

## ORIGINAL ARTICLE

# Transcriptomic and microRNAomic profiling reveals multi-faceted mechanisms to cope with phosphate stress in a dinoflagellate

Xinguo Shi<sup>1</sup>, Xin Lin<sup>1</sup>, Ling Li<sup>1</sup>, Meizhen Li<sup>1</sup>, Brian Palenik<sup>2</sup> and Senjie Lin<sup>1,3</sup>

<sup>1</sup>State Key Laboratory of Marine Environmental Science and College of Ocean and Earth Sciences, Xiamen University, Xiamen, China; <sup>2</sup>Scripps Institution of Oceanography, University of California, San Diego, CA, USA and <sup>3</sup>Department of Marine Sciences, University of Connecticut, Groton, CT, USA

Although gene regulation can occur at both transcriptional and epigenetic (microRNA) levels, combined transcriptomic and microRNAomic responses to environmental stress are still largely unexplored for marine plankton. Here, we conducted transcriptome and microRNAome sequencing for *Prorocentrum donghaiense* to understand the molecular mechanisms by which this dinoflagellate copes with phosphorus (P) deficiency. Under P-depleted conditions, G1/S specific cyclin gene was markedly downregulated, consistent with growth inhibition, and genes related to dissolved organic phosphorus (DOP) hydrolysis, carbon fixation, nitrate assimilation, glycolysis, and cellular motility were upregulated. The elevated expression of ATP-generating genes (for example, rhodopsin) and ATP-consuming genes suggests some metabolic reconfiguration towards accelerated ATP recycling under P deficiency. MicroRNAome sequencing revealed 17 microRNAs, potentially regulating 3268 protein-coding genes. Functional enrichment analysis of these microRNA-targeted genes predicted decreases in sulfatide (sulfolipid) catabolism under P deficiency. Strikingly, we detected a significant increase in sulfolipid sulfatide content (but not in sulphoquinovosyldiacylglycerol content) and its biosynthesis gene expression, indicating a different sulfolipid-substituting-phospholipid mechanism in this dinoflagellate than other phytoplankters studied previously. Taken together, our integrative transcriptomic and microRNAomic analyses show that enhanced DOP utilization, accelerated ATP cycling and repressed sulfolipid degradation constitute a comprehensive strategy to cope with P deficiency in a model dinoflagellate.

The ISME Journal advance online publication, 26 May 2017; doi:10.1038/ismej.2017.81

## Introduction

Marine phytoplankton experience environmental fluctuations in light intensity, nutrient availability, temperature and pH. In the future ocean, warming and intensified stratification may bring greater nutrient stress. Phosphorus (P) is an essential nutrient for phytoplankton growth as it is required for cellular structures (membranes, DNA, RNA), metabolism (nucleotides, NADH, NADPH), storage of energy (ATP), cell signaling (cAMP, IP3), and biochemical regulation (protein phosphorylation) (Karl, 2014; Lin *et al.*, 2016). In many parts of the ocean, dissolved inorganic phosphorus (DIP), the chemical form of P that can be taken up and assimilated directly, can be diminishingly scarce (Karl, 2014; Lin *et al.*, 2016). P can be a limiting

factor for phytoplankton growth even in eutrophic coastal waters where excess nitrogen is introduced due to human activities (Lin *et al.*, 2016). Various strategies have been identified in phytoplankton to cope with low P availability, including changing P transporters to increase uptake (Perry, 1976), switching to dissolved organic phosphorus (DOP) for P source (Dyhrman and Palenik, 1997; Dyhrman *et al.*, 2007; Karl, 2014), remodeling cell membranes (Shemi *et al.*, 2016), lowering P quota, bypassing P consuming process in glycolysis reactions (Wurch *et al.*, 2011; Dyhrman *et al.*, 2012), and recycling lipid P (Dyhrman *et al.*, 2012). The realization of these strategies may involve regulation of gene expression at multiple levels, ranging from transcriptional, post-transcriptional, to post-translational. While transcriptional regulation has been extensively studied, post-transcriptional regulation (for example, via microRNA) of phytoplankton responses to P stress has not been explored even though its existence has been reported (O'Rourke *et al.*, 2013). This study was aimed to address this gap of research by investigating genome-wide responses to P stress in a dinoflagellate.

Correspondence: S Lin, Department of Marine Sciences, University of Connecticut, 1080 Shennecossett Road, Groton, CT 06340, USA.

E-mail: senjie.lin@uconn.edu

Received 17 November 2016; revised 18 April 2017; accepted 25 April 2017

Dinoflagellates generally have limited transcriptional regulation (Lin, 2011), as demonstrated in field work (Alexander *et al.*, 2015) and culture studies (Moustafa *et al.*, 2010). Recent discoveries of microRNAs in dinoflagellates (Baumgarten *et al.*, 2013; Lin *et al.*, 2015) suggest that this form of post-transcriptional regulation may be a major mechanism controlling gene expression in dinoflagellates. We conducted transcriptome and microRNAome sequencing for *Prorocentrum donghaiense*, a model dinoflagellate species prone to form blooms in low P environments (Lu *et al.*, 2005), grown under P-replete and P-deficient conditions. Our data provide the most comprehensive insights to date of how a dinoflagellate copes with P deficiency with some novel mechanisms, meanwhile demonstrating the power of combined transcriptomic and microRNAomic analysis in gaining a panoramic view of phytoplankton responses to phosphorus stress.

## Materials and methods

### *Cultures and physiological measurements*

*P. donghaiense* cultures were grown in 5 l seawater medium at 20 °C under a photon flux of  $100 \mu\text{E m}^{-2} \text{s}^{-1}$  with a 14:10 h light:dark cycle. P-replete cultures were grown in L1 medium (without silicate) with  $36 \mu\text{M}$  phosphate concentration. The P-depleted cultures were acclimated to  $1 \mu\text{M}$  of phosphate with other nutrients being the same as in L1 medium. Each condition was set up in triplicate. All cultures were sampled daily for cell count and measurement of physiological parameters. Samples were collected on day 3 for RNA-seq, microRNA sequencing, qRT-PCR, western blot and LC-MS analyses and again on day 6 for LC-MS analysis for both the P-replete and P-depleted cultures. On both days, phosphate concentrations in the P-depleted cultures were essentially undetectable while that in the P-replete cultures were above  $15 \mu\text{M}$ .

The cell concentration was measured using a Sedgwick-Rafter counting chamber under a microscope. Chlorophyll *a* was measured fluorometrically. Fv/Fm ratio as an indicator of photosynthetic efficiency was measured with a pulse amplitude modulated (PAM) fluorometer (Walz, Effelrich, Germany). Cell diameter was estimated by Z1 particle counter (Beckman Coulter, Brea, VA, USA). DIP concentration in the growth media was measured using a standard molybdate method. Alkaline phosphatase activity (APA) was determined as reported previously (Lin *et al.*, 2012).

### *RNA isolation and sequencing*

About  $10^7$  cells were collected on the third day of the experiment and homogenized in TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) to extract RNA as previously reported (Lin *et al.*, 2010). Ethanol precipitation instead of RNA binding column was used in this isolation, which enables us to analyze both mRNA

and miRNA from one sample (Mraz *et al.*, 2009). The RNA concentration was assessed using RNA 6000 Nano Lab Chip Kit on Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The RNA integrity number of the samples was all above 6.0. Then these isolated RNA were used for sequencing and qRT-PCR analysis. Equal amounts of RNA masses that were extracted from the triplicate samples were pooled for RNA-seq and microRNA sequencing.

A total of 3  $\mu\text{g}$  RNA was used to generate paired-end RNA-seq library using Illumina TruSeq™ RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). Small RNA libraries were generated using NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB, Ipswich, MA, USA) with the same amount of total RNA mass as used in RNA-seq. The libraries were purified (AMPure XP system) and quantified using the Agilent Bioanalyzer 2100 system and then sequenced on an Illumina HiSeq 2000 platform. The raw data have been deposited to NCBI in SRA (SRX2554989, SRX2555036, SRX2557351 and SRX2557352).

### *Bioinformatic analyses*

Raw reads for RNA-Seq and microRNA sequencing were cleaned up by trimming adapter sequences, removing poly-N containing reads ( $\geq 10\%$  of nucleotides in each read) and filtering low quality reads (Q-score  $\leq 20$  for  $\geq 50\%$  of nucleotides in each read) using a proprietary program (Tian *et al.*, 2012). All the downstream analyses were based on clean data with high quality.

Transcriptome assembly was performed using Trinity platform (v2012-10-05; min\_kmer\_cov = 2; prior normalization and other parameters set default) (Haas *et al.*, 2013) with the two sets of sequence reads as the input. The assemblies were blasted against GeneBank non-redundant (nr) and Uniprot (Swiss-Prot) with a cutoff e-value of  $1e-5$ . The Gene Ontology (GO) functional classes and biochemical pathway for each transcript were assigned based on the GO terms annotated to their corresponding homologs in the UniProt database and Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Differentially expressed genes were estimated using edgeR (V3.16.5) in the Bioconductor package, a program that has often been used for non-replicated samples. Based on the instruction of edgeR software, typical values for the Biological Coefficient of Variation (BCV) for data sets arising from well-controlled experiments are 0.4 for human data, 0.1 for data on genetically identical model organisms or 0.01 for technical replicates (Robinson *et al.*, 2010). We selected the value of 0.1 to calculate dispersion, in which the deduced significant differentially expressed gene number matched well with our qPCR results. Reads for each gene were normalized using 'TMM' in edgeR, and significance of differentially expressed genes was assigned with edgeR's 'exactTest' function following previous reports (Marchetti *et al.*, 2012; Gong *et al.*, 2016). Genes with fold changes  $> 2$  and FDR  $< 0.05$  (adjusted

*P*-value, determined by the Benjamini and Hochberg multiple-testing correction implemented in the 'p.adjust' method of R) were defined as differentially expressed genes.

For our microRNA sequencing data set, reads less than 16 nt in length were removed. Sequence tags originating from non-coding RNAs (rRNA, scRNA, snoRNA, snRNA and tRNA) deposited in Rfam 10.1 database were also excluded. To identify miRNAs from the remaining data set, the criteria previously used for dinoflagellates (Baumgarten *et al.*, 2013; Lin *et al.*, 2015) were used: (1) Both the guide (miRNA) and the passenger (miRNA\*) sequences of the precursor were mapped by clean reads; (2) there were at least 2 nucleotide 3' overhang in the miRNA or miRNA\* duplex; and (3) the minimum fold energy (MFE) of the hairpin structure is less than  $-25 \text{ kcal mol}^{-1}$ . Bowtie program (v0.12.9, -v 0 -k 1) (Langmead *et al.*, 2009) was used to map small RNA tags to *P. donghaiense* transcript sequences. The mapped small RNA tags were used to look for known miRNA existing in miRBase21.0. The remaining unannotated sRNA sequences were analyzed by miREvo (v1.1, -M -p 10 -g 50 000) (Wen *et al.*, 2012) and quantified by mirdeep2 (v2.0.0.5 quantifier.pl, -p -m -r -y -g 0) (Friedlander *et al.*, 2012) to predict potential novel miRNAs through exploring hairpin structure, dicer cleavage sites and the minimum free energy. To predict target mRNAs of these miRNAs, the 3'UTR sequence of each unigene transcript was used to predict animal type targets using miRanda (Enright *et al.*, 2003) with stringent criteria (miRanda-3.3a, -sc 140 -en -15 -scale 4 -strict). The expression level of miRNA was normalized by TPM (transcript per million) with previously published criteria (Zhou *et al.*, 2010): Normalized expression = mapped readcount/total reads\*1 000 000. After normalization, the miRNA expression profiles between the two microRNA libraries (P-replete and P-depleted) were compared. Differential expression analysis of two conditions was performed using the DEGseq R package (Wang *et al.*, 2010). The *P*-value was adjusted using the *q*-value (Storey, 2003). The differentially expressed miRNAs were defined with *q*-value < 0.01 and  $|\log_2(\text{foldchange})| > 1$ . GO enrichment and KEGG pathways were analyzed for the predicted target genes of the differentially expressed miRNAs using DAVID (Huang *et al.*, 2009) with UNIPROT ACCESSION terms as previously reported (Rosic *et al.*, 2014). All the transcripts of *P. donghaiense* with Swiss-prot annotation were used as background gene list for the enrichment analyses. Corrected Benjamini *P*-value < 0.05 as cutoff was conducted to yield statistically significant gene enrichment for particular pathways or GO categories.

#### Mass spectrometry analysis of lipids

Total lipid was extracted from about  $10^7$  cells using chloroform: methanol (2:1, w/w) as reported (Bligh

and Dyer, 1959). Total lipids generated from the triplicate cultures were combined respectively for the P-replete and P-depleted groups to achieve a sufficient sample size for machine loading. Lipids analysis was performed with an Agilent 1290 high-performance liquid chromatograph system equipped with an Agilent 6490 triple quadruple electrospray ionization mass spectrometer. Precursor ion of target lipid was identified as previously described (Li *et al.*, 2014). The specific head group of target lipid was identified using product ion mode based on sulfatides and SQDG standards (Avanti Polar Lipids, Alabaster, AL, USA). Lipid extracts were separated at 40 °C on an extended C18 column (1.8  $\mu\text{m}$ ,  $2.1 \times 150 \text{ mm}$ ; Agilent) for negative mode using  $[\text{M-H}]^-$  form (Hsu *et al.*, 1998). The mobile phase were 85% methanol (A) and isopropanol (B) containing 0.025%  $\text{NH}_4\text{OH}$ ; the LC gradients were as follows: 0 min, 95% A and 5% B; 10 min, 85% A and 15% B; 16 min, 50% A and 50% B; 18 min, 45% A and 55% B; 33 min, 35% A and 65% B; 38 min, 20% A and 80% B; 43 min, 85% A and 15% B; 48 min, 95% A and 5% B. The flow rate was  $0.2 \text{ ml min}^{-1}$ . Quantitation was titrated using external calibration curves of sulfatides and SQDG standards, and normalized to the recovery of the internal standard.

#### Gene and protein quantification and phylogenetic analysis

Reverse transcription quantitative PCR (RT-qPCR), western blot and phylogenetic analysis were performed as previously reported (Shi *et al.*, 2015) to determine the transcriptional or translational levels or classification of protein-coding genes, respectively. The primers used for qPCR are listed in Supplementary Table 8. For microRNA RT-qPCR, TransScript miRNA RT Enzyme mix (TransGen, Beijing, AQ202-01) was used to facilitate cDNA synthesis. Then the RT product was used as the template for qPCR. All qPCR reactions were performed on a CFX96 Real-time PCR System (BioRad, Hercules, CA, USA) using iQTM SYBR Green Supermix in 96-well plates according to manufacturer's recommendations. Calmodulin (Shi *et al.*, 2015) and U6 snRNA were used as references to calibrate the expression of the mRNA and miRNAs, respectively. Three technical and three biological replicates were conducted for both P-replete and P-depleted culture samples. The comparative threshold ( $2^{-\Delta\Delta\text{Ct}}$ ) method was used to assess the relative expression levels (Livak and Schmittgen, 2001).

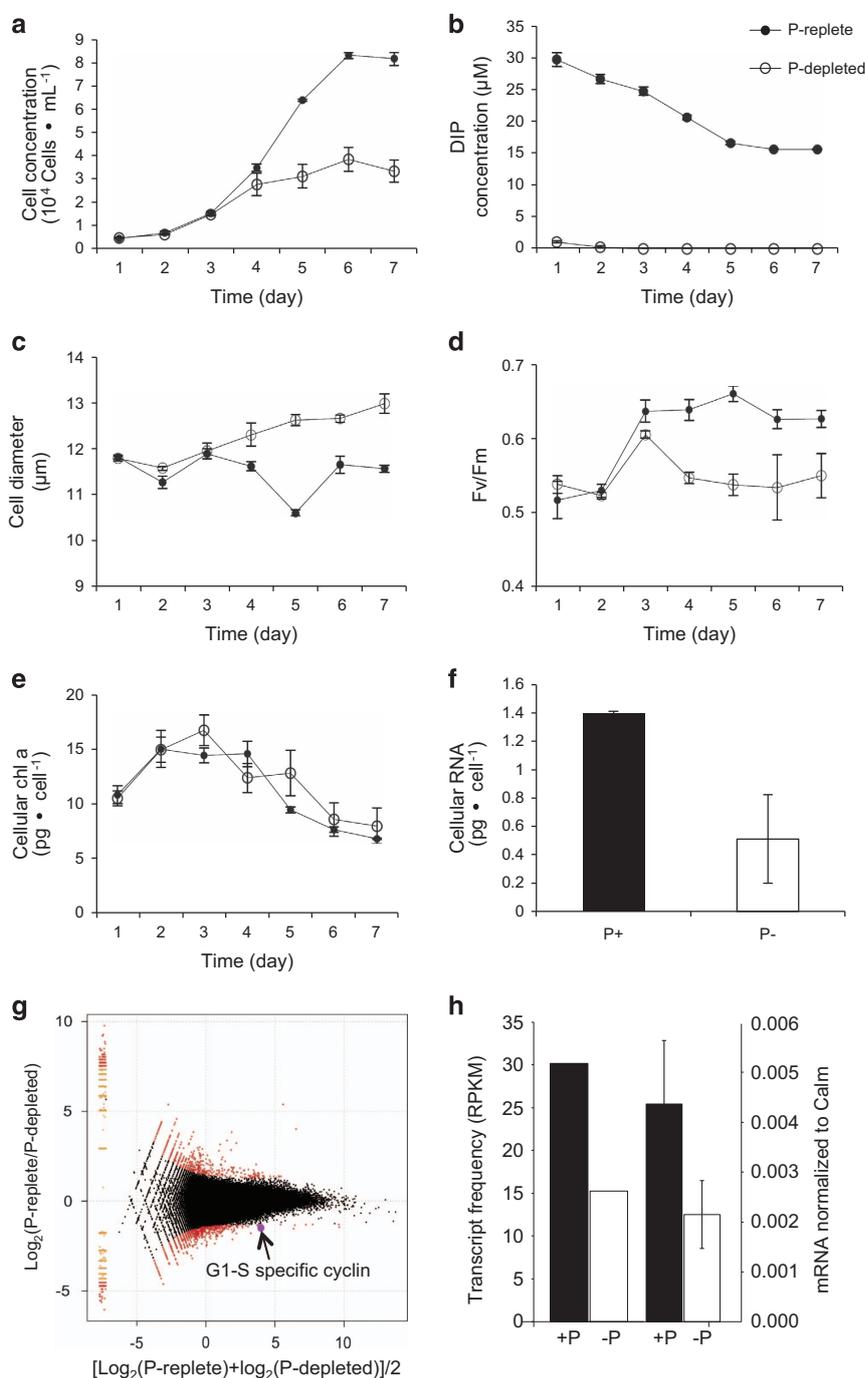
## Results and discussion

#### General physiological and molecular responses to P deficiency

Compared with the P-replete cultures, the P-depleted cultures grew slower and entered stationary phase earlier, with day 3 as the diverging point between the

two groups (Figure 1a). DIP in the medium was barely detectable in the P-depleted cultures since day 1 but remained above 15  $\mu\text{M}$  throughout the experiment in the P-replete cultures (Figure 1b), and

accordingly the P-depleted cultures exhibited markedly higher alkaline phosphatase (AP) activity as measured from day 4 (Supplementary Figure 1). This is consistent with the relationship between AP



**Figure 1** Physiological parameters and molecular responses of *P. donghaiense* under P-replete and P-depleted conditions. (a) Cell concentration. (b) DIP concentration in culture media. (c) Average cell diameter. (d) Cellular chlorophyll *a* content. (e) Photosynthetic efficiency as indicated by Fv/Fm; (f) cellular RNA content on day 3. (g) Comparison scatter plot of RNA-seq analyses of *P. donghaiense* in P-depleted vs P-replete treatments. Each point represents an identified unigenes. Plotted are the fold change ratio and the average of read counts in the P-depleted and P-replete libraries for each gene. The significantly differentially expressed genes are highlighted in red. Points below  $y=0$  line indicate genes upregulated under the P-depleted condition, while points above  $y=0$  line indicate genes downregulated under the P-depleted condition. G1-S specific cyclin is highlighted by bigger purple point and an arrow. (h) Abundances of G1-S specific cyclin transcript under the P-replete and the P-depleted conditions obtained from RNA-seq transcriptomic (bars and Y-axis on the left, RPKM: Reads Per Kilobases per Millionreads) and qRT-PCR (bars and Y-axis on the right) analyses. Error bars indicate  $\pm$  s.d. of biological triplicates; in cases where no error bars are visible, it is because the s.d. values are very small.

activity and cell growth under P-stressed previously reported for this species (Ou *et al.*, 2010). Therefore, day 3 represented the transition from intracellular P storage consumption to P depletion. Furthermore, the P-depleted cultures exhibited a larger cell size (Figure 1c), suggesting persistent photosynthesis and cell growth under P deficiency, consistently with the unchanged cellular chlorophyll *a* (Figure 1d) and upregulation of both Rubisco and ferredoxin thioredoxin reductase (Table 1), despite a lower quantum efficiency of photosystem II (Fv/Fm) (Figure 1e).

Samples were collected on day 3 from these P-replete and P-depleted cultures, and transcriptomes were sequenced to obtain insights into how cell responded to the early stage of P stress. On this day, cellular RNA content in the P-depleted cultures was only about 36.7% of that in the P-replete cultures (Figure 1f). *De novo* assembly of the RNA-seq reads generated 193 643 unigenes (practically defined as the longest sequence of Trinity transcript cluster) with an average length of 701 bp and N50 of 1209 bp (Supplementary Table 1). About 40% of these unigenes had significant matches to documented genes in public databases (Supplementary Table 2), with their GO term, COG, and KOG pathway diversity annotated (Supplementary Figure 2). A total of 1657 unigenes (0.86%) were differentially expressed (Figure 1g), among which 596 and 1064 genes were up- and downregulated respectively under P-depleted condition with >2 fold changes and FDR < 0.05. A setback exists in this result due to the lack of biological replicates in our sequencing; however, a method previously adopted (EdgeR) for non-replicated samples was used in our data analysis, and the differential expression patterns of some of these genes were verified by RT-qPCR results. Of the 1657 unigenes, 342 could be annotated, including 240 up- and 102 downregulated. The functional classification of these genes indicates

their involvement in various metabolic pathways (Figure 2), including nitrogen uptake, photosynthesis, cell cycle, glycolysis, amino acid and protein metabolism, protein synthesis (ribosomal proteins), intracellular transport (a group of dynein heavy chains), signal transduction and methyltransferase activity (Supplementary Data sets 1 and 2). The expression trends were verified for a selection of genes using RT-qPCR (Table 1). Consistent with the growth inhibition under P deficiency (Figure 1a), G1/S checkpoint-specific cyclin (Supplementary Data set 2), a protein critical for regulating cell cycle progression, showed significant downregulation both in the RNA-seq and RT-qPCR data sets (Figures 1g and h). This contrasts sharply with another dinoflagellate, *Alexandrium catenella*, in which growth inhibition and cell cycle arrest in G1 phase by P deficiency is accompanied by upregulation of a cell cycle inhibitory protein (cdc protein 20/fizzy) instead (Zhang *et al.*, 2014).

For miRNA identification, due to the lack of genome data for *P. donhaiense*, the method previously reported for predicting miRNAs from genomes (Tarver *et al.*, 2012) was not applicable. So we used the criteria proposed for dinoflagellate (Baumgarten *et al.*, 2013; Lin *et al.*, 2015). Considering that microalgae can possess microRNA shorter than 18 nt (Lopez-Gomollon *et al.*, 2014), and the 3' overhang and the pre-miRNA double-stranded stem need to have a minimal length of 16 bp (Starega-Roslan *et al.*, 2011), 16 nt or longer reads were used for further analysis. As a result, from the 12 303 754 and 12 784 951 filtered reads from P-replete and P-depleted samples respectively, 17 mature microRNAs were identified, 7 being known and 10 novel microRNAs. The lengths of these miRNAs ranged from 16 to 22 nt, with a peak at 20 nt (Supplementary Figure 3). This contrasted with typical miRNAs in plants and animals with length peaks at 21 and 22 nt,

**Table 1** Genes and fold changes of their expression under P deficiency compared with under P-replete condition from RNA-seq transcriptome and microRNAome profiling, RT-qPCR and western blot

Gene ID	Annotation <sup>a</sup>	E-value	RNA-seq	qPCR	microRNA	Western blot
comp95052_c0	Alkaline phosphatase	0	1.41	1.93 <sup>b</sup>	ND	NA
comp52959_c0	N-acylneuraminate-9-phosphatase-like	1.20E-33	1.27	1.06	ND	NA
comp101639_c0	High-affinity phosphate transporter	2.79E-70	1.32	1.04	ND	NA
comp98084_c2	Ferredoxin thioredoxin reductase variable alpha chain	5.76E-15	2.52 <sup>b</sup>	1.68 <sup>b</sup>	ND	NA
comp376390_c0	Ciliary dynein heavy chain	2.07E-120	5.25 <sup>b</sup>	1.5	ND	NA
comp122586_c0	Chloroplast 50S ribosomal protein L9	1.05E-15	3.04 <sup>b</sup>	1.62 <sup>b</sup>	ND	NA
comp101619_c0	Ribulose 1,5-bisphosphate carboxylase oxygenase large subunit, partial	3.47E-09	1.65	1.92 <sup>b</sup>	ND	NA
comp101143_c2	Rhodopsin 1, partial	1.10E-50	1.02	1.04	ND	1.8
comp1675839_c0	N-sulphoglucosamine sulphohydrolase	1.83E-53	1.01	NA	-2.44 <sup>b, d</sup>	NA
comp95787_c2	Extracellular sulfatase Sulf-1	7.34E-12	1.24	NA	2.45 <sup>b, c</sup>	NA
comp101603_c0	Arylsulfatase A	2.53E-72	-1.06	NA	2.45 <sup>b, c</sup>	NA
comp201041_c0	Arylsulfatase B	3.20E-169	1.13	NA	2.45 <sup>b, c</sup>	NA

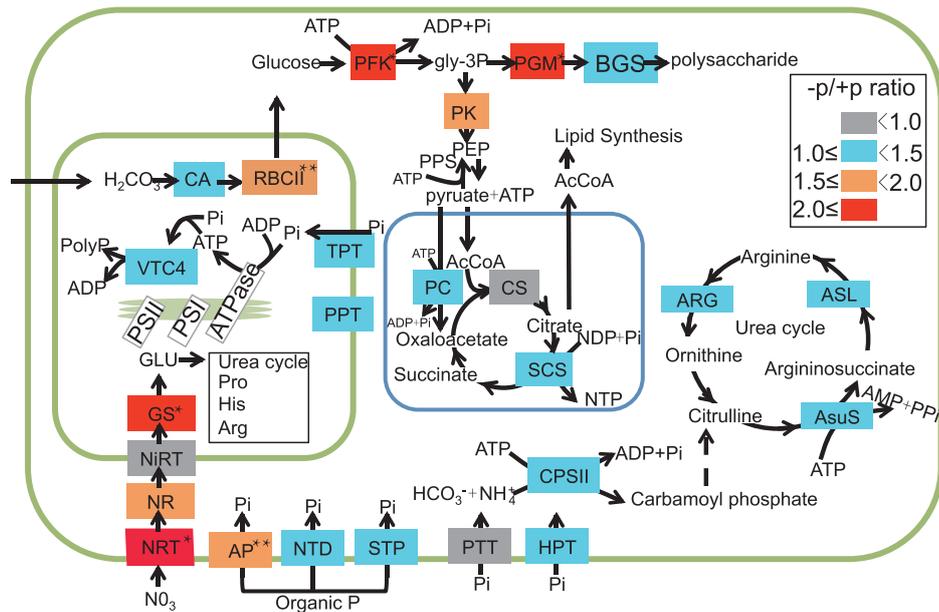
Abbreviations: NA, not analyzed; ND, not detected.

<sup>a</sup>Annotation based on NCBI nr or Swiss-Prot database.

<sup>b</sup>Fold change value with statistical significance ( $P < 0.05$ ).

<sup>c</sup>microRNA ID: Pdh-miR14; the fold change of Pdh-miR14 was estimated by microRNAome.

<sup>d</sup>microRNA ID: Pdh-miR1-5p; the fold change of Pdh-miR1-5p was estimated by qPCR.



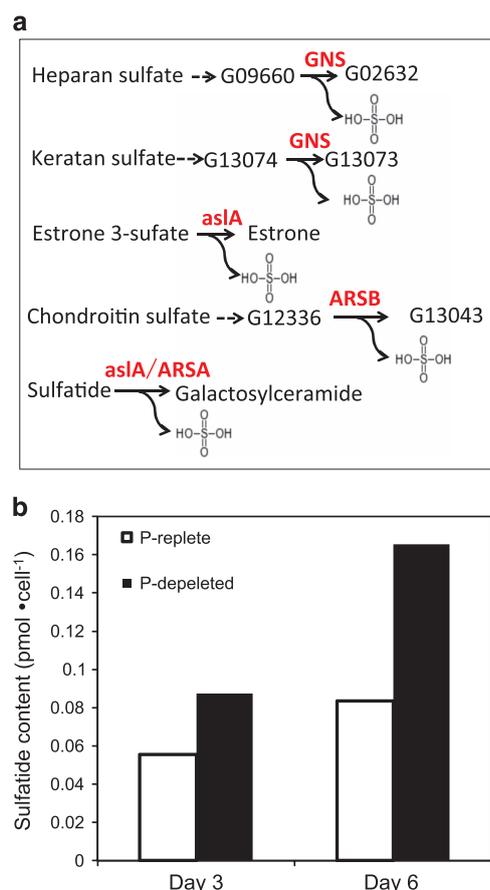
**Figure 2** *P. donghaiense* metabolic pathways responding to phosphate limitation. GS, glutamate synthase; CS, citrate synthase; SCS, Succinyl coenzyme A synthetase; NiRT, nitrite transporter; NR, nitrite reductase; NTD, 5'-Nucleotidase; STP, Serine/threonine phosphatase; TPT, plastidic triose-phosphate/phosphate translocator; VTC4, Vacuolar transporter chaperone 4; PPT, phosphoenolpyruvate/phosphate translocator; PC, pyruvate carboxylase; Acetyl-CoA, Acetyl coenzyme A; CPSII, Carbamoyl phosphate synthetase II; AsuS, argininosuccinate synthase; ASL, argininosuccinate lyase; ARG, arginase; PTT, phosphate transporter; HPT, phosphate transporter; CA, Carbonic anhydrase; RBCII, Form II RuBisCO; NRT, Nitrate transporter; AP, alkaline phosphatase; PFK, Phosphofructokinase; PGM, Phosphoglucumutase; BGS, Callose synthase; PK, Pyruvate kinase; PEP, Phosphoenolpyruvate; PPS, PEP synthetase. Color filling of gene name box depicts the fold change of expression levels of the gene based on RNA-seq data; \*Indicates significant up- or downregulation based on RNA-seq data; \*\*Indicating significantly up- or downregulation based on qRT-PCR.

respectively, and agreed with non-canonical miRNAs identified in diatoms and haptophytes (Lopez-Gomollon *et al.*, 2014). Target gene analysis against the transcriptome described above yielded 3268 protein-coding genes as the targets of those microRNAs with stringent criteria (Supplementary Figure 4). Fourteen of these mature microRNAs showed differential expression between P-replete and P-depleted conditions, including 3 down- and 11 upregulated in the P-depleted culture group, and RT-qPCR for 5 of these verified their trends of expression (Supplementary Table 3). Most of the total microRNA target genes (3092 or 94.61%) were targets of these 14 differentially expressed microRNAs. DAVID enrichment analysis (Huang *et al.*, 2009) showed that these target genes were significantly enriched in gene families of RNA recognition motif 2, Cyclic nucleotide-binding domain, EF-hand domain pair, ion transport protein, sulfatase, ABC transporter and ABC-2 type transporter (Supplementary Table 4) A similar enrichment prediction of these functions were also detected using Interproscan (Supplementary Table 5). There was 8.93-fold enrichment in the RNA recognition motif 2-containing protein family. The RNA recognition motif 2 is one of the most abundant eukaryotic protein domains involved in post-transcriptional regulation of gene expression (Maris *et al.*, 2005). All of the eight RNA recognition motif 2-containing genes were targets of upregulated microRNA (Supplementary Table 6). Because the general

function of a microRNA is binding to mRNA to yield mRNA degradation or translation repression (Bartel, 2009), the downregulation of the microRNAs putatively regulating this gene family would predict reduced microRNA-independent post-transcriptional regulation under P limitation. Another interesting enriched protein family was sulfatase. Previous report indicated that sulfate metabolism has an important role under P limitation in phytoplankton (Van Mooy *et al.*, 2006, 2009). According to PFAM enrichment result in the current study, 14 microRNA targets related to sulfatase metabolism were enriched. Twelve of these 14 unigenes were targeted by upregulated microRNA, and the other two was targeted by downregulated and unchanged microRNA separately (Supplementary Table 7). The main function of this group of sulfatases is to break down sulfatides and release sulfate (Figure 3). All of these findings suggest that post-translationally regulated sulfatase metabolism potentially has a pivotal role when *P. donghaiense* experienced phosphate limitation.

#### *DIP and DOP utilization regulated both transcriptionally and by microRNA*

From the transcriptome data, we identified eight low-affinity and one high-affinity phosphate transporters, which may allow the species to tune its P uptake depending on changing phosphate availability (Lin *et al.*, 2016). Besides, a common strategy to



**Figure 3** microRNA target genes that regulate sulfate metabolism and cellular sulfatide content in the P-replete and P-depleted *P. donghaiense* cultures. (a) microRNA target genes that regulate sulfate metabolism. Target genes of upregulated microRNA are shown in red. GNS, N-acetylglucosamine-6-sulfatase; asIA, arylsulfatase; ARSA, arylsulfatase A; ARSB, arylsulfatase B. Dotted arrow indicates multiple steps. G-numbers are glycan number from KEGG. (b) Cellular sulfatide content of P-replete and P-depleted *P. donghaiense* cells on day 3 and day 6.

cope with DIP deficiency in phytoplankton is to utilize DOP (Karl, 2014), which is dominated by phosphoesters of C-O-P ester bonds (Dyhrman *et al.*, 2007). Phosphoesters are typically hydrolyzed by phosphatases to release phosphate (Lin *et al.*, 2016). In our transcriptome data set, AP (alkaline phosphatase) and a NLI (nuclear interacting factor) interacting factor-like phosphatase gene (Supplementary Data set 1, NIFs) was found to be significantly upregulated under P limitation. For AP, both qRT-PCR (Table 1) and enzyme activity assays (Supplementary Figure 1) showed remarkable upregulation (1.93 and 1.7 folds respectively, with  $P < 0.05$ , ANOVA) under P deficiency. Other DOP utilization related genes, such as STP (serine/threonine-protein phosphatase) and 5'-nucleotidase (NTD), were also identified. All of this suggests that *P. donghaiense* is able to utilize various types of phosphoester substrates, including nucleic acids, ATP and lipids as a source of phosphate.

We found that some of the phosphorus metabolism-related GO terms were targets of differentially expressed microRNAs (Supplementary Data set 3), impacting a complex metabolic network (Supplementary Figure 3). These GO terms can be sorted into four P related groups, phosphoester hydrolysis, phospholipase, phosphate-group transfer, and phosphorus-containing anhydrides hydrolyase, which accounted for 43, 3, 43 and 20 microRNA target genes respectively. All the detected 17 microRNAs were predicted to be involved in regulating genes related to DOP metabolism listed above, indicating a broad post-transcriptional regulation existing in this dinoflagellate for DOP utilization. In accordance, these phosphorus metabolism-related genes showed no transcriptional regulation under the contrasting P conditions from our RNA-seq data set.

It is notable that the 16 dynein heavy chain (DHC) genes detected were substantially upregulated ( $> 2$  fold,  $FDR < 0.05$ ) at the transcriptional level under P limitation. DHC has two basic domains—an ATP hydrolysis site domain located at the N terminus and a ATP-sensitive track-binding site at the C terminus of the protein (Asai and Wilkes, 2004). The ATP hydrolysis site domain usually contains a P-loop motif, a catalytic ATP-binding site (Asai and Koonce, 2001; Asai and Wilkes, 2004). Cytoplasmic dyneins are involved in organelle movement toward the center of the cell (minus end of the microtubules) whereas the axonemal dyneins drive flagellum movement. Phylogenetic analysis of DHC indicated that *P. donghaiense* possess two types of cytoplasmic (cyto I and II) and several axonemal forms (outer arm subunits  $-\alpha$   $-\beta$  and  $-\gamma$  and inner arm subunit  $1\beta$ ) (Supplementary Figure 6). All the six subgroups exhibited higher expression levels in P-depleted cultures, suggesting promotion of intracellular trafficking or cell motility by P deficiency. Increased cellular motility would facilitate prey exploration to supplement P-nutrition (Lin *et al.*, 2016).

#### Changes in ATP cycling under P-depleted condition

Under P-limited conditions, two groups of ATP-acquisition and ATP-metabolism pathways were found to change expression levels in concert. One group of such pathways was ATP-consuming process or apparatus, including PFK (phosphofructokinase, Supplementary Data set 1) in glycolysis, dynein heavy chain, and ribosomes. PFK is an important enzyme in glycolysis (Dyhrman, 2016) and its modification under P deficiency is believed to be common in higher plants, green algae and diatoms (Theodorou *et al.*, 1991; Dyhrman *et al.*, 2012; O'Rourke *et al.*, 2013). Our data showed its 3.15-fold higher transcript abundance with statistical significance under P deficiency (Figure 2) (Supplementary Data set 1). This is consistent with the previous finding in green algae that this ATP-consuming activity is usually activated by low

concentrations of Pi (inorganic phosphate) but inhibited by high concentrations of Pi (Botha and Turpin, 1990). Besides, the metabolic process of the 16 DHC described above are also ATP consuming, as they contain an ATPase domain, and were significantly upregulated in the P-depleted cultures. Furthermore, 40 ribosomal proteins were significantly upregulated under P deficiency based on Swiss-Prot annotation (Supplementary Data set 1), and not one was found to be downregulated (Supplementary Data set 2). As protein synthesis is an ATP-consuming process, which take up about 75% of a cell's total energy budget (Lane and Martin, 2010), the upregulation of ribosomal proteins (and hence protein synthesis activity) can potentially add significantly to the demand for ATP.

Several ATP synthesis pathways were also identified as upregulated under P limitation. Of these, proton pump-type rhodopsin identified here resembles bacterial homologs that functions as a light-driven proton pump to generate an outward proton gradient to activate ATP synthesis (Martinez *et al.*, 2007). Since its discovery in dinoflagellates (Lin *et al.*, 2010), proton pump rhodopsin has been reported to supplement energy needs in coping with nutritional deficiency in dinoflagellates and diatoms (Marchetti *et al.*, 2012; Guo *et al.*, 2014). In *P. donghaiense*, this gene appears to compensate for the light-limited photosynthesis in dim environment (Shi *et al.*, 2015). Both RNA-seq and RT-qPCR results showed that the transcriptional level of this rhodopsin changed little between P-replete and P-deplete conditions (Table 1). However, the Western blot assay for this protein showed that P-depleted cultures expressed 1.31-fold (when normalized to amount of total protein used) or 1.80-fold (when normalized to GAPDH as reference) (Table 1) more rhodopsin than the P-replete cultures ( $P < 0.05$ , ANOVA). This suggests that this rhodopsin may serve as a supplementary energy (ATP) generating mechanism under P deficiency. In accordance, five of the eight detected ATP synthase subunits were highly expressed in the P-depleted cultures (Supplementary Figure 7), among which, subunit b showed a remarkable (2.07-fold) upregulation and gamma subunit a moderate (1.46-fold) upregulation.

#### *Increased sulfatide content and decreased degradation under P limitation*

Studies have shown in phytoplankters such as *Aureococcus anophagefferens* and *Prochlorococcus* spp., P deficiency led to increased sulfolipids relative to phospholipids, due to partial replacement of phospholipids by sulfonated lipids to reduce cellular P quanta (Van Mooy *et al.*, 2006, 2009; Wurch *et al.*, 2011). In these organisms, UNP-sulfoquinovose synthase was upregulated under P deficiency to increase sulphoquinovosyldiacylglycerol (SQDG) sulfolipids. Using HPLC-MS analysis, we also detected SQDG in *P. donghaiense*, but with

no consistent increasing trend in its abundance under the P-depleted condition (Supplementary Figure 8). Besides, the SQDG synthase gene was not detected in the transcriptome data set, indicating a low transcriptional level. However, our HPLC-MS analysis revealed a sulfatide product ion spectrum at  $m/z$  764 (Supplementary Figure 9) and the HOSO<sub>3</sub><sup>-</sup> ion of sulfatide at  $m/z$  97 (Hsu *et al.*, 1998). Our quantification result indicated that a different type of sulfolipid than SQDG, sulfatides, increased by 1.57 and 1.98 fold on per cell basis in the exponential (day 3) and stationary (day 6) growth phases respectively under P limitation (Figure 3b). If normalized to cell surface area, the fold change decreased slightly to 1.56 and 1.65 respectively. The pooling of the triplicate samples for the HPLC-MS analysis has prevented us from having statistical test for the significance of the difference between the two P conditions; however, the consistent pattern on both sampling events and for both normalization methods suggests that the difference we observed reflects the actual difference between the two P conditions.

Notably, seven of the 14 target genes of dynamic microRNAs were assigned KO pathways of sulfur metabolism, with all their regulating microRNAs being upregulated under P deficiency (Supplementary Tables 3 and 7), predicting decreases in their encoded proteins. Five of the seven genes functionally promoted reactions to release sulfate (Figure 3; Supplementary Table 3). One of these is involved in sulfatide degradation (Table 1), suggesting reduced sulfolipid degradation under P stress. Other sulfate metabolic pathways, including those for heparin sulfate, keratin sulfate, estrone 3-sulfate and chondroitin sulfate, were also targets of the upregulated microRNAs. These results are all generally consistent with the findings in *Synechococcus* and *Aureococcus* regarding substituting sulfate for PO<sub>4</sub><sup>3-</sup> in lipids under P deficiency (Van Mooy *et al.*, 2006). However, if verified in future research with more extensive sampling, these results would suggest a distinct mechanism in the dinoflagellate. In contrast to the upregulated sulfolipid (SQDG) synthesis in P-limited cyanobacteria (Van Mooy *et al.*, 2006), *P. donghaiense* would seem to reduce degradation of sulfolipids (sulfatides) under P deficiency.

## Concluding remarks

Our study reveals metabolic responses to P stress in the dinoflagellate *P. donghaiense* that are regulated at transcriptional, microRNA-mediated post-transcriptional, and translational levels. The results demonstrate that the integrated microRNAomic-transcriptomic (ideally proteomic too) approach is powerful for gaining a comprehensive understanding of the molecular mechanisms by which a dinoflagellate responds to varying P conditions and

potentially other environmental conditions. Application of this approach to more species and ultimately to natural assemblages of phytoplankton, ideally with replicate samples sequenced, will illuminate the potentially different strategies in different lineages of phytoplankton to cope with P stress, global warming, and other environmental challenges.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgements

The project was supported by the National Natural Science Foundation of China grants #41330959, 31661143029, 41606121, China Postdoctoral Science Foundation Grant 2016M602070, and National Key Research and Development Program of China grant 2016YFA0601202. Thanks are due to Ms. Brittany Sprecher for her kind assistance in English language.

## References

- Alexander H, Rouco M, Haley ST, Wilson ST, Karl DM, Dyrhman ST. (2015). Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *Proc Natl Acad Sci USA* **112**: E5972–E5979.
- Asai DJ, Koonce MP. (2001). The dynein heavy chain: structure, mechanics and evolution. *Trends Cell Biol* **11**: 196–202.
- Asai DJ, Wilkes DE. (2004). The dynein heavy chain family. *J Eukaryot Microbiol* **51**: 23–29.
- Bartel DP. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**: 215–233.
- Baumgarten S, Bayer T, Aranda M, Liew YJ, Carr A, Micklem G et al. (2013). Integrating microRNA and mRNA expression profiling in *Symbiodinium microadriaticum*, a dinoflagellate symbiont of reef-building corals. *BMC Genomics* **14**: 704.
- Bligh EG, Dyer WJ. (1959). A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**: 911–917.
- Botha FC, Turpin DH. (1990). Molecular, kinetic, and immunological properties of the 6-phosphofructokinase from the green alga *Selenastrum minutum*: activation during biosynthetic carbon flow. *Plant Physiol* **93**: 871–879.
- Dyrhman ST, Palenik BP. (1997). The identification and purification of a cell surface alkaline phosphatase from the dinoflagellate *Prorocentrum minimum* (Dinophyceae). *J Phycol* **33**: 602–612.
- Dyrhman ST, Ammerman JW, Van Mooy BA. (2007). Microbes and the marine phosphorus cycle.
- Dyrhman ST, Jenkins BD, Ryneerson TA, Saito MA, Mercier ML, Alexander H et al. (2012). The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS One* **7**: e33768.
- Dyrhman ST. (2016). Nutrients and their acquisition: phosphorus physiology in microalgae. *The Physiology of Microalgae*. Springer, pp 155–183.
- Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. (2003). MicroRNA targets in drosophila. *Genome Biol* **5**: R1.
- Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res* **40**: 37–52.
- Gong W, Browne J, Hall N, Schrueth D, Paerl H, Marchetti A. (2016). Molecular insights into a dinoflagellate bloom. *ISME J* **11**: 439–452.
- Guo Z, Zhang H, Lin S. (2014). Light-promoted rhodopsin expression and starvation survival in the marine dinoflagellate *Oxyrrhis marina*. *PLoS One* **9**: e114941.
- Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* **8**: 1494–1512.
- Hsu F-F, Bohrer A, Turk J. (1998). Electrospray ionization tandem mass spectrometric analysis of sulfatide: determination of fragmentation patterns and characterization of molecular species expressed in brain and in pancreatic islets. *Biochim Biophys Acta* **1392**: 202–216.
- Huang DW, Sherman BT, Lempicki RA. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57.
- Karl DM. (2014). Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. *Ann Rev Mar Sci* **6**: 279–337.
- Lane N, Martin W. (2010). The energetics of genome complexity. *Nature* **467**: 929–934.
- Langmead B, Trapnell C, Pop M, Salzberg SL. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Li J, Han D, Wang D, Ning K, Jia J, Wei L et al. (2014). Choreography of transcriptomes and lipidomes of nannochloropsis reveals the mechanisms of oil synthesis in microalgae. *Plant Cell* **26**: 1645–1665.
- Lin S, Zhang H, Zhuang Y, Tran B, Gill J. (2010). Spliced leader-based metatranscriptomic analyses lead to recognition of hidden genomic features in dinoflagellates. *Proc Natl Acad Sci USA* **107**: 20033–20038.
- Lin S. (2011). Genomic understanding of dinoflagellates. *Res Microbiol* **162**: 551–569.
- Lin S, Cheng S, Song B, Zhong X, Lin X, Li W et al. (2015). The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* **350**: 691–694.
- Lin S, Litaker RW, Sunda WG. (2016). Phosphorus physiological ecology and molecular mechanisms in marine phytoplankton. *J Phycol* **52**: 10–36.
- Lin X, Zhang H, Huang B, Lin S. (2012). Alkaline phosphatase gene sequence characteristics and transcriptional regulation by phosphate limitation in *Karenia brevis* (Dinophyceae). *Harmful Algae* **17**: 14–24.
- Livak KJ, Schmittgen TD. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* **25**: 402–408.
- Lopez-Gomollon S, Beckers M, Rathjen T, Moxon S, Maumus F, Mohorianu I et al. (2014). Global discovery and characterization of small non-coding RNAs in marine microalgae. *BMC Genomics* **15**: 697.

- Lu D, Goebel J, Qi Y, Zou J, Han X, Gao Y *et al.* (2005). Morphological and genetic study of *Prorocentrum donghaiense* Lu from the East China Sea, and comparison with some related *Prorocentrum* species. *Harmful Algae* **4**: 493–505.
- Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT *et al.* (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci USA* **109**: E317–E325.
- Maris C, Dominguez C, Allain FH. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**: 2118–2131.
- Martinez A, Bradley AS, Waldbauer JR, Summons RE, DeLong EF. (2007). Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proc Natl Acad Sci USA* **104**: 5590–5595.
- Moustafa A, Evans AN, Kulis DM, Hackett JD, Erdner DL, Anderson DM *et al.* (2010). Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. *PLoS One* **5**: e9688.
- Mraz M, Malinova K, Mayer J, Pospisilova S. (2009). MicroRNA isolation and stability in stored RNA samples. *Biochem Biophys Res Commun* **390**: 1–4.
- O'Rourke JA, Yang SS, Miller SS, Bucciarelli B, Liu J, Rydeen A *et al.* (2013). An RNA-Seq transcriptome analysis of orthophosphate-deficient white lupin reveals novel insights into phosphorus acclimation in plants. *Plant Physiol* **161**: 705–724.
- Ou L, Huang B, Hong H, Qi Y, Lu S. (2010). Comparative alkaline phosphatase characteristics of the algal bloom dinoflagellates *Prorocentrum donghaiense* and *Alexandrium catenella*, and the diatom *Skeletonema costatum*. *J Phycol* **46**: 260–265.
- Perry M. (1976). Phosphate utilization by an oceanic diatom in phosphorus-limited chemostat culture and in the oligotrophic waters of the central North Pacific. *Limnol Oceanogr* **21**: 88–107.
- Robinson MD, McCarthy DJ, Smyth GK. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- Rosic N, Ling EY, Chan CK, Lee HC, Kaniewska P, Edwards D *et al.* (2014). Unfolding the secrets of coral-algal symbiosis. *ISME J* **9**: 844–856.
- Shemi A, Schatz D, Fredricks HF, Van Mooy BA, Porat Z, Vardi A. (2016). Phosphorus starvation induces membrane remodeling and recycling in *Emiliana huxleyi*. *New Phytol* **211**: 886–898.
- Shi X, Li L, Guo C, Lin X, Li M, Lin S. (2015). Rhodopsin gene expression is regulated by the light dark cycle, light spectrum and light intensity in the dinoflagellate *Prorocentrum*. *Front Microbiol* **6**: 555.
- Starega-Roslan J, Koscianska E, Kozlowski P, Krzyzosiak WJ. (2011). The role of the precursor structure in the biogenesis of microRNA. *Cell Mol Life Sci* **68**: 2859.
- Storey JD. (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann Stat* **31**: 2013–2035.
- Tarver JE, Donoghue PC, Peterson KJ. (2012). Do miRNAs have a deep evolutionary history? *Bioessays* **34**: 857–866.
- Theodorou ME, Elrifi IR, Turpin DH, Plaxton WC. (1991). Effects of phosphorus limitation on respiratory metabolism in the green alga *Selenastrum minutum*. *Plant Physiol* **95**: 1089–1095.
- Tian SL, Zhou SL, Zhou GY, Li RQ. (2012). Method and device for processing sequencing data. Patent CN102831330A.
- Van Mooy BA, 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.
- Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblí *et al.* (2009). Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* **458**: 69–72.
- Wang L, Feng Z, Wang X, Zhang X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**: 136–138.
- Wen M, Shen Y, Shi S, Tang T. (2012). miREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinformatics* **13**: 140.
- Wurch LL, Bertrand EM, Saito MA, Van Mooy BA, Dyhrman ST. (2011). Proteome changes driven by phosphorus deficiency and recovery in the brown tide-forming alga *Aureococcus anophagefferens*. *PLoS One* **6**: e28949.
- Zhang C, Lin S, Huang L, Lu W, Li M, Liu S. (2014). Suppression subtraction hybridization analysis revealed regulation of some cell cycle and toxin genes in *Alexandrium catenella* by phosphate limitation. *Harmful Algae* **39**: 26–39.
- Zhou L, Chen J, Li Z, Li X, Hu X, Huang Y *et al.* (2010). Integrated profiling of microRNAs and mRNAs: microRNAs located on Xq27.3 associate with clear cell renal cell carcinoma. *PLoS One* **5**: e15224.

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