The Plant Cell, Vol. 28: 616–628, March 2016, www.plantcell.org

BREAKTHROUGH REPORT

Diatom Phytochromes Reveal the Existence of Far-Red-Light-Based Sensing in the Ocean[OPEN]

Antonio Emidio Fortunato,[a,1] Marianne Jaubert,[a,1] Gen Enomoto,[b] Jean-Pierre Bouly,[a] Raffaella Raniello,[c] Michael Thaler,[a] Shruti Malviya,[d,2] Juliana Silva Bernardes,[a] Fabrice Rappaport,[e,3] Bernard Gentili,[f] Marie J.J. Huysman,[g,h,i] Alessandra Carbone,[a,j] Chris Bowler,[d] Maurizio Ribera d'Alcalà,[c,4] Masahiko Ikeuchi,[b] and Angela Falciatore[a,4]

[a] Sorbonne Universités, UPMC, Institut de Biologie Paris-Seine, CNRS, Laboratoire de Biologie Computationnelle et Quantitative UMR 7238, 75006 Paris, France
[b] Department of Life Sciences (Biology), Graduate School of Arts and Sciences, The University of Tokyo, Meguro, Tokyo 153-8902, Japan
[c] Stazione Zoologica Anton Dohrn, 80121 Naples, Italy
[d] Ecole Normale Supérieure, PSL Research University, Institut de Biologie de l'Ecole Normale Supérieure, CNRS UMR 8197, INSERM U1024, F-75005 Paris, France
[e] Institut de Biologie Physico-Chimique, UMR 7141 CNRS-UPMC, 75005 Paris, France
[f] Sorbonne Universités, UPMC Univ-Paris 6, CNRS, UMR 7093, Laboratoire d'Océanologie de Villefranche, F-06230 Villefranche/mer, France
[g] Protistology and Aquatic Ecology, Department of Biology, Ghent University, B-9000 Gent, Belgium
[h] Department of Plant Systems Biology, VIB, B-9052 Gent, Belgium
[i] Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium
[j] Institut Universitaire de France, 75005 Paris, France

ORCID IDs: 0000-0002-3173-5742 (A.E.F.); 0000-0001-5419-5796 (M.J.); 0000-0002-3680-2043 (R.R.); 0000-0002-5492-9961 (M.R.d.A.)

The absorption of visible light in aquatic environments has led to the common assumption that aquatic organisms sense and adapt to penetrative blue/green light wavelengths but show little or no response to the more attenuated red/far-red wavelengths. Here, we show that two marine diatom species, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*, possess a bona fide red/far-red light sensing phytochrome (DPH) that uses biliverdin as a chromophore and displays accentuated red-shifted absorbance peaks compared with other characterized plant and algal phytochromes. Exposure to both red and far-red light causes changes in gene expression in *P. tricornutum*, and the responses to far-red light disappear in *DPH* knockout cells, demonstrating that *P. tricornutum* DPH mediates far-red light signaling. The identification of *DPH* genes in diverse diatom species widely distributed along the water column further emphasizes the ecological significance of far-red light sensing, raising questions about the sources of far-red light. Our analyses indicate that, although far-red wavelengths from sunlight are only detectable at the ocean surface, chlorophyll fluorescence and Raman scattering can generate red/far-red photons in deeper layers. This study opens up novel perspectives on phytochrome-mediated far-red light signaling in the ocean and on the light sensing and adaptive capabilities of marine phototrophs.

INTRODUCTION

Phytochromes (PHY) represent a major class of photoreceptors first discovered in terrestrial plants as red and far-red (R-FR) light sensors controlling key adaptive and developmental processes (Franklin and Quail, 2010). Subsequently, these photoreceptors have been identified in several prokaryotic and eukaryotic organisms, from phototrophs to heterotrophs (Rodriguez-Romero et al., 2010; Auldridge and Forest, 2011).

PHY spectral characteristics are determined by covalent interaction with a heme-derived linear tetrapyrrole chromophore in the N-terminal photosensory module (PSM) of the apoprotein. The canonical PSM typically consists of conserved consecutive PAS (Period/Arnt/Single-minded), GAF (cGMP phosphodiesterase/adenylyl cyclase/FhlA), and PHY (Phy specific) domains, which undergo reversible conformational changes upon irradiation, from a Pr to a Pfr form and vice versa. The photoswitchable PSM controls the activity of the C-terminal output module (OPM) involved in signal propagation. Compared with the PSM, larger diversification is found in the OPM, which can contain PAS, ATPase (H), histidine kinase (KD), diguanylate cyclase (GGDEF),

[1] These authors contributed equally to this work.
[2] Current address: CSIR-National Institute of Oceanography, Dona Paula, Goa 403 004, India.
[3] Deceased.
[4] Address correspondence to maurizio@szn.it or angela.falciatore@upmc.fr.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Angela Falciatore (angela.falciatore@upmc.fr).
[OPEN]Articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.15.00928

phosphodiesterase (EAL), and response receiver (REC) domains in different numbers and combinations, suggesting divergence in the signal propagation cascade (Rockwell et al., 2006).

Aquatic environments are less transparent to R-FR irradiation than the atmosphere (Mobley, 1994). This makes the recent discovery of PHYs in a growing number of aquatic organisms a puzzling issue and opens novel perspectives on the possible aquatic origin of PHYs. Found in numerous algal lineages that originated after cyanobacterial endosymbiosis (Streptophyta, Prasinophyta, and Glaucophyta), PHYs have also been identified in some algae derived from secondary endosymbiosis (Cryptophyta and Heterokonts) (Falciatore and Bowler, 2005; Keeling, 2013; Duanmu et al., 2014; Rockwell et al., 2014b; Li et al., 2015). The characterization of representative PHYs from multicellular and unicellular marine phototrophs (phytoplankton) revealed a wide spectral variation among them. Several of these PHYs exhibit a R-FR photocycle, while others sense orange, green, and even blue light wavelengths (Duanmu et al., 2014; Rockwell et al., 2014b). Similarly to proteorhodopsin-based sensing (Man et al., 2003), this variability has been interpreted to be a result of selective pressure to tune the photoreceptor absorbance to the spectral characteristics of light in marine environments.

In this study, we extended analysis of the phytochrome photoreceptor class to marine diatoms, one of the most prominent and diversified groups of photosynthetic secondary endosymbionts in the ocean (Field et al., 1998; Kooistra et al., 2007; Moustafa et al., 2009). Able to actively colonize a wide range of ecological niches, including subsurface layers and turbulent environments (Margalef, 1978), diatoms experience a wide range of light conditions. Because of water mixing or sinking, diatoms can also be displaced to depths where light is limiting for growth (Smetacek, 1985). They can remain there for long periods of time or can be carried back to the photic zone, where a fast readjustment of light-dependent processes is likely to be important.

Action spectra specificity for expression of a gene encoding a Fucoxanthin chlorophyll *a/c* binding protein first suggested the existence of phytochrome-based R and FR light sensing in the diatom *Thalassiosira weissflogii* (Leblanc et al., 1999). Subsequent genome sequencing of the centric diatom *Thalassiosira pseudonana* (Armbrust et al., 2004) and the pennate diatom *Phaeodactylum tricornutum* (Bowler et al., 2008) confirmed the presence of diatom PHY-like sequences, DPH, previously reported as Dph (Montsant et al., 2007). Here, we characterize DPHs from these two model species (Tp-DPH and Pt-DPH) and investigate the functional responses of Pt-DPH to different light conditions, revealing that DPHs can function as an R-FR light-sensitive molecule controlling FR light responses. We further identified *DPH* in the genome and transcriptome of diverse diatom species, leading us to propose novel hypotheses regarding the signal(s) to which DPH might respond in marine environments.

RESULTS

P. tricornutum and *T. pseudonana* DPHs Are R-FR Light-Reversible Chromoproteins

Domain prediction analysis of *P. tricornutum* and *T. pseudonana* DPHs revealed an overall protein architecture very close to the bacterial two-component histidine kinase PHYs (BphPs), with a conserved OPM organization constituted by the H, KD, and REC sensor domains (Figure 1A). Notably, canonical GAF and PHY domains were identified in the PSM of both protein sequences, but the N-terminal PAS domain sequence was predicted only in Tp-DPH (Figure 1A; Supplemental Data Set 1). In spite of this difference, both DPHs possess a Cys residue in the N-terminal extremity, at the same position as the Cys residue that mediates chromophore binding in BphPs (Figure 1A; Supplemental Figure 1) (Lamparter et al., 2004; Wagner et al., 2005). This conservation suggests that Biliverdin IXα (BV) might constitute the heme-derived chromophore of DPHs.

To assess their biochemical properties, the Pt-DPH full-length and PSM, named Pt-DPH-FL and Pt-DPH-PSM, respectively, as well as the Tp-DPH-PSM were expressed in *Escherichia coli* together with a *Synechocystis PCC6803* heme oxygenase (HO1) producing BV. Purified recombinant DPHs were obtained as soluble holoproteins covalently bound to BV, as confirmed by zinc staining analysis (Figure 1B, insets). Pt-DPH-PSM recapitulated the absorption properties of Pt-DPH-FL. For both of them, illumination with 757-nm LEDs induces within 10 s the photoconversion of Pt-DPH-PSM as well as Pt-DPH-FL into the typical Pr form, maximally absorbing at 700 nm. Illumination with 639-nm LEDs produces a Pfr-enriched mixture peaking at 700 and 750 nm, and corresponding to the remaining Pr form and the photoconverted Pfr form, respectively (Figure 1B). Tp-DPH-PSM exhibited similar spectral properties, with absorption maxima of the Pr form at 686 nm and Pfr at 764 nm (Figure 1B). Notably, no dark relaxation was observed using Pt-DPH-FL, and only ~15% of Tp-DPH-PSM Pfr reverted to Pr over 24 h (Supplemental Figures 2A and 2B).

The identification of several heme oxygenase encoding genes in the *T. pseudonana* and *P. tricornutum* genomes (Supplemental Data Set 2) makes BV a strong DPH chromophore candidate. Putative bilin reductase encoding genes (*pebA* and *pebB-like* genes) that could catalyze a further reduction of BV were also found (Supplemental Data Set 2). Nevertheless, they putatively encode enzymes more closely related to phycoerythrobilin (PEB) synthesis than to ferredoxin-dependent bilin reductases, which convert BV to phytochromobilin (PΦB) in cyanobacteria and plants, respectively (Rockwell et al., 2014a). We therefore tested the possible binding of different chromophores by expressing Pt-DPH-PSM with *ho1* in *E. coli* along with the *Arabidopsis thaliana HY2* or *Synechococcus pebA-pebB* bilin reductase genes producing PΦB and PEB, respectively. The native absorption spectra (Supplemental Figures 2C and 2D), difference spectra, and acidic urea-denatured difference spectra (Figures 1C and 1D) of the different purified Pt-DPH-PSM were almost identical to BV-bound Pt-DPH-PSM. Together, these analyses reveal that both DPHs are bona fide BV binding phytochromes, undergoing a R-FR light reversible photoconversion.

Pt-*DPH* Expression Is Light Regulated

To explore DPH function, we first monitored *DPH* expression by RT-qPCR in *P. tricornutum* cells grown under 12-h-ligh/12-h-dark cycles or maintained in prolonged darkness for 48 h. This revealed that Pt-*DPH* transcripts exhibited a pronounced diurnal rhythmic expression pattern, with an accumulation at the end of the light period, as observed for Tp-*DPH* (Ashworth et al., 2013) (Figure 2A). In

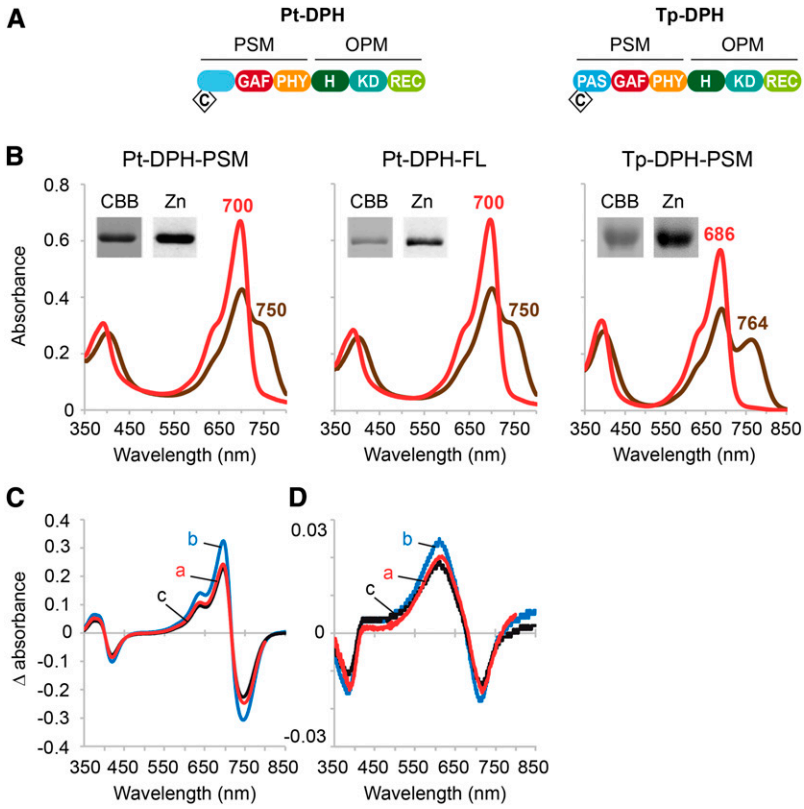

Figure 1. Biochemical Analysis of Pt-DPH and Tp-DPH.

(A) Schematic domain organization of Pt-DPH and Tp-DPH proteins. PAS, Period-Arnt-Sim domain; GAF, cyclic di-GMP phosophodiesterase/adenyl cyclase/Fhla domain; H, Histidine Kinase A Domain; KD, HATPase c domain; REC, Response receiver domain; C, cysteine residue putatively binding biliverdin IXα (BV). In Pt-DPH, the PAS domain was not predicted.

(B) Absorption spectra of recombinant Pt-DPH-PSM, Pt-DPH-FL, and Tp-DPH-PSM purified from BV-producing *E. coli* after irradiation with LED emitting light at 757 and 765 nm (red line) or 639 and 690 nm (brown line) for Pt-DPH and Tp-DPH, respectively. In the inset, corresponding purified proteins visualized by Coomassie blue staining (CBB) or zinc-induced fluorescence (Zn).

(C) and **(D)** Different absorption spectra of native **(C)** and acidic urea denatured **(D)** recombinant Pt-DPH-PSM purified from *E. coli* expressing BV (a), together with phytochromobilin (PØB) (b) or PEB (c) biosynthesis enzymes, and irradiated as above.

contrast, in prolonged darkness, Pt-*DPH* mRNA levels remained at a constant low level (Figure 2A). To determine whether the changes in transcript levels affect Pt-DPH protein content, we generated Pt-DPH-HA transgenic lines in which HA-tagged DPH is under the control of its own promoter and terminator sequences (Pt-DPH_{Pro}:Pt-*DPH-HA*) and that recapitulate endogenous Pt-*DPH* mRNA expression patterns (Supplemental Figures 3A and 3B). In these lines, Pt-DPH-HA protein levels gradually decrease during the light period and reaccumulate during the dark period (Figure 2B), revealing that mRNA and protein levels display opposite trends under diurnal conditions. Interestingly, when the dark treatment was prolonged, Pt-DPH-HA levels also progressively decreased (Figure 2B). This expression pattern, with Pt-DPH-HA maximal accumulation at the end of the dark period and possible subsequent degradation, points to a potential role of Pt-DPH during dark to light transitions.

R-FR-Light-Mediated Gene Expression in *P. tricornutum*

To explore the potential functional impact of Pt-DPH in response to R and FR light, we compared *P. tricornutum* genome-wide gene expression patterns upon shifting cells from dark to R (660 nm) or FR (765 nm) light (5 $\mu mol \cdot m^{-2} \cdot s^{-1}$) or maintaining them in darkness for 30 min. The reference time point (T0) collected just before exposure to light therefore corresponds to high Pt-DPH accumulation (Figure 2B). Transcripts from light- and dark-treated cells were compared with those from T0 (Figure 3A). Hierarchical clustering identified 282 genes significantly responding to 30 min of R and/or FR light exposure, which were grouped into three main classes: 80 FR-responsive genes (Class 1), 177 R-responsive genes (Class 2), and 22 genes responsive to both R and FR light (Class 3) (Figure 3B). Most of the Class 1 genes encode proteins of unknown function (55%) or are involved in transcription and signaling (25%). A high proportion of unknown genes were also found in Class 2 (48%), together with putative proteins implicated in nitrogen or carbon metabolism (36%). Interestingly, the large majority of Class 3 genes (75%) are involved in tetrapyrrole biosynthesis pathways (Supplemental Data Set 3).

To address the signaling pathways triggering R or FR light responses, we further analyzed the responsiveness of

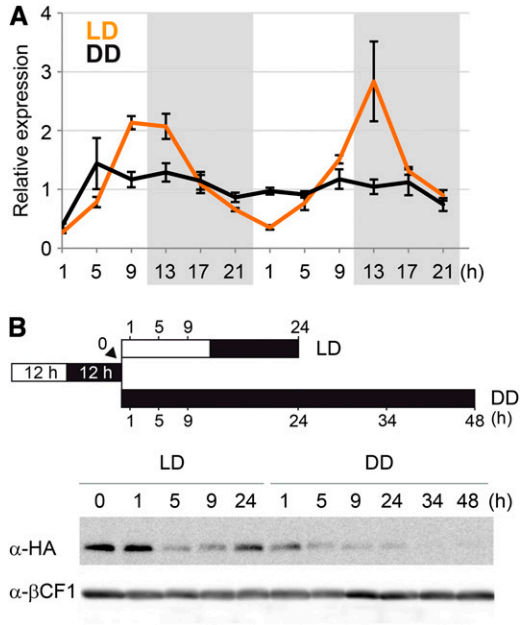

Figure 2. Pt-DPH Expression in Different Light Conditions.

(A) RT-qPCR analysis of Pt-*DPH* mRNAs in *P. tricornutum* cells grown in diurnal 12-h-light/12-h-dark cycles (LD) and in continuous dark (DD) for 2 d. *RPS* and *TBP* were used for normalization (n = 3; black bars = ±SD).
(B) Immunoblot analysis of Pt-DPH-HA protein in LD and DD in a transgenic line expressing Pt-DPH_{Pro}:Pt-*DPH*-HA, using the antihemagglutinin antibody (α-HA) and the α-βCF1 antibody, as loading control. The sampling time points are indicated.

representative Class 1 and Class 2 genes to a wider range of R and FR wavelengths (660, 688, 732, 740, and 765 nm, 5 $\mu mol \cdot m^{-2} \cdot s^{-1}$). We tested genes putatively encoding a transcription factor (HSF4.6a) and a molecular chaperone (DNAJ) among Class 1 genes, and an acetylornithine transaminase (ACOAT) and a glutamine synthase III (GLNAIII) among Class 2. RT-qPCR analyses revealed a wavelength-dependent induction for all tested genes. Notably, an action spectrum of the Class 1 genes shows a progressive induction under longer wavelengths (max of induction at 765 nm), closely matching the absolute absorption spectrum of Pt-DPH Pfr form (Figure 3C). Their expression levels progressively decreased at shorter wavelengths, and no induction at all was seen at 688 nm, which is the Qy band of chlorophyll absorption. Treatment with the photosystem II inhibitor DCMU prior to FR light exposure had no effect on the expression of these genes (Figure 3D), which further rules out any regulation of their expression by feedback from photosynthesis. Any photosystem I-mediated regulation can also be excluded given the absence of absorption of photosystem I complexes above 710 nm (Veith and Büchel, 2007; Ikeda et al., 2008). On the contrary, the Class 2 genes are maximally induced by the shorter wavelengths and their expression profiles do not match the Pt-DPH spectra of any of the two forms (Figure 3E). However, the induction globally overlaps with the wavelength range seen in the absorption spectrum from whole cells (Figure 3E) and is dramatically impaired when cells are treated with DCMU (Figure 3F).

These analyses reveal the existence of at least two signaling pathways mediating R and FR gene expression responses, the latter one being potentially linked to Pt-DPH activity.

Pt-DPH Is Required for FR Light Signaling

We used TALE nuclease genome editing to generate several independent Pt-*DPH* knockout lines (Daboussi et al., 2014). Two different biallelic mutants (KO1 and KO2) and a transgenic line bearing the wild-type Pt-*DPH* sequence for both alleles (transformation control) were used to assess the involvement of Pt-DPH in the regulation of R-FR-light-mediated gene expression (Figure 4A; Supplemental Figure 3C). We monitored the expression of the same genes tested above, as well as two more genes encoding putative histidine kinase proteins, HISK and GAF-HISK for Class 1 and two genes putatively encoding a chromate transporter (CHRT) and zeaxanthin epoxidase (ZEP1) for Class 2 (Figures 4B and 4C). This analysis showed that the FR-mediated induction of the Class 1 genes was maintained in both wild-type and control lines but was completely abolished in the two Pt-*DPH* loss-of-function mutants (Figure 4B). Expression of Class 2 genes by red light was not affected in the KO lines (Figure 4C).

Given that Pt-DPH can undergo a R-FR photocycle, we further tested whether it could act as a light-regulated kinase in vitro, as already observed for several phytochrome photoreceptors (Yeh and Lagarias, 1998; Karniol and Vierstra, 2003; Giraud et al., 2005). Incubation of recombinant holo-Pt-DPH-FL in the presence of [ϒ-^{32}P]ATP revealed rapid protein autophosphorylation upon FR light illumination, whereas only a basal signal could be detected from R-light-irradiated Pt-DPH-FL (Figure 4D).

Together, these data demonstrate that FR light regulation of gene expression in *P. tricornutum* involves Pt-DPH, possibly through the activation of a phosphorylation cascade.

DPH Is Present in Diverse Diatom Species and Expressed at Different Ocean Depths

To address the potential extent of R-FR light sensing in the diverse diatom group, we investigated the presence of *DPH*-like sequences in a total of 89 diatom species from available genomic and transcriptomic data (Supplemental Data Set 4). We identified 25 *DPH* sequences in 19 centric and pennate diatoms. While only one *DPH* was identified in the centric species, up to four genes were found in some pennate species (e.g., *Seminavis robusta* and *Amphora coffeaeformis*). Extended phylogenetic analyses grouped these sequences in the Heterokont PHY clade, highlighting distinct subgroups of DPHs in pennate and centric diatoms, and of PHYs in brown algae (*Ectocarpus siliculosus*) and brown algal viruses (Figure 4A). As previously shown, the Heterokont PHYs appear to have a distinct origin from PHYs identified in Archeaplastida and Cryptophyta (Figure 4A) (Duanmu et al., 2014; Li et al., 2015), and they form a sister clade to the fungal phytochromes (Fphs). Heterokont and fungal PHYs cluster apart from the bacterial BphPs, despite sharing common features with an OPM made up of the H, KD, and REC domains and a conserved N-terminal Cys residue at the position shown to bind BV in several BphPs (Figure 1; Supplemental Figure 1).

An in-depth comparative analysis of DPH sequences further revealed interesting novel features. Three DPHs found in

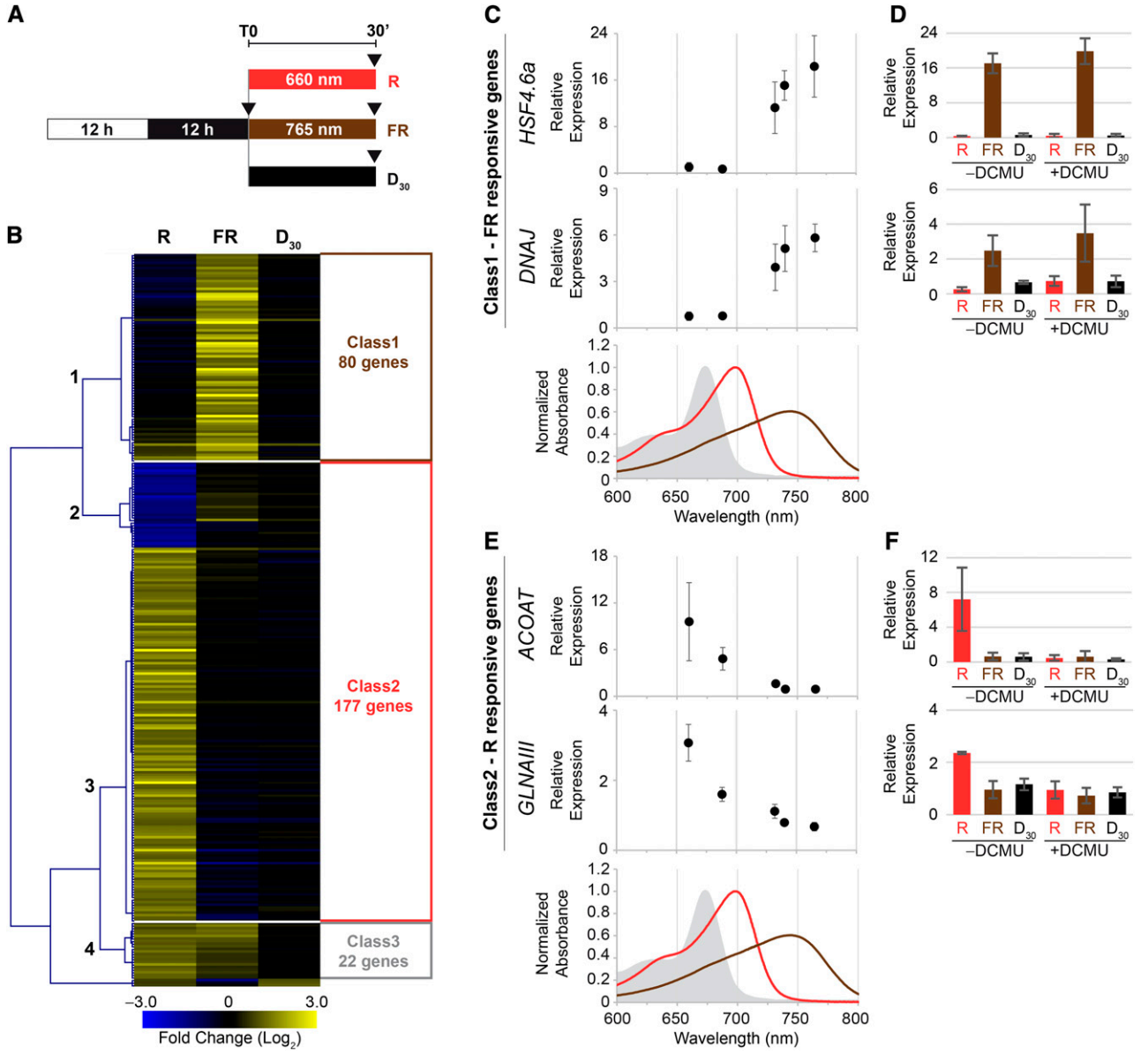

Figure 3. R-FR-Light-Mediated Gene Expression Changes.

(A) Experimental setup used for microarray analysis: *P. tricornutum* cells were collected at the end of the 12-h-dark period (T0) and following 30 min of irradiation with 660 nm (R) or 765 nm (FR) LED lights (5 $\mu mol \cdot m^{-2} \cdot s^{-1}$) or an additional dark treatment (D_{30}).
(B) Hierarchical clustering analysis of the 282 genes (Supplemental Data Set 3) obtained from the microarray analysis from R, FR, and D_{30} conditions, with respect to the T0. The dendrogram on the left defined four major groups assigned to three classes of light regulated genes based on the expression profiles.
(C) and **(E)** Wavelength dependence of Class 1 (*HSF4.6a* and *DNAJ*) **(C)** and Class 2 (*ACOAT* and *GLNAIII*) **(E)** gene expression induction performed by RT-qPCR on cells irradiated for 30 min with different LED lights (660, 688, 732, 740, or 765 nm, 5 $\mu mol \cdot m^{-2} \cdot s^{-1}$). The graphs below the RT-qPCR data represent the absorption spectra of the pure Pt-DPH Pr (red line) and Pfr (brown line) forms, and the whole-cell absorption spectra of *P. tricornutum* (filled gray).
(D) and **(F)** RT-qPCR analyses of the *HSF4.6a* and *DNAJ* genes **(D)** and *ACOAT* and *GLNAIII* genes **(F)**, performed on wild-type cells irradiated as described in **(A)**, in presence or absence of the photosystem II inhibitor DCMU. *RPS* was used as normalizer gene and values were relativized to T0 ($n = 3$; black bars = ±SD).

A. coffeaeformis showed an additional Cys residue in the GAF domain, at a similar position to the chromophore binding site in plant and cyanobacterial (Cphs) PHYs as well as in *E. siliculosus* PHYs (Supplemental Figure 1) (Rockwell et al., 2014b). Moreover, domain prediction analyses with different tools (Supplemental Data Set 1) revealed a possible diversification of the DPH N-terminal part that typically contains a PAS domain in other eukaryotic PHYs. Pfam and SMART did not detect conserved PAS domains in any of the DPHs from pennate diatom species and HHBlits predicted with high confidence the PAS only in three DPHs

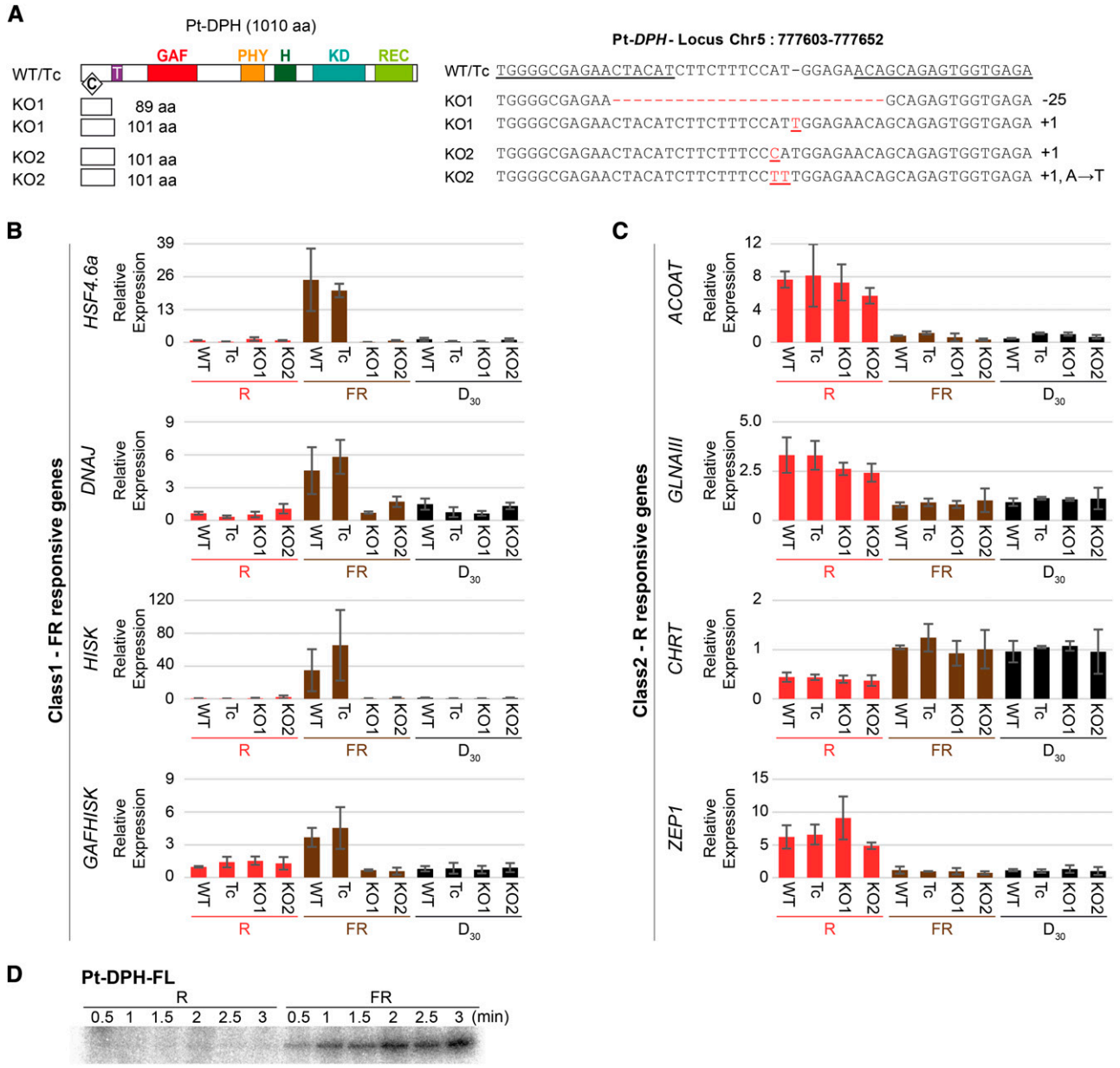

Figure 4. Characterization of Pt-*DPH* Knockout Lines.

(A) Scheme of the translation products of Pt-*DPH* in the wild-type and transformation control (Tc), and in two knockout lines (KO1 and 2). Sequences retrieved from KO1 and KO2 lines shown in red, for the different alleles, the mutations in the Pt-*DPH* locus compared with the wild type, and on the right the nucleotide insertions (+), deletions (−), and/or transversions.
(B) and **(C)** RT-qPCR analysis of Class1 (*HSF4.6a*, *DNAJ*, *HISK*, and *GAFHisK*) **(B)** and Class 2 genes (*ACOAT*, *GLNAIII*, *CHRT*, and *ZEP1*) **(C)** on wild-type cells, Tc, KO1, and 2 lines, exposed for 30 min to 660 nm (R, red) or 765 nm (FR, brown) LED lights (5 µmol·m^{-2}·s^{-1}) or maintained in darkness (D_{30}, black). Expression data were normalized using the T0 point (cf. Figure 3A) and the *RPS* gene as normalizer (n = 3; black bars = ±SD).
(D) Autophosphorylation assay of recombinant holo-Pt-DPH-FL incubated with [ϒ-^{32}P]ATP under 639-nm (R) or 757-nm (FR) LED lights up to 3 min, revealed by SDS-PAGE and autoradiography.

(*S. robusta* DPH4 and *A. coffeaeformis* DPH1 and DPH2). However, secondary structure analysis performed by JPred (Supplemental Data Set 1) identified patterns similar to those found in the PAS domain of *Deinococcus radiodurans* BphP, for which the crystal structure has been determined (Wagner et al., 2005).

Regarding the vertical distribution of DPHs in the water column, there are several records of diatoms actively growing in the subsurface layer (Crombet et al., 2011) or in highly mixed water columns (Backhaus et al., 2003; Smetacek et al., 2012) where cells may be displaced at depth and remain there for long periods of

time. We queried the MAREDAT (marine ecosystem data) database, which contains information about diatom abundance (Leblanc et al., 2012; Buitenhuis et al., 2013) to explore the distribution of marine diatom species bearing the *DPH* gene (Supplemental Data Set 4). The search identified two diatoms expressing DPH, *Thalassiosira rotula* and *Skeletonema costatum* (Figure 5B), and revealed their wide vertical distribution along the water column, including deep-water layers.

Estimation of R-FR Light Sources for DPH Modulation along the Water Column

To assess the identity and extent of light signals that may modulate DPH activity in the subsurface layers, we calculated the vertical distribution of underwater R and FR scalar irradiance (610- to 800-nm band), which derives from attenuated sunlight, Raman scattering, and phytoplankton chlorophyll fluorescence (Figure 5C) and, for comparison, attenuated sunlight only. We incorporated Raman and chlorophyll fluorescence because they each transfer part of the energy from the more penetrating bands (e.g., blue-green) to the R-FR bands (Mobley, 1994; Ragni and Ribera d'Alcala, 2004). Three case studies representing different phytoplankton abundance scenarios in different oceanic contexts were investigated: (1) marine station alt_0033 displaying low phytoplankton at surface and a pronounced deep chlorophyll maximum (DCM) at 62 m (1.6 mg/m^3 at DCM); (2) station Mc1027 with a pronounced surface chlorophyll maximum (0 to 5 m) (average 8.2 mg/m^3); and (3) station STB-7 with a DCM deeper than 100 m (0.1 mg/m^3 at DCM). For all the sites considered, this analysis indicated that direct sunlight in the R-FR band becomes a minor contributor to irradiance below 15 to 20 m, whereas total R-FR irradiance is greater than $1 \cdot 10^{-2}$ $\mu E \cdot m^{-2} \cdot s^{-1}$ above 35 or 75 m depth and greater than $1 \cdot 10^{-3}$ $\mu E \cdot m^{-2} \cdot s^{-1}$ above 75 or 100 m, for the least transparent (Mc1027) and the most transparent (STB-7) station, respectively (Figure 5C). We subsequently computed the relative molar absorptivities of Pt-DPH Pr and Pt-DPH Pfr at different depths by convoluting the underwater scalar irradiance spectrum with the absorption spectra of the two forms. The data show that that ratios between the absorptivities of Pt-DPH Pr and Pt-DPH Pfr is always larger than 1 and converges to values around 2.2 (Figure 5D). High photosynthetic biomass does not change the ratio of light captured by DPH. This is because chlorophyll fluorescence provides a similar contribution to R and FR light and because the selective absorption of R light by chlorophyll, even at the high cell densities of the selected sites corresponding to an algal bloom, does not affect the ratio significantly. Taken together, the data suggest that below the first meter (Figure 5D), the regulation of Pt-DPH activity cannot be based on modulation of the R/FR ratio.

DISCUSSION

Diatom Phytochrome Mediates FR Light Signaling

Recent characterization of phytochromes from taxonomically diverse algae unveiled significant structural diversity in these proteins and wide spectral variations that are not restricted to R and FR light (Duanmu et al., 2014; Rockwell et al., 2014b). Our study in diatoms establishes the presence of this important photoreceptor class in prominent marine species and further enhances our knowledge of phytochrome diversity. Because our investigation mostly relies on the analysis of transcriptomic data and on a limited set of complete genome sequences, we cannot provide a complete overview of the conservation of DPH throughout the diatoms. Nevertheless, the presence of multiple DPH members in some pennate species and in *E. siliculosus* suggests that gene duplication events contributed to shape the Heterokont PHYs.

The analysis of DPHs also unveils several distinct features of the diatom proteins compared with the other eukaryotic PHYs, such as the presence of a second cysteine potentially involved in chromophore binding in the PSM, and a diversification of the N-terminal PHY sequence. Although we cannot rule out the possibility of a PAS domain folding, the reported N-terminal sequence diversification is of particular interest given the deep conservation of the PHY N-terminal module across eukaryotes. Analysis of additional PHY sequences in metagenomic and metatranscriptomic data from phytoplankton (e.g., *Tara Oceans*) (Bork et al., 2015) will likely provide new insights about DPH evolution and a better understanding of their distribution and expression in specific environments.

The Tp-DPH and Pt-DPH proteins characterized in this work are photoreversible R-FR-light-sensitive proteins (Figure 1). Detailed analyses of gene expression in *P. tricornutum* revealed specific cellular responses to R and FR light and establish a role for DPH in FR, but not R light signaling (Figures 2 and 3). Red light induced changes in expression of ~200 genes likely triggering a reorganization of nitrogen and carbon metabolism (Jungandreas et al., 2014). R-light-mediated regulation of the tested genes was abolished by DCMU, as was observed previously using higher R light fluences (Valle et al., 2014), but was not impaired in Pt-*DPH* knockout cells, positing that regulation of R-responding genes is linked to the chloroplast and/or cellular metabolic status. In contrast, a role for Pt-DPH in FR signaling was first suggested by the identification of a significant number of genes (~80) specifically induced by FR light. The comparison of Pt-DPH absorption spectra with the action spectra for gene expression showed that maximal induction of selected FR-responsive genes occurs at wavebands specific to the Pfr form. Analyses of these genes in transcriptome data sets from *P. tricornutum* cells exposed to other wavebands (blue, green, red, and white) (Nymark et al., 2009; Chauton et al., 2013; Valle et al., 2014) showed either no expression changes or opposite changes compared with those induced by FR. The loss of responsiveness of Class 1 genes in knockout cells further demonstrated the involvement of DPH in FR light signaling (Figure 3). Finally, autophosphorylation assays also suggest that Pt-DPH mediates FR light signal propagation through a phosphorylation cascade. While we cannot exclude that FR light might be perceived by another still unknown light absorbing molecule, our data make DPH a strong FR light photoreceptor candidate in *P. tricornutum* and presumably also in other diatoms.

Perspectives on Phytochrome-Mediated Underwater Light Perception

Our study incites a reevaluation of the ecological and functional significance of R and FR light signals underwater. While green and blue light rapidly become predominant with depth, they also

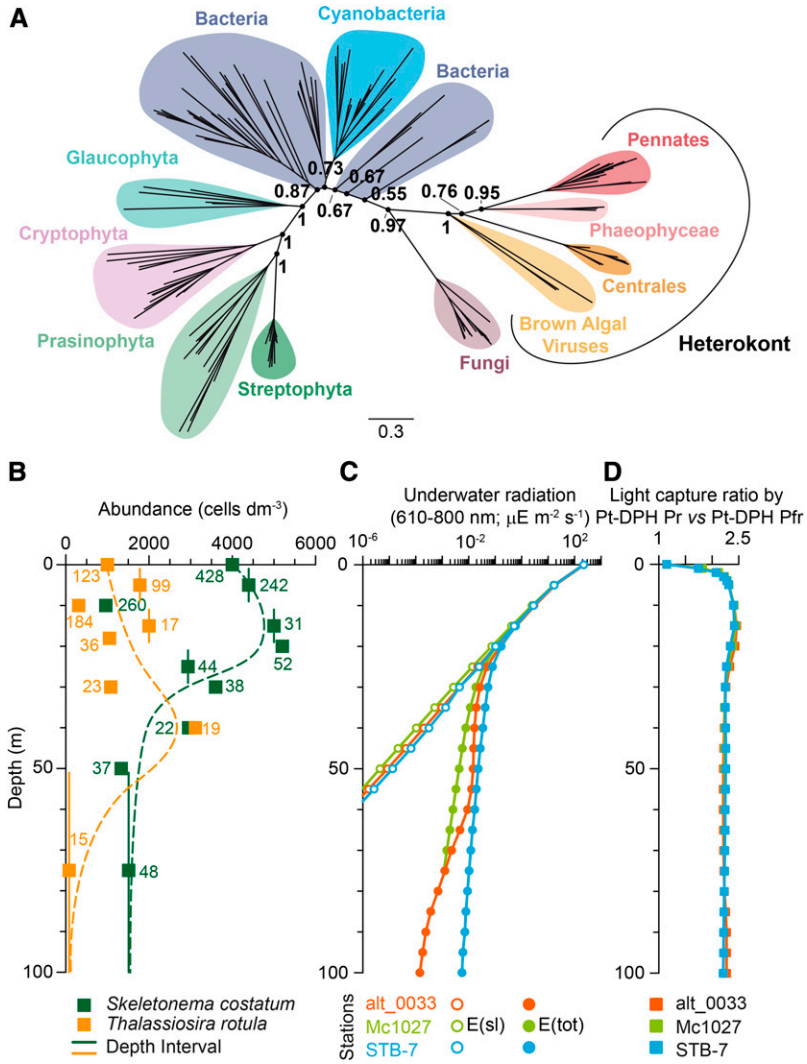

Figure 5. Exploration of DPH Abundance and Distribution and of the Light Sources Possibly Triggering DPH in the Marine Environment.

(A) Phylogenetic analysis of PHY protein superfamily using the GAF-PHY region. The tree is midpoint-rooted and was obtained applying the Bayesian inference method. The posterior probability value support is shown only for the basal nodes (black dots).

(B) Vertical distribution of *T. rotula* (yellow) and *S. costatum* (green) from the MAREDAT database. Each data point is the median of the abundance values (cells/dm^3) for single depth (squares) or depth intervals centered on the median depth (vertical lines). The number of observations for each data point is provided. Dashed lines show interpolation of the data points.

(C) Vertical profiles of the available scalar irradiance in the wavelength range of 610 to 800 nm. Open circles, irradiance derived from the attenuated sunlight E(sl) only; closed circles, total irradiance E(tot), deriving from sunlight, Raman scattering, and phytoplankton chlorophyll. Marine station codes are reported at the bottom.

(D) Ratios between the relative molar absorptivities of Pt-DPH Pr and Pt-DPH Pfr at different depths in the same spectral range.

generate, via Raman scattering and chlorophyll fluorescence, a flux of photons that partially compensate the attenuation of R-FR wavebands from sunlight (Figure 5B). Taking into account these parameters, our results indicate that such light fluxes may potentially trigger DPH responses. Compared with terrestrial environments, the intensity of the R and FR light within most of the photic zone is lower than under a canopy, where the lowest fluence rate values are in the order of few $\mu E \cdot m^{-2} \cdot s^{-1}$ (~0.5% incident solar radiation) (De Castro, 2000) and are comparable with those detectable under 9 mm of sandy soil, ($\sim 3 \cdot 10^{-2}$ $\mu E \cdot m^{-2} \cdot s^{-1}$) (Bliss and Smith, 1985). Such intensities are nonetheless well above the minimal fluences required to trigger very low fluence responses in plants, which are in the range of $\sim 1 \cdot 10^{-4}$ $\mu E \cdot m^{-2} \cdot s^{-1}$ (Furuya and Schafer, 1996). We hypothesize that perception of low light levels might be favored by the binding of BV by DPH. Other eukaryotic PHYs in oxygenic phototrophs use more reduced bilin chromophores that shift their absorption properties toward shorter wavelengths (Rockwell et al., 2014a). The displacement of the DPH absorption spectrum toward FR wavebands, beyond the spectral window of photosynthetic pigments,

may enhance DPH specificity, possibly providing an adaptive value in the poor FR light marine environments.

Our study highlights that, in addition to the generally lower light intensities in aquatic compared with terrestrial environments, there is a relatively constant R/FR light ratio along the water column. Due to Raman and chlorophyll fluorescence contributions, R and FR light intensities covary with a ubiquitous excess of the R component (Figure 5D). Consequently, DPH-mediated responses are unlikely to be regulated by R-FR or FR-R photoconversion (Figure 5D). However, the DPH-dependent FR responses reported here suggest that an equilibrium between the Pr/Pfr forms is generated following dark to light transition, and that could be maintained under constant R/FR ratios and in the absence of efficient dark reversion. Tight control of *P. tricornutum* DPH protein content, likely resulting from both transcriptional and posttranscriptional regulation (Figure 2B), may contribute to Pt-DPH regulation and to restrict its activity to the early light period. Similarly to land plants, DPH could therefore contribute to the photoperiodic regulation of cellular events and metabolism (Wang and Wang, 2015) by activating specific signaling cascades.

Given the strong attenuation of R-FR light with depth, the activity of DPH might be restricted to the surface layer, thus potentially acting as a surface detector. However, other functions can also be envisaged below the ocean surface, which take into account the important contribution of chlorophyll fluorescence to R-FR irradiation at deeper depths (Figure 5C). In one scenario, DPH could act as a biotic signal detector sensing elevated chloroplast fluorescence from nearby phytoplankton cells, notably during algal blooms. Alternatively, at the low phytoplankton cell densities typically found in the ocean (Figure 5C), DPH might rather sense chlorophyll autofluorescence of the diatom cell itself, a prevalent source of R-FR light at depths greater than 15 to 20 m. In this second scenario, DPH could act as a sensor of chloroplast activity or chlorophyll concentration and may work together with other light absorbing molecules. Which signal is the prevalent trigger of DPH responses (external versus internal signals) may depend on DPH subcellular localization, on R/FR photoreceptor sensitivity, on the size and density of neighboring cells as well as on their chlorophyll/chloroplast content and activity, which are all questions left open for future investigation. Partial answers to these questions will likely derive from the characterization of the FR-specific genes regulated by DPH. Most of these genes encode pioneer proteins, potentially involved in the regulation of processes relevant for diatom survival and/or proliferation in specific light environments. DPH-regulated traits could therefore confer benefits to diatom growth and contribute, at least in part, to their ecological success in the ocean.

Finally, the discovery of DPH-mediated FR responses opens new perspectives on the light sensing abilities of an ecologically relevant group of phytoplankton and stimulates novel hypotheses about the possible function of this light sensor, among which its potential for being a biotic signal detector. Further characterization of DPH photochemical and biochemical properties are expected to provide completely novel insights about its physiological action, and the more than 50% genes of unknown function, specifically induced by FR in *P. tricornutum*, constitute a promising entry point for future molecular studies aimed at solving the aquatic phytochrome conundrum.

METHODS

Culture Conditions

The wild-type *Phaeodactylum tricornutum* (Pt1 8.6; CCMP2561) cells and transgenic lines were grown in f/2 medium (Guillard, 1975) at 18°C under a 12-h-light/12-h-dark regime using 50 μmol photons$\cdot m^{-2}\cdot s^{-1}$ white light (Philips TL-D De Luxe Pro 950). For gene expression studies, cells were specifically illuminated for 30 min at low irradiance (5 $\mu E\cdot m^{-2}\cdot s^{-1}$) with a series of narrow-spectrum LED light sources (λ_{max}: 660, 688, 732, 740, or 765 nm, and half-bandwidth below 17 nm for all the wavelengths). DCMU treatment was performed as described (Huysman et al., 2013). All the experiments were done with cells in exponential phase of growth (1 to $2\cdot 10^6$ cells/mL).

Expression, Purification, and Spectral Analysis of DPH Proteins

We refer to the diatom phytochrome gene, mRNA, and protein (Pt-*DPH*, Pt-*DPH* mRNA, and Pt-DPH, respectively) following the nomenclature reported by Quail et al. (1994). Pt-*DPH* and Tp-*DPH* gene sequences were synthesized by GenScript with codons optimized for *Escherichia coli* expression. Sequences were cloned into pET28a (Novagen) in NcoI/XhoI sites for the full-length Pt-*DPH* gene and NcoI/SalI for the PSM of Pt-*DPH* and Tp-*DPH*, and expressed as C-terminal His6 tagged proteins in *E. coli* BL21 strain containing either the pKT270 plasmid for BV IXa, pKT272 for PΦB, or pKT278 for PEB production (Mukougawa et al., 2006). Recombinant proteins were expressed with 0.1 mM (Pt-DPH) or 0.5 mM (Tp-DPH) isopropyl β-D-1-thiogalactopyranoside at 18°C overnight and purified and analyzed as described (Enomoto et al., 2012). Light-emitting diodes were used for R (639 nm for Pt-DPH and 690 nm for Tp-DPH) and FR (757 nm for Pt-DPH and 765 nm for Tp-DPH) light irradiations (half-bandwidths below 24 nm for the different wavelengths and fluences of ~20 and ~3 µmol photons$\cdot m^{-2}\cdot s^{-1}$ for R and FR irradiation, respectively). The different phytochromes were maximally photoconverted within the 10 s of illumination. Absorption spectra were recorded using UV-2600PC (Shimadzu) or UvikonXS (Secomam) spectrophotometers. Difference spectra were calculated as "FR-irradiated" minus "R-irradiated" spectra. Proteins denaturation was done in acidic urea (pH 2) (Enomoto et al., 2012). Pt-DPH-FL and Tp-DPH-PSM dark relaxation was followed during 24 h at room temperature starting from DPH purified under room lighting. Autophosphorylation assay was performed with ~16.8 µM Pt-DPH-FL holoprotein as described (Hirose et al., 2008), under irradiation with R or FR light up to 3 min.

Construction of Pt-Dph Reporter Lines

The Pt-DPH_{Pro}:Pt-*DPH-HA* construct was obtained through a double-joint PCR approach (Yu et al., 2004) using the Phusion High-Fidelity DNA polymerase (Thermo Fisher). The Pt-*DPH* coding sequence was obtained after PCR amplification from cDNA using the oligos Phy_DraI-Fw and Phy_XhoI_Rv, inserted into the pENTR1A vector (Invitrogen) in *DraI*/*Xho*I, and recombined with the pDEST-C-HA vector (Siaut et al., 2007). The Pt-*DPH-HA* fragment was amplified from the pDEST-Pt-*DPH-HA* plasmid using CodPromPhy_Fw and PhyHA_AT_Rv oligos and then fused to the Pt-*DPH* 5′ upstream region amplified from the genomic DNA with the PhyProm_Fw2 and the PromCodPhy_Rv primer pairs, and the Pt-*DPH* terminator with the oligos Pt-PhyAt_Fw and Pt-PhyAt_Rv0, using the PhyProm_Fw3 and Pt-PhyAt_Rv1 primers (primer sequences are listed in Supplemental Table 1). The final fragment was cloned into the TOPO-TA vector (Invitrogen). Diatom transgenic lines were obtained by cotransformation with nourseothricin resistance plasmid as described (De Riso et al., 2009).

Construction of Pt-*DPH* Knockout Lines by TALEN-Mediated Genome Editing

A pair of TALE nucleases was designed by Cellectis Bioresearch to bind upstream and downstream the targeted cleavage site of Pt-*DPH* gene

(TGGGGCGAGAACTACAT and the sequence corresponding to ACAGCAGAGTGGTGAGA on the opposite strand, respectively; from +258 to +307 relatively to the start codon) and transformed in *P. tricornutum*. Three transgenic lines containing both TALENs were screened by T7 endonuclease assay (Daboussi et al., 2014). Amplification and sequencing from individual subclones allowed us to isolate two independent lines containing biallelic mutations together with another clone containing the wild-type sequence, used as a control.

RNA Extraction and Gene Expression Analysis

Total RNA was extracted as described (Huysman et al., 2013). For RT-qPCR, 500 ng of total RNA was reverse-transcribed using the QuantiTect reverse transcription kit (Qiagen) and the reaction accomplished with 12.5 ng cDNA as template, following the SsoAdvanced Universal SYBR Green Supermix instructions (Bio-Rad), in a CFX 96 real-time detection system (Bio-Rad). Primer efficiencies were determined as described (Pfaffl, 2001). *RPS* and/or *TBP* were used as reference genes and normalization performed as described (Siaut et al., 2007).

For microarray analyses, total RNA was processed with the RNA cleanup protocol of the RNeasy Mini Kit (Qiagen). Two hundred nanograms of total RNA was labeled with fluorescent Cy3 and Cy5 dyes using the Low Input Quick Amp Labeling Kit, Two-Color (Agilent). Two biologically independent experiments were tested for each condition, using dye switch, and hybridized on 8x60K *P. tricornutum* whole-genome 60-mer oligonucleotide microarrays (Agilent Technologies) following the Agilent protocol. The microarray chip (Platform GPL21013) was designed based on the *P. tricornutum* genome, version Phatr2 (http://genome.jgi.doe.gov/Phatr2/Phatr2.home.html). Slides were scanned using a 2-micron Agilent microarray scanner and the resulting images analyzed using the Agilent Feature Extraction 11.5 software. Microarray normalization was performed by global lowess using Goulphar (Lemoine et al., 2006). For each gene, the Cy5/Cy3 ratios corresponding to the different probes were averaged and mean of the biological replicates was calculated. Statistical significance of the variations in gene expression was determined using the MeV 4.9 version of SAM. The two independent replicates of the D_{30} conditions were used as control group in SAM by performing a two-class unpaired analysis to test the independent replicates of the R and FR conditions, separately. The analysis was performed using all six possible unique permutations, S0 (Tusher et al., 2001) and K-nearest neighbors imputer (10 neighbors) as imputation engine. For gene selection, the delta value was set to obtain a false discovery rate $<10^{-4}$% and absolute $\log_2$ (fold change) value >1. Hierarchical clustering analysis were performed using MeV 4.9 (Eisen et al., 1998). Four major clusters were identified following the dendrogram obtained from the clustering algorithm and class assignment determined based on the expression profile.

Immunoblot Analysis

Proteins were extracted as described (Coesel et al., 2009), loaded onto 7.5% SDS-PAGE gels (Laemmli, 1970), and transferred to nitrocellulose membranes (Sigma-Aldrich) in CAPS buffer (Fisher). Membranes were blocked with TBS-T 5% milk and incubated overnight at 4°C with α-HA (Covance) and α-βCF1 (β-subunit of chloroplastic ATPase; Agrisera) as loading control. Following incubation with the horseradish peroxidase-conjugated secondary antibodies (Promega), proteins were revealed with Clarity reagents (Bio-Rad) and an Image Quant LAS4000 camera (GE Healthcare).

Data Mining, Protein Domain Identification, and Plastid Localization

Phytochrome sequences were collected from already characterized PHYs, TBLAST-N searches on public NCBI database with *Arabidopsis thaliana* PHYB and Pt-DPH protein sequences as queries, and local BLAST-P searches onto the assembled transcriptomic database of the MMETSP (Marine Microbial Eukaryote Transcriptome Sequencing Project) (Keeling et al., 2014) using the Pt-DPH GAF-PHY domain sequence as query. The *Seminavis robusta* phytochrome sequence data were produced by the U.S. Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) in collaboration with the user community. Newly identified sequences (E-value $<10^{-5}$) were analyzed for domain composition using Pfam (http://pfam.xfam.org/) and SMART (http://smart.embl-heidelberg.de/). HHBlits (Remmert et al., 2012) was used to determine PAS domain divergence by performing profile-profile comparison, taking into account secondary structure information (PSIPRED; Jones, 1999). HHblits was used to build profiles for PAS_2 (PF08446, from Pfam seed sequences) and each DPH protein (sequences upstream of the predicted GAF domain [PF01590]). HHblits scores were computed with default parameters. Secondary structures were predicted by JPred (Drozdetskiy et al., 2015). Chromophore binding site prediction was done by aligning PHY sequences with ClustalW (Larkin et al., 2007) and manually corrected.

Phylogenetic Analysis of the Phytochrome Family Prediction

A preliminary database comprising 224 protein sequences was created and sequences independently trimmed to the GAF-PHY protein domains (Supplemental Data Set 5). Protein alignment, preliminary phylogenies, and estimation of the best amino acid model given the data (LG +I+G+F; Akaike criterion) was performed as described (Oliveri et al., 2014). Genealogical relationships among protein sequences were inferred using MrBayes 3.2.3 (Ronquist et al., 2012) on the CIPRES Science Gateway (Miller et al., 2010). The analysis was performed with two independent runs and two million generations per run. After a burn-in of the 25% of the total trees per run, the remaining trees were used to reconstruct a consensus tree and to get posterior probabilities for node supports. The final tree was edited in FigTree 1.4 (http://tree.bio.ed.ac.uk/software/Figuretree/).

DPH Ecological Distribution Analysis

Diatom species possessing DPH sequences were searched in the MAREDAT database (Leblanc et al., 2012; Buitenhuis et al., 2013). A total of 544 and 1274 entries about distribution information were recovered for *Thalassiosira rotula* and *Skeletonema costatum*, respectively. Data were grouped by depth, when corresponding to the standard sampling depths (0, 10, 20, 30, 40, and 50 m), by depth interval when derived from samples in between standard depths, e.g., 1 to 9 m, 11 to 19 m, etc., or by depth ranges when poorly sampled, e.g., 51 to 100 m.

Identity of Light Signals Possibly Driving DPH

Underwater scalar irradiance with a spectral resolution of 5 nm and each 5-m depth was derived by the Hydrolight model (Mobley et al., 1993; Mobley and Sundman, 2013), which allows us to compute the contributions to E of the solar irradiance, the Raman scattering, and the chlorophyll a fluorescence, independently. Simulations were run using in situ chlorophyll *a* profiles and clear sky irradiance at Solar Zenith Angle of 30° as inputs. Station data are as follows: alt_0033, 41.75°N, 10.50°E, 08/08/2004, MC1027, 40.81°N, 14.25°E, 18/09/2012; STB-7, 22.03°S, 120.21°W, 09/11/2004.

The relative molar absorptivities of Pt-DPH Pr or Pt-DPH Pfr were computed by convoluting total E at different depths with the spectra of Pt-DPH Pr and Pt-DPH Pfr normalized by Pt-DPH Pr maximum. The absolute absorption spectrum of Pt-DPH Pfr was calculated integrating the method suggested by Giraud et al. (2005), i.e., measuring spectra obtained under steady state illumination at two different wavelengths, 620 and 705 nm (halogen 150-W lamp, filtered through interference filters, half-bandwidth 6 nm) in our case, and the spectrum obtained after illumination at 780 nm (halogen 150-W lamp, filtered through cut-on filter, half-bandwidth 6 nm),

considering that it corresponds to 100% of Pr form. The absorption spectrum of *P. tricornutum* was obtained as described (Ragni and Ribera d'Alcala, 2007).

Accession Numbers

Sequence data from this article have been deposited in the GenBank data library under the following accession numbers: KT175537, codon modified Pt-*DPH*; and KT175538, codon modified Tp-*DPH*. Accession numbers of the sequences used for phylogenetic analysis are in Supplemental Data Set 2. The microarray study is MIAME compliant (Gene Expression Omnibus accession number GSE73915).

Supplemental Data

Supplemental Figure 1. Alignment of PAS-GAF domains of representative PHYs.

Supplemental Figure 2. Dark reversion and chromophore binding experiments.

Supplemental Figure 3. Molecular characterization of Pt-DPH-HA and Pt-*DPH* knockout lines.

Supplemental Table 1. Primer list.

Supplemental Data Set 1. PAS domain prediction in DPH proteins.

Supplemental Data Set 2. Putative enzymes implicated in chromophore biosynthesis.

Supplemental Data Set 3. Microarray data of Figure 3B.

Supplemental Data Set 4. List of the proteins used in the phylogenetic analysis

Supplemental Data Set 5. Sequence alignment.

ACKNOWLEDGMENTS

We thank A. Verméglio and B. Bailleul for discussion, F. Devaux and J. Mehrej for support with microarray analysis, W. Vyverman for sharing unpublished sequence data of *S. robusta* DPHs, E. Vellucci for support with the Hydrolight simulations, and M. Siaut for supporting the project at the early stage. This work was funded by a French ATIP CNRS grant (2009), by a Human Frontier Science Program Organization research grant (RGY0082/2010), by a Gordon and Betty Moore Foundation grant (GBMF 4966), and by grants from the Marie-Curie ITN CALIPSO (ITN 2013 GA 607607) and Marie Curie ITN AccliPhot (ITN 2012 GA 316427) to A.F.; by the Italian Program RITMARE to M.R.d.A.; by the French ''Investissements d'Avenir'' program CALSIMLAB (ANR-11-LABX-0037-0, program ANR-11-IDEX-0004-02) funds from the Institut Universitaire de France to A.C.; by the French "Initiative d'Excellence" program (LabEx DYNAMO, ANR-11-LABX-0011) to F.R.; and by the ERC ''Diatomite'' project, the French ''Investissements d'Avenir'' program MEMO LIFE (ANR-10-LABX-54), and PSL Research University (ANR-11-IDEX-0001-02) to C.B. This article is dedicated to the memory of Fabrice Rappaport.

AUTHOR CONTRIBUTIONS

A.F. and C.B. discovered diatom phytochromes. A.F., M.J., J.-P.B., A.E.F., M.R.d.A., and M.I. designed experiments and wrote the article. A.E.F., M.J., M.T., J.-P.B., and R.R. performed molecular studies. M.J.J.H. provided the *S. robusta* DPH sequence information. G.E. performed Pt-DPH spectral analysis. C.B. and S.M. explored *DPH* distribution. J.S.B. and A.C. analyzed the PHY domain. M.R.d.A., B.G., and F.R. analyzed underwater light signals. All authors discussed results and revised and approved the manuscript.

Received November 3, 2015; revised February 16, 2016; accepted February 29, 2016; published March 3, 2016.

REFERENCES

Armbrust, E.V., et al. (2004). The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science **306:** 79–86.

Ashworth, J., Coesel, S., Lee, A., Armbrust, E.V., Orellana, M.V., and Baliga, N.S. (2013). Genome-wide diel growth state transitions in the diatom *Thalassiosira pseudonana*. Proc. Natl. Acad. Sci. USA **110:** 7518–7523.

Auldridge, M.E., and Forest, K.T. (2011). Bacterial phytochromes: more than meets the light. Crit. Rev. Biochem. Mol. Biol. **46:** 67–88.

Backhaus, J.O., Hegseth, E.N., Wehde, H., Irigoien, X., Hatten, K., and Logemann, K. (2003). Convection and primary production in winter. Mar. Ecol. Prog. Ser. **251:** 1–14.

Bliss, D., and Smith, H. (1985). Penetration of light into soil and its role in the control of seed germination. Plant Cell Environ. **8:** 475–483.

Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E., and Wincker, P. (2015). Tara Oceans. Tara Oceans studies plankton at planetary scale. Introduction. Science **348:** 873.

Bowler, C., et al. (2008). The Phaeodactylum genome reveals the evolutionary history of diatom genomes. Nature **456:** 239–244.

Buitenhuis, E.T., et al. (2013). MAREDAT: towards a world atlas of MARine Ecosystem DATa. Earth Syst. Sci. Data **5:** 227–239.

Chauton, M.S., Winge, P., Brembu, T., Vadstein, O., and Bones, A.M. (2013). Gene regulation of carbon fixation, storage, and utilization in the diatom *Phaeodactylum tricornutum* acclimated to light/dark cycles. Plant Physiol. **161:** 1034–1048.

Coesel, S., Mangogna, M., Ishikawa, T., Heijde, M., Rogato, A., Finazzi, G., Todo, T., Bowler, C., and Falciatore, A. (2009). Diatom PtCPF1 is a new cryptochrome/photolyase family member with DNA repair and transcription regulation activity. EMBO Rep. **10:** 655–661.

Crombet, Y., Leblanc, K., Quéguiner, B., Moutin, T., Rimmelin, P., Ras, J., Claustre, H., Leblond, N., Oriol, L., and Pujo-Pay, M. (2011). Deep silicon maxima in the stratified oligotrophic Mediterranean Sea. Biogeosciences **8:** 459–475.

Daboussi, F., et al. (2014). Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. Nat. Commun. **5:** 3831.

De Castro, F. (2000). Light spectral composition in a tropical forest: measurements and model. Tree Physiol. **20:** 49–56.

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., and Falciatore, A. (2009). Gene silencing in the marine diatom *Phaeodactylum tricornutum*. Nucleic Acids Res. **37:** e96.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Res. **43:** W389–W394.

Duanmu, D., et al. (2014). Marine algae and land plants share conserved phytochrome signaling systems. Proc. Natl. Acad. Sci. USA **111:** 15827–15832.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA **95:** 14863–14868.

Enomoto, G., Hirose, Y., Narikawa, R., and Ikeuchi, M. (2012). Thiol-based photocycle of the blue and teal light-sensing cyanobacteriochrome Tlr1999. Biochemistry **51:** 3050–3058.

Falciatore, A., and Bowler, C. (2005). The evolution and function of blue and red light photoreceptors. Curr. Top. Dev. Biol. **68:** 317–350.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. Science **281:** 237–240.

Franklin, K.A., and Quail, P.H. (2010). Phytochrome functions in Arabidopsis development. J. Exp. Bot. **61:** 11–24.

Furuya, M., and Schafer, E. (1996). Photoperception and signalling of induction reactions by different phytochromes. Trends Plant Sci. **1:** 301–307.

Giraud, E., Zappa, S., Vuillet, L., Adriano, J.M., Hannibal, L., Fardoux, J., Berthomieu, C., Bouyer, P., Pignol, D., and Verméglio, A. (2005). A new type of bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in *Rhodopseudomonas palustris*. J. Biol. Chem. **280:** 32389–32397.

Guillard, R.R.L. (1975). Culture of phytoplankton for feeding marine invertebrates. In Culture of Marine Invertebrate Animals, W.L. Smith and M.H. Chanley, eds (New York: Plenum Press), pp. 26–60.

Hirose, Y., Shimada, T., Narikawa, R., Katayama, M., and Ikeuchi, M. (2008). Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein. Proc. Natl. Acad. Sci. USA **105:** 9528–9533.

Huysman, M.J., et al. (2013). AUREOCHROME1a-mediated induction of the diatom-specific cyclin dsCYC2 controls the onset of cell division in diatoms (*Phaeodactylum tricornutum*). Plant Cell **25:** 215–228.

Ikeda, Y., Komura, M., Watanabe, M., Minami, C., Koike, H., Itoh, S., Kashino, Y., and Satoh, K. (2008). Photosystem I complexes associated with fucoxanthin-chlorophyll-binding proteins from a marine centric diatom, *Chaetoceros gracilis*. Biochim. Biophys. Acta **1777:** 351–361.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292:** 195–202.

Jungandreas, A., Schellenberger Costa, B., Jakob, T., von Bergen, M., Baumann, S., and Wilhelm, C. (2014). The acclimation of *Phaeodactylum tricornutum* to blue and red light does not influence the photosynthetic light reaction but strongly disturbs the carbon allocation pattern. PLoS One **9:** e99727.

Karniol, B., and Vierstra, R.D. (2003). The pair of bacteriophytochromes from *Agrobacterium tumefaciens* are histidine kinases with opposing photobiological properties. Proc. Natl. Acad. Sci. USA **100:** 2807–2812.

Keeling, P.J. (2013). The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. Annu. Rev. Plant Biol. **64:** 583–607.

Keeling, P.J., et al. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biol. **12:** e1001889.

Kooistra, W.H.C.F., Gersonde, R., Medlin, L.K., and Mann, D.G. (2007). The origin and evolution of the diatoms: their adaptation to a planktonic existence. In Evolution of Primary Producers in the Sea, P.G. Falkowski and A.H. Knoll, eds (Amsterdam, Heidelberg: Elsevier, Academic Press), pp. 207–249.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680–685.

Lamparter, T., Carrascal, M., Michael, N., Martinez, E., Rottwinkel, G., and Abian, J. (2004). The biliverdin chromophore binds covalently to a conserved cysteine residue in the N-terminus of Agrobacterium phytochrome Agp1. Biochemistry **43:** 3659–3669.

Larkin, M.A., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics **23:** 2947–2948.

Leblanc, C., Falciatore, A., Watanabe, M., and Bowler, C. (1999). Semi-quantitative RT-PCR analysis of photoregulated gene expression in marine diatoms. Plant Mol. Biol. **40:** 1031–1044.

Leblanc, K., et al. (2012). A global diatom database – abundance, biovolume and biomass in the world ocean. Earth Syst. Sci. Data **4:** 149–165.

Lemoine, S., Combes, F., Servant, N., and Le Crom, S. (2006). Goulphar: rapid access and expertise for standard two-color microarray normalization methods. BMC Bioinformatics **7:** 467.

Li, F.W., Melkonian, M., Rothfels, C.J., Villarreal, J.C., Stevenson, D.W., Graham, S.W., Wong, G.K., Pryer, K.M., and Mathews, S. (2015). Phytochrome diversity in green plants and the origin of canonical plant phytochromes. Nat. Commun. **6:** 7852.

Man, D., Wang, W., Sabehi, G., Aravind, L., Post, A.F., Massana, R., Spudich, E.N., Spudich, J.L., and Béjà, O. (2003). Diversification and spectral tuning in marine proteorhodopsins. EMBO J. **22:** 1725–1731.

Margalef, R. (1978). Life forms of phytoplankton as survival alternatives in an unstable environment. Oceanol. Acta **1:** 493–509.

Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE) (New Orleans, LA: Institute of Electrical and Electronics Engineers), pp. 1–8.

Mobley, C.D. (1994). Light and Water: Radiative Transfer in Natural Waters. (San Diego, CA: Academic Press).

Mobley, C.D., Gentili, B., Gordon, H.R., Jin, Z., Kattawar, G.W., Morel, A., Reinersman, P., Stamnes, K., and Stavn, R.H. (1993). Comparison of numerical models for computing underwater light fields. Appl. Opt. **32:** 7484–7504.

Mobley, C.D., and Sundman, L.K. (2013). Hydrolight 5.2. Ecolight 5.2. (Bellevue, WA Sequoia Scientific).

Montsant, A., et al. (2007). Identification and comparative genomic analysis of signaling and regulatory components in the diatom *Thalassiosira pseudonana*. J. Phycol. **43:** 585–604.

Moustafa, A., Beszteri, B., Maier, U.G., Bowler, C., Valentin, K., and Bhattacharya, D. (2009). Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science **324:** 1724–1726.

Mukougawa, K., Kanamoto, H., Kobayashi, T., Yokota, A., and Kohchi, T. (2006). Metabolic engineering to produce phytochromes with phytochromobilin, phycocyanobilin, or phycoerythrobilin chromophore in *Escherichia coli*. FEBS Lett. **580:** 1333–1338.

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

Oliveri, P., Fortunato, A.E., Petrone, L., Ishikawa-Fujiwara, T., Kobayashi, Y., Todo, T., Antonova, O., Arboleda, E., Zantke, J., Tessmar-Raible, K., and Falciatore, A. (2014). The Cryptochrome/Photolyase family in aquatic organisms. Mar. Genomics **14:** 23–37.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29:** e45.

Quail, P.H., Briggs, W.R., Chory, J., Hangarter, R.P., Harberd, N.P., Kendrick, R.E., Koornneef, M., Parks, B., Sharrock, R.A., Schafer, E., Thompson, W.F., and Whitelam, G.C. (1994). Spotlight on phytochrome nomenclature. Plant Cell **6:** 468–471.

Ragni, M., and Ribera d'Alcala, M. (2004). Light as an information carrier underwater. J. Plankton Res. **26:** 433–443.

Ragni, M., and Ribera d'Alcala, M. (2007). Circadian variability in the photobiology of *Phaeodactylum tricornutum*: pigment content. J. Plankton Res. **29:** 141–156.

Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat. Methods **9:** 173–175.

Rockwell, N.C., Duanmu, D., Martin, S.S., Bachy, C., Price, D.C., Bhattacharya, D., Worden, A.Z., and Lagarias, J.C. (2014b). Eukaryotic algal phytochromes span the visible spectrum. Proc. Natl. Acad. Sci. USA **111:** 3871–3876.

Rockwell, N.C., Lagarias, J.C., and Bhattacharya, D. (2014a). Primary endosymbiosis and the evolution of light and oxygen sensing in photosynthetic eukaryotes. Front. Ecol. Evol. **2:** 66.

Rockwell, N.C., Su, Y.S., and Lagarias, J.C. (2006). Phytochrome structure and signaling mechanisms. Annu. Rev. Plant Biol. **57:** 837–858.

Rodriguez-Romero, J., Hedtke, M., Kastner, C., Müller, S., and Fischer, R. (2010). Fungi, hidden in soil or up in the air: light makes a difference. Annu. Rev. Microbiol. **64:** 585–610.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. **61:** 539–542.

Siaut, M., Heijde, M., Mangogna, M., Montsant, A., Coesel, S., Allen, A., Manfredonia, A., Falciatore, A., and Bowler, C. (2007). Molecular toolbox for studying diatom biology in Phaeodactylum tricornutum. Gene **406:** 23–35.

Smetacek, V., et al. (2012). Deep carbon export from a Southern Ocean iron-fertilized diatom bloom. Nature **487:** 313–319.

Smetacek, V.S. (1985). Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. Mar. Biol. **84:** 239–251.

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

Valle, K.C., Nymark, M., Aamot, I., Hancke, K., Winge, P., Andresen, K., Johnsen, G., Brembu, T., and Bones, A.M. (2014). System responses to equal doses of photosynthetically usable radiation of blue, green, and red light in the marine diatom *Phaeodactylum tricornutum*. PLoS One **9:** e114211.

Veith, T., and Büchel, C. (2007). The monomeric photosystem I-complex of the diatom *Phaeodactylum tricornutum* binds specific fucoxanthin chlorophyll proteins (FCPs) as light-harvesting complexes. Biochim. Biophys. Acta **1767:** 1428–1435.

Wagner, J.R., Brunzelle, J.S., Forest, K.T., and Vierstra, R.D. (2005). A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. Nature **438:** 325–331.

Wang, H., and Wang, H. (2015). Phytochrome signaling: time to tighten up the loose ends. Mol. Plant **8:** 540–551.

Yeh, K.C., and Lagarias, J.C. (1998). Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry. Proc. Natl. Acad. Sci. USA **95:** 13976–13981.

Yu, J.H., Hamari, Z., Han, K.H., Seo, J.A., Reyes-Domínguez, Y., and Scazzocchio, C. (2004). Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. Fungal Genet. Biol. **41:** 973–981.