The Myb-like transcription factor Phosphorus Starvation Response (PtPSR) controls conditional P acquisition and remodeling in marine microalgae

Amit Kumar Sharma, Alice Mühlroth, Juliette Jouhet, Eric Maréchal, Leila Alipanah, Ralph Kissen, Tore Brembu, Atle Bones, Per Winge

Research

New Phytologist

The Myb-like transcription factor phosphorus starvation response (PtPSR) controls conditional P acquisition and remodelling in marine microalgae

Amit Kumar Sharma[1], Alice Mühlroth[1], Juliette Jouhet[2], Eric Maréchal[2], Leila Alipanah[1], Ralph Kissen[1], Tore Brembu[3], Atle M. Bones[1] and Per Winge[1]

[1]Cell, Molecular Biology and Genomics Group, Department of Biology, Norwegian University of Science and Technology, 7491, Trondheim, Norway; [2]Laboratoire de Physiologie Cellulaire Végétale, Centre National de la Recherche Scientifique, Commissariat à l'Energie Atomique, Institut National de la Recherche Agronomique, Université Grenoble Alpes, 38000, Grenoble, France; [3]Department of Biotechnology and Food Science, Norwegian University of Science and Technology, 7491, Trondheim, Norway

Authors for correspondence:
Per Winge
Tel: +47 73 59 62 29
Email: per.winge@ntnu.no

Atle M. Bones
Tel: +47 73 59 86 92
Email: atle.m.bones@ntnu.no

Received: *15 April 2019*
Accepted: *29 September 2019*

New Phytologist (2020) **225:** 2380–2395
doi: 10.1111/nph.16248

Key words: alkaline phosphatase (AP), diatom, lipid remodelling, Myb transcription factor, phosphate stress response, phosphorus, reshuffling of phosphate containing compounds.

Summary

- Phosphorus (P) is one of the limiting macronutrients for algal growth in marine environments. Microalgae have developed adaptation mechanisms to P limitation that involve remodelling of internal phosphate resources and accumulation of lipids.
- Here, we used *in silico* analyses to identify the P-stress regulator PtPSR (*Phaeodactylum tricornutum* phosphorus starvation response) in the diatom *P. tricornutum*. *ptpsr* mutant lines were generated using gene editing and characterised by various molecular, genetics and biochemical tools.
- PtPSR belongs to a clade of Myb transcription factors that are conserved in stramenopiles and distantly related to plant P-stress regulators. PtPSR bound specifically to a conserved *cis*-regulatory element found in the regulatory region of P-stress-induced genes. *ptpsr* knockout mutants showed reduction in cell growth under P limitation. P-stress responses were impaired in *ptpsr* mutants compared with wild-type, including reduced induction of P-stress response genes, near to complete loss of alkaline phosphatase activity and reduced phospholipid degradation.
- We conclude that PtPSR is a key transcription factor influencing P scavenging, phospholipid remodelling and cell growth in adaptation to P stress in diatoms.

Introduction

Phosphorus (P) is one of the most important macronutrients for algae and is a limiting factor of marine primary production in some parts of the world's oceans (Sarthou *et al.*, 2005; Elser *et al.*, 2007). It can be found in its inorganic form (P_i, orthophosphate) or in dissolved organic phosphorus (DOP) (Van Mooy *et al.*, 2009). Recent studies have revealed underlying common molecular mechanisms of stramenopiles to deal with P stress that can be divided into global and specific P-stress responses (Dyhrman *et al.*, 2012; Cruz de Carvalho *et al.*, 2016; Cañavate *et al.*, 2017; Mühlroth *et al.*, 2017; Alipanah *et al.*, 2018). In microalgae, specific P-stress responses involve P transportation, P scavenging and degradation of alternative P sources such as DOP (Dyhrman, 2016). Global P-stress responses include cell division arrest, reduction of the RNA pool and redirection of main carbon fluxes towards carbon storage molecules such as chrysolaminarin or triacylglycerol (TAG, Mühlroth *et al.*, 2013; Brembu *et al.*, 2017). Even though a few molecular P-stress mechanisms are known, key transcriptional regulators have not been identified in stramenopiles.

Transcriptome and lipidome studies in the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*, as well as the eustigmatophyte *Nannochloropsis oceanica*, have provided detailed insights of gene regulation patterns during P limitation (Dyhrman *et al.*, 2012; Cruz de Carvalho *et al.*, 2016; Mühlroth *et al.*, 2017; Alipanah *et al.*, 2018). P acquisition is activated by transcriptional induction of genes encoding high and low affinity P_i transporters, phosphodiesterases and phosphatases (Donald *et al.*, 1997; Shimogawara *et al.*, 1999). Phosphodiesters such as phospholipids (PLs) and nucleic acids are degraded by extracellular and intracellular glycerophosphoryl diester phosphodiesterases (GDPDs) and nucleotidases, which are poorly studied in stramenopiles (Dyhrman *et al.*, 2012; Baldwin, 2013). A recent study in *N. oceanica* showed that expression of cellular and extracellular GDPDs was induced as part of an early P-stress response, suggesting that degradation of membrane-bound PLs was likely part of the P-stress-inducible PL-recycling scheme (Mühlroth

et al., 2017). This PL-recycling scheme includes degradation of PLs, reduction of PL biosynthesis and synthesis of P-free lipid classes (Martin *et al.*, 2011; Cañavate *et al.*, 2017). Phosphatases involved in P acquisition include phosphomonoesterases (PMEs) and alkaline phosphatases (APs) degrading phosphomonoesters, enabling the cell to obtain orthophosphate by degrading intracellular and extracellular DOP sources (Flynn *et al.*, 1986; Theodorou *et al.*, 1991; Mühlroth *et al.*, 2017; Alipanah *et al.*, 2018). The P-stress-inducible APs are the most commonly encountered phosphatases and are found in a wide range of phytoplankton. AP activity measurements are therefore used as a biomarker for P limitation in the marine environment and single cultures (Dyhrman *et al.*, 2012; Lin *et al.*, 2012; Feng *et al.*, 2015; Cañavate *et al.*, 2017). Several putative APs identified in diatoms have either bacterial origin or are similar to the characterised low P-inducible APs of the green alga *Chlamydomonas reinhardtii* (Dyhrman *et al.*, 2012; Alipanah *et al.*, 2018). APs in *C. reinhardtii* are either membrane bound, which allows AP detection in intact concentrated cells or extracellular, allowing AP activity measurements in the filtrate (Quisel *et al.*, 1996; Xu, 2006; Li *et al.*, 2012; Dyhrman, 2016).

Long-term P-stress-induced processes classically involve transcription factors (TFs) that are often lineage-specific and key players in gene expression. However, apart from *in silico* studies, little information is known regarding nutrient stress-induced TFs in stramenopiles (Rayko *et al.*, 2010; Matthijs *et al.*, 2016, 2017; Thiriet-Rupert *et al.*, 2016). The myeloblastosis (Myb) family of TFs is widespread in eukaryotic organisms and is a large, functionally diverse family (Dubos *et al.*, 2010; Rayko *et al.*, 2010). The Myb-related TFs PHOSPHORUS STARVATION RESPONSE 1 (PSR1) and PHOSPHATE STARVATION RESPONSE1 (PHR1) were found to act as important regulators of P-stress responses in *C. reinhardtii* and *Arabidopsis thaliana*, respectively (Wykoff *et al.*, 1999; Rubio *et al.*, 2001). PSR1 and PHR1 control the expression of P transporters, enzymes involved in DOP degradation and P scavenging, and P starvation-specific secreted polypeptides under low P_i conditions (Dubos *et al.*, 2010). Myb TFs are divided into several classes depending on the adjacent repeats of the Myb DNA-binding domain (Dubos *et al.*, 2010; Chiou & Lin, 2011). Phylogenetic analysis showed that PHR1 and PSR1 share the same Myb DNA-binding domain with a SHLQKYR motif located at its C-terminal part (Wykoff *et al.*, 1999; Rubio *et al.*, 2001; Thiriet-Rupert *et al.*, 2016). In addition to the Myb DNA-binding domain, PHR1 and PSR1 also share similarities in a C-terminal, predicted coiled-coil (CC) domain, which is a glutamine-rich sequence potentially involved in protein–protein interactions (Dubos *et al.*, 2010; Chiou & Lin, 2011). Under low P_i, PSR1 and PHR1 form homodimers, probably through their CC domain, that bind an imperfect palindromic DNA sequence termed the PHR1-binding sequence (P1BS) present in the promoters of P_i starvation-responsive genes (Wykoff *et al.*, 1999; Rubio *et al.*, 2001).

The marine pennate diatom *P. tricornutum* has become a model species for molecular studies in diatoms, due to its well characterised genome and development of advanced molecular engineering techniques (Falciatore *et al.*, 1999; Bowler *et al.*, 2008; Karas *et al.*, 2015; Nymark *et al.*, 2016). The recently established CRISPR/Cas9 gene editing system in *P. tricornutum* opens up possibilities for efficient reverse-genetic pipelines to identify key players in metabolic pathways and transcriptional regulation (Nymark *et al.*, 2016; Sharma *et al.*, 2018). In the present study, we used available transcriptome data to identify the P-stress-induced Myb-related TF PtPSR (*P. tricornutum* phosphorus starvation response). *ptpsr* mutants generated by gene editing were subjected to chemical, biochemical, physiological and transcriptional analyses to investigate its role as a regulator of P limitation responses.

Materials and Methods

Growth conditions

Axenic *P. tricornutum* wild-type cells (PtWT) (Bohlin clone Pt1 8.6, CCMP632) were grown in liquid culture as described previously under constant light (100 µmol photons m^{-2} s^{-1}), at 20°C with shaking (150 rpm) (Nymark *et al.*, 2009). Before experiments, the cells (PtWT and mutants) were acclimatised by cultivation in continuous light under shaking (150 rpm) for 2–3 wk; cells were diluted 10-fold every second day to maintain exponential growth. The experimental cultures were inoculated to *c.* 1×10^5 cells ml^{-1} in either f/2 medium (36.38 µM PO_4^{3-}, +P cells) or f/2 medium with reduced P concentration (0.6 µM PO_4^{3-}, −P cells). The mutant lines (*ptpsr 1_3.4, ptpsr 1_3.10, ptpsr 2_2.2, ptpsr 2_2.3*) and PtWT were grown in four biological replicates under both −P and +P conditions. For validation of the AP activity experiment, RNA samples at 48 h were taken to confirm reproducibility of the experimental setup under P limitation (data not shown). For cell counts, a NovoCyte flow cytometer (ACEA Biosciences, San Diego, CA, USA) was used with light excitation at 488 nm, a forward angle light scatter (FSC, 0 ± 13), and the fluorescence filter FL3 with absorption at 670 nm as a threshold. Cultures were regularly tested for axenity by plating on agar plates with f/2 medium and 1 g l^{-1} peptone followed by incubation for 3–5 d at room temperature. Axenic growth was confirmed throughout the experiments.

Vector construction and sgRNA design

The CRISPR/Cas9 vector pKsdiaCas9_sgRNA (Addgene ID: 74923, Watertown, MA, USA) and pPtPuc3_Cas9_sgRNA (Addgene ID: 109219) were used to produce InDels in *Phatr2_47256* encoding PtPSR in *P. tricornutum* (Nymark *et al.*, 2016; Sharma *et al.*, 2018). The CRISPR/Cas9 vector, pKsdiaCas9_sgRNA, is delivered via biolistic method and plasmid is integrated into the host genomic DNA, while pPtPuc3_Cas9_sgRNA is delivered via a bacterial conjugation method and remains as an episome inside the cells (Nymark *et al.*, 2016; Sharma *et al.*, 2018; Slattery *et al.*, 2018). A custom-made Perl script that search for unique PAM sites with low homology to other genomic loci that might produce off-target effect was used to design single guide RNA (sgRNA) adapters for targeting two regions in *PtPSR* gene. *PtPSR* PAM 1 (target site 1)

was located 5′ of the Myb DNA-binding domain so that out-of-frame InDels would produce a nonfunctional protein lacking the Myb domain, and *PtPSR* PAM 2 (target site 2) targeted the conserved domain 2. The adapter sequences (followed by PAM sequence: 5′-NGG-3′) were selected as described in Nymark *et al.* (2016) by using an in-house program based on a Perl script. The adapters targeting *PtPSR* were ligated into the pKSdiaCas9_sgRNA and pPtPuc3_Cas9_sgRNA plasmids as described previously by Nymark *et al.* (2016, 2017). *PtPSR* PAM1 lines were generated using pKS_diaCas9_sgRNA plasmids (Nymark *et al.*, 2016) and *PtPSR* PAM2 lines using pPtPuc3_diaCas9_sgRNA with transformation via biolistic bombardment and bacterial conjugation, respectively (Sharma *et al.*, 2018).

Transformation

Phaeodactylum tricornutum cells were transformed as described previously (Falciatore *et al.*, 1999; Nymark *et al.*, 2017). For delivery by conjugation to *P. tricornutum*, cells were transformed using *Escherichia coli* (DH10β) as previously described (Sharma *et al.*, 2018; Slattery *et al.*, 2018). *E. coli* cells with conjugative pT-MOB plasmid and pPtPuc3_diaCas9_sgRNA plasmid were grown at 37C to OD_{600} 0.8–1.0, harvested by centrifugation (10 min, 3000 ***g***) and resuspended in 500 µl SOC media (Karas *et al.*, 2015). *Phaeodactylum tricornutum* cells were grown for 4 d on 50% f/2-Si agar plates and then scraped off, resuspended and adjusted to 5×10^8 cells ml^{-1}. For the conjugation, 200 µl each of *P. tricornutum* and *E. coli* cells were mixed and plated on 50% seawater (SW) f/2 with 5% LB agar plates and incubated in darkness at 30°C for 90 min. Plates were transferred to constant light (100 µmol photons m^{-2} s^{-1}) at 20°C. After 2 d, *P. tricornutum* cells were transferred to selection plates. Colonies were selected from selection plates after 2 wk and transferred to f/2 liquid medium containing 50 µg ml^{-1} zeocin or on new selection plates.

Screening primary *ptpsr* mutant colonies and isolation of *ptpsr* mutant lines

Phaeodactylum tricornutum mutations at PAM1 and PAM2 sites were identified by high-resolution melting (HRM) analysis followed by Sanger sequencing of PCR amplicons using primers from Supporting Information Table S1, as described by Nymark *et al.* (2016, 2017). Pure mutant lines were obtained as previously described (Nymark *et al.*, 2016; Sharma *et al.*, 2018). Primary mutant colonies were diluted (*c.* 200 cells ml^{-1}), and 250 µl of cells were transferred to selection plates (50% SW f/2, 100 µg ml^{-1} zeocin, 1% (w/v) agar). Mutants grown on selection plates were picked and checked for target mutations as described above. Clean biallelic clones were identified by HRM and the PCR product was cloned into the pCR4-TOPO TA vector (Invitrogen). At least six PCR clones from each colony were sequenced to get an overview of the different InDels present within a colony. Allelic InDels and mutations were verified using single nucleotide polymorphisms (SNPs) observed in the gene (Sharma *et al.*, 2018).

Alkaline phosphatase assay

Phaeodactylum tricornutum cells were harvested for cellular AP activity (cAP) measurements by centrifugation (15 min, 4000 ***g*** and 1 min, 16 000 ***g***). The supernatant was removed and replaced by fresh seawater to obtain a cell concentration of 7×10^6 cells ml^{-1} 0.5 ml concentrated cell suspension was added to 0.5 ml Tris-HCl–pNPP solution (100 mM Tris-HCl and 10 mM *p*-nitrophenyl phosphate (pNPP) were suspended in sterile seawater and adjusted to pH 9.5). After 2 h incubation in dimmed light at room temperature, presence of pNP (*p*-nitrophenol) was measured by absorption at 405 nm. As a blank, 0.5 ml concentrated cell suspension was added to 0.5 ml Tris-HCl–pNPP solution immediately before measuring. A dilution series of pNP standard was measured at 405 nm and together with the cell count used to determine the AP activity (mM conversion of pNP per 10^6 cells). External AP activity (eAP) was measured using the supernatant of centrifuged cell samples and measured as described above.

Phosphorus determination

Samples for dissolved P_i were collected as supernatant after centrifugation of algal cells and stored at −20°C until analysis. Pi determination was done according to the standard colorimetric method by the reaction of phosphomolybdic acid at 880 nm with the Analytical Flow Solution IV analyser (Hansen & Koroleff, 1999). For cellular P, defined volumes of the cell cultures (10–50 ml) were filtered on a *c.* 1-µm Whatman GF/C filter and stored at −20°C until analysis following Norwegian standard NS6878 (2004).

RNA isolation and quantitative real-time PCR

Depending on cell density, 100–250 ml of four biological replicates of each mutant line were quickly harvested through vacuum filtration on 0.65-µm Durapore membrane filters (DVPP; Merck Millipore, Darmstadt, Germany) at time points 2, 24 or 48 h. Cells were washed from filters using 1 ml of the same growth medium and then centrifuged at 16 000 ***g*** for 1 min at 4°C. The supernatant was discarded, and the pellet was flash frozen in liquid nitrogen and stored at −80°C. Total RNA was isolated using the Spectrum Plant Total RNA Extraction Kit (Sigma-Aldrich) and quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, NC, USA). Reverse transcription of total RNA was performed from 1000 ng RNA using the QuantiTect Reverse Transcription Kit (Qiagen) and oligo(T) and random primers. Four biological replicates from each line, time point and treatment were used to perform quantitative real-time PCR (qRT-PCR) on a LightCycler 480 using the LightCycler 480 SYBR green I master kit (Roche Applied Science, Indianapolis, IN, USA). The standard method included a preincubation for 5 min at 95°C, followed by 50 cycles of amplification consisting of 10 s at 95°C, 10 s at 58°C and 15 s at 72°C. Both genes of interests (GOIs) and housekeeping genes used for qRT-PCR were chosen based on the microarray data from previous P limitation experiments in *P. tricornutum* and *T. pseudonana* (Table S2) (Dyhrman

et al., 2012; Alipanah *et al.*, 2018). LinReg PCR 11.1 software was used to determine primer efficiency, and qBasePlus software for identifying the Ct value (Ramakers *et al.*, 2003; Ruijter *et al.*, 2009). Means of the individual biological samples ($n = 4$) were calculated and used to determine the ratio and the respective $\log_2$ value, using the qBasePlus software (Biogazelle, www.qbaseplus.com) and one-way ANOVA analysis. For *ptpsr* mutants, regulation of selected genes at each time point was calculated using PtWT as reference under the same conditions.

Phylogenetic analysis of Myb DNA-binding domain

Protein alignments were generated using Mega 7.0.26 (Kumar *et al.*, 2016). A number of substitution matrices were evaluated (Mega 7.0.26) and the best one selected. Maximum likelihood (ML) trees were created with Mega 7.0.26 using the LG model as the amino acid substitution model, with gamma-distributed mutation rates among amino acid sites and complete deletion of gaps and missing data. Mega 7.0.26 was also used to create bootstrapped neighbour-joining (N-J) (Saitou & Nei, 1987) trees using the Jones–Taylor–Thornton (JTT) model as the amino acid substitution model. In total, 1000 nonparametric bootstrap inferences were executed both for ML and N-J trees. The trees were refined using Affinity Designer (Serif (Europe) Ltd, Nottingham, UK).

PtPSR recognition motif screen

The Multiple EM for Motif Elicitation (MEME) web server was used for identifying potential transcription factor binding sites in promoters from 84 selected genes, using the 0-order model to normalise for biased distribution of bases (Bailey *et al.*, 2006). The promoters were selected from genes that showed a specific and robust induction under P-limited conditions, that is the GEO dataset GSE66063 (Alipanah *et al.*, 2018).

Electrophoresis mobility shift assay

Full-length *PtPSR* was cloned into pBADM-20(+) (Addgene) to generate a fusion protein with an N-terminal His-tag. The recombinant protein was expressed in *E. coli* BL21-CodonPlus (DE3)-RIPL (Stratagene, La Jolla, CA, USA), purified on HisTrap column (GE Healthcare, Pittsburgh, PA, USA) and treated with Tobacco Etch Virus (TEV) protease to cleave off the tag. For the Electrophoresis Mobility Shift Assay (EMSA), three 51 bp DNA oligos (Integrated DNA Technologies, Leuven, Belgium) were used: one containing three PtPSR binding motifs and labelled with IRDye700 ('probe'), a similar unlabelled one ('competitor') and one unlabelled where the PtPSR binding motifs were mutated ('mutated competitor'). Double-stranded probes, DNA-binding reactions and EMSA were performed as described by Nakata & Ohme-Takagi (2014) using the Odyssey Infrared EMSA Kit (Li-Cor Biosciences, Lincoln, NE, USA).

SDS-PAGE

Tris-glycine SDS-PAGE mini gels were run using Precision Plus Protein Dual Color Standards (Bio-Rad) as the molecular weight marker and stained with SimplyBlue SafeStain (Life Technologies).

Lipid contents and composition

Lipid contents and class was determined as described in Mühlroth *et al.* (2017). Algae cultures were harvested by centrifugation (15 180 ***g***, 10 min) and subsequent filtration. The filter was resuspended in 0.5 ml seawater, centrifuged (20 000 ***g***, 1 min) and the pellets were stored at −80°C. Lipids were extracted and quantified by mass spectrometric analysis according to Jouhet *et al.* (2017).

Results

Identification and characterisation of PtPSR and phylogenetic analysis of the Myb DNA-binding domain

Myb TFs in plants and green algae are known components of a central regulatory system controlling transcriptional responses to P limitation (Jyoti *et al.*, 2019). When we investigated the transcriptional response to P limitation in genes encoding TFs of the diatom *P. tricornutum*, using a publicly available dataset (Alipanah *et al.*, 2018), the most responsive TF at 48 and 72 h after P deprivation was a transcript encoding a Myb-like TF annotated as PtMyb1R_SHAQKYF5 (*Phatr2_47256*), here re-named PtPSR. *PtPSR* transcript induction was more than four times stronger compared with other Myb-like genes (Fig. 1a). We also analysed a transcriptome dataset of P limitation in the eustigmatophyte *N. oceanica* (Mühlroth *et al.*, 2017). Two *N. oceanica* transcripts encoding Myb-like TFs closely related to *PtPSR* (No_10770 and No_8993 in the genome assembly by Vieler *et al.* (2012)) displayed a similar early induction under P limitation, suggesting that the role of this TF is conserved among stramenopiles (Fig. S1). *Nannochloropsis* spp. harbour 35 Myb and Myb-related TF genes (15 R2R3 Myb, 8 R1R2R3 Myb and 12 Myb related); three of these are related to PtPSR (Hu *et al.*, 2014). PtPSR belongs to a family of nine Myb-like TFs in *P. tricornutum* that is characterised by a single Myb DNA-binding domain-containing a SHAQKYF-like motif (Rayko *et al.*, 2010). SHAQKYF-like Mybs in diatoms can be further divided into three classes. Class I SHAQKYF-like Mybs are related to *Arabidopsis* ARR and PHR1-like proteins, class II are weakly related to Arabidopsis CCA1 and LHY proteins while class III are diatom-specific (Fig. S2; Rayko *et al.*, 2010). PtPSR (PtMyb1R_SHAQKYF5) belongs to class I together with PtMyb1R_SHAQKYF8 (*Phatr2_50328*), which showed lower induction than PtPSR at 72 h under P limitation (Fig. 1a). PtPSR encodes a 575 amino acids (AA) long protein with four conserved domains (CDs) of unknown function and the Myb DNA-binding domain at 219 to 275 AA (Fig. 1b). A putative nuclear localisation signal (NLS, GGKSSRSSKVRRKGK, amino acid residues 198–212) predicted by Localizer 1.0 (Sperschneider *et al.*, 2017) is present upstream of the Myb DNA-binding domain. The CD2 domain is conserved in diatoms and contains a predicted helical structure. The CD3 domain showed similarity

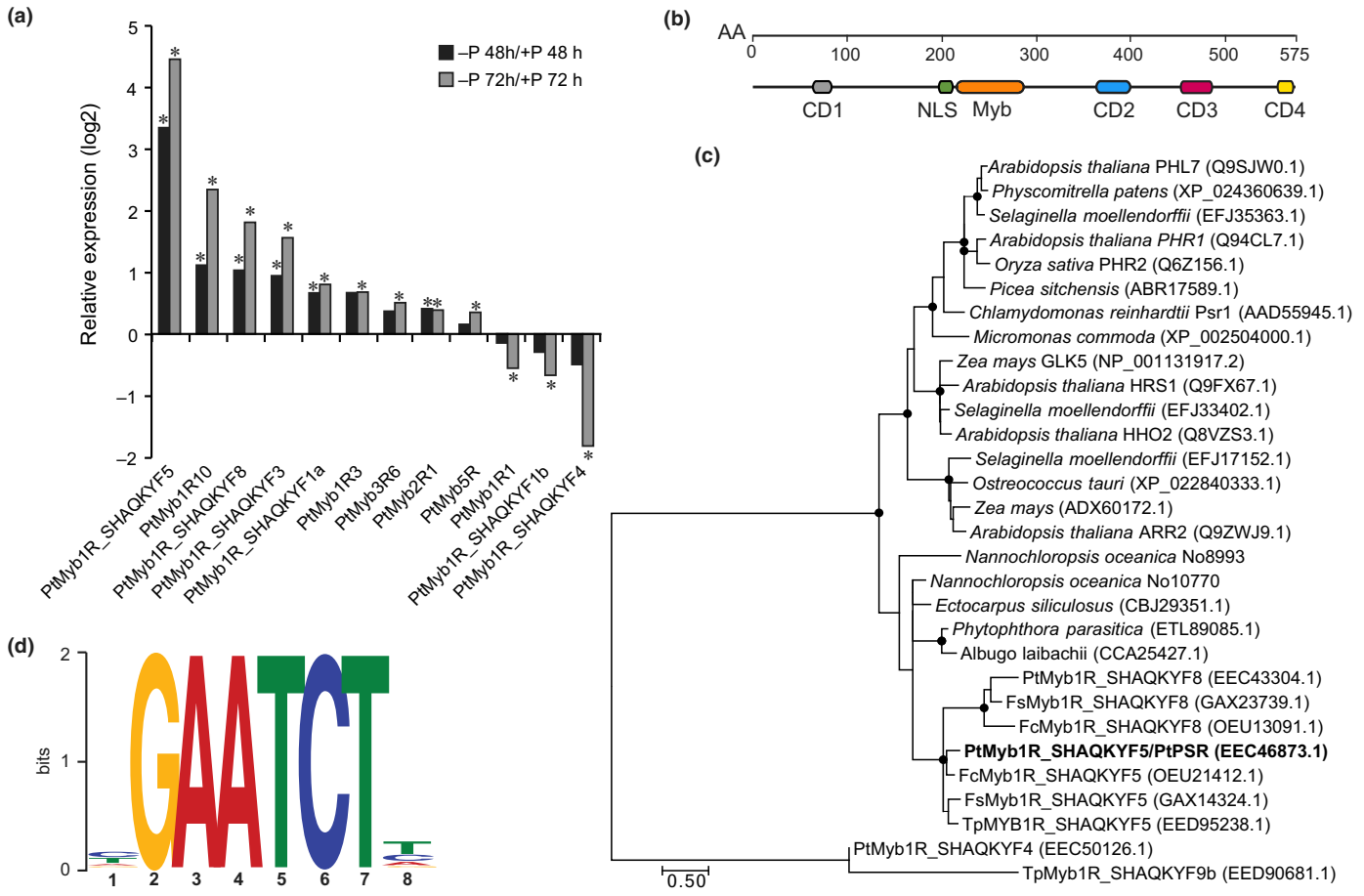


Fig. 1 The Myb transcription factor *PtPSR* is strongly induced at the transcript level by P limitation. (a) Transcriptional responses of all *Phaeodactylum tricornutum* Myb TFs significantly regulated (*, $P < 0.01$) by P depletion after 48 and 72 h. The ratios were $\log_2$ transformed. Raw data from Alipanah *et al*. (2018). Myb nomenclature is taken from Rayko *et al*. (2010). (b) Domain organisation of *PtPSR*. CD, conserved domain; NLS, nuclear localisation signal. (c) Unrooted maximum likelihood (ML) tree based on a protein alignment of *PtPSR* and related class I SHAQKYF TFs. The overall topologies for ML and Neighbour-Joining (NJ) trees are the same. Black circles indicate nodes supported by ML and NJ bootstrap values (1000× sampling, > 60%). (d) 100% conserved *PtPSR* recognition motif (expect value (*e*) = 4.8×10^{-61}) based on the promotor sequence of 84 P-stress-inducible genes (Supporting Information Fig. S1) analysed with MEME (Bailey *et al.*, 2006).

in several diatoms and was identified as a putative CC domain using both COILS and Parcoil2 CC prediction programs (Lupas *et al.*, 1991; McDonnell *et al.*, 2006). Similar CC domains are found in Myb-CC TFs in plants including PHR1 in Arabidopsis. However, no clear function could be identified for the CDs. A phylogenetic analysis of the Myb domain indicated that Subgroup I SHAQKYF-like Mybs are distributed throughout stramenopiles (Fig. 1c).

In order to identify genes exhibiting a P limitation-specific transcript response, we collected publicly available transcriptome datasets for response to silicon limitation (Sapriel *et al.*, 2009), cadmium exposure (Brembu *et al.*, 2011), N limitation (Alipanah *et al.*, 2015) and P limitation (Alipanah *et al.*, 2018). The promoter regions of the 84 genes strongest induced by 72 h of P stress in the dataset from Alipanah *et al.* (2018), but little affected or not in the other datasets (Fig. S3) were analysed for motifs using the MEME web server (Bailey *et al.*, 2006). Our analysis revealed a highly conserved motif containing 8-bp sequence 5′-YGAATCTH-3′ (Fig. 1d) within 81 of these genes, with up to six copies of the motif in the promoter region of a single gene, such as the alkaline phosphatase AP1 *Phatr2_49678* (Fig. 2). This enriched motif was similar to the recognition motifs of Arabidopsis P-stress-inducible Myb TFs, HYPERSENSITIVITY TO LOW PHOSPHATE-ELICITED PRIMARY ROOT SHORTENING1 (HRS1, At1g13300, *P*-value: $1.1 \times e^{-02}$) and its homologues HRS1 HOMOLOG2 (HHO2, At1g68670, *P*-value: $4.7 \times e^{-04}$) and HRS1 HOMOLOG6 (HHO6, At1g49560, *P*-value: $2.6 \times e^{-04}$) (Fig. S4), suggesting that this could be the recognition motif for PtPSR, which itself has four of these conserved motifs in its promoter. Transcription factor binding sites are known to be conserved through evolution in both plants and animals (Nitta *et al.*, 2015).

We performed an EMSA analysis to verify the binding of recombinant PtPSR to the predicted recognition motif using IRDye-labelled DNA probes containing three GAATCT motifs (Fig. S5a). PtPSR binds specifically to the GAATCT sequences,

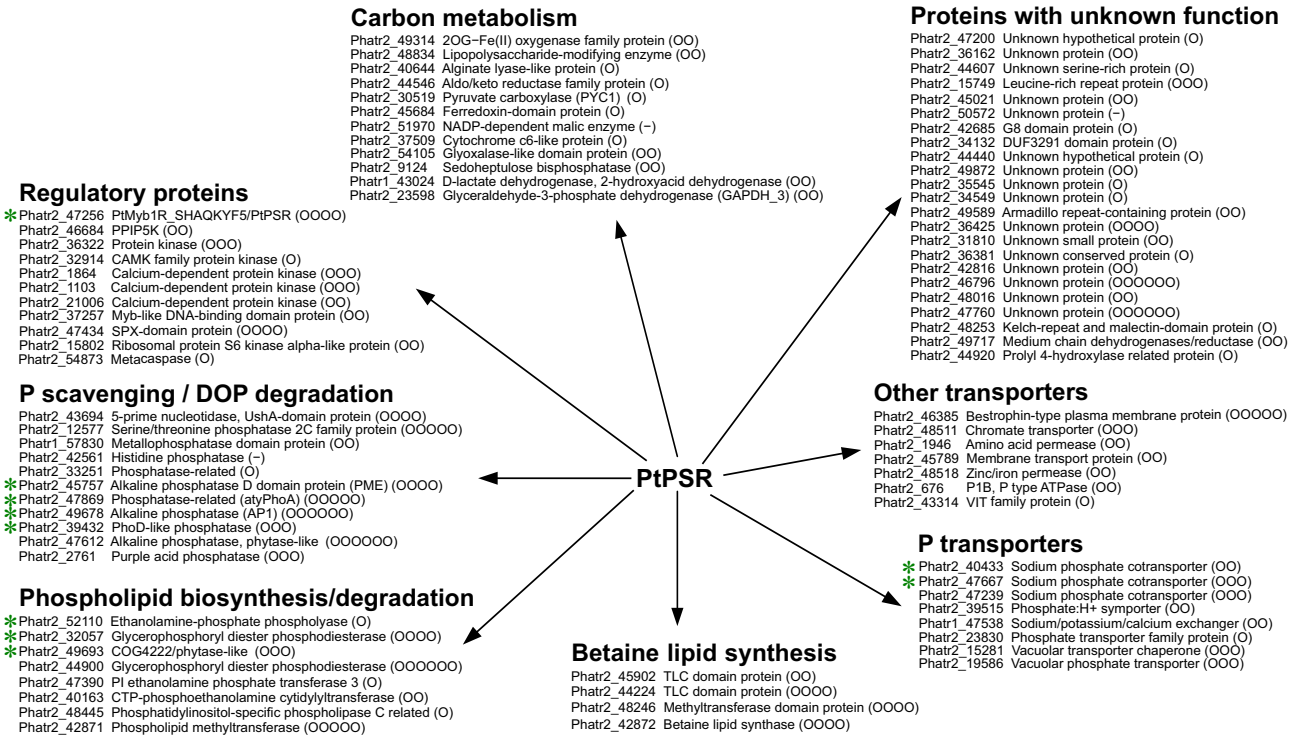


Fig. 2 Strongly P stress-inducible genes contain *PtPSR* recognition motifs. Eighty-one of the 84 strongest upregulated genes under P stress contained putative *PtPSR* recognition motifs in promotor region in *Phaeodactylum tricornutum*. Transcriptome data were derived from microarray data (Sapriel *et al.*, 2009; Brembu *et al.*, 2011; Alipanah *et al.*, 2015; Alipanah *et al.*, 2018). (O), number of *PtPSR* recognition motif within the promotor region of the gene; (−), genes without *PtPSR* recognition motif; green asterisks, genes validated by qRT-PCR.

as signal was lost from the PtPSR/DNA complex when the unlabelled competitor DNA probe was added but not when an unlabelled DNA probe with mutated motifs was added (Fig. S5b). Also, the unrelated protein *A. thaliana* NSP1 (Kong *et al.*, 2012), expressed and purified under similar conditions as PtPSR, did not bind the probes. For some reason, PtPSR bound to the probe did not migrate into the native gel, while migration of PtPSR was not prevented on SDS-PAGE (Fig. S5c).

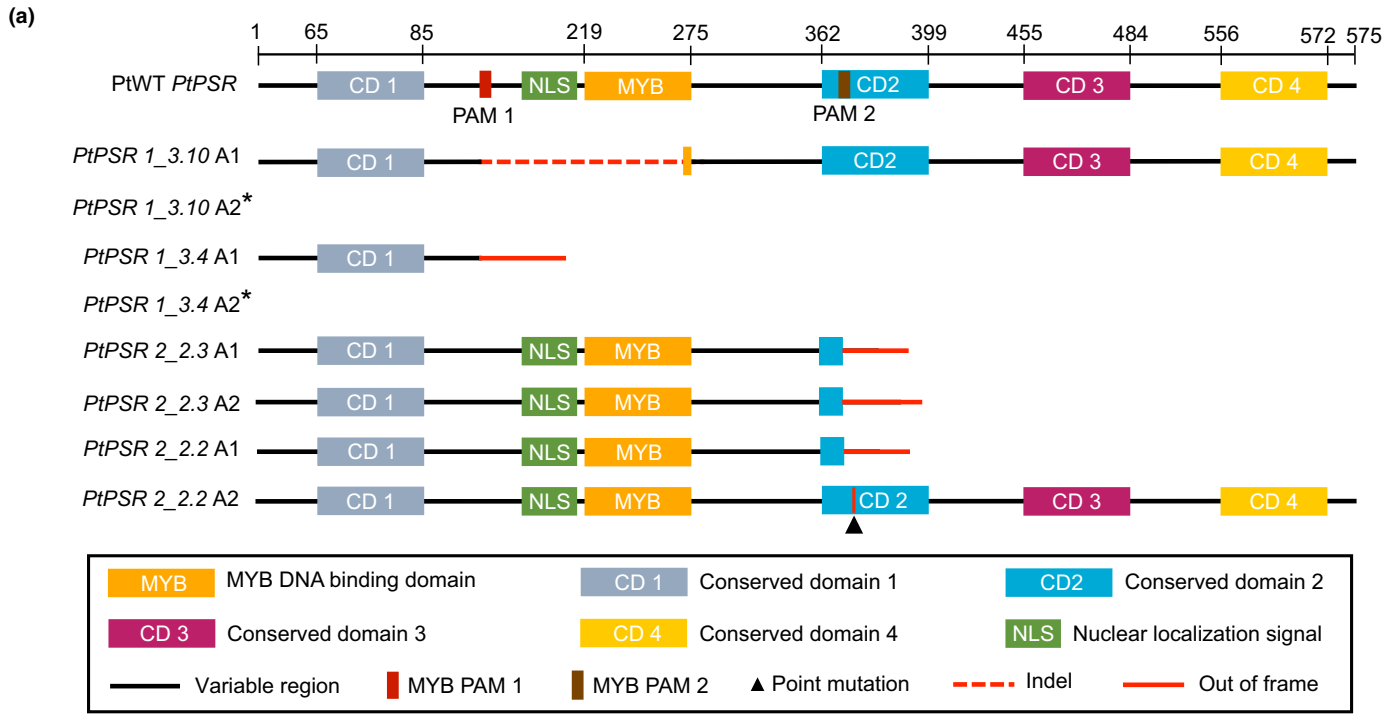
Generation of *ptpsr* mutants

ptpsr mutant cell lines were generated using CRISPR/Cas9-based gene editing. Two sgRNAs were designed targeting the N-terminus between the CD1 and Myb domains (PAM1) and the CD2 domain (PAM2), respectively (Fig. 3a). Among the transformants obtained, two mutants for each target site were further analysed. Both of the PAM1 mutants, ptpsr *1_3.4 and ptpsr 1_3.10,* carried large deletions in allele 1 and rearrangements in allele 2, resulting in a complete loss of the Myb DNA-binding domain (Figs 3a,b, S6). None of the PAM1 lines harboured wild-type alleles. A 504 bp deletion and 2 bp insertion in *ptpsr 1_3.4* introduced a premature stop codon due to out-of-frame InDels. A 537 bp in-frame deletion of *ptpsr 1_3.10* resulted in deletion of the Myb domain while all other CDs remained intact (Fig. 3c). To test for off-target effects, loci showing the highest probability for off-target by CRISPR/Cas9 (4 bp mismatch against sgRNA) were sequenced (Table S3). The sequencing data showed no mutation at the predicted off-target regions (one for each sgRNA) tested in *P. tricornutum* (data not shown). In the PAM2 mutant *ptpsr 2_2.3*, editing of both alleles of *PtPSR* resulted in the loss of CDs except CD1 and the Myb DNA-binding domain, due to a 14 bp deletion in allele 1 and a 1 bp insertion in allele 2. *ptpsr 2_2.2* had a 14 bp deletion in allele 1 resulting in an out-of-frame mutation. However, allele 2 of *ptpsr 2_2.2* contained a point mutation in CD2, leading to an exchange of a 100% conserved leucine to phenylalanine (L373F). In conclusion, the *PtPSR* PAM2 mutants had an active and probably functional Myb DNA-binding domain, but lacked CD3 and CD4, whereas the *PtPSR* PAM1 mutants exhibited an inactive Myb DNA-binding domain.

Effect of P limitation on *ptpsr* mutants

ptpsr mutants and PtWT were cultivated under P-deplete (−P) and replete (+P) conditions. Cell growth and cellular P were monitored to observe the effect of mutations in the *PtPSR* gene. AP activity and the dynamic regulation of P-stress response genes were analysed.

P household and P transport in ptpsr mutants Cell growth and specific growth rate (μ) presented in Fig. 4(a) and 4(b), respectively, showed no significant difference between *ptpsr* mutants and PtWT under +P conditions. After 48 h of −P cultivation, all mutant lines showed reduced μ compared with PtWT; however,



(b)

```
                                  PAM1    20 bp target site 1
PtWT PtPSR:      5´ GC--TACCCTCTACACATCGGCAGCAACACGGAATC//CATGCAGACGTCACCAGCGAACGGATCAAGAGTCATTTACAA 3´

PtPSR 1_3.10 A1:    GC--TAC----------------------------//----------------------------------TTACAA -537 bp
PtPSR 1_3.10 A2:    ---------------------------------//CATGCAGACGTCACCAGCGAACGGATCAAGAGTCATTTACAA rearrangement

PtPSR 1_3.4 A1:     GCGC-----------------------------//--CATGCAGACGTCACCAGCGAACGGATCAAGAGTCATTTACAA +2 bp,-504 bp
PtPSR 1_3.4 A2:     ----------------------------ACGGAA//CATGCAGACGTCACCAGCGAACGGATCAAGAGTCATTTACAA rearrangement

                                                  20 bp target site 2      PAM2
PtWT PtPSR:      5´ AGAAAAAGAGAAAAATTCTCCGTTGGGGACGTCCATGAGTTATC-TCATGGGCCTTTTTTTCTCGTTGCGGCAACAACTG 3´

PtPSR 2_2.3 A1:     AGAAAAAGAGAAAAATTCTCCGTTGGGGACGTCCATGC------------CTTTTTTTCTCGTTGCGGCAACAACTG -14 bp
PtPSR 2_2.3 A2:     AGAAAAAGAGAAAAATTCTCCGTTGGGGACGTCCATGAGTTATCCTCATGGGCCTTTTTTTCTCGTTGCGGCAACAACTG +1 bp

PtPSR 2_2.2 A1:     AGAAAAAGAGAAAAATTCTCCGTTGGGGACGTCCATGC------------CTTTTTTTCTCGTTGCGGCAACAACTG -14 bp
PtPSR 2_2.2 A2:     AGAAAAAGAGAAAAATTCTCCGTTGGGGACGTCCATGAGTTATT-TCATGGGCCTTTTTTTCTCGTTGCGGCAACAACTG C>T
```

(c)

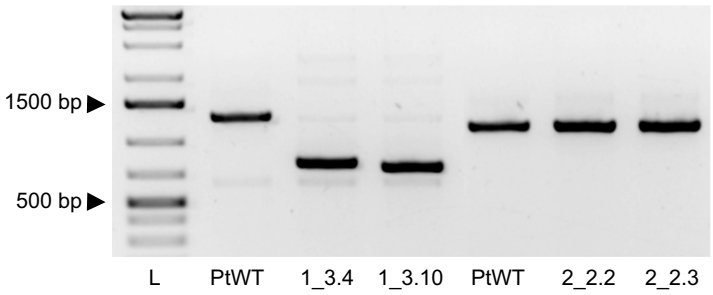


Fig. 3 Alignment of *Phaeodactylum tricornutum* wild-type (PtWT) and *ptpsr* mutant lines. (a) Representation of putative conserved domains in the *PtPSR* protein in comparison to designed mutants. The block size of the conserved domain does not represent the actual size of conserved domains. PAM sites and results of mutants are indicated for the specific mutants. (b) Partial gene sequence alignment of PtWT *PtPSR* compared with the *PtPSR* mutants (*ptpsr 1_3.4, ptpsr 1_3.10, ptpsr 2_2.2, ptpsr 2_2.3*). The 20-bp target sites 1 and 2 are indicated in red and green, respectively, with their corresponding PAM sequence in blue. PAM site 1 (CCT) refer to the reverse strand. The size of each deletion/insertion and the name of the allele are indicated at the right and left, respectively. Rearrangements in allele 2 of the *PtPSR PAM1* mutant lines were identified by high throughput sequencing (*). For a more detailed presentation of the rearrangements see Supporting Information Fig. S6. (c) PCR-amplified DNA fragments of the target sites.

μ was significantly reduced only in *ptpsr 1_3.4* and *1_3.10* (Fig. 4b). Cellular P (cP) was reduced in all mutant lines except *ptpsr 2_2.2* after 48 h in +P conditions (Fig. 4c). This trend was strengthened by –P cultivation; all mutant lines (except *ptpsr 1_3.10*) showed significantly reduced cP after 24 and 48 h (Fig. 4c). The cP content of *ptpsr 1_3.4* under −P conditions was 1/7 of PtWT cells at 48 h. Dissolved P (dP) in the medium showed variation between PtWT and mutants under +P and –P conditions, but no clear pattern was observed (Fig. 4d).

An expression analysis was performed on *PtPSR* in PtWT and the mutant lines, using PCR primers that were not affected by the mutations. *PtPSR* expression in PtWT was quickly induced by –P conditions. After 2 h, *PtPSR* expression in PtWT was 10-fold higher in –P compared with +P cultures (Fig. 5a). *PtPSR* expression in the mutant lines in –P conditions was moderately affected compared with PtWT –P conditions; however, at 48 h *PtPSR* was downregulated in *ptpsr 2_2.3* and *ptpsr 1_3.4* compared with PtWT –P conditions (Fig. 5b). To further investigate the transcriptional effect of PtPSR inactivation, we analysed several genes belonging to the P-stress response-specific gene set that carried the 100% conserved PtPSR core recognition motif (GAATCT) in their promoters (Figs 2, S3). As shown in Fig. 5(a), all genes displayed P limitation-induced expression in PtWT at the three time points analysed.

Alkaline phosphatase enzyme activity and phosphomonoesterase gene expression The survey of P starvation-induced genes containing GAATCT motifs included several phosphatases and genes of lipid metabolism, all of which were upregulated in −P conditions in wild-type (Fig. 5a). Four genes involved in phosphomonoester degradation and AP activity were analysed: *AP1* (*Phatr2_49678*), *atypical PhoA (atyPhoA, Phatr2_47869*) (Li *et al.*, 2018; Lin *et al.*, 2015), *phosphomonoesterase* (*PME, Phatr2_45757*) and *PhoD* (*Phatr2_39432*). These four genes showed a strong gradual increase in their gene expression in PtWT cells during cultivation under P limitation compared with +P conditions (Figs 5a, S7a). By contrast, expression of these genes in the mutant lines was generally downregulated in −P cultures at all time points compared with PtWT (Fig. 5c–f). *ptpsr 1_3.4* showed the lowest AP activity and highest downregulation

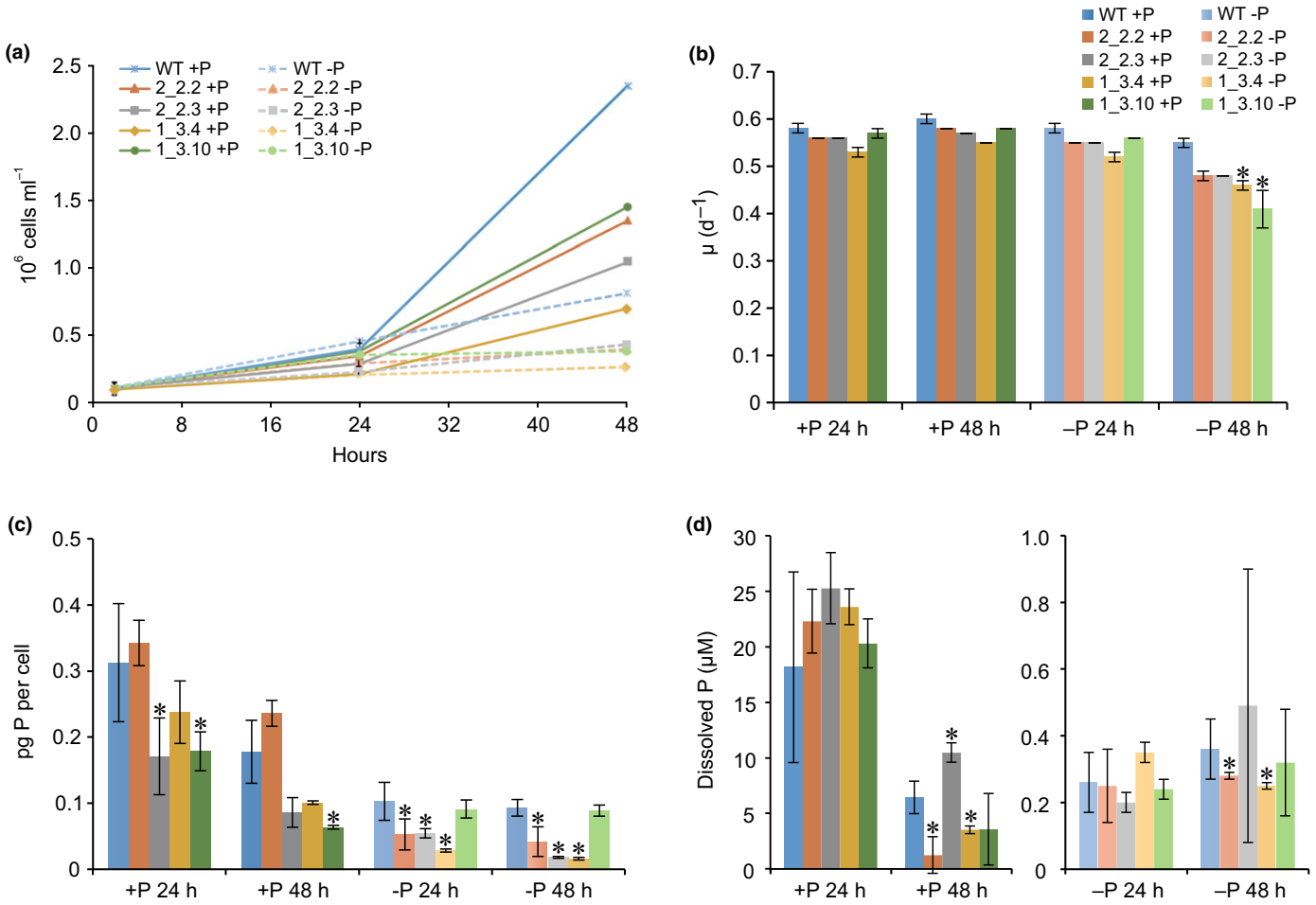


Fig. 4 Growth, cellular and dissolved phosphorus of *Phaeodactylum tricornutum* wild-type (PtWT) and *ptpsr* mutants in −P and +P conditions. (a) Time analysis of cell count (cells ml^{-1}) of +P (closed symbols) and −P (open symbols) cell lines. (b) Specific growth rate μ (d^{-1}). (c) Cellular P (pg P per cell) in +P and −P conditions. (d) Dissolved P (μM) after 24 and 48 h of cultivation under +P and −P conditions. Asterisks indicate significant difference (*, $P < 0.05$) compared with PtWT in the same conditions. Data are expressed as means ± SD ($n = 4$).

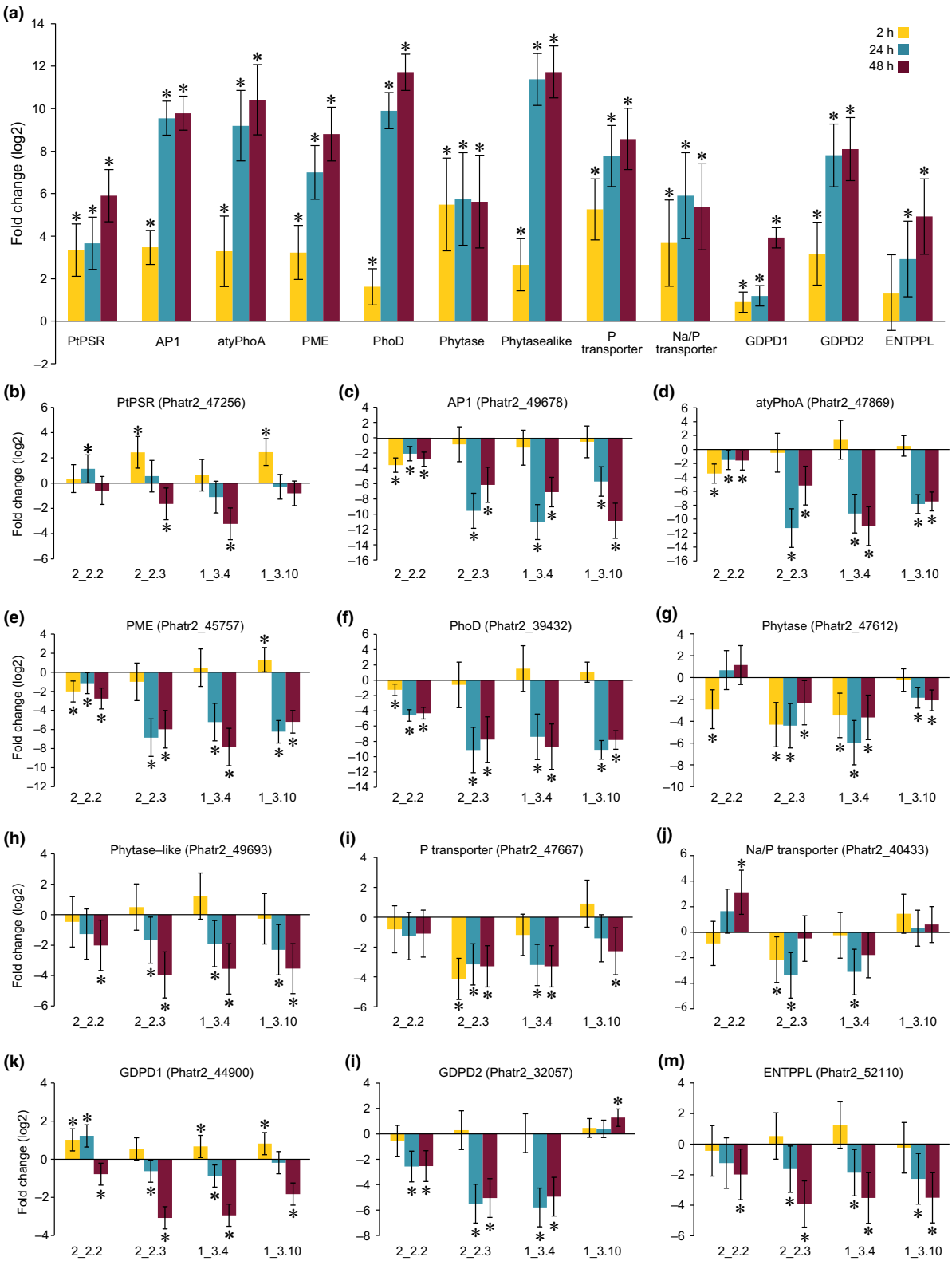


Fig. 5 Relative mRNA expression of P-stressed induced genes in *Phaeodactylum tricornutum* wild-type (PtWT) and *ptpsr* mutants under phosphate-deplete conditions. (a) The histograms shows fold change (log_2 transformed) of PtWt after 2, 24 or 48 h of growth under phosphate-deplete conditions compared with 2 h growth under phosphate-replete conditions. (b–m) All *PtPSR* mutants (*ptpsr 1_3.4*, *ptpsr 1_3.10*, *ptpsr 2_2.2*, *ptpsr 2_2.3*) compared with PtWT cells at same time points all under phosphate-deplete conditions. Statistically significant upregulation (*) relates to the *ptpsr* mutants compared with PtWT at the same time point and is based on ANOVA test. Each data point represents the mean value ($n = 4$) and error bars represent 95% confidence intervals. AP1, alkaline phosphatase 1; atyPhoA, atypical PhoA; ENTPPL, phosphoethanolamine phospholyase; GDPD1/2, glycerophosphoryldiester phosphodiesterases 1/2; PhoD, alkaline phosphatase D; PME, phosphomonoesterase.

of genes related to phosphomonoester degradation compared with PtWT under P limitation. Therefore, the most significant effects on the regulation of PME- and AP-encoding genes were observed in mutants with an impaired Myb DNA-binding domain and CC domain.

AP external (eAP) and cellular (cAP) enzyme activities were detected in the mutant lines and correlated to the expression levels of genes involved in phosphomonoester degradation. At +P conditions eAP and cAP enzyme activities were very low in PtWT and below detection limit in all mutants (Table 1). Under P limitation, however, the AP activities in PtWT increased more than 220 times, whereas *ptpsr 1_3.4*, *ptpsr 1_3.10* and *ptpsr 2_2.3* showed no detectable activities and *ptpsr 2_2.2* significantly lower activities (Table 1).

We also analysed two genes encoding a phytase (myo-inositol hexaphosphate phosphohydrolase) (*Phatr2_47612*) and a phytase-like (*Phatr2_49693*) protein, which could be involved in phytate degradation and/or phosphate scavenging. The phytase was induced at all time points in PtWT, but showed significantly lower expression in mutants except for *ptpsr* 2_2.2 (Fig. 5a,g). The phytase-like gene expression was strongly induced after 24 h in PtWT in P-deplete conditions (Fig. 5a). By contrast, expression of this gene in the *ptpsr* mutants was significantly reduced compared with PtWT at 24 and 48 h in P-deplete conditions (Fig. 5h). Inorganic P is taken up by the cell via P transporters. The low affinity Pi co-transporter (*Phatr2_47667*) was induced under P stress at all time points in PtWT and significantly reduced in *ptpsr* mutants compared with PtWT except *ptpsr 2_2.2* (Fig. 5i). Expression of the low affinity P_i co-transporter, Na+/P transporter gene (*Phatr2_40433*, Fig. 5j) followed the same trend but was less consistent.

Lipid class regulation in ptpsr mutants Two GDPDs (GDPD1 (*Phatr2_44900*) and GDPD2 (*Phatr2_32057*)) and one phosphoethanolamine phospholipase (ENTPPL, *Phatr2_52110*) are coupled to PL degradation. In PtWT cells there was a gradual increase of expression for genes encoding PL-degrading enzymes under both replete and deplete conditions compared with 2 h +P conditions (Fig. 5a). *ptpsr* mutant lines showed a downregulation of these genes compared with PtWT under similar conditions (Figs 5k–m, S7k–m).

Lipid class profiling showed that *ptpsr 1_3.4* and *ptpsr 1_3.10* had significantly higher PL as well as total lipid content compared with PtWT under both P-replete and P-deplete conditions (Figs 6, S8). P limitation resulted in a complete loss of PLs in PtWT, while the P-free lipid classes digalactosyldiacylglycerol (DGDG), diacylglyceride (DAG) and TAG increased at 72 h. PL levels in all *ptpsr* mutant lines except *ptpsr 2_2.2* were reduced 2–3 times under P limitation compared with nonlimited conditions but did not reach zero as in PtWT. P-free lipid classes in *ptpsr 1_3.4, 1_3.10* and *2_2.3* mutants were comparatively higher than PtWT in P-limited conditions at 72 h.

Table 1 Alkaline phosphatase activity of *ptpsr* mutants and *Phaeodactylum tricornutum* wild-type (PtWT).

PtPSR cell lines and conditions		eAP activity (mM *p*NP per 10^6 cells) 48 h	cAP activity (mM *p*NP per 10^6 cells) 48 h
+P	PtWT	0.8 ± 1.5	0.3 ± 0.4
	2_2.2	nd*	nd*
	2_2.3	nd*	nd*
	1_3.4	nd*	nd*
	1_3.10	nd*	nd*
−P	PtWT	217.5 ± 27.1	71.1 ± 13.9
	2_2.2	38.4 ± 0*	2.5 ± 0*
	2_2.3	nd*	nd*
	1_3.4	nd*	nd*
	1_3.10	1.0 ± 0*	nd*

External and cellular alkaline phosphatase (eAP and cAP, respectively) assay was taken after 48 h of cultivation under both conditions (mM pNP/10^6 cells). nd, not detected; *Significantly differentially regulated ($P < 0.05$) compared with PtWT. '±' represents SD.

Discussion

PtPSR, a Myb-like TF playing an essential role in the conditional adaptation of microalgae to P limitation, has been identified and functionally characterised. Downstream genes containing the PtPSR recognition motif include genes encoding proteins linked to P scavenging/acquisition, PL remodelling, transport, carbon metabolism and cellular regulation (Fig. 2). EMSA analyses verified specific binding of PtPSR to the recognition motif. *ptpsr* knockout mutants with loss of the Myb DNA-binding domain showed generally reduced expression of P-stress response genes, low cP content, highly reduced AP activity and reduced growth under P limitation (Table 2; Figs 4, 5). The helical CD2 domain, mutated in *ptpsr 2_2.2*, is also important for the P-stress response in *P. tricornutum*, as reflected in lower gene expression levels of P-stress-inducible genes (Fig. 5).

PtPSR identification and characterisation

The Myb-related P-stress-inducible TFs PHR1 in *A. thaliana* and PSR1 in *C. reinhardtii* contain within the Myb DNA-binding domain the conserved motif SHLQKYR, which is identical to the motif found in PtPSR and its orthologues in other diatoms. Myb proteins having the SHLQKYR motif are members of the class I SHAQKYF Myb proteins and are common in plants and algae. *A. thaliana* encodes more than 30 Mybs with the SHLQKYR motif, many of them known to be involved in regulating P starvation-induced genes. Phylogenetic studies here and in Rayko *et al.* (2010) showed a taxonomic distribution of the TFs related to P-stress response in green algae (*C. reinhardtii*), plants (*A. thaliana*) and stramenopiles that points to a common lineage of green algae-like SHAQKYF-like TFs (Fig. 1c). Rayko *et al.* (2010) showed through *in silico* studies that PtPSR belongs to the Myb-related proteins (MybR1) and is closely related to *T. pseudonana* TpMyb1R_SHAQKYF5 and *P. tricornutum* PtMyb1R_SHAQKYF8 (Phatr2_50328). In transcription studies of *T. pseudonana* and *N. oceanica*, the most significant upregulated TFs under P limitation were *TpMyb1R_SHAQKYF5* and the PtPSR-related TF *NannoCCMP1779_10770*, respectively

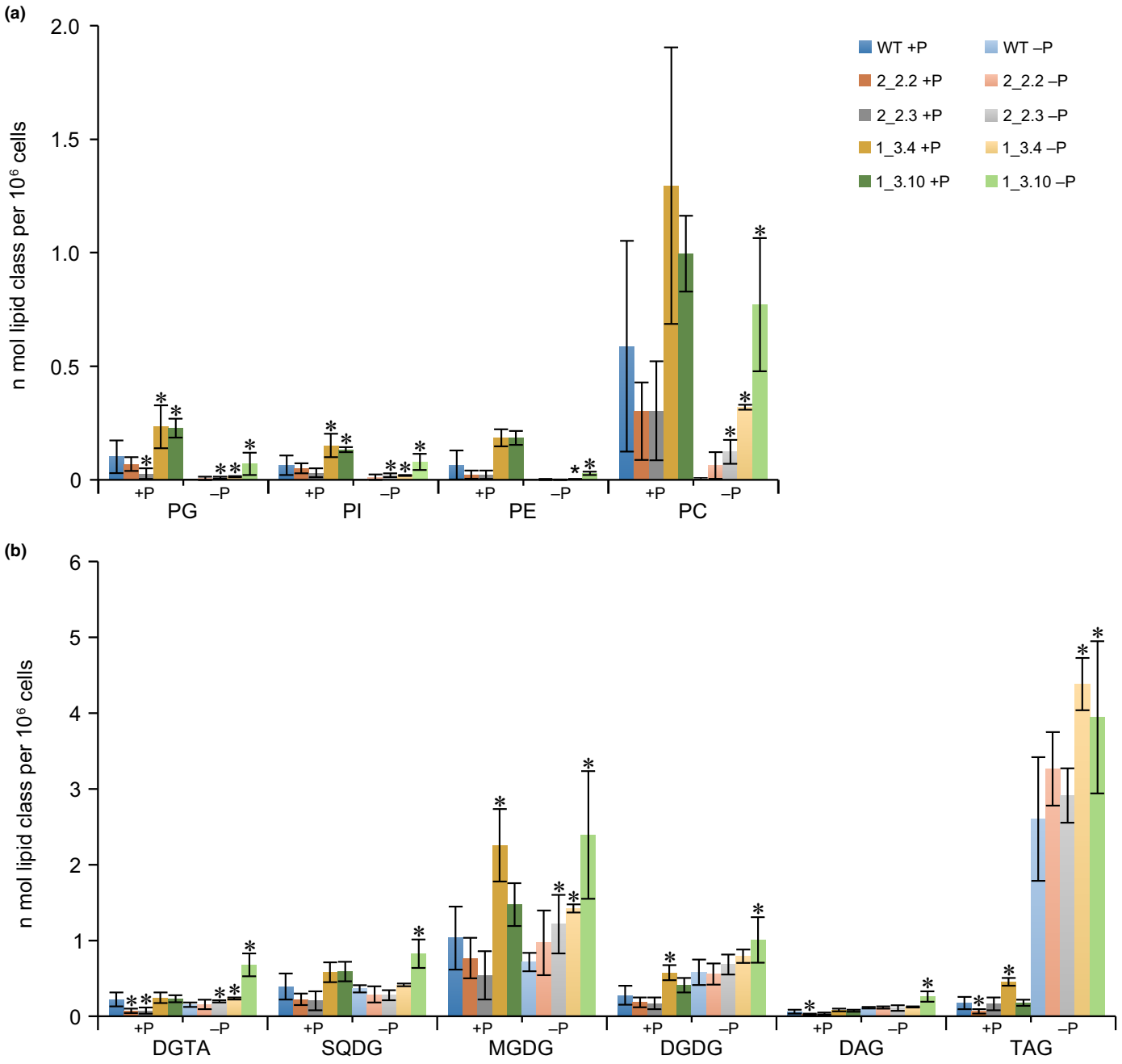


Fig. 6 Lipid class content of *Phaeodactylum tricornutum* wild-type (PtWT) and *ptpsr* mutants under +P and −P conditions at 72 h. Targeted lipid class profiling is based on linear trap quadrupole (LTQ) identification and LC-MS/MS. (a) Phospholipids. (b) Nonphosphorus glycerolipids. DAG, diacylglyceride; DGDG, digalactosyldiacylglycerol; DGTA, diacylglycerolhydroxymethyltrimethylalanine; MGDG, monogalactosyldiacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; SQDG, sulfoquinovosyl diacylglycerols; TAG, triacylglycerides. An asterisk denotes significant ($P < 0.05$) P effect on lipid classes of *ptpsr* mutants compared with PtWT either in −P or +P conditions. The mean of four biological replicates from the third independent experiment (qPCR data showed the same gene expression), when cells were grown until 72 h, ± SE are shown.

(Dyhrman *et al.*, 2012; Mühlroth *et al.*, 2017, Fig. S1). These findings indicate a similar regulation mechanism in *T. pseudonana* and *N. oceanica* under P stress as we found in *P. tricornutum*.

The MEME algorithm identified a highly enriched, conserved 8-bp sequence (5′-YGAATCTH-3′) present in the promoters of almost all P-stress-induced genes analysed (Figs 1d, 2), with high similarity to the binding sites of Myb-related TFs HRS1, HHO2 and HHO6 in Arabidopsis (Fig. S4; Rubio *et al.*, 2001; Medici *et al.*, 2015; Nagarajan *et al.*, 2016). That the GAATCT motif is indeed a PtPSR recognition motif was confirmed by EMSA analyses (Fig. S5). Large-scale studies confirm that TFs with related DNA-binding domains tended to bind very similar DNA sequences, resulting in analogous regulation in distantly related species (Weirauch *et al.*, 2014). Both HHO2 and HRS1 were shown in

Table 2 Characteristics of *PtPSR* mutants and *Phaeodactylum tricornutum* wild-type (PtWT) under P limitation.

Mutant line vs PtWT under P limitation	*PtPSR 1_3.4*	*PtPSR 1_3.10*	*PtPSR 2_2.3*	*PtPSR 2_2.2*
Myb DNA-binding domain	Inactive	Inactive	Intact	Intact
Allele 1 : CD2–CD4	X	√	X	X
Allele2: CD2–CD4	X	X	X	√(PM in CD2)
Specific growth rate (% of PtWT)	80%	74%	87%	87%
Cellular phosphorus (% of PtWT)	22%	22%	22%	110%
AP activity (% of PtWT)	nd	nd	nd	4–17.5%
Gene expression related to AP activity (% of PtWT)	0.2–2%	0.1–2%	0.1–2%	4–13%
Gene expression related to phospholipid degradation (% of PtWT)	*GDPDs* 0.3–3%, *ENTPPL* 2%	*GDPDs* 0.3–1%, *ENTPPL* 4%	*GDPDs* 1–2%, *ENTPPL* 1%	*GDPDs* 1–13%, *ENTPPL* 12%
Gene expression of low affinity P_i transporter (% of PtWT)	11%	27%	9%	67%
Gene expression of high affinity Na^+/P_i transporter (% of PtWT)	80%	106%	61%	200%

Percentage of enzymatic activity or gene expression of the mutant lines compared with the PtWT under P limitation. AP, alkaline phosphatase; CD, conserved domain; ENTPPL, phosphoethanolamine phospholyase; GDPD, glycerophosphoryldiester phosphodiesterase; nd, not detected; PM, point mutation. Presence (√) or absence (X) of conserved domains (CD2–CD4) in the encoded proteins. Gene names in italics.

microarray analyses and functional studies to be involved in the regulation of P_i acquisition and signalling under P limitation (Misson *et al.*, 2005; Medici *et al.*, 2015; Nagarajan *et al.*, 2016). The *PtPSR* promoter contains four PtPSR recognition motifs; thus, *PtPSR* expression may be regulated via a positive feedback loop. The reduced *PtPSR* gene expression level observed in *ptpsr* knockout mutants compared with PtWT support this view. However, *PtPSR* expression after 2 h of P limitation is more induced in the mutant lines compared with PtWT, suggesting that other TFs are involved in the initial induction of *PtPSR* expression during P limitation.

PtPSR mutations and their trends

The roles of the CDs of PtPSR are currently unknown. PtPSR CD1, CD2 and CD4 domains showed low similarities to CDs of other Myb-related TFs. CD3 contained a putative CC motif that was also seen in PHR1 and PSR1 and suggested to be involved in protein–protein interactions (Rayko *et al.*, 2010).

The analyses of *ptpsr* mutants indicate that PtPSR acts as a positive regulator of P-stress-inducible genes. Mutants defective in the Myb DNA-binding domain (*ptpsr 1_3.4* and *ptpsr 1_3.10*) were unable to induce these P-stress-regulated genes in a similar way as in the PtWT. Mutants where the CD2 domain was mutated (*ptpsr 2_2.2* and *ptpsr 2_2.3,* intact Myb DNA-binding domain) showed the same responses, but had a slightly reduced transcript induction for some target genes. The truncated PtPSR *1_3.4* and *1_3.10* protein, with loss of Myb DNA-binding domain function, was likely unable to enter the nucleus due to loss of the NLS and to bind to the promoter sequences, reducing activation of P-stress-inducible genes (Figs 3, 4). *ptpsr 1_3.4* had a severe reduction of growth rate and cP content compared with PtWT. All mutant lines, except *ptpsr 2_2.2*, showed lower cP content compared with PtWT under both −P and +P conditions, indicating that PtPSR is central in regulating P acquisition and scavenging both under normal and P-limited growth conditions (Figs 2, 4). In the *ptpsr 1_3.10* mutant line, which lacks the Myb DNA-binding domain but has intact CDs, most of the P-stress-inducible genes showed significant downregulation compared with wild-type, particularly at 24 and 48 h. The *ptpsr 2_2.3* mutant line has an intact Myb DNA-binding domain and CD1, but no CD2-4. Therefore, the PtPSR protein expressed in *ptpsr 2_2.3* can bind to P-stress-inducible genes but may not be able to interact with other key regulatory proteins. Together these findings indicate that other TFs, such as PtMyb1R_SHAQKYF8, which has a CD2 domain similar to PtPSR, may partially compensate for the lack of PtPSR. The *ptpsr 2_2.2* mutant line underlined the importance of CD2, although its function is still unknown. *ptpsr* 2_2.2 contains one complete copy of PtPSR with an L373F substitution in the CD2 domain and one copy with deletion of CD2-4, which led to reduction of AP activities and gene expression levels of P-stress response genes compared with PtWT under P limitation (Table 2; Fig. 5). The observed phenotypes of *ptpsr 2_2.2* indicated that the replacement of the conserved Leu with a Phe in the CD2 domain render it partially functional.

P acquisition and scavenging in *ptpsr* mutants

Our findings showed rapid transcriptional induction of proteins involved in P acquisition, such as P_i transporters, during −P cultivation. It was also seen that the complex regulation of P_i transporters is partly regulated by PtPSR. Reduced cP content in *ptpsr* mutant lines indicated low P_i supply, which supports a reduced intracellular and extracellular P_i transport. The reduced cP may also explain the low cell division rate for some of the mutants (Fig. 4a). The low affinity P_i co-transporter (*Phatr2_47667*) was induced by PtPSR under P stress. The postulated high affinity Na^+/P_i co-transporter (*Phatr2_40433*), related to human SLC34A1, showed variable differential regulation in the *ptpsr* mutant lines when compared with PtWT (Fig. 5i–j). In addition to PtPSR, other TFs are likely involved in regulation of P supply to the cell. Posttranscriptional regulation might also have a role, as shown for the low affinity Na^+-dependent P_i transporter DsSPT1 in the green alga *Dunaliella salina* (Moseley *et al.*, 2006; Li *et al.*, 2012).

We identified genes that were involved in AP activity regulated by PtPSR as an immediate response to low P_i concentration. *AP1*, *atyPhoA*, *PME* and *PhoD*, encoding putative phosphomonoester degrading enzymes, showed strongly reduced transcript induction under P limitation in *ptpsr* mutant lines (Fig. 5c–f). In line with these results, the *ptpsr* mutant lines displayed strongly reduced or no AP activity compared with PtWT (Table 1).

AP activity is an essential survival mechanism under low P_i in several algae (Dyhrman, 2016). *In silico* studies of AP abundance showed a wide range of algal species containing several AP sequences per genome with similarities to identified bacterial AP forms (Lin *et al.*, 2012). Our results show that *AP1, atyPhoA, PME* and *PhoD* are regulated by PtPSR and are potentially involved in eAP and cAP activities to degrade DOP under P limitation. From plants it is known that SPX domain-containing proteins control a set of processes for Pi homeostasis by fine tuning Pi transportation and Pi sensing (Secco *et al.*, 2012; Wang, 2004). Whether the SPX proteins in diatoms have related functions remains to be determined, but one gene encoding a SPX protein (*Phatr2 _47434*) may be regulated by PtPSR (Fig. S3).

Phospholipid degradation in *P. tricornutum ptpsr* mutants

Phospholipid degradation and betaine and TAG lipid synthesis are common P-stress responses in stramenopiles, and have been observed in *P. tricornutum* (Abida *et al.*, 2015; Cañavate *et al.*, 2017). The lipid profile of PtWT at 72 h of P limitation showed such common P-stress response as depletion of PL and increase in TAG, DAG and DGDG compared with PtWT grown under normal conditions (Fig. 6). However, the amount of DGTA, SQDG and MGDG were almost similar in −P and +P conditions, and could indicate that the lipid class remodelling was in its early stage in our study. Previous studies in *P. tricornutum* have shown a significant increase of P-free lipids classes such as SQDG and DGTA after 7–10 d of P limitation (Abida *et al.*, 2015; Cañavate *et al.*, 2017). These lipid classes are important for PL replacement and lipid droplet formation under nutrient stress (Lupette *et al.*, 2019). Our findings and previous studies have shown that the dynamics of lipid class changes in *P. tricornutum* under P stress are complex and have to be investigated in depth in future studies.

Our lipid analysis showed a higher level of PLs in the *ptpsr 1_3.4* and *1_3.10* mutants compared with PtWT under both P-replete and P-deplete conditions, indicating reduced PL degradation. PLs are phosphodiesters and are degraded extracellularly and intracellularly by GDPDs and lipases; not much information is known about these enzymes in stramenopiles, but several putative annotated lipases and GDPDs are transcriptionally induced under P stress (Dyhrman *et al.*, 2012; Yamaguchi *et al.*, 2014; Mühlroth *et al.*, 2017; Alipanah *et al.*, 2018). We have shown that *ENTPPL*, *GDPD2* and *GDPD1* are under the regulation of PtPSR (Fig. 5k–m). Extracellular GDPD2 may hydrolyse glycerolphosphodiesters into alcohols and *sn*-glycerol-3-phosphates that are further degraded to inorganic P, while GDPD1 (with unknown compartmentalisation) could be involved in intracellular PL degradation producing precursors for lipid accumulation. ENTPPL degrades the PL biosynthesis intermediate phosphoethanolamine to acetaldehyde and release P_i. Membrane-bound PL degradation has been suggested to be a part of the P-stress-inducible PL-recycling scheme and P-acquiring mechanism in stramenopiles (Mühlroth *et al.*, 2017; Alipanah *et al.*, 2018). Our findings support the hypothesis that PL degradation is a P-stress-inducible pathway that, in part, is regulated by PtPSR.

Previous studies have proposed that the PL-recycling scheme includes degradation of PLs, reduction of PL biosynthesis and synthesis of P-free lipid classes (Martin *et al.*, 2011; Cañavate *et al.*, 2017; Mühlroth *et al.*, 2017). Whether PtPSR is involved in the regulation of genes involved in TAG and betaine lipid accumulation is still unclear. However, *ptpsr 1_3.4, 1_3.10* and *2_2.3* showed significantly increased TAG, galactolipid and DGTA content compared with PtWT, indicating either PtPSR regulation of genes involved in the synthesis of betaine lipid, TAG and galactolipid classes or an indirect effect of reduced PL degradation. Bioinformatic data showed that the betaine lipid synthesis gene (*Phatr2_42872*) contains putative PtPSR recognition motifs (Fig. 2), supporting the idea that PtPSR may regulate P-free lipid biosynthesis in later time points.

In summary, we have identified one of the master regulators of specific P-stress responses in the diatom *P. tricornutum*, the Myb-related TF PtPSR and its binding motif. Phylogenetic analysis showed that PtPSR-like Mybs are evolutionarily conserved in stramenopiles. Functional studies indicate a central role for PtPSR in regulation of extra- and intracellular P acquisition, scavenging and phospholipid remodelling during acclimation to P-deplete conditions.

Acknowledgements

We sincerely thank Torfinn Sparstad, Trude Johansen and Kjersti Andresen for excellent technical assistance. We would also like to thank Valérie Gros for technical help with the lipid analysis. Financial support from the HAVBRUK2 programme (project 267474/E40) of the Norwegian Research Council to AMB is gratefully acknowledged. EM and JJ are supported by the French National Research Agency (Oceanomics ANR-11-BTBR-0008, GlycoAlps ANR-15-IDEX-02 and GRAL 10-LABX-0049).

Author contributions

PW, TB and AMB conceived the study. AKS and PW designed target sequences and developed the mutant lines. AKS, AM and LA performed experiments. JJ and EM performed lipid analyses. RK and AKS performed EMSA experiments. AMB provided resources. AKS and AM drafted the paper together with AMB, PW and TB. All authors reviewed and approved the manuscript.

ORCID

Tore Brembu https://orcid.org/0000-0001-5622-8037
Per Winge https://orcid.org/0000-0003-4380-1951

References

Abida H, Dolch L-J, Meï C, Villanova V, Conte M, Block MA, Finazzi G, Bastien O, Tirichine L, Bowler C *et al.* 2015. Membrane glycerolipid remodeling triggered by nitrogen and phosphorus starvation in *Phaeodactylum tricornutum*. *Plant Physiology* **167**: 118–136.

Alipanah L, Rohloff J, Winge P, Bones AM, Brembu T. 2015. Whole-cell response to nitrogen deprivation in the diatom *Phaeodactylum tricornutum*. *Journal of Experimental Botany* **66**: 6281–6296.

Alipanah L, Winge P, Rohloff J, Najafi J, Brembu T, Bones AM. 2018. An integrated analysis of phosphorus depletion responses in the diatom *Phaeodactylum tricornutum*. *PLoS ONE* **13**: e0193335.

Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research* **34**: W369–W373.

Baldwin DS. 2013. Organic phosphorus in the aquatic environment. *Environmental Chemistry* **10**: 439–454.

Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP *et al.* 2008. The Phaeodactylum genome reveals the dynamic nature and multi-lineage evolutionary history of diatom genomes. *Nature* **456**: 239–244.

Brembu T, Jørstad M, Winge P, Valle KC, Bones AM. 2011. Genome-wide profiling of responses to cadmium in the diatom *Phaeodactylum tricornutum*. *Environmental Science and Technology* **45**: 7640–7647.

Brembu T, Mühlroth A, Alipanah L, Bones AM. 2017. The effects of phosphorus limitation on carbon metabolism in diatoms. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **372**: 20160406.

Cañavate JP, Armada I, Hachero-Cruzado I. 2017. Aspects of phosphorus physiology associated with phosphate-induced polar lipid remodelling in marine microalgae. *Journal of Plant Physiology* **214**: 28–38.

Chiou T-J, Lin S-I. 2011. Signaling network in sensing phosphate availability in plants. *Annual Review of Plant Biology* **62**: 185–206.

Cruz de Carvalho MH, Sun HX, Bowler C, Chua NH. 2016. Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations. *New Phytologist* **210**: 497–510.

Donald KM, Scanlan DJ, Carr NG, Mann NH, Joint I. 1997. Comparative phosphorus nutrition of the marine cyanobacterium *Synechococcus WH7803* and the marine diatom *Thalassiosira weissflogii*. *Journal of Plankton Research* **19**: 1793–1813.

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. 2010. MYB transcription factors in Arabidopsis. *Trends in Plant Science* **15**: 573–581.

Dyhrman ST. 2016. Nutrients and their acquisition: phosphorus physiology in microalgae. In: Borowitzka M, Beardall J, Raven J, eds. *The physiology of microalgae*. Dordrecht, the Netherlands: Springer, 155–183.

Dyhrman ST, Jenkins BD, Rynearson TA, Saito MA, Mercier ML, Alexander H, Whitney LAP, Drzewianowski A, Bulygin VV, Bertrand EM *et al.* 2012. The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS ONE* **7**: e33768.

Elser JJ, Bracken MES, Cleland EE, Gruner DS, Harpole WS, Hillebrand H, Ngai JT, Seabloom EW, Shurin JB, Smith JE. 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters* **10**: 1135–1142.

Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C. 1999. Transformation of nonselectable reporter genes in marine diatoms. *Marine Biotechnology* **1**: 239–251.

Feng TY, Yang ZK, Zheng JW, Xie Y, Li DW, Murugan SB, Yang WD, Liu JS, Li HY. 2015. Examination of metabolic responses to phosphorus limitation via proteomic analyses in the marine diatom *Phaeodactylum tricornutum*. *Scientific Reports* **5**: 10373.

Flynn KJ, Öpik H, Syrett PJ. 1986. Localization of the alkaline phosphatase and 5′-nucleotidase activities of the diatom *Phaeodactylum tricornutum*. *Microbiology* **132**: 289–298.

Hansen HP, Koroleff F. 1999. Determination of nutrients. In: Grasshoff K, Kremling K, Ehrhardt M, eds. *Methods of seawater analysis, 3rd edn*. Weinheim, Germany: Wiley-VCH, 159–228.

Hu J, Wang D, Li J, Jing G, Ning K, Xu J. 2014. Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae Nannochloropsis. *Scientific Reports* **4**: 5454.

Jouhet J, Lupette J, Clerc O, Magneschi L, Bedhomme M, Collin S, Roy S, Maréchal E, Rébeillé F. 2017. LC-MS/MS versus TLC plus GC methods: consistency of glycerolipid and fatty acid profiles in microalgae and higher plant cells and effect of a nitrogen starvation. *PLoS ONE* **12**: e0182423.

Jyoti A, Kaushik S, Srivastava VK, Datta M, Kumar S, Yugandhar P, Kothari SL, Rai V, Jain A. 2019. The potential application of genome editing by using CRISPR/Cas9, and its engineered and ortholog variants for studying the transcription factors involved in the maintenance of phosphate homeostasis in model plants. *Seminars in Cell & Developmental Biology*. doi: 10.1016/j.semcdb.2019.03.010.

Karas BJ, Diner RE, Lefebvre SC, McQuaid J, Phillips APR, Noddings CM, Brunson JK, Valas RE, Deerinck TJ, Jablanovic J *et al.* 2015. Designer diatom episomes delivered by bacterial conjugation. *Nature Communications* **6**: 6925.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**: 1870–1874.

Kong XY, Kissen R, Bones AM. 2012. Characterization of recombinant nitrile-specifier proteins (NSPs) of *Arabidopsis thaliana*: dependency on Fe (II) ions and the effect of glucosinolate substrate and reaction conditions. *Phytochemistry* **84**: 7–17.

Li SH, Xia BB, Zhang C, Cao J, Bai LH. 2012. Cloning and characterization of a phosphate transporter gene in *Dunaliella salina*. *Journal of Basic Microbiology* **52**: 429–436.

Li T, Guo C, Zhang Y, Wang C, Lin X, Lin S. 2018. Identification and expression analysis of an atypical alkaline phosphatase in *Emiliania huxleyi*. *Frontiers in Microbiology* **9**: 2156.

Lin X, Wang L, Shi X, Lin S. 2015. Rapidly diverging evolution of an atypical alkaline phosphatase (PhoA(aty)) in marine phytoplankton: insights from dinoflagellate alkaline phosphatases. *Frontiers in Microbiology* **6**: 868.

Lin X, Zhang H, Cui Y, Lin S. 2012. High sequence variability, diverse subcellular localizations, and ecological implications of alkaline phosphatase in dinoflagellates and other eukaryotic phytoplankton. *Frontiers in Microbiology* **3**: 235.

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. *Science* **252**: 1162–1164.

Lupette J, Jaussaud A, Seddiki K, Morabito C, Brugière S, Schaller H, Kuntz M, Putaux J-L, Jouneau P-H, Rébeillé F. 2019. The architecture of lipid droplets in the diatom *Phaeodactylum tricornutum*. *Algal Research* **38**: 101415.

Martin P, Van Mooy BA, Heithoff A, Dyhrman ST. 2011. Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom *Thalassiosira pseudonana*. *ISME Journal* **5**: 1057–1060.

Matthijs M, Fabris M, Broos S, Vyverman W, Goossens A. 2016. Profiling of the early nitrogen stress response in the diatom *Phaeodactylum tricornutum* reveals a novel family of RING-domain transcription factors. *Plant Physiology* **170**: 489–498.

Matthijs M, Fabris M, Obata T, Foubert I, Franco-Zorrilla JM, Solano R, Fernie AR, Vyverman W, Goossens A. 2017. The transcription factor bZIP14 regulates the TCA cycle in the diatom *Phaeodactylum tricornutum*. *EMBO Journal* **36**: 1559–1576.

McDonnell AV, Jiang T, Keating AE, Berger B. 2006. Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* **22**: 356–358.

Medici A, Marshall-Colon A, Ronzier E, Szponarski W, Wang R, Gojon A, Crawford NM, Ruffel S, Coruzzi GM, Krouk G. 2015. AtNIGT1/HRS1 integrates nitrate and phosphate signals at the Arabidopsis root tip. *Nature Communications* **6**: 6274.

Misson J, Raghothama KG, Jain A, Jouhet J, Block MA, Bligny R, Ortet P, Creff A, Somerville S, Rolland N *et al.* 2005. A genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined plant

responses to phosphate deprivation. *Proceedings of the National Academy of Sciences, USA* **102**: 11934–11939.

Moseley JL, Chang CW, Grossman AR. 2006. Genome-based approaches to understanding phosphorus deprivation responses and PSR1 control in *Chlamydomonas reinhardtii*. *Eukaryotic Cell* **5**: 26–44.

Mühlroth A, Li K, Røkke G, Winge P, Olsen Y, Hohmann-Marriott MF, Vadstein O, Bones AM. 2013. Pathways of lipid metabolism in marine algae, co-expression network, bottlenecks and candidate genes for enhanced production of EPA and DHA in species of chromista. *Marine Drugs* **11**: 4662–4697.

Mühlroth A, Winge P, El Assimi A, Jouhet J, Maréchal E, Hohmann-Marriott MF, Vadstein O, Bones AM. 2017. Mechanisms of phosphorus acquisition and lipid class remodelling under P limitation in a marine microalga. *Plant Physiology* **175**: 1543–1559.

Nagarajan VK, Satheesh V, Poling MD, Raghothama KG, Jain A. 2016. Arabidopsis MYB-Related HHO2 exerts a regulatory influence on a subset of root traits and genes governing phosphate homeostasis. *Plant and Cell Physiology* **57**: 1142–1152.

Norwegian Standard NS-EN ISO 6878. 2004. *Water quality – determination of phosphorus – ammonium molybdate spectrometric method*. Oslo, Norway: Norwegian Standards Association.

Nakata M, Ohme-Takagi M. 2014. Electrophoresis mobility shift assay. *Bio-Protocol* **4**: e1099.

Nitta KR, Jolma A, Yin Y, Morgunova E, Kivioja T, Akhtar J, Hens K, Toivonen J, Deplancke B, Furlong EEM. 2015. Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. *eLife* **4**: e04837.

Nymark M, Sharma A, Hafskjold M, Sparstad T, Bones A, Winge P. 2017. CRISPR/Cas9 gene editing in the marine diatom *Phaeodactylum tricornutum*. *Bio-Protocol* **7**: e2442.

Nymark M, Sharma AK, Sparstad T, Bones AM, Winge P. 2016. A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports* **6**: 24951.

Nymark M, Valle KC, Brembu T, Hancke K, Winge P, Andresen K, Johnsen G, Bones AM. 2009. An integrated analysis of molecular acclimation to high light in the marine diatom *Phaeodactylum tricornutum*. *PLoS ONE* **4**: e7743.

Quisel JD, Wykoff DD, Grossman AR. 1996. Biochemical characterization of the extracellular phosphatases produced by phosphorus-deprived *Chlamydomonas reinhardtii*. *Plant Physiology* **111**: 839–848.

Ramakers C, Ruijter JM, Lekanne Deprez RH, Moorman AFM. 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**: 62–66.

Rayko E, Maumus F, Maheswari U, Jabbari K, Bowler C. 2010. Transcription factor families inferred from genome sequences of photosynthetic stramenopiles. *New Phytologist* **188**: 52–66.

Rubio V, Linhares F, Solano R, Martín AC, Iglesias J, Leyva A, Paz-Ares J. 2001. A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes & Development* **15**: 2122–2133.

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB, Moorman AFM. 2009. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**: e45.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406–425.

Sapriel G, Quinet M, Heijde M, Jourdren L, Tanty V, Luo G, Le Crom S, Lopez PJ. 2009. Genome-wide transcriptome analyses of silicon metabolism in *Phaeodactylum tricornutum* reveal the multilevel regulation of silicic acid transporters. *PLoS ONE* **4**: e7458.

Sarthou G, Timmermans KR, Blain S, Tréguer P. 2005. Growth physiology and fate of diatoms in the ocean: a review. *Journal of Sea Research* **53**: 25–42.

Secco D, Wang C, Arpat BA, Wang Z, Poirier Y, Tyerman SD, Wu P, Shou H, Whelan J. 2012. The emerging importance of the SPX domain-containing proteins in phosphate homeostasis. *New Phytologist* **193**: 842–851.

Sharma AK, Nymark M, Sparstad T, Bones AM, Winge P. 2018. Transgene-free genome editing in marine algae by bacterial conjugation – comparison with biolistic CRISPR/Cas9 transformation. *Scientific Reports* **8**: 14401.

Shimogawara K, Wykoff DD, Usuda H, Grossman AR. 1999. *Chlamydomonas reinhardtii* mutants abnormal in their responses to phosphorus deprivation. *Plant Physiology* **120**: 685–694.

Slattery SS, Diamond A, Wang H, Therrien JA, Lant JT, Jazey T, Lee K, Klassen Z, Desgagné-Penix I, Karas BJ *et al.* 2018. An expanded plasmid-based genetic toolbox enables Cas9 genome editing and stable maintenance of synthetic pathways in *Phaeodactylum tricornutum*. *ACS Synthetic Biology* **7**: 328–338.

Sperschneider J, Catanzariti AM, Deboer K, Petre B, Gardiner DM, Singh KB, Dodds PN, Taylor JM. 2017. LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific Reports* **7**: 44598.

Theodorou ME, Elrifi IR, Turpin DH, Plaxton WC. 1991. Effects of phosphorus limitation on respiratory metabolism in the green alga *Selenastrum minutum*. *Plant Physiology* **95**: 1089–1095.

Thiriet-Rupert S, Carrier G, Chénais B, Trottier C, Bougaran G, Cadoret JP, Schoefs B, Saint-Jean B. 2016. Transcription factors in microalgae: genome-wide prediction and comparative analysis. *BMC Genomics* **17**: 282.

Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblížek M, Lomas MW, Mincer TJ, Moore LR, Moutin T *et al.* 2009. Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* **458**: 69–72.

Vieler A, Wu G, Tsai C-H, Bullard B, Cornish AJ, Harvey C, Reca I-B, Thornburg C, Achawanantakun R, Buehl CJ. 2012. Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genetics* **8**: e1003064.

Wang Y. 2004. Structure and expression profile of the Arabidopsis PHO1 gene family indicates a broad role in inorganic phosphate homeostasis. *Plant Physiology* **135**: 400–411.

Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K *et al.* 2014. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**: 1431–1443.

Wykoff DD, Grossman AR, Weeks DP, Usuda H, Shimogawara K. 1999. Psr1, a nuclear localized protein that regulates phosphorus metabolism in Chlamydomonas. *Proceedings of the National Academy of Sciences, USA* **96**: 15336–15341.

Xu J. 2006. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Molecular Ecology* **15**: 1713–1731.

Yamaguchi H, Arisaka H, Otsuka N, Tomaru Y. 2014. Utilization of phosphate diesters by phosphodiesterase-producing marine diatoms. *Journal of Plankton Research* **36**: 281–285.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Transcriptional induction of two PtPSR-like Myb TFs by P limitation in Nannochloropsis oceanica.

Fig. S2 Phylogenetic relationship of diatom SHAQKYF-like motif-containing Mybs.

Fig. S3 Regulated genes under stress conditions containing identified PtPSR recognition sites in promoter region in Phaeodactylum tricornutum.

Fig. S4 Comparison of the PtPSR recognition motif to binding sites of Arabidopsis P-stress inducible Myb-related TFs.

Fig. S5 EMSA analysis of PtPSR binding to the specific motif.

Fig. S6 A schematic description of the deletions and rearrangements in ptpsr PAM1 mutants deduced from RNA-seq data.

Fig. S7 Relative mRNA expression of P-stress induced genes in PtWT and ptpsr mutants under phosphate-replete conditions.

Fig. S8 Total lipid content of PtWT and ptpsr mutants in +P and −P conditions at 72 h.

Table S1 Primers for gDNA amplification of mutants and High-resolution melting (HRM) primers.

Table S2 Primer pairs used for qRT-PCR.

Table S3 Off-target locus for PAM1 and PAM2. Mismatch bases are underlined.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.