**
SYNW1815	*pstS*: ABC transporter, substrate binding protein, phosphate	-0.48	-	-	-	**1.54**
SYNW2507	*pstS*: ABC transporter, substrate binding protein, phosphate	-0.76	-	-	-0.22	0.90
SYNW2224	*som:* Possible porin	nd	-	-	-0.63	**3.30**
SYNW1270	*pstC*: Putative phosphate ABC transporter	nd	-	-	-0.23	**1.84**
SYNW1271	*pstA*: Putative phosphate ABC transporter	nd	-	-	-	**1.31**
SYNW1272	*pstB*: Putative phosphate ABC transporter	nd	-	-	-0.47	**1.37**

1. Log_2 fold difference between mutant and wild type
2. Output of SAM analysis for comparison of gene expression in the *ptrA* mutant relative to wild type during early P stress
3. not applicable in insertion mutant
4. no significant difference in gene expression
5. data from Tetu *et al.,* 2009



genes of interest down-regulated in the *ptrA* mutant relative to wild type

1. *psiP1* (SYNW0165)
2. *phoA* (SYNW0196)
3. Glycosyltransferases, LPS and cell wall biogenesis related genes (SYNW0424-SYNW0456)
4. possible phytase (SYNW0762)
5. PTOX
6. *swmB* cluster including genes of unknown function
7. Phycobilisome gene cluster (5 genes)
8. *phoA* (SYNW2390)
9. Zn^{2+} ABC transport system, cyanate lyase, ferredoxin-nitrite reductase

genes of interest up-regulated in the *ptrA* mutant relative to wild type

11. ATP synthase subunit C
12. RNA polymerase β' subunit
13. *ptrA* (partial transcript)
14. ribonucleotide reductase
15. SYNW1511,SYNW1512, putative sulfate transporter
16. cytochrome b_6f
17. ribosomal protein gene cluster (16 genes)
18. *psbA* PSII D1 protein

figure 3

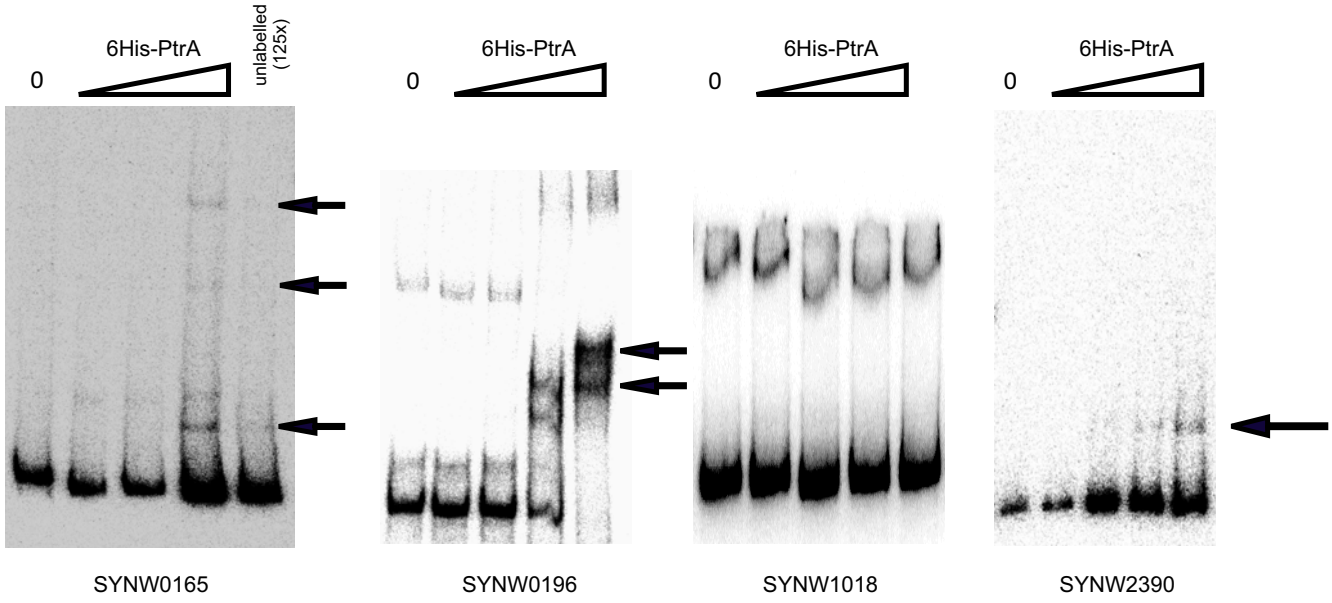
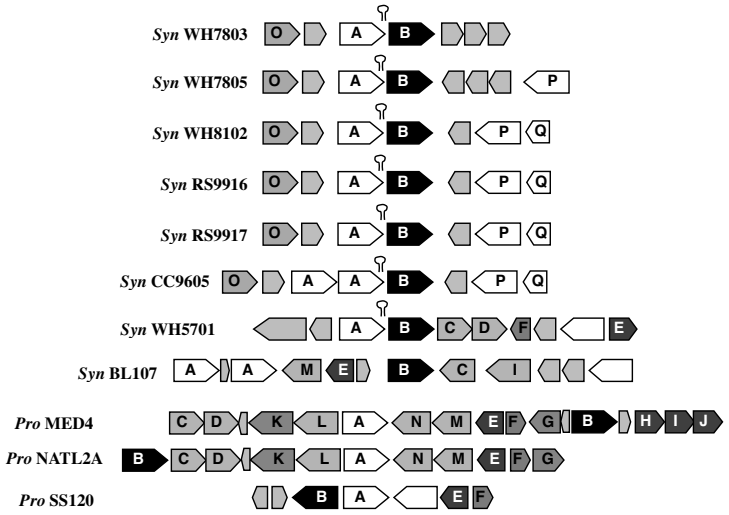

0
6His-PtrA
unlabelled (125x)
SYNW0165
0
6His-PtrA
SYNW0196
0
6His-PtrA
SYNW1018
0
6His-PtrA
SYNW2390

figure 4.



code	gene	description
A	*pstS*	periplasmic phosphate binding protein
B	***ptrA***	**CRP family transcriptional regulator**
C	*phoB*	Two component response regulator for phosphate
D	*phoR*	Two component histidine kinase for phosphate
E	*gap3*	glyceraldehyde-3-phosphate dehydrogenase
F	*arsR*	possible transcriptional regulator ArsR family
G	*arsC*	arsenate reductase
HIJ	*pstCAB*	ABC transport system for phosphate
K	-	putative secreted phosphatase
L	*som*	porin homolog
M	*proP*	permease of the major facilitator superfamily
N	*chr1*	chromate transporter, CHR family
O	*dnaQ*	DNA polymerase III, epsilon subunit
P	*msrA*	peptide methionine sulfoxide reductase
Q	*rpaC*	putative regulator

hypothetical or conserved hypothetical gene

predicted transcription terminator

figure 5.