



**HAL**  
open science

# Light sensing in the Ocean: studying diatom phytochrome photoreceptors

Carole Duchêne

► **To cite this version:**

Carole Duchêne. Light sensing in the Ocean: studying diatom phytochrome photoreceptors. Morphogenesis. Sorbonne Université, 2022. English. NNT : 2022SORUS164 . tel-04145175

**HAL Id: tel-04145175**

**<https://theses.hal.science/tel-04145175>**

Submitted on 29 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sorbonne Université

École doctorale Complexité du Vivant

*Laboratoire biologie du chloroplaste et perception de la lumière chez les micro-algues*

## **Light sensing in the Ocean : studying diatom phytochrome photoreceptors**

Par Carole Duchêne

Thèse de doctorat de Biologie

Dirigée par Dr. Angela Falciatore

Co-encadrante Dr. JAUBERT Marianne

Présentée et soutenue publiquement le 28 juin 2022

Devant un jury composé de :

GROSSMAN Arthur	Prof. Dr. Carnegie Institution for Science Stanford University	Rapporteur
PETROUTSOS Dimitris	Chercheur CNRS	Rapporteur
BHAYA Devaki	Prof. Dr. Carnegie Institution for Science Stanford University	Examinatrice
BABIN Marcel	Professeur d'Université, Université de Laval	Examinateur
CARBONE Alessandra	Professeur d'Université, Sorbonne Université	Examinatrice
FANKHAUSER Christian	Professeur d'Université, Université de Lausanne	Examinateur
FALCIATORE Angela	Directrice de recherche CNRS	Directrice de thèse
JAUBERT Marianne	Maître de conférences Sorbonne Université	Co-encadrante de thèse

*Choose a job you love, and you will never have to work a day in your life.*

*Confucius*

## ACKNOWLEDGMENTS

---

This thesis is obviously a team work, and I would like to thank all the people who surrounded me, supervised and supported me during the whole duration of my thesis. It has been a pleasure to develop this research project here.

I would like to start by thanking Angela, my thesis director, for giving me the opportunity to join this project (and many others!). Her enthusiasm for science and her passion for diatoms are a driving force for our research.

I thank Marianne for her patience, her rigor and her calm in all circumstances. Thank you for the freedom that was given to me in this thesis work. It was a pleasure to work with you.

I would like to underline the involvement of Jean-Pierre Bouly, who, even if he is not officially the supervisor of my thesis, has clearly been indispensable in the research that is presented here. Thank you also for the discussions and the hypotheses that always carry us further (and for the beers too).

I salute Alessandro, we started our theses (and masters!) together and we finish them two months apart. Thank you for your presence, your style and your ideals.

I would like to thank my colleagues of the UMR7238, where I started my PhD experience and of the UMR7141: the basement community, including an office only of students; thanks to Sandrine and Soizic who make this lab run, and for your enthusiasm for microscopy. Thank you to all those with whom I ate OUTSIDE every lunchtime and in all weathers during these last two years. The pandemic has slowed our social interactions, but this daily break has remained convivial. Thank you for the lab meetings of several hours, because these discussions are extremely formative.

Thank you to the whole UMR7141 for this research atmosphere and this dynamic so enriching. Thank you for the lab meetings of several hours, because these discussions are extremely formative.

Thank you all for putting up with my whistling and singing.

I would like to warmly thank Dr. Chris Bowler, for your interest in this project, even in the Arctic ice. I would like to thank you and Prof. Marcel Babin for the opportunity to live this experience. I want also to thank Dr. Juan Pierella Karlusich for his advice and patience.

I want to thank Dr. Maurizio Ribera d'Alcalà for his enthusiasm about my results, and for the discussions that followed.

Finally, I would like to thank all members of the jury for accepting to examine my work.

## TABLE OF CONTENTS

<b>Introduction</b> .....	9
Diatom roles in the biogeochemical cycles of carbon and silica .....	10
Diatom origin and evolution.....	11
Life in the ocean: characteristics of the diatom environment.....	15
Cellular aspects of diatom life and their regulation by light.....	26
Light sensing.....	32
Model species and genetic tools.....	50
Objective of the Thesis .....	55
<b>Chapter1: Photoreceptors repertoire and ditribution in Ochrophyta</b> .....	59
Abstract.....	61
Introduction.....	62
Results and Discussion.....	66
Conclusions .....	87
Material and methods:.....	89
Bibliography .....	92
Supplementary Figures.....	104
Supplementary Data.....	116
<b>Chapter 2: Marine diatom phytochromes are sensors of depth and phytoplankton concentration through sensing of underwater light variations</b> .....	119
Abstract.....	121
Introduction.....	122
Results .....	125
Discussion .....	141
Material & Methods .....	146
References .....	152
Supplementary Figures.....	160
Supplementary Tables.....	172
Supplementary Data.....	173
Supplemental Method.....	174
<b>Chapter 3: DPH in bentic diatom species</b> .....	181
Results .....	183
Discussion .....	190
Material and methods.....	195

Supplementary Figure.....	197
<b>Chapter 4: DPH sinalization.....</b>	<b>199</b>
Results .....	200
Conclusion .....	215
Material and methods.....	215
<b>Discussion .....</b>	<b>221</b>
<b>Bibliography .....</b>	<b>231</b>
<b>Annexes .....</b>	<b>251</b>

## LIST OF FIGURES AND TABLES

<b>Introduction</b> .....	9
Figure 1. Diatom evolution .....	12
Figure 2. Diversity of diatoms .....	14
Figure 3. Environmental context for phytoplankton life in the water column .....	17
Figure 4. Example of vertical profile of oxygen and light in two types of sediments .....	19
Figure 5. Ice associated algal communities and their dynamics .....	21
Figure 6. Examples of underwater light fields in different aquatic environments .....	23
Figure 7. Light field sediments and in ice.....	25
Figure 8. Diatom life cycle.....	27
Figure 9. The photosynthetic light reactions in diatoms.....	30
Figure 10. Sensing of light cues in diatoms .....	39
Figure 11. Phylogenetic analysis phytochrome photoreceptor .....	42
Figure 12. Signaling mechanisms in phytochromes.....	45
Figure 13. Algal phytochromes absorption spectra. ....	49
Table 1. Sequenced genomes of centric diatoms .....	51
Table 2. Sequenced genomes of pennate diatoms.....	52
<b>Chapter1: Photoreceptors repertoire and ditribution in Ochrophyta</b> .....	59
Figure 1. Photoreceptors repertoire in Stramenopiles .....	67
Figure 2. LOV-based photoreceptors in Ochrophyta .....	72
Figure 3. Ochrophyta CPF family protein tree. ....	78
Figure 4. Phylogeny of the rhodopsins.....	83
Figure 5. Phylogeny of the phytochromes .....	85
Table 1. Details of the parameters and reference sequences used in our photoreceptor search..	90
Figure S1. Stramenopile species tree.....	104
Figure S2. Bioinformatic analysis of the putative LOV-containing photoreceptors .....	105
Figure S3. LOV-based photoreceptors families in Ochrophyta.....	106
Figure S4. Analysis of the aureochrome and helmchromes .....	107
Figure S5. Analysis of the CPF family in Ochrophyta .....	108
Figure S6. Zoom on the CPD branch of the CPF tree.....	109
Figure S7. Zoom on the Cry-DASH branch of the CPF tree .....	110
Figure S8. The 6-4 photolyases.....	111
Figure S9. The plant-Cry like proteins .....	112
Figure S10. Sequence analysis of the CryP sub-families.....	113
Figure S11. Repartition of rhodopsins and phytochromes in Stramenopiles .....	114
Figure S12. Analysis of the rhodopsin photoreceptor types .....	115
Figure S13. Alignment of selected phytochrome sequences .....	116

<b>Chapter 2: Marine diatom phytochromes are sensors of depth and phytoplankton concentration through sensing of underwater light variations</b> .....	119
Figure 1. PtDPH activity reporter system and experimental conditions.....	126
Figure 2. Action and inhibition spectra for PtDPH-mediated responses .....	128
Figure 3. Modelling of PtDPH activity <i>in vivo</i> .....	130
Figure 4. PtDPH activity in modeled environmental fields .....	133
Figure 5. PtDPH activity in <i>in situ</i> measured field fields.....	134
Figure 6. DPH distribution is linked to latitude, temperature and optical parameters.....	136
Figure 7. Spectral properties of DPH from various diatoms.....	139
Table 1. Spectral and photochemical properties of different recombinant DPH.....	140
Figure S1. Experimental setup for the action and inhibition spectra .....	160
Figure S2. Control lines show no induction of YFP in the action spectra.....	160
Figure S3. Expression analysis of PtDPH-regulated and PtDPH-non regulated genes upon far-red and blue irradiation. <i>P. tricornutum</i> WT and DPH KO mutant.....	161
Figure S4. Light modeling examples.....	162
Figure S5. DPH activity in different modelled environmental condition .....	163
Figure S6. DPH activity in <i>in situ</i> measured environmental light fields over a year time series	164
Figure S7. Biogeography of DPH and Aureochromes (Aureo) from centric diatoms .....	164
Figure S8. Metagenome Assembled Genomes (MAGs) geographical distribution of centric diatoms.....	165
Figure S9. DPH relative abundance compared to Aureochrome relative abundance.....	166
Figure S10. Correlation and individual generalized additive models analysis for DPH and Aureochromes relative abundance .....	167
Figure S11. Complex generalized additive models explaining DPH relative abundance with a combination of environmental parameters for each filter size .....	168
Figure S12. Alignment of a portion of the PAS and GAF domains of selected DPH and reference phytochrome sequences.....	168
Figure S13. Purified recombinant DPH .....	169
Figure S14. TpDPH restores blue and far-red-dependent YFP induction in PtDPH KO line.....	169
Figure S15. Projection of DPHs in modeled light fields.....	170
Figure S16. Method for DPH and diatom Aureochrome gene search .....	171
Table S1. Species from which DPH have been characterized in this study. ....	172
Table S2. Sequences of primers used in this study .....	172
Table S3. SSN parameters for DPH and diatom Aureochrome search in Tara Oceans data .....	173
<b>Chapter 3: DPH in benthic diatom species</b> .....	181
Table 1. Pennate diatom species with phytochrome(s). ....	183
Figure 1. Absorption spectra and difference spectra between the Pr and Pfr forms for 4 <i>A. coffeaeformis</i> phytochromes.....	184
Figure 2. Fluorescence emission spectra at 77K of <i>P. tricornutum</i> .....	187
Figure 3. DPH controls F710 amplitude in response to increased FR/R ratio .....	188



Figure 4. Kinetics of induction of F710.....	188
Figure 5. 7K fluorescence spectra of centric diatom species.....	198
Figure 6. 7K fluorescence spectra of different pennate diatom species.....	190
Figure S1. DPH controls F710 amplitude in response to increased FR/R ratio.....	196
Table S1. Primers used to clone <i>A. coffeaeformis</i> DPHs.....	197
<b>Chapter 4: DPH Signalization .....</b>	<b>199</b>
Figure 1. Phytochrome chromophores biosynthesis pathways.....	201
Table 1. Characteristic of the candidate Heme oxygenase in <i>Phaeodactylum tricornutum</i> .....	202
Figure 2. Spectral properties of TweiDPH expressed SynHO1).....	203
Figure 3. Scheme of the vector structure allowing to co-express <i>T. weissflogii</i> DPH photosensory module.....	204
Table 2. Characteristic of the candidate BVR in <i>Phaeodactylum tricornutum</i> .....	205
Figure 4. Spectral properties of TweiDPH expressed with SynHO1 and different PtBVR s.....	205
Figure 5. DPH localization. Venus-PtDPH-expressing <i>P. tricornutum</i> cells .....	206
Figure 6. DPH localization. PtDPH-NeonGreen-expressing <i>P. tricornutum</i> cells.....	207
Figure 7. Venus localization. Venus-expressing <i>P. tricornutum</i> cells .....	207
Figure 8. PtCPF1 localization. Venus-PtCPF1-expressing <i>P. tricornutum</i> cells .....	208
Figure 9. Elements from two-component systems .....	209
Figure 10. DPH-OPM interacts with PtHpt in Yeast Two_Hybrid.....	210
Table 3. Proteins with Response regulators domains in <i>P.tricornutum</i> genome .....	212
Figure 10. DPH-OPM interacts with PtHpt in Yeast Two_Hybrid.....	210
Figure 11. Regulation of ProHsf4.6a by light, cell concentration and agitation.....	214
Figure 12. Proposition of DPH signaling mechanisms in <i>P. tricornutum</i> .....	215
Table S1. Primers data used in this study .....	215



# INTRODUCTION

Diatoms are unicellular photosynthetic eukaryotes that can be found in all wet environments, from freshwater to marine ecosystems, including brackish waters, and extended to sea ice and terrestrial soil (Armbrust, 2009). One of their characteristics is the presence of a silicified cell wall (the “frustule”) surrounding the cell, giving them a diversity of beautiful shapes (Babenko et al., 2022). Although microscopic, diatoms are major primary producers in the ocean and have tremendous roles in the cycle of carbon, silica and nutrients.

In the last 20 years, the availability of genomic information and genetic resources in model species such as *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, started to unravel the molecular secret of diatoms and unlock diatom potential for biotechnology application.

As photosynthetic organisms, light is for diatoms both a source of energy and of information about their surrounding environment. Light in the open ocean is blue-enriched due to the absorption of red and longer wavelengths by water; in more turbid waters (coastal regions for example), blue is absorbed by other components and the light field is enriched in green light. Light shapes different aspects of diatom life, such as photosynthetic acclimation, cell cycle control, sexual reproduction and movement. These traits are controlled by light-driven processes such as photosynthesis and photoperception, through specialized proteins that absorb light. Genomic studies revealed that diatoms possess different types of photoreceptors: blue light sensing Aureochromes and cryptochrome-like proteins, putative green-sensing rhodopsins, but also red/far-red sensing phytochromes, which is surprising given the light environment diatom live in.

Recent functional investigation showed that diatoms phytochromes (DPH) are bona fide red/far-red photoreceptors *in vitro*, and regulate gene expression in response to far-red light in the model diatom *Phaeodactylum tricornutum*. Despite the availability of DPH knockout in *P. tricornutum*, the physiological function of this photoreceptor is still unknown.

My PhD work was designed to get insights into the role and significance of DPH for diatom life in the marine environment. To address this question, different research approaches were developed. By combining bioinformatics, mathematical modeling of environmental light and DPH response to it, experimental approaches of genetics and biochemistry on molecular model diatoms, this work has provided a much complete view of DPH action in diatoms.

## DIATOM ROLES IN THE BIOGEOCHEMICAL CYCLES OF CARBON AND SILICA

70% of our planet's surface is covered by ocean. Marine ecosystems contribute to about 50 Gt of net carbon fixation (GtC) per year, which is equivalent to the contribution of terrestrial ecosystems (Field et al., 1998). If the contribution of land and sea in the global primary production is roughly equal, primary produced biomass is drastically different (Bar-On et al., 2018). Indeed, primary producers represent 450 Gt C on land, sustaining 20GtC of consumers biomass, while in the marine environment there is 1 GtC of primary producers and 5 GtC of consumers (Bar-On et al., 2018). This reversed pyramid is explained by the high turnover rate of fixed carbon in the marine environment, and high efficiency in the transfer to higher trophic levels. Indeed, marine fixed carbon can fuel higher trophic levels with high efficiency (10% of the energy is transferred to higher trophic levels in the ocean, compared to 1% in terrestrial environments (Trebilco et al., 2013), or exported to the depth, where it can be re-mineralized as CO<sub>2</sub> or be buried in the sediments for geological times if reaches the bathypelagic layer.

Among phytoplankton (marine primary producers), diatoms play an important role. They are considered as responsible for 40% of marine primary production (Nelson et al., 1995)

and constitute about 40% of the particulate organic carbon that can be exported to the depth (Jin et al., 2006). This is due to their high sinking rate, which is linked to the diatom's large size and silicified cell wall, and to their tendency to form aggregates (Sarhou et al., 2005). However, large variations in cell size and silicification (Si/C ratio) exist both in different diatom species and in different environmental conditions, and all diatoms do not contribute to the carbon export in the same way (reviewed in Tréguer et al., 2017). Numerical models predicting carbon export in the ocean have to take into account this heterogeneity and the wide diatom diversity in the ocean (Malviya et al., 2016). These models are complex and difficult to accurately parameterize, and the exact contribution of diatoms to biological pump (export of carbon to the depth) remains to be quantified (Tréguer et al., 2017). The role of nanoplanktonic diatoms, which are difficult to observe due to their small size (<5µm), is gaining interest, and there is evidence that although small, these diatoms contribute to the export of carbon (Leblanc et al., 2018).

As the main silicified organism in the ocean, diatoms also play a tremendous role in the cycle of silica (Si), which is linked to the cycle of carbon and nitrogen in the ocean (Tréguer and De La Rocha, 2013). Diatoms fix 240 Tmoles per year of dissolved Si in their frustule, of which 6.5 Tmoles are buried into the sediments (Struyf et al., 2009).

Diatoms as primary producers thus play a tremendous role in the marine trophic chain, and their silica cell wall gives them important roles in the ocean biogeochemical cycles.

## DIATOM ORIGIN AND EVOLUTION

Diatoms belong to the Stramenopile, Alveolata, Rhizaria (SAR) eukaryotic lineage (Burki et al., 2020). The branching of SAR compared to other major eukaryotic lineages is still uncertain (Fig. 1). Inside the SAR supergroup, diatoms belong to the Stramenopile group, and more precisely to the Ochrophyta clade, which are photosynthetic Stramenopiles. In addition to diatoms, this clade includes a large diversity of ecologically important organisms, especially in marine environments: multicellular brown algae, which are the main primary producers in tidal and sub-tidal regions (next to the shores) (Bringloe et al., 2020), Chrysophytes, that show a variability of trophic modes and can be important

grazers or important primary producers depending on the species and the environment (Olefeld et al., 2018), or Pelagophytes, which can be important components of the picophytoplankton (size <math>< 2\mu\text{m}</math>) in tropical regions (Guérin et al., 2021).

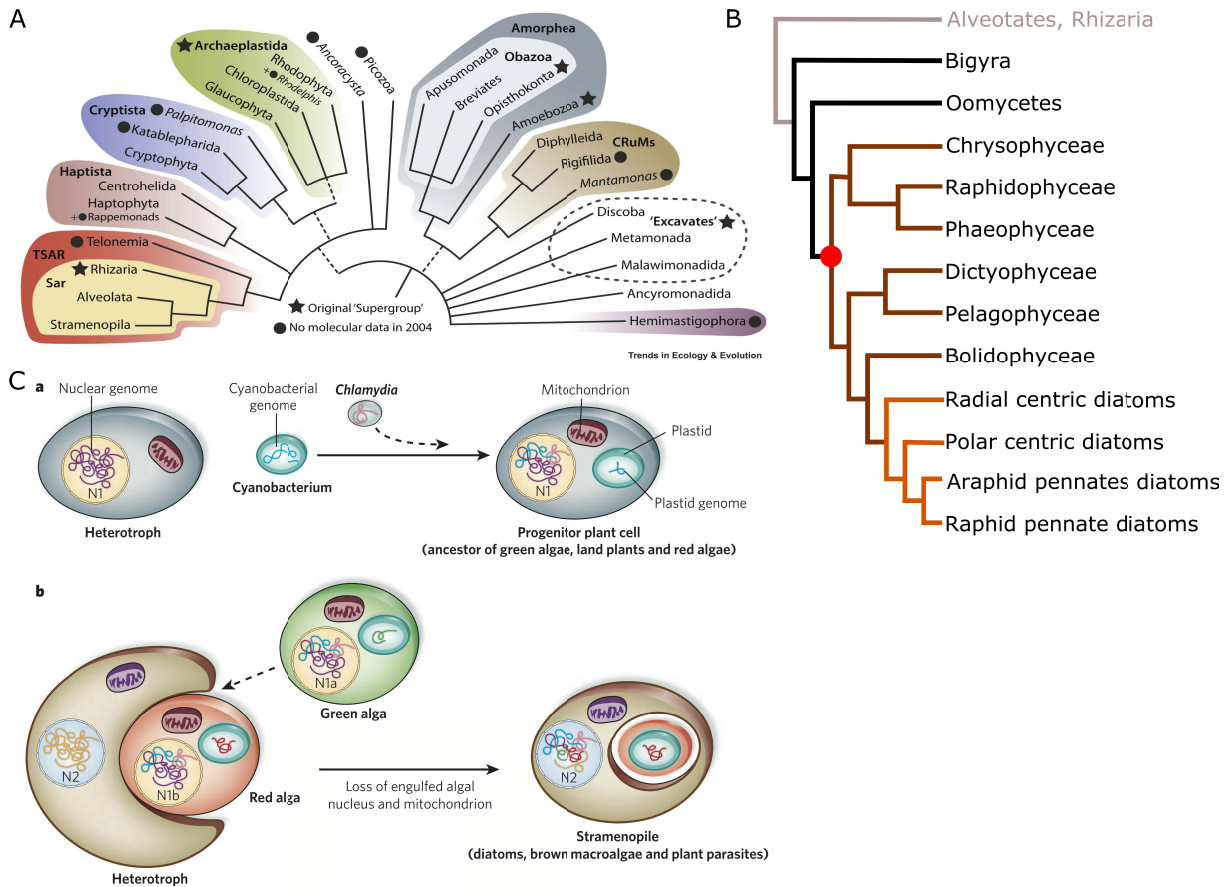


Figure 1. Diatom evolution. A, position of Stramenopiles in the Eukaryotic tree of life (from (Burki et al., 2020)). B, Close up on the Stramenopile lineage, and position of Ochrophyta (brown branches) and diatoms within Ochrophyta (Orange branches). The red dot indicates the acquisition of photosynthesis in the ancestor of Ochrophyta. Adapted from (Azuma et al., 2022; Dorrell et al., 2022). C, The complex history of endosymbiosis in diatom. a. Primary endosymbiosis and b. Higher order endosymbiosis, giving rise to a plastid of red algal origin; Genes from green algae are found in the nuclear genome, suggesting previous association with a green alga. From (Armbrust, 2009)

Primary endosymbiosis occurred about 900 million years ago with the association of a cyanobacteria and a eukaryotic host, giving rise to the Archaeplastidia (Shih and Matzke, 2013). Subsequently, secondary endosymbiosis event, i.e the engulfment and stable association of a photosynthetic eukaryote from the Archaeplastidia lineage by a (heterotroph) eukaryote, independently occurred multiple times during the evolution, with several examples of plastids of green algae origins (Jackson et al., 2018) and red algae

origin (Keeling, 2013). Higher order endosymbiosis events also exist especially in dinoflagellates (Alveolate). Different eukaryotic lineages carry a plastid of red algal origin, such as Cryptophytes, Haptophytes and Ochrophytes. Plastid phylogenies placed red-origin plastids of these algae as a single clade, which led to the idea that secondary endosymbiosis of a red algae chloroplast is a single event that led to the different eukaryotic lineages with red alga chloroplasts (Cryptophytes, Haptophytes and Ochrophytes; this is the chromalveolate hypothesis) (Keeling, 2013). However, large-scale hosts phylogenies place the different eukaryotic hosts on different eukaryotic branches, and the support for this hypothesis has diminished (Burki et al., 2016; Burki et al., 2020). One hypothesis could be that the red-alga derived plastid from different eukaryotic branches is the result of independent symbiosis of closely related red alga. Alternative hypotheses to conciliate plastid and host phylogenies proposed that a single secondary endosymbiotic event (a cryptophyte associating with a red alga) was followed by higher order endosymbiosis events (the ancestor of Ochrophyta adopting a Cryptophyte, followed by the ancestor of haptophyte adopting an early Ochrophyte, although the order is not clearly established): this is the “Cryptophyte first” hypothesis (Stiller et al., 2014).

In Ochrophyta, the early events leading to plastid acquisition are not clearly established, but there is a common agreement that these algae are the result from a single endosymbiosis event, giving rise to their plastid of ultimate red alga origin (red dot in Fig.1B) (Stiller et al., 2014; Burki et al., 2016; Dorrell et al., 2017; Strassert et al., 2021; Azuma et al., 2022; Dorrell et al., 2022).

Diatoms, and Ochrophyta in general, possess plastid-localized proteins, which are encoded by genes of red algal origin in the nucleus; these are supposed to result from endosymbiotic gene transfer from the red alga symbiont to the host. In addition, diatoms possess a number of genes (about 2% of *Phaeodactylum tricornutum* genes) of green algal origin, addressed or not to the chloroplast. This could be the remains of an association with a green algae preceding the establishment of the endosymbiosis with the red algae, that gave rise to the actual chloroplast (Moustafa et al., 2009). Diatoms also have genes of bacterial origin (2.3% in *Phaeodactylum tricornutum*), supposed to be the result of horizontal gene transfer. Diatom genomes are thus a mosaic of genes from different

origins, which might explain their ecological success. For example: the LHCX proteins, involved in photoprotection (see also below, photosynthesis section), are supposed of green algal origin; ISIP2a, a membrane protein allowing iron uptake, is also related to green alga proteins; finally, diatom living in sea ice possess ice-binding proteins of bacterial origin (Dorrell et al., 2022).

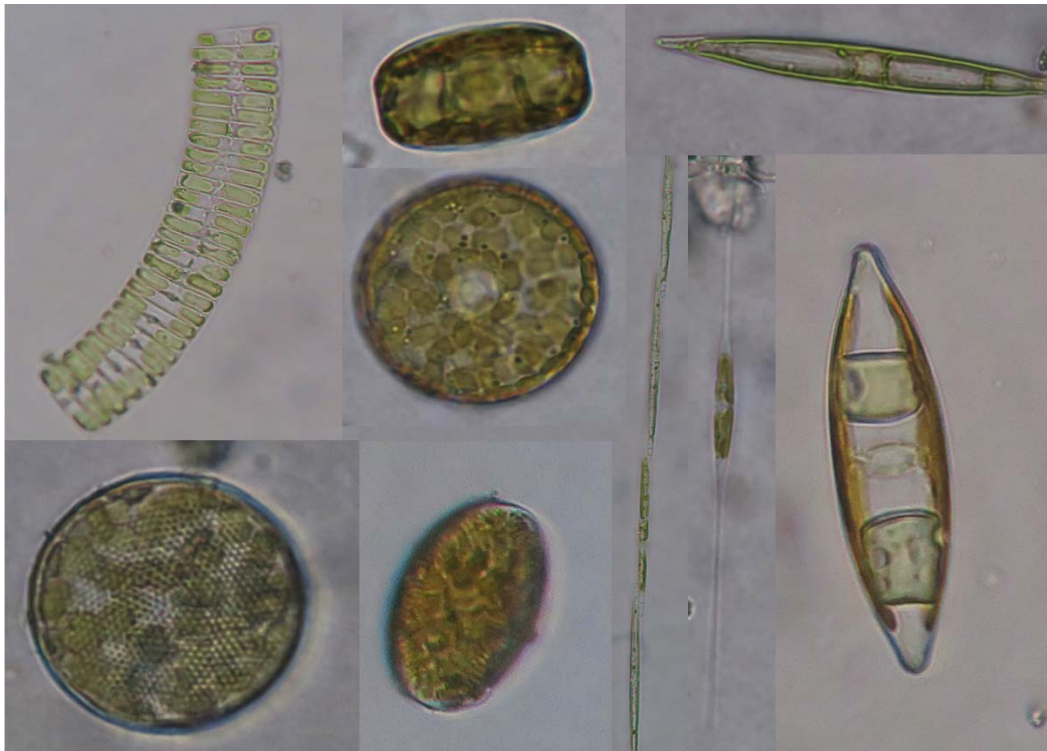


Figure 2. Diversity of diatoms. Diatoms sampled during the Dark Edge cruise in the Arctic in October 2021 (coordinated by Prof. M. Babin). Upper left, *Fragilariopsis* sp. chains (pennate species), middle centric species, probably *Porosira* (upper 2) and *Thalassiosira* species. Right, pennate diatoms, probably *Navicula* (upper and right), the chain forming *Pseudo-Nitzschia* and the needle-like *Cylindrotheca* (middle).

Diatoms are historically separated into 2 groups based on their shape: the centric diatoms, which are cylindrical, and the more elongated pennate diatoms (Fig. 2). Centric species are further divided into radial and polar centric, and pennates into araphid and raphid based on the presence of the raphe, a structure along the frustule that enables gliding movements on a surface. However, this classification does not reflect diatom evolution (See Figure 1) (Medlin, 2016). Thanks to their silica frustule, diatom fossil records can be traced back, and with the combined use of molecular and fossil data, diatom important evolutionary transitions can be dated. Diatoms are clearly found in



fossil records from 180 million years ago (Ma) in the Jurassic period, but some report earlier rise of diatoms after the Permian-Triassic mass extinction crisis (250Ma) (Sims et al., 2006; Sorhannus, 2007). First diatoms were radial centric diatoms, and diversified during the Cretaceous. Multipolar centric, then araphid Pennate diatoms appear during the Cretaceous; the apparition of raphid penates is dated from the late Cretaceous. This group diversified rapidly and nowadays, raphid pennate species outnumber araphid pennate and centric species combined (Nakov et al., 2018).

Diatoms are today a major lineage of photosynthetic eukaryotes, and they are considered to be the most species-rich clade of algae (Kooistra et al., 2007; Mann and Vanormelingen, 2013; Nakov et al., 2018). They come in a diversity of size and beautiful shapes (Fig. 2) and occupy a wide range of ecological niches in the contemporary ocean (Malviya et al., 2016).

## LIFE IN THE OCEAN: CHARACTERISTICS OF THE DIATOM ENVIRONMENT

Phytoplankton growth is essentially controlled by light and nutrient availability (nitrogen, phosphorous, silica and iron, but also vitamin B12), plus eventually by CO<sub>2</sub> concentration and temperature. Diatoms have colonized very different environmental niches and can display different adaptive strategies.

### DIATOMS IN THE WATER COLUMN: PLANKTONIC LIFESTYLE

Diatoms can have a planktonic lifestyle, i.e. drifting in fresh or salted water. I will describe here the global framework to understand phytoplankton life in the oceans. In the water column, light and nutrients usually show opposite gradients with depth: light is high at the surface and decreases exponentially with depth; nutrients are low at the surface and high at depth (Fig. 3A). The “euphotic zone” defines the upper ocean layer that is illuminated (usually defined as the water layer where light intensity is higher than 1% of the surface light intensity) and where photosynthesis can take place.

In parallel, temperature is usually high at the surface and decreases at depth (Fig. 3). In a water column with no or little turbulence, the nutrients in the photic zone will be depleted

by phytoplankton growth and exported to the depth due to dead cells sinking. Phytoplankton growth is then nutrient-limited, and this can result in a “deep chlorophyll maximum” (DCM), with a higher phytoplankton concentration at a depth reflecting a trade off between light and nutrient requirements (Fig. 3A) (Mann and Lazier, 2006). This situation is common all year long in Tropical regions and in late summer in temperate waters. In temperate regions, seasonal variation will change the vertical structure of the water column. In winter, the upper oceanic layer is mixed to an important depth due to high turbulence (wind at the surface, storms) and nutrients from deeper waters are brought back to the surface (Fig. 3B). In this upper mixed layer, the physico-chemical parameters of the water are homogenous (temperature, density, nutrient concentration) (Mann and Lazier, 2006). However, light is in general lower in winter, for astronomic reasons (shorter photoperiod and lower sun angles, so a lower light intensity) and if the mixed layer is deeper than the photic zone, cells are periodically taken to the darker deep waters out of the photic zone. Phytoplankton growth is therefore considered to be light-limited. In spring, light limitation is alleviated either by the increase in day length and light intensity and/or by the mixed layer thinning that allows cells to stay longer in the photic zone. Nutrients are still high for the winter mixing events, and usually sustain spring blooms. Later in the season, increase in temperature will cause thermal stratification of the water column. In the photic zone, nutrients are quickly used for phytoplankton growth; some of it may be recycled within photic zone through grazers and cell lysis, but sinking of dead cells (eventually in aggregates and fecal pellets from grazers) will export nutrients to the deep and deplete the upper layer. Low mixing prevents the replenishment of nutrient stock at the surface and subsurface or deep chlorophyll maximum can appear in these conditions (Fig 3B) (Mann and Lazier, 2006). Finally, in autumn, the deepening of the mixer layer can sustain fall bloom in some regions.

Some marine environments show specificities that differentiate them from the global model presented above. In upwelling region, wind-induced currents cause nutrient-rich water from the deep to arise to the surface (Mann and Lazier, 2006). Coastal regions are

further influenced by tides and freshwater run-off from land, which also show seasonal rhythms (equinox tides, river discharge increase in spring).

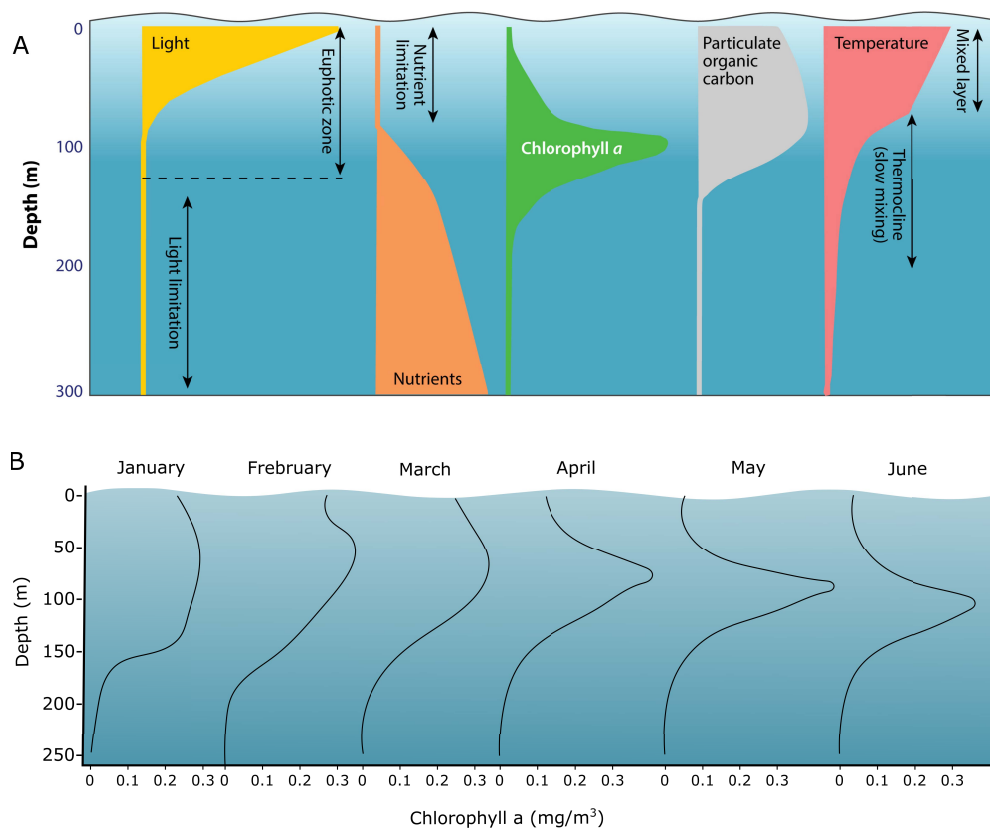


Figure 3. Environmental context for phytoplankton life in the water column. A. Biotic and abiotic parameters across the water column in stratified waters. Light is high at the surface, while nutrients are high at depth, leading to the formation of the maximum of chlorophyll and phytoplankton concentration at depth (DCM). Drawn by (Pierella Karlusich et al., 2020) with data from Bermuda Atlantic Time Series station. B. Seasonal variation of the chlorophyll profile with depth. In January, deep mixing brings enough nutrients to the photic zone to sustain phytoplankton growth. From winter to spring, increased stratification leads to the formation of a DCM, explained in A. Drawn from (Sauzède et al., 2015), data from the same station as A.

Several studies showed that species succession during the season is remarkably conserved from year to year (Caracciolo et al., 2021; Longobardi et al., 2022). Both papers concluded that time of year per se, i.e. astronomical parameters such as photoperiod (Longobardi et al., 2022) and light intensity (Caracciolo et al., 2021), are primary determinants of phytoplankton community succession. Internal factors such as biotic interactions also control self-organization patterns and annual oscillations (Caracciolo et al., 2021).

Diatoms are usually considered to thrive in high nutrient, highly mixed environments (i.e. “spring-like” waters) (Margalef, 1978). Indeed, they often dominate the spring bloom and they are thought to be well adapted to these highly variable environments (especially to light variations with mixing (Lavaud et al., 2007; Lavaud and Lepetit, 2013)). However, diatoms are found in all marine regions, including tropical regions (Malviya et al., 2016). A recent review highlighted possible adaptation strategies to stratified water columns in diatoms (Kemp and Villareal, 2018). Indeed, diatom can be a major contributor of the DCM in stratified waters by adapting to low light levels (“shade flora” (Kemp et al., 2000; Goldman and McGillicuddy, 2003)), but also by regulating their buoyancy to perform vertical migration to acquire nutrient at depth and return to the near surface at a speed up to 7m/h (Moore and Villareal, 1996; Singler and Villareal, 2005). In addition, diatoms are able to store nutrients, including nitrate, in their vacuoles, allowing them to store nutrients during sporadic mixing events and to divide at higher rates than their competitors in nutrient scarce conditions (Dortch et al., 1984; Behrenfeld et al., 2021). Diatoms also possess a complete urea cycle, allowing them to rapidly recover from nutrient starvation (Allen et al., 2011). Finally, diatom-diazotroph association fuel diatom and phytoplankton community with newly fixed nitrogen, and alleviates nitrogen limitation (Singler and Villareal, 2005; Foster et al., 2011).

---

## DIATOMS IN MARINE SEDIMENTS: BENTHIC ORGANISMS

The benthic zone is the bottom layer of a water body, including the sediment surface and its first centimeters. More generally, we call “benthic” any algae growing on a substrate (organic or not). This can include very different environments, as tidal mudflats, surf beaches, coral reefs, coastal sediments or organic surfaces such as sea turtle shells or macroalgae blades. Benthic diatoms play major roles in this ecosystem: they are primary producers at the base of the trophic chain, they mediate exchanges at the sediment-water interface, and they have major roles in stabilizing the sediments through the secretion of Extracellular Polymeric Substances (EPS) (Underwood and Kromkamp, 1999; Middelburg et al., 2000; Serôdio and Catarino, 2000; Underwood and Paterson, 2003).

Sediments are very peculiar environments compared to the water column. Light is attenuated by the sediments at the  $\mu\text{m}$  scale (compared to meters in the water column), and there are also steep gradients of temperature, nutrients, salinity, and if the microorganism population is dense, gradients of  $\text{O}_2$ , Dissolved Inorganic Carbon (DIC) and pH (Fig.4). In addition, temporal variation of these parameters with tides and day/night cycles can be extreme in the tidal zones (Cartaxana et al., 2016b; Cartaxana et al., 2016a; Marques da Silva et al., 2017).

Diatoms have colonized benthic environments and can dominate the microbenthos (i.e benthic microalgal community) in some regions, for example intertidal mudflats. Some species can shift between benthic and planktonic lifestyle and are called tychoplanktonic. Most of the diatom benthic species are pennate and are extremely diverse (An et al., 2020). Two types of benthic diatoms can be distinguished: epipsammic diatoms, that live attached to sediment grain and are not motile (or very little), as opposed to epipellic diatoms, free-living forms that can migrate daily over several millimeter of sediment. These are mainly raphid pennates that can glide on sediments thanks to their raphe.

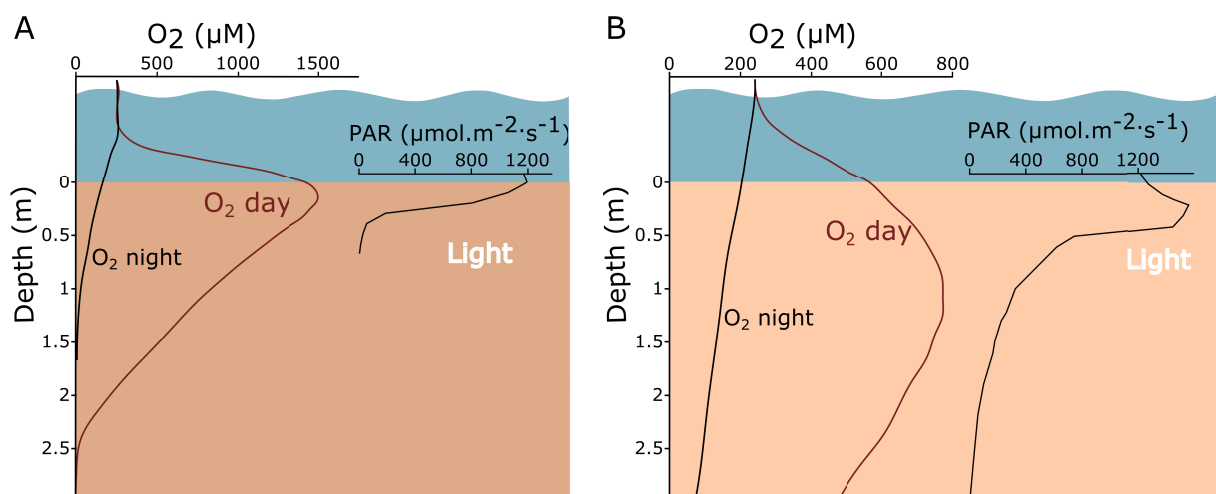


Figure 4. Example of vertical profile of oxygen and light in two types of sediments (A. mud and B. coarse sand) containing diatoms (A. epipellic, B epipellic and epipsammic species). The photic zone in the muddy sediments is very thin, and was densely populated. In the sandy sediment, the photic zone is deeper, and we observe an increase in light intensity in the first 0.1mm at the near surface due to high scattering and low absorption From (Cartaxana et al., 2016b)

Epipellic diatoms show impressive migratory behaviors to cope with these environmental constraints. They usually dominate in fine, soft sediments and silt habitat, and they are

able to move toward light, or away from it, to find optimal light conditions to perform photosynthesis (Fig. 4A). In dense photosynthetic communities, photosynthesis activity depletes inorganic carbon stocks and micromigration can allow alternate between high light/low DIC surface and low light/high DIC subsurface environment, with continuous replacements of the cells at the surface (Consalvey et al., 2004; Vieira et al., 2016; Marques da Silva et al., 2017). Deeper migratory patterns (at the mm scale) are also controlled by diurnal light rhythms (up during the light phase) and tidal rhythms (up during low tide, burial before tidal inundation). These rhythms can persist in continuous condition, highlighting an endogenous control of the timing of this behavior (Palmer and Round, 1967; Barnett et al., 2020). However, in the dark diatoms have to adapt their metabolism: they can grow heterotrophically (Lewin and Lewin, 1960) and in anoxic sediments, they can also perform nitrate respiration based on intracellular stored nitrates (Kamp et al., 2011; Merz et al., 2021).

In sediments with larger grain size, epipellic and epipsammic species can co-exist (Fig. 4B). Light penetrates deeper (see also next paragraph) and epipsammic diatoms (that do not migrate) adapt to high light through physiological photoprotective mechanisms, as opposed to “behavioral” mechanisms such as downward migration of epipellic diatoms (Cartaxana et al., 2011; Barnett et al., 2015). When photosynthetic organisms density is high, as occurs in biofilms, the community can self-organize in layers (niche differentiation). Benthic diatoms usually compose the uppermost layer, with green algae and cyanobacteria underneath. Anoxygenic phototrophs such as green and purple sulfur bacteria can be found below the cyanobacteria, where  $O_2$  is depleted and  $H_2S$  concentrations are high (Stal et al., 1985; Stal et al., 2019).

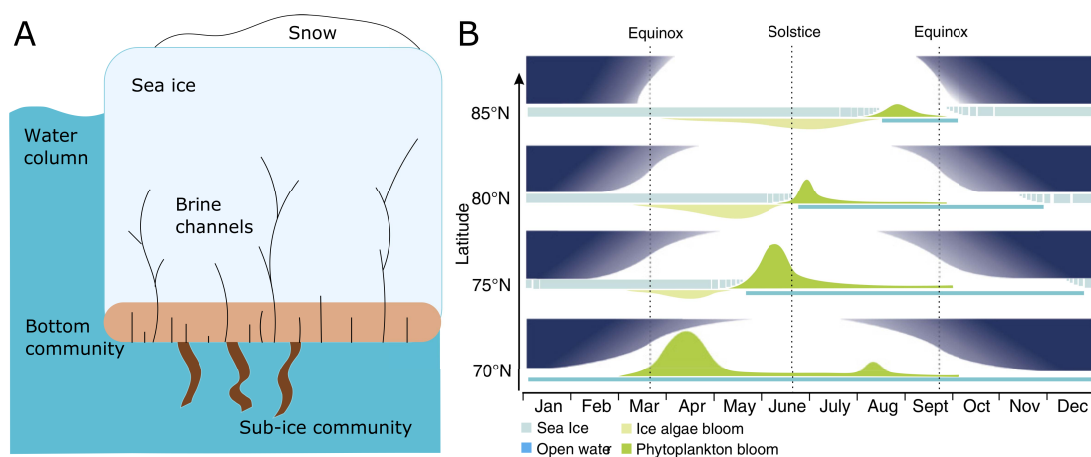


Figure 5. A. Ice associated algal communities B. Time of phytoplankton blooms in the Arctic ecosystem at different latitudes, from (Daase et al., 2021)

One other peculiar environments colonized by diatoms is sea ice (Fig5). Diatoms (especially pennate diatoms) dominate the sea ice community and can represent up to 90% of the algal cell abundance (Arrigo, 2016). Sea ice algae are usually found in the bottom 20cm of the ice sheet, where sufficient nutrients are available from the seawater below. Most common diatom genera in the bottom sea-ice community are *Nitzschia*, *Fragilariopsis* and *Navicula*. The sea ice interior is characterized by low nutrient and high salinity in brine pockets and channels, and thus less favorable to algal growth. Pennate diatom aggregates and chains of the centric diatom *Melosira* can also colonize the underside of the sea ice. The polar algae communities are subjected to drastic environmental changes along the seasons. In autumn, the ice starts to form when temperature fall below  $-1.76^{\circ}\text{C}$  and traps protists into the sea ice (Rózańska et al., 2008). Some algae can form cyst and resting stages/spores to survive the adverse conditions within sea ice. During winter, the polar night and the thickness of sea ice limit light availability for alga, and there is a selection mechanisms causing the overall diversity of the ice community to decrease. The return of the sun and lengthening of the days trigger the sea ice algal bloom, when low-light adapted algae within the sea-ice start to proliferate (Rozanska et al., 2009). At this time of the year, sea-ice bloom is often the sole source of carbon for heterotrophs. Further temperature increase causes the ice to melt, and algae are released to the water column (Leu et al., 2015). Different diatom types are thought to be adapted to the different stages of seasonal progression (Croteau et al., 2021). As in sediments, some pennate diatoms are motile and adjust their position in the sea ice to

optimize their growth (Aumack et al., 2014). Diatoms use EPS secretion and ice-binding proteins to protect themselves from low temperature and adhere and migrate on ice.

---

## THE PECULIARITY OF LIGHT IN AQUATIC ENVIRONEMENTS

The light field is defined both by light intensity and light spectral quality. In terrestrial environments, the light field depends on cloud cover, time of day, time of year, latitude and shade (notably by other photosynthetic organisms). In aquatic environments, an additional dimension has to be considered: depth.

### LIGHT IN THE WATERCOLUMN

As already mentioned before, light intensity decreases with depth in the water column. The main physical processes determining the underwater light field are absorption and scattering, which will both reduce the light intensity. Major light-absorbing components in aquatic environments are water, Colored Dissolved Organic Matter (CDOM) and particulate matter as inanimate particles and phytoplankton (Kirk, 2011).

Water molecules themselves absorb light and can contribute to an important part of light absorption coefficient of natural waters (it can represent 68% of the absorption of Photosynthetically Active Radiations, PAR). The absorption spectra of pure water (Fig. 6) show high absorption in the long wavelength part of the light spectrum (above 500 nm). Salts present in seawater have negligible effect on light attenuation. CDOM, also called gilvin or yellow substance, derives from the decomposition of organic matter. In coastal regions, CDOM is a major component of light absorption coefficient, while it is less concentrated in clear oceanic waters. Source of CDOM is terrestrial organic matter (or resuspended organic matter from sediments). CDOM absorbs mainly in the blue, with an exponential decrease towards longer wavelengths (Fig. 6b and c).



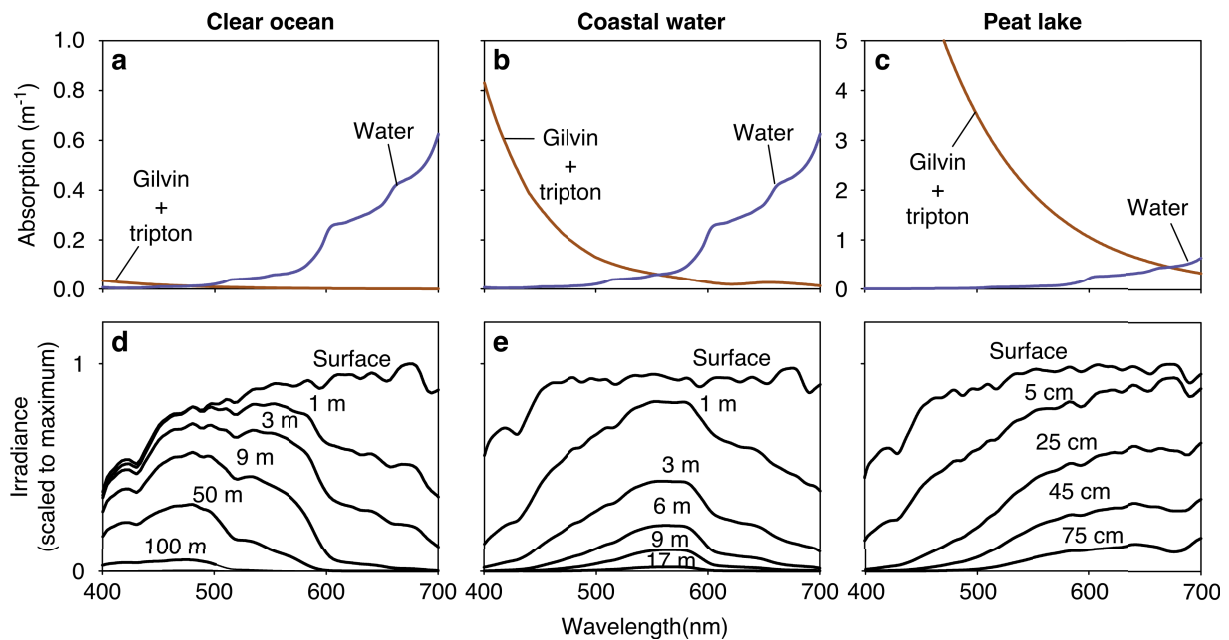


Figure 6. Examples of underwater light fields in different aquatic environments and decomposition of their absorption spectra. A In clear oceanic waters (Pacific Ocean, station ALOHA), light penetrate deep (see spectra at 100m) and is blue enriched due to absorption by water. B, in coastal waters from the Baltic sea, absorption by CDOM and particles (gilvin+tripton) is higher, which leads low light penetration, but with higher penetration of green light. C. Finally, in the Groot Moost lake, a peat lake in the Netherlands, turbidity is so high that light is attenuated in the first meter, and the light field is red-enriched.(Figure from (Stomp et al., 2007)).

Inanimate particulate matter is also called tripton or non-algal particles. It is composed of mineral particles and dead phytoplankton cells, and its absorption spectrum globally has the same shape as CDOM (Fig. 6). Photosynthetic pigments from living phytoplankton cells absorb light at specific wavelengths. Total phytoplankton absorption (both in shape and value) depends on the species composition of the phytoplankton community, their concentration, but also on their physiological state (cell size and intracellular pigment concentration). In general, phytoplankton absorption spectra show a strong absorption in blue and second peak in red corresponding to the chlorophylls (Kirk, 2011).

Scattering is the deviation of a photon from its original path following interaction with a particle or a molecule (Kirk, 2011). In aquatic environments, interaction with particles will increase the optical path for photons, and thus increase their probability of being absorbed; this will intensify the vertical attenuation of light. Scattering in natural waters

is roughly neutral towards wavelengths, with a slight attenuation towards long wavelengths.

Marine waters can be classified based on their optical properties (Morel et al., 2006). Case I waters are oceanic waters, where optical properties are directly linked to phytoplankton abundance. Coastal waters with limited terrigenous inputs can also be considered as case I waters. At low phytoplankton concentration, i.e. in oligotrophic waters, absorption of long wavelengths by water will result in blue-enriched light field at depth (Fig. 6a). Case II waters are waters for which resuspended sediments and dissolved organic matter (for example from river discharge) are important contributors to the optical properties. Absorption in the blue is very high in these waters, resulting in green-shifted light fields. (Fig. 6b). In some extremely turbid ecosystems, absorption of blue and red can lead to a red-enriched light field (Fig. 6c).

#### LIGHT IN SEDIMENTS AND BIOFILMS

Less is known on light attenuation in marine sediments compared to light in the water column. The main protagonists involved in light attenuation are mineral particles (sand, silt, mud) and organic matter (living particulate and dissolved) (Cartaxana et al., 2016b). Here, water absorption is negligible compared to the other actors, but water is an important component as wet sediments scatter light less than dry sediments and light will penetrate deeper in wet sand (Kühl and Jorgensen, 1994). Both absorption and scattering are more intense than in the water column due to the high density of sediment particles and algae living in it, and the depth of the “photic zone” is limited to a few mm (Fig. 4).

The size of the mineral particles is crucial in sediment light attenuation (Kühl et al., 1994), through effect on scattering. Indeed, small particle size scatter light more compared to coarse grains. Light attenuation by sand is higher in the blue (maximum around 450nm), and decreases toward longer wavelengths. Presence of organic matter and microalgae will increase light attenuation in the blue range, with also a peak around 680nm due to chlorophyll absorption (Haardt H, Nielsen Gae, 1980; Kühl and Jorgensen, 1994). In biofilms, i.e. very dense communities of microorganisms, the layers of different organisms will also change the attenuation coefficient with depth due to their different pigment

composition (Fig. 7A) (Kühl and Jorgensen, 1992), for example chlorophylls and bacteriochlorophylls.

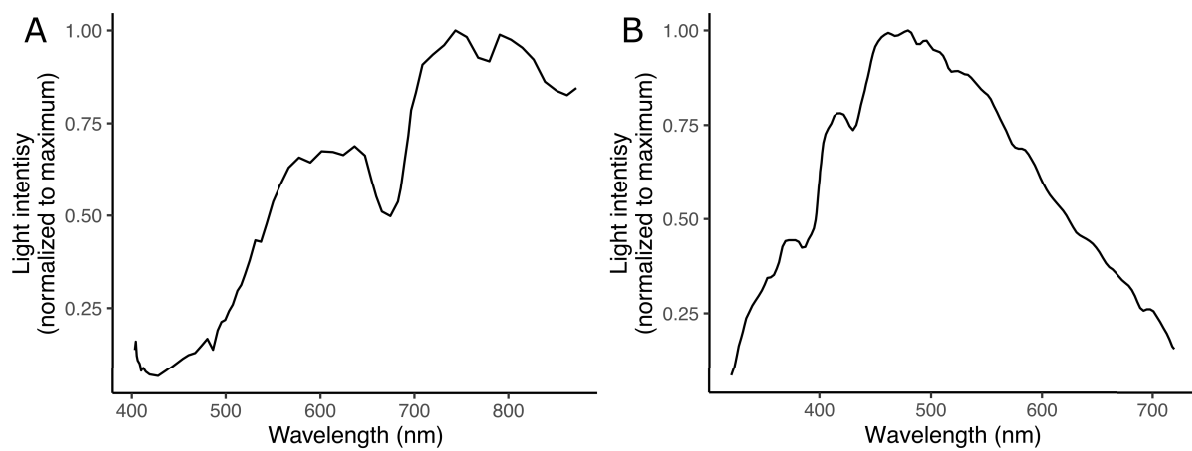


Figure 7. Light field sediments and in ice. A. Light under 1mm of coarse-grained coastal sediment occupied by diatoms and cyanobacteria. Note that light is more absorbed at 680nm due the chlorophyll a. From (Lassen et al., 1992). B Light under sea-ice. Maximum light penetration in the ice is at 460nm. Data from (Lund-Hansen et al., 2020).

Intense scattering in sediments and in biofilms will also create a “subsurface irradiance maxima”. Indeed, scalar irradiance, i.e. light intensity integrated from all directions, is higher than at surface in the first 0.5 millimeter below the sediments and can reach up to 280% of the surface (downwelling) irradiance in fine sand (Kühl et al., 1994). One other difference with the water column environment is the exposure to high UV light. If both the sea surface and the sediment surface are exposed to direct sunlight, the cells are diluted in the full volume of the mixer layer in the ocean, while they are concentrated to the surface and first mm in the sediment. Algae in sediment are thus exposed to much higher doses of UV light (Garcia-Pichel and Bebout, 1996).

As water, ice absorbs mostly in red part of the light spectra, resulting in blue-enriched light fields (Fig. 7B)(Ehn and Mundy, 2013). However, in the ice algae layer, the algae density is so high that absorption in the blue becomes important, and the light field can shift towards green. In addition, ice is a highly scattering media resulting in reduced light intensity.

## LIFE CYCLE

All diatom species sequenced so far are thought to be diploid in their vegetative states and divide mitotically during the major part of their life cycle (Falciatore et al., 2020). Diatoms have a “bloom and burst” life cycle, with periods of intense replication when the environmental conditions are favorable. The bloom's dynamics are not fully understood, and the bloom's demise can be due to different parameters (not mutually exclusive), for example nutrient depletion, grazing pressure, or pathogen attack (Behrenfeld et al., 2021).

Some diatoms can form resting stages to survive unfavorable environmental conditions and resume growth when better conditions are met. The cues that trigger resting stage formation are not clear for all species, but high cell density and nitrogen limitation are suspected (McQuoid and Hobson, 1996; Pelusi et al., 2020). Two types of resting stages can exist: resting spores, which are morphologically different from vegetative growing cells, and resting cells. Both have a reduced metabolism, and resting cells are able to successfully germinate after centuries (*Skeletonema marinoi* (Harnstrom et al., 2011)) and even millennia (*Chaetoceros muelleri* (Sanyal et al., 2021)) in sediments. Resting stages are important in diatom species succession, and resting stage stocks accumulated in sediments can form a reservoir of viable cells, seeding blooms in spring in coastal regions (Eilertsen et al., 1995)

Some aspects of the diatom life cycles are shaped by their silica cell wall, the frustule. The frustule is made of two pieces that fit together like a Petri dish and encapsulate the cell. The two pieces, also called theca, are slightly different in size: the larger one (like the Petri dish lid) is called the epitheca, and the smaller one is the hypotheca. Due to this rigid silica cell wall, daughter cells following mitotic division are different (see Fig. 8). One inherited the epitheca and will result in cell size comparable with the mother cell, while the cell that inherited the small hypotheca will form a new hypotheca inside the maternally inherited one (using it as a template), resulting in smaller cell size. This results in size reduction with mitotic division, which can end with cell death if the cell size becomes too small to

allow survival. In some diatoms, sexual reproduction is the only stage where size restoration can occur (Bilcke et al., 2022) (Fig. 8).

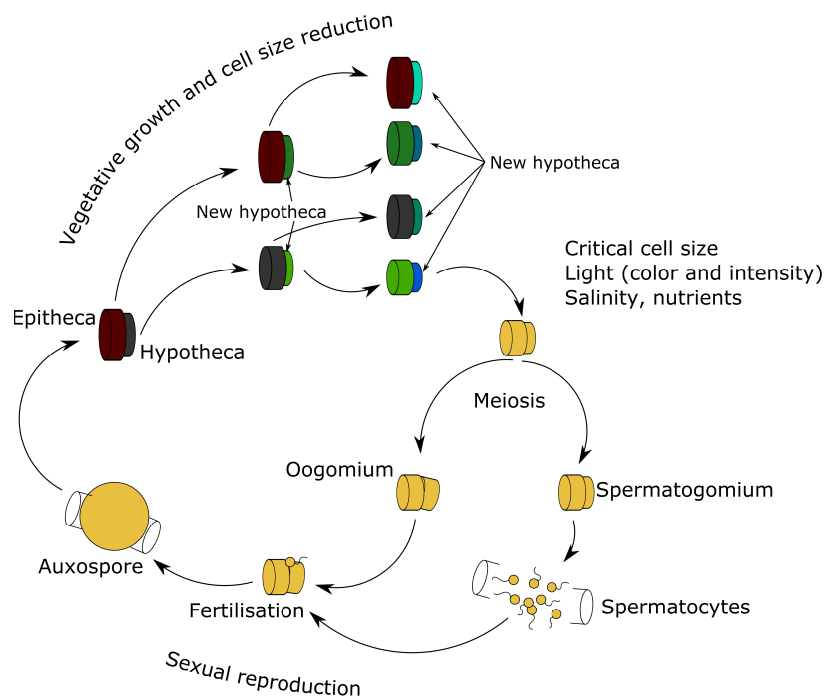


Figure 8. Diatom life cycle. Clockwise from the red diatom: during mitotic division, each daughter cell inherits a theca and build a new hypotheca. This results in cell size reduction. When critical cell size is reached, or in response to other cues, sexual reproduction can occur, restoring cell size (yellow cells: example of the sexual reproduction in *Thalassiosira weissflogii*, redrawn from (Bilcke et al., 2022)).

Sexual reproduction can be triggered when critical cell size is reached, but also in response to environmental cues (nutrient, salinity, pheromones, light) (Amato, 2010; Godhe et al., 2014; Moeys et al., 2016; Basu et al., 2017; Moore et al., 2017). Compared to the diploid part of diatom life cycle, which can last years, sexual reproduction is short and has rarely been observed in nature. Centric and pennate species show differences in their sexual cycles (which are arguments to classify them too). In general, centric species are homothallic, i.e. clonal cells can generate large oogonia or small, motile spermatocytes (Fig. 8). Pennate are usually heterothallic, i.e. two mating type are required to perform sexual reproduction (Bilcke et al., 2022). In raphid pennate, compatible mating type cells must pair to form the gametangia that will undergo meiosis and gametogenesis. Here, the two gametes mating types are indistinguishable (isogamy). In all diatom groups, fertilization results in a large cell called auxospore that will resume mitotic division and vegetative life phase. The model diatom *P. tricornutum* and *T. pseudonana* are presumed

to predominantly reproduce asexually in the lab and are rarely used to study sexual reproduction (one exception is the study by Moore et al., 2017). Other species have arisen as model species for sexual reproduction, mainly the pennate species *Seminavis robusta* and *Pseudo-nitzschia multistriata*.

Numerous studies reported the effect of light on the diatom life cycle. Blue light is especially efficient in controlling mitotic cell cycle progression in *P. tricornutum* through blue light photoreceptor-mediated control of cyclin expression (Huysman et al., 2013). Resting spore germination is sensitive to photoperiod (Eilertsen et al., 1995) and light quality, especially to blue light (*Skeletonema costatum*, *Thalassiosira minima*, and *Chaetoceros sp.* (Shikata et al., 2009)) and blue and red light through photosynthesis activity in *Leptocylindrus danicus* (Shikata et al., 2011). Finally, sexual reproduction is controlled by red and far-red light in the diatom *Stephanopyxis palmeria* in a reversible manner (induction by red light, repression by far-red light applied after red light) (Ren et al., 2008). Red light also triggers sexual reproduction in the pennate diatom *Haslea ostreicola* (Mouget et al., 2009), but in *Seminavis robusta*, blue light is the more efficient waveband (Bilcke et al., 2021). Light intensity also matters: in *Seminavis robusta*, higher reproduction rate occurs at the lowest intensity tested (4  $\mu\text{mol photons/m}^2/\text{s}$ , Bilcke et al., 2021) and in *Thalassiosira weissflogii* a period of darkness and/or low light are necessary to trigger sexual reproduction (Armbrust, 1990).

---

## PHOTOSYNTHESIS

As mentioned earlier, there is a common agreement that diatom chloroplasts arose from red algae. Most of the components of oxygenic photosynthesis are conserved in diatoms, but they also possess unique features linked to their complex evolutionary history (Büchel et al., 2022).

The chloroplasts possess a network of membranes containing the photosynthetic apparatus: the thylakoids. Photosynthesis takes place in 2 phases: the “light reactions”, where light energy is harvested and transformed into chemical energy, and the “biochemical reactions”, when this chemical energy is used to fix carbon ( $\text{CO}_2$ ). The

photosynthetic chain in the thylakoids carries out the light reactions (Fig. 9). Protein complexes are embedded in the thylakoid membrane: photosystems (PS) I and II, cytochrome b6f complexes and ATP synthase. Photosystems are composed of core reaction centers containing a special pair of chlorophyll a and Light harvesting complexes (LHC) containing pigments. Light is absorbed by pigments in the LHC or directly by the 2 chlorophylls in the reaction centers and light energy transferred to the reaction centers where charge separation occurs. This fuels a flow of electrons in the photosynthetic chain from water ( $H_2O$ , with production of  $O_2$ ) to  $NADP^+$ , producing reducing power in the form of NADPH. The electron flow also allows the build-up a proton gradient across the thylakoid membrane. The return of  $H^+$  from the lumen to the stroma fuels the production of ATP by the ATP synthase.

Compared to the green algae and plants, diatom-specificity contributing to the light reactions are notably the spatial organization of protein complexes in the plastid and the protein and pigment content (Büchel et al., 2022). Diatom chloroplasts are embedded in 4 membranes, and their thylakoids are organized in stacks of 3 thylakoids (Fig. 9), and the different complexes are heterogeneously localized: PSI are rather in the stroma-facing membranes, while PSII are enriched in the core of the stacked membranes, facing other thylakoid membranes. ATP synthase is in the outer, stroma-facing membranes. This organization can change with light conditions (see also below).

The light harvesting complexes of diatoms differ from other eukaryotes (Büchel et al., 2022). LHC proteins can be grouped in 3 main families: LHCF, the major LHC group; LHCR, associated with PSI and related to red alga LCFs; and LHCX, that are of supposed green alga origin and are involved in photoprotection. In addition to chlorophyll a, LHCs bind chlorophyll c and fucoxanthins. This carotenoid pigment is specific for Ochrophyta and gives these algae their brownish color.

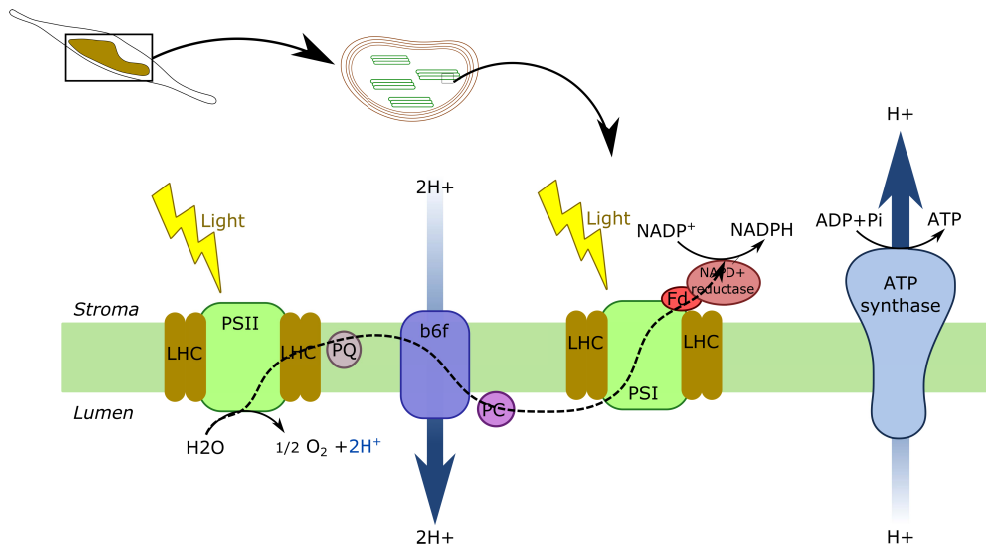


Figure 9. The photosynthetic light reactions in diatoms (the dotted line illustrates the linear flow of electron across the photosynthetic complexes).

Excess light energy can damage photosystems, especially the PSII and therefore photosynthetic organism evolved sophisticated photoprotection mechanisms. In addition to efficient PSII repair systems, diatoms can also quickly adjust PSII activity and dissipate excess energy: this can be detected *in vivo* by a decrease in PSII fluorescence, and is called Non Photochemical Quenching (NPQ) (Lepetit et al., 2022). At least part of the NPQ capacity is related to a xanthophyll cycle involving the conversion of diadinoxanthin to diatoxanthin pigment under high light. This xanthophyll cycle is different from the one found in plants, both in terms of pigments (plants use violaxanthin, antheraxanthin and zeaxanthin pigments) and enzymatic reactions (two-step process in plants and only one in diatom). LHCX antenna are also involved in NPQ regulation, but the exact mechanisms is unknown (Bailleul et al., 2010).

Acclimation to both light intensity and light quality occurs in diatoms. High light acclimation is depicted by high maximal photosynthetic rate with high photoprotective mechanisms (high NPQ) (Schellenberger Costa et al., 2013). In *P. tricornutum*, blue wavelengths are sufficient to trigger this phenomenon, through aureochromes and cryptochromes photoreceptors (Juhás et al., 2014; Mann et al., 2017) which control expression of light-harvesting proteins. Few studies suggest that chromatic acclimation, i.e. acclimation to light quality, can also occurs and identified specific photosynthetic



hallmarks for cells grown in red light (Fujita and Ohki, 2004; Herbstová et al., 2015; Herbstová et al., 2017). In *P. tricornutum* and *Nitzschia closterium*, cells grown in red light showed a fluorescence spectra peak in FR light, associated with the presence of a specialized LHC in *P. tricornutum*, Lhcf15 (for more details, see chapter 3). A re-organization of the photosynthetic complexes in the thylakoid membrane was also described (Bína et al., 2016), with a stronger spatial heterogeneity. Schellenberg Costa and coauthors (Schellenberger Costa et al., 2013) compared chloroplast membrane proteins of *P. tricornutum* cells grown under blue or red light and identified proteins specific for the different light conditions (10 upregulated in blue light and linked to high light acclimation, 4 upregulated in red light including Lhcf15). Transcriptomic responses to changes in light quality, although at high intensity, go in the same direction, with an induction of many LHCF and LHCR genes in green and red, while blue light induced LHCX genes and genes involved PSII repair (Valle et al., 2014).

Fixation of CO<sub>2</sub> by Rubisco in the Calvin-Benson-Bassham cycle is the biochemical phase of photosynthesis. However, in aquatic environments, CO<sub>2</sub> concentrations are very low, as CO<sub>2</sub> dissolved in H<sub>2</sub>CO<sub>2</sub>, which dissociates in HCO<sub>3</sub><sup>-</sup>. The majority Dissolved Inorganic Carbon (DIC) in water is in the form of HCO<sub>3</sub><sup>-</sup>, which cannot cross membranes, while CO<sub>2</sub> can. Diatoms have developed efficient Carbon Concentration Mechanisms (CCM) to cope with the low CO<sub>2</sub> concentration in water (Kroth and Matsuda, 2022). These algae possess several HCO<sub>3</sub><sup>-</sup> transporters, some of which can pump HCO<sub>3</sub><sup>-</sup> inside the cell, and carbonic anhydrases converting HCO<sub>3</sub><sup>-</sup> to CO<sub>2</sub> and vice-versa in the different compartments. Some of these enzymes are thought to be secreted and convert HCO<sub>3</sub><sup>-</sup> to CO<sub>2</sub> at the cell surface. CO<sub>2</sub> passively crosses the plasma membrane and is converted to HCO<sub>3</sub><sup>-</sup> in the cytosol, to which we should add HCO<sub>3</sub><sup>-</sup> actively pumped at the plasma membrane. HCO<sub>3</sub><sup>-</sup> is then transported inside the chloroplast by unidentified transporters, to the stroma where it is transformed back into CO<sub>2</sub> in the vicinity of RubisCO. This CCM is an important mechanism for diatoms in the wild, as mining surface seawater metagenomes and metatranscriptomes revealed an important number of carbonic anhydrase transcripts (Pierella Karlusich et al., 2021). Some studies showed that some diatoms possess genes homologs to genes involved in the C<sub>4</sub> mechanisms from plant, but functional studies of these genes did not confirm the occurrence of this process (Haimovich-Dayana et al., 2013),

and mining environmental data showed that these genes rarely co-occur, limiting the possible involvement of C4 in the environment (Pierella Karlusich et al., 2021).

---

## DIATOM MOTILITY

Pennate raphid diatoms are able to glide on surface at high speed (up to 25µm/s). Not much is known about the mechanisms of diatom motility, but it involves secretion of mucilage through the raphe (Poulsen et al., 2022). Current models involve the secretion of mucilage via the Golgi and transport of vesicles to the plasma membrane. Extracellular mucilage is linked via transmembrane proteins to myosin motors, themselves connected to actin filaments. Myosin action builds a force along the actin filaments, which pushes the cell in the opposite direction.

Diatom movement is tightly controlled, especially by light. We already underlined diatom movements in sediments. In natural communities, blue light is the most efficient waveband to trigger positive phototaxy (Barnett et al., 2020). Numerous studies on monocultures in laboratory conditions showed the effects of different wavebands: cells are in general more attracted by blue light (*Craticula sp*, *Pinnularia sp*, *Nitzschia sp* (Cohn and Weitzell, 1996); *Nitzschia communis* (Nultsch, 1971), *Nitzschia perminuta* (McLachlan et al., 2009)) and repelled by red (Nultsch, 1971; Cohn and Weitzell, 1996), but this is intensity- and strain-dependent. In addition, changes in illumination can change the direction or the speed of the movement. Finally, the tip of the cell seems to be the site of light perception (Cohn et al., 1999), and further signaling events involve calcium release from intracellular socks (McLachlan et al., 2012).

## LIGHT SENSING

Light is an essential source of information for organisms on Earth. All organisms (photosynthetic or not) possess specialized proteins called photoreceptors that perceive light and mediate the response to this environmental cue. These proteins usually bind a pigment (also called chromophore) that enables them to absorb light at specific

wavelengths (Möglich et al., 2010). Upon light absorption, changes in the chromophore conformation impact protein conformation, changing its activity (protein-protein interaction, phosphorylation or other enzymatic activity for example). This will activate specific signaling cascade(s) leading to the biological responses.

A photoreceptor typically displays several protein domains or modules that mediate different part of the photoreceptors function: light-sensing domains and signal transduction effector domains. Different classes of light-sensing domains are known, with different chromophores and light sensing properties. Three types of blue-light photoreceptors bind flavin chromophores: photoreceptors with a Light-Oxygen Voltage (LOV) domain, which bind FMN, sensors of Blue Light Using FAD (BLUF) and cryptochromes, which bind FAD. Other photoreceptors detect blue light via the 4-hydroxycinnamic chromophore and are called Photoactive Yellow Protein (PYP), while the UVB photoreceptor UVR8 perceives light through absorption by tryptophan residues. Rhodopsins perceive blue and green lights through a retinal chromophore, and finally phytochromes perceive red and far-red lights with a bilin chromophore. The combination of the same type of photosensing domain with different effector domains leads to protein with different functions. As example (see also below), the LOV-based photosensors are found in combination with a variety of effector domains across the three kingdoms of life (Glantz et al., 2016).

Land plants and green algae possess examples of 5 of the above mentioned photoreceptor types (no homologs of PYP or BLUF have been reported in Viridiplantae so far). Cryptochromes, which are photoreceptors known across all eukaryotes, are found in land plant and green algae. UVR8 photosensors are thought to be restricted to the green lineage. LOV domains exist in combination with a serine/threonine kinase domain (phototropins), F-box domain and Kelch repeat (Zeitlupe family), PAS domain (PAS/LOV in land plants), Histidine kinase domain (LOV-HK from *Ostreococcus* and other mameliophyceae), but also in combination with other photosensory domain like phytochrome photosensory module (neochromes ferns and mosses). Different types of rhodopsins also exist in the green algae *Chlamydomonas reinhardtii*, while the phytochromes photoreceptors are conserved in land plants but not in all green algae.

Exploration of diatom genomes has led to the discovery of many putative photoreceptors in these organisms (Jaubert et al., 2017). Some are classical photoreceptors known from bacteria, land plant or animals, while others show unique domain combinations. Photoreceptor types in diatoms include cryptochromes, LOV-based aureochromes, rhodopsins and phytochromes. No homologs of the plant phototropins have been identified in diatoms. Although first identified in *P. tricornutum* (Bowler et al., 2008; Djouani-Tahri et al., 2011), close examination of the putative LOV-HK homologs in diatoms showed that they lack the critical amino acids for interaction with the chromophore. Diatoms have genes with similarity to UVR8, but these lack the critical tryptophans involved in light sensing and are thus not considered as UVB photoreceptors (Fernández et al., 2016).

In the following section, I will describe the current knowledge on the diatom photoreceptors, with a particular focus on phytochromes, which represent the main topic of my PhD. Some paragraphs (Cryptochromes, Aureochromes and Rhodopsins section, marked with \*) have been extracted from a recently review published in the “Molecular Life of Diatom book”, that I co-authored (Jaubert et al, 2022). The full chapter is available at the end of this manuscript, in the annex section. Additional information on new diatom photoreceptors and their distribution are described in the result session (Chapter 1).

---

## DIATOM PHOTORECEPTORS

### **Aureochromes \***

The aureochromes (AUREOs) are unique blue light photoreceptors that possess both an FMN-binding light-oxygen-and-voltage (LOV) domain (Crosson et al. 2003), and a bZIP domain typical for bZIP transcription factors (TFs) (Dröge-Laser et al. 2018) (Fig. 1). The blue light phototropin photoreceptors of green algae and plants also possess two LOV domains, but utilize a serine/threonine (Ser/Thr) kinase domain for signal transduction instead (Christie 2007). AUREOs were originally discovered in the xanthophyte alga *Vaucheria frigida* in 2007 (Takahashi et al. 2007). Their name refers to “aurum” (Latin for gold), because of the golden-brownish colour of most stramenopiles. Takahashi et al. (2007) identified two orthologs in *V. frigida*, VfAUREO1, and VfAUREO2. Using an RNA

interference approach to silence these two genes individually, they demonstrated that both AUREOs are involved in the regulation of photomorphogenic responses. Meanwhile, AUREOs have been identified in other stramenopiles (Ishikawa et al. 2009; Jungandreas et al. 2014) and in a raphidophyte (Ji et al. 2017), but not in non-photosynthetic oomycetes (Kroth et al. 2017). AUREOs are not present in red algae, which are considered to represent the endosymbiotic ancestors of stramenopile plastids (Archibald 2015). This indicates that the ancestral AUREO gene with its unique combination of LOV and bZIP domains may have been provided either by the putative host cell of the secondary endosymbiosis event, or that it evolved very early within the stramenopiles, possibly via domain shuffling (Di Roberto and Peisajovich 2014).

Stramenopile AUREOs differ in their structures. While AUREO1 proteins possess the typical LOV domain, AUREO2 proteins have a mutation within the LOV domain, which prevents non-covalent binding of flavin needed for light absorbance in the blue range. The reason is a steric hindrance from a methionine residue within the binding cavity (Banerjee et al. 2016a). AUREO2, therefore, is not a real photoreceptor, but could still be involved in light regulation, e.g. by forming a dimer with a light-sensing AUREO1 protein. In reciprocal experiments, genetic modification of AUREO2 from *P. tricornutum*-restored flavin binding (Serif 2017), while introducing a point mutation at the same site in PtAUREO1a, led to loss of flavin binding (Banerjee et al. 2016a). Based on this distinction and on phylogenetic analyses, aureochromes in diatoms and other organisms have been classified as either AUREO1-type (a/b/c etc.) or as AUREO2 (Schellenberger Costa et al. 2013b). The algae studied so far all possess one AUREO2 protein and one or more AUREO1 isoforms (Table 2).

Blue light absorption, both in aureochromes and phototropins, causes the formation of an adduct between the flavin and a nearby cysteine within a few microseconds, starting the signalling cascade (Toyooka et al. 2011; Kerruth et al. 2014). The domain topology of AUREOs is inverted as compared to most other characterized LOV proteins because the sensory domain of AUREOs is at the C-terminus of the receptor. The J $\alpha$  helix of the AUREO-LOV domain allosterically regulates the fold of the N-terminally flanking A' $\alpha$  helix (Herman and Kottke 2015). Subsequent unfolding of the A' $\alpha$  helix exposes a high affinity dimerization site and enables the formation of the light state dimer of LOV (Herman et al.

2013; Herman and Kottke 2015; Heintz and Schlichting 2016). Indeed, in full-length AUREO, the J $\alpha$  helix plays a crucial role in the formation of the light state of the receptor (Banerjee et al. 2016b). The activation of the LOV domain results in a loss of the helical secondary structure of the bZIP domain (Banerjee et al. 2016b), indicating that there is a direct communication between the two domains. If a DNA binding site is available, the helical fold of the bZIP domain is increased by light (Banerjee et al. 2016b), resulting in rigidification of the domain (Heintz and Schlichting 2016). In a complementary approach, Tian and colleagues showed the importance of a previously overlooked C $\alpha$  helix in promoting the conformational protein changes (Tian et al. 2020). bZIP domains have a general tendency to dimerize, and are only capable of binding DNA as dimers (Tateyama et al. 2018). For some AUREOs, dimerization and DNA binding have been shown to be induced by blue light (Hisatomi and Furuya 2015; Banerjee et al. 2016b; Nakatani and Hisatomi 2018; Nakajima et al. 2021). PtAUREO1a occurs in the dark as a dimer/higher oligomer (Banerjee et al. 2016b) or as a monomer in equilibrium with a dimer (Heintz and Schlichting 2016). Light induces the dimerization of LOV domains and the association of the monomers (Kobayashi et al. 2020; Goett-Zink et al. 2020), which is the rate-limiting step in the process of DNA binding (Akiyama et al. 2016). Based on these properties, AUREOs could potentially be used as an optogenetic tool, for instance, to increase protein stability under blue light conditions (Hepp et al. 2020) or directly as a light-driven gene switch.

AUREO function has been extensively characterized in *P. tricornutum*, especially for PtAUREO1a that acts as a key regulator of the diatom cell cycle (Table 1; Chapter "Cellular Hallmarks and Regulation of the Diatom Cell Cycle"). Evidence from silenced and knockout lines of *P. tricornutum* further indicates that PtAUREO1a and PtAUREO1b might be involved in regulation of photoacclimation (Table 1; Schellenberger Costa et al. 2013b; Serif et al. 2017; Mann et al. 2017; Madhuri et al. 2019). PtAUREO1c might be a high light sensor in vivo because it recovers faster and is much less sensitive to light than PtAUREO1a (Bannister et al. 2019). Recent studies indicate that AUREOs may have a large impact on the cells. Changes in transcriptomes in response to a shift from red to blue light were analyzed in wild-type *P. tricornutum* cells, in PtAUREO1a knockout and in PtAUREO1a complemented lines (Mann et al. 2020). Wild-type cells react within minutes

by up- or down-regulating 75% of the genes, while this massive change in gene expression is mostly inhibited in PtAUREO1a knockout strains (Mann et al. 2020). PtAUREO1a, therefore, must have a specific function in cellular regulation that cannot be complemented by other AUREOs. This raises the question of how a single photoreceptor can affect such a large number of genes. Possibly, a cascade of TF transduces the initial response of PtAUREO1a to blue light. This is supported by findings that the transcript abundance of a large number of diatom TFs and photoreceptors (Rayko et al. 2010) is strongly and rapidly affected by blue light in wild-type cells, but not in PtAUREO1a knockout mutants (Mann et al. 2020). When common gene expression patterns are analyzed, the aureochromes are placed in different clusters (Ait-Mohamed et al. 2020), supporting the idea that AUREO1 isoforms may have different roles. There also is some evidence that AUREOs might be involved in regulation of the diel cycle and the expression of some AUREOs follows a different diurnal pattern (Banerjee et al. 2016b). The diurnal rhythmic expression of PtAUREO1a and 1c can still be detected when the cells are kept in the dark, while PtAUREO1b expression appears to be light activated, and PtAUREO2 oscillates only weakly throughout the day. The recently discovered diatom clock component RITMO1/PtbHLH1a (Annunziata et al. 2019) is strongly induced by blue light in wild-type cells, but not in the PtAUREO1a mutant, indicating that PtAUREO1a might be involved in triggering the diatom clock in response to blue light.

### **Cryptochrome/Photolyase Family \***

The cryptochrome/photolyase family proteins (CPFs) are widespread blue light-absorbing flavoproteins with similar primary sequences, but very diverse functions. Most CPFs non-covalently bind a flavin adenine dinucleotide (FAD) cofactor as a specific prosthetic chromophore. Other chromophores such as 5,10-methenyltetrahydrofolic acid (MTHF), 8-hydroxy-5-deazaflavin, or flavin-mononucleotide (FMN) may also be associated with some CPFs as light antennae (Essen et al. 2017). CPF members have a characteristic conserved photolyase-related (PHR) domain, but amino and carboxy terminal extensions are highly variable in both length and primary sequence. Photolyases (PL) are blue light-activated enzymes repairing UV-induced DNA lesions, such as

cyclobutane pyrimidine dimer (CPD PL) or (6-4) pyrimidine-pyrimidone photoproducts. The cryptochromes (CRY) of plants and some insects do not have any DNA repair activity, but are either blue light photoreceptors or light-independent components of the central circadian oscillator (Chaves et al. 2011). The activation of light-dependent CRY has been well studied. The photoperception process starts with a very fast photoinduced reduction of FAD. This redox- and light-dependent change at the core of the protein leads to conformational changes, allowing specific interactions with other protein partners. Globally, both light- dependent and -independent CRY, are involved in transcriptional regulation, respectively, by inhibiting transcriptional repressors or by inactivating transcription.

Phylogenetic relationships in the CPF family helped to identify five major super- classes (sc) which do not necessarily converge functionally. For instance, sc1 includes (6-4) PLs but also light-dependent animal CRYs and light-independent CRYs involved in the transcription/translation feedback loop of the circadian clock. Class I and III CPD photolyases group also together phylogenetically with proteins with different functions such as light dependent plant photoreceptor CRY and plant- like photoreceptor CRY in sc3, but are separate from class II CPD photolyases which are all found in sc4. The last two super-classes (sc2 and sc5) include all CRY-DASH (named after *Drosophila*, *Arabidopsis*, *Synechocystis*, Human) and the proteobacterial PL/CRYs (Fortunato et al. 2015; Ozturk 2017). A surprising result from genome sequence analyses was that diatoms do not possess canonical plant CRY photoreceptors, even though blue light is preponderant in the ocean and CRY regulate so many physiological processes in plants. However, further genomic and functional investigations in diatoms revealed novel CPF variants, including the animal-like CPF1 and plant-like CRYs (CryP) (see below) and several DASH CRYs (Table 2). The biological function of Cry DASH is not yet clearly defined in diatoms or indeed in other organisms, but several studies suggest that they might have a single-stranded DNA CPD PL activity, a signalling role, or be involved in the regulation of metabolism, consistent with their organellar localization (Kleine et al. 2003; Froehlich et al. 2010).

Only two diatom CPFs have been characterized in detail to date: CPF1 and CryP of *P. tricornutum* (Fig. 1b). Initially discovered in diatoms, these proteins are extensively



represented in metatranscriptomic data derived from marine environments (Coesel et al. 2021). CPF1 belongs to sc1, is localized in the nucleus, and has (6-4) PL DNA repair activity both in vitro (Coesel et al. 2009) and in vivo (De Riso et al. 2009). However, like plant CRY, CPF1 regulates the transcription of several genes acting in pathways modulating photoprotection, cell division, nutrient assimilation, etc., under blue light (Coesel et al. 2009), so is likely to be also a photoreceptor. CPF1 is also able to bind mammalian CLOCK protein in heterologous cells where it partially represses the CLOCK/BMAL1 heterodimer (Coesel et al. 2009) in the positive loop of the circadian clock (Kume et al. 1999).

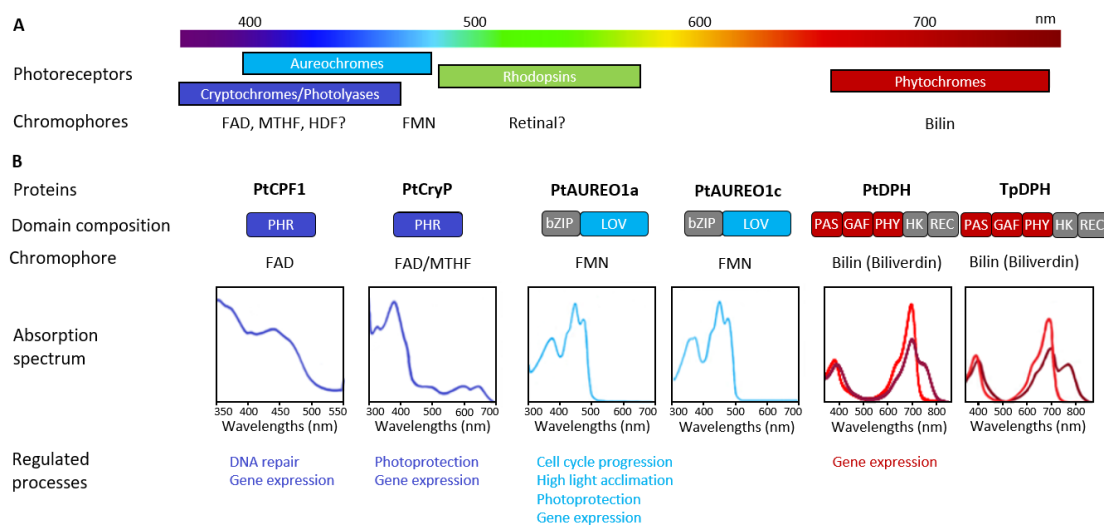


Figure 10 (Figure 1 from the book chapter). Sensing of light cues in diatoms. (a) Photoreceptor classes identified in diatom genomes. Cryptochrome-Photolyase, Aureochrome, Rhodopsin, and Phytochrome families are positioned schematically according to the range of wavelengths they usually absorb. The nature of the chromophores is indicated below the respective photoreceptor boxes, with question marks indicating putative chromophore. (b) Summary of characterized photoreceptors from *P. tricornutum* (Pt) and *T. pseudonana* (Tp). Domain architecture, chromophore, absorption spectrum, and regulated processes are indicated for: cryptochrome photolyase 1 (PtCPF1) (Coesel et al. 2009); plant-like Cry (PtCryP) (Juhás et al. 2014); aureochromes PtAUREO1a (Heintz and Schlichting 2016) and PtAUREO1c (Bannister et al. 2019) and diatom phytochromes (PtDPH and TpDPH) (Fortunato et al. 2016). Domains involved in the light sensing are indicated in blue or red

CryP belongs to the group of Plant-like photoreceptors CRY (sc3). Recombinant *P. tricornutum* CryP, produced in *E. coli*, binds FAD and MTHF chromophores (Juhás et al. 2014). Moreover, FAD photoreduction, a common mechanism in CPFs that bind FAD, is also active in vitro, suggesting that activation of CryP is light-dependent. Originally, CryP

was identified as a blue light regulator of light harvesting proteins directly involved in photoprotection (LHCX) (Juhas et al. 2014). However, subsequent analyses of transcriptional responses to illumination by blue light after prolonged dark, revealed that gene expression was already deregulated in the dark in CryP knockdown lines compared to wild type (König et al. 2017). Therefore, despite the presence of chromophores, CryP may not be a major blue light photoreceptor but rather a component involved in the global modulation of transcription, requiring other blue light photoreceptors to signal the light responses.

### **Rhodopsins \***

These light-sensing integral membrane proteins found in Archaea, bacteria, and eukaryotes share a topology of seven transmembrane alpha helices within which a retinal chromophore is covalently bound (Ernst et al. 2014). Rhodopsins exhibit a wide variety of spectral tuning in the blue-green part of the spectrum depending on the nature of a few influential amino acids interacting with the retinal (Man 2003; Ernst et al. 2014). Rhodopsins function as light-driven ion pumps, ion channels, or light sensors (Grote et al. 2014). The discovery that H<sup>+</sup>-pump rhodopsin converts light into ATP has challenged the assumed monopoly of photosynthesis as a phototrophy-enabling mechanism (Béjà et al. 2001; Finkel et al. 2013). Recently, a distinct group of microbial rhodopsins, the heliorhodopsins, has been identified after analysing environmental genomic samples (Pushkarev et al. 2018). Heliorhodopsins do not have the capacity for light-triggered ion transport but they do have a long photocycle, suggesting that they could act as signalling photoreceptors (Pushkarev et al. 2018). H<sup>+</sup>-pump rhodopsins and heliorhodopsins are present in bacteria, Archaea, and algae and are highly represented in environmental genomic data (Pushkarev et al. 2018; Coesel et al. 2021).

Diatom rhodopsin-like sequences falling into the H<sup>+</sup>-pump group, based on conservation of key amino acids and phylogeny, were first identified in the transcriptome of *Pseudonitzschia granii* (Marchetti et al. 2012), then in those of other species (Marchetti et al. 2015), and in the genome of *F. cylindrus* (Mock et al. 2017) (Table 2). Because *P. granii* rhodopsin-like transcripts are highly abundant in low iron conditions, it has been

hypothesized that this proteorhodopsin-like protein could be involved in energy production under conditions of iron deficiency that affect photosynthesis (Marchetti et al. 2012). It is noteworthy that a gene homologous to heliorhodopsin has been identified in the genomes of *P. tricornutum* and other pennate diatoms, but not in centric ones (Pushkarev et al. 2018) (Table 2). G-protein-coupled receptor rhodopsin-like genes, homologs of receptors that transduce a wide range of stimuli including light, hormones, volatile molecules, glycoproteins, nucleotides, and chemokines in eukaryotes (Costanzi et al. 2009), have also been identified in various diatom genomes (Port et al. 2013), with evidence of expansion of this gene family in some species (Osuna-Cruz et al. 2020). However, information about the spectral and functional properties of these and other diatom rhodopsins is still lacking, so their function as light sensors remains to be established.

## PHYTOCHROMES

Phytochromes are known for more than 60 years as master regulators of plant photomorphogenesis, where they control germination, de-etiolation, shade avoidance, stomatal development, entrainment of the circadian clock, and flowering (non exhaustive list!) in response to light and to the ratio of red to far-red wavelengths (Franklin and Quail, 2010). These photoreceptors were subsequently found in other organisms, photosynthetic or not: prokaryotes such as cyanobacteria and anoxygenic photosynthetic bacteria, but also non-photosynthetic bacteria; photosynthetic eukaryotes such as glaucophyta and prasinophyte (primary endosymbiosis) or cryptophytes and ochrophytes (secondary endosymbiosis), and heterotrophic eukaryotes such as Fungi (Rockwell and Lagarias, 2020). Evolutionary, phytochromes are thought to originate from bacteria (Rockwell and Lagarias, 2020), and transferred to eukaryotes at least twice independently. Indeed, phytochrome phylogenetic studies showed that plant phytochromes form a large and robust clade, sister to prasinophyte, glaucophyte and cryptophyte phytochromes (“Eukaryotic phytochrome type 1” in Fig 11) while Fungi and Stramenopile phytochromes form an independent clade (“Eukaryotic phytochrome type 2” in Fig 11). Phytochromes from the eukaryotic type 1 clade have been proposed to arise from endosymbiotic gene transfer from cyanobacteria (Kooß and Lamparter, 2017), but other studies resolve cyanobacterial phytochrome as an independent clade within

bacterial phytochromes (Duanmu et al., 2014; Li et al., 2015a). Within the eukaryotic type 1 clade, glaucophyte phytochromes are the first to diverge, followed by cryptophyte phytochromes; however, cryptophyte algae are not part of the archaeplastidia themselves (Burki et al., 2020). This incongruence between phytochrome tree and species tree has not been resolved. In the other eukaryotic branch, Fungi and Stramenopile phytochromes are sister clades, but the origin of these phytochrome has not been determined. Horizontal gene transfer, deletion but also duplication and diversification have shaped the evolutionary history of this photoreceptor (Rockwell and Lagarias, 2020).

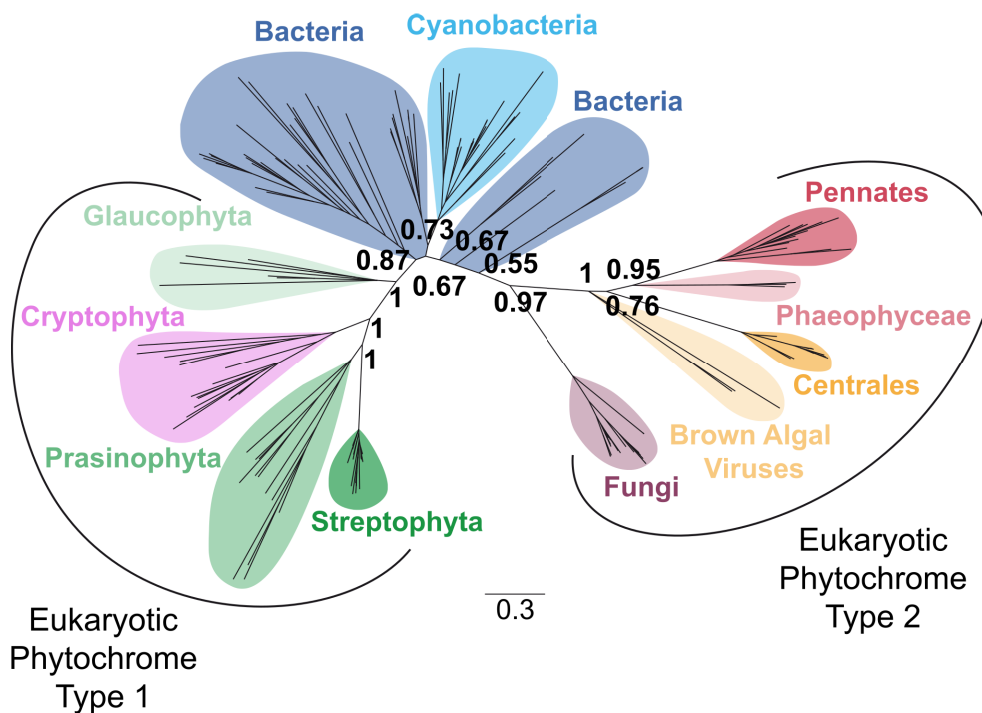


Figure 11: Phylogenetic analysis phytochrome photoreceptor (based on the GAF and PHY proteins domains). Eukaryotic Type 1 phytochrome contains phytochromes from Viridiplantae and Cryptophyte, while eukaryotic type 2 contains Fungi and Stramenopile phytochromes. These include phytochromes from brown algae (Phaeophyceae), Diatom (Pennate and Centrales) and brown algal viruses. From (Fortunato et al., 2016).

These photoreceptors share a common structure, with a conserved Photosensory Core Module (PCM) in the first part of the molecule, composed of PAS, GAF and PHY domains, and a more variable Output Module (OM) (Fig. 12). The chromophore is a linear tetrapyrrole derived from heme (bilin), which is covalently bound to a cysteine either in the PAS or in the GAF domain (Fig. 12). Different chromophore can be used: biliverdin (in

proteobacterial and fungal phytochromes for examples (Bhoo et al., 2001; Karniol et al., 2005; Blumenstein et al., 2005) or more reduced bilins, such as phytochromobilin (plants) or phycocyanobilin (streptophyte algae and cyanobacteria (Rockwell et al., 2014; Hughes et al., 1999). Heme oxygenase is the enzyme necessary to produce BV from heme, while other enzymes (Bilin Reductases) are required to produce the other chromophores. Light absorption triggers photoisomerization of the bilin around a double bond, from E to Z conformation or vice versa. Phytochromes exist in two states, a red absorbing form Pr with the chromophore in the Z conformation (in canonical plant phytochromes), and a far-red absorbing form Pfr, bound to chromophore in E conformation (Rüdiger and Thümmler, 1994). Absorption of red (R) light by Pr will trigger the Pr → Pfr reaction, and absorption of far-red (FR) by Pfr will trigger the Pfr → Pr reaction (Fig. 12A). In the absence of light, some phytochromes will revert to its most stable form: this is called the dark or thermal reversion (Mancinelli, 1994). The dark state can be Pr (in canonical phytochromes) or Pfr (bathyphytochromes). The equilibrium between the two forms depends on the light environment, presence and binding of partners, temperature and dark reversion rate. In addition, phytochromes work as dimers *in vivo* (Klose et al., 2015; Brockmann et al., 1987)

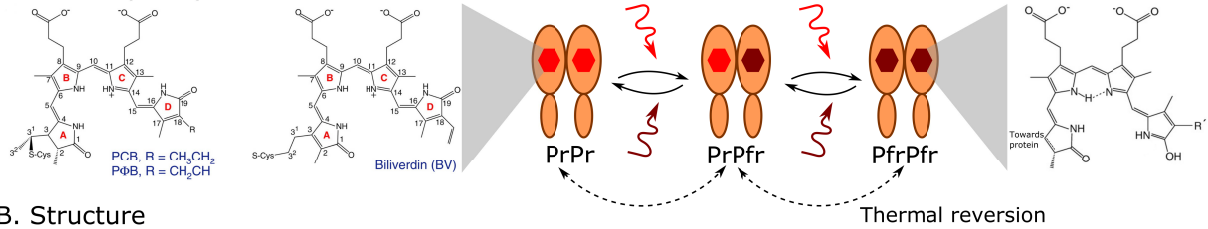
### **Plant phytochromes**

Land plants possess variable copy number of phytochrome genes, which regulate different aspects of morphogenesis both in redundant and specific manners (Franklin and Quail, 2010; Legris et al., 2019). Seed plants possess 3 conserved types of phytochromes: PhyA, PhyB and PhyC. In the model plant *Arabidopsis thaliana*, subsequent duplication of PhyB gave rise to PhyE, found in all Angiosperms, and PhyD, which is specific for Brassicaceae (Mathews and McBreen, 2008). Phytochrome responses were first classified as Type I, represented by PhyA responses, and Type II. Type I responses are also called Very Low Fluence Responses, and function as a light- but not wavelength-sensitive response (Shinomura et al., 1996). Indeed, PhyA rapidly converts from Pr to Pfr upon light illumination of any wavelength, but is very unstable and degraded; this response is irreversible. PhyA also mediates High Irradiance Responses, which require long exposure to high light intensities. Type II responses, or Low Fluence Responses, are the classical phytochrome responses: they are induced by a pulse of red light, show R/FR reversibility,

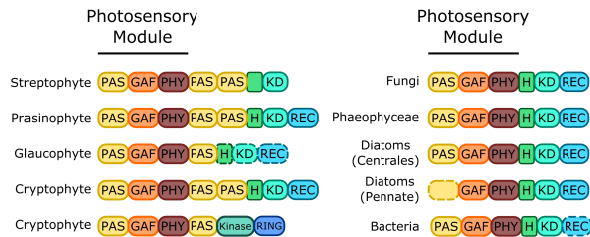
sensitivity to the R/FR ratio and follow the reciprocity law, i.e the amplitude of response observed is a function of the total amount of photon received (Time of exposure\*Light intensity) (Mancinelli, 1994). In addition, the rate of photoconversion and dark reversion is affected by ambient temperature, and phytochromes are involved in temperature sensing in plants (Jung et al., 2016; Legris et al., 2016). Structural and biochemical differences (thermal reversion) could explain the differences between phytochrome responses (Legris et al., 2019; Burgie et al., 2021).

Plant phytochromes bind phytochromoblin as chromophore with a cysteine in their GAF domain, and their output module is composed of 2 PAS domains and a Histidine-Kinase Related domain (HRKD) (Fig. 12B). Upon light exposure, phytochrome is shuttled to the nucleus and interacts with Phytochrome Interacting Factors (PIF), which are bHLH transcription factors. PIF will be phosphorylated, ubiquitinated and addressed to the proteasome for degradation (See Fig. 12 for more details on the phytochrome signaling cascade in plants). Whether all the phytochrome HRKD has a serine/threonine phosphorylation activity is debated (Boylan and Quail, 1996; Shin et al., 2016; Ni et al., 2017; Li et al., 2022).

## A. General principles

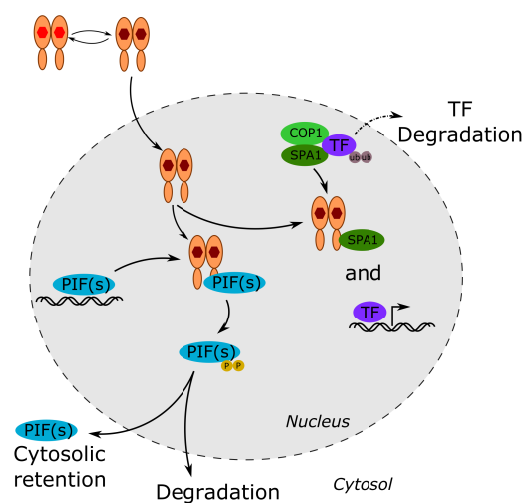


## B. Structure

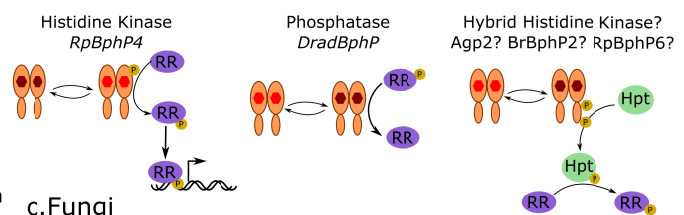


## C. Signalling cascades

### a. Plant



### b. Bacteria



### c. Fungi

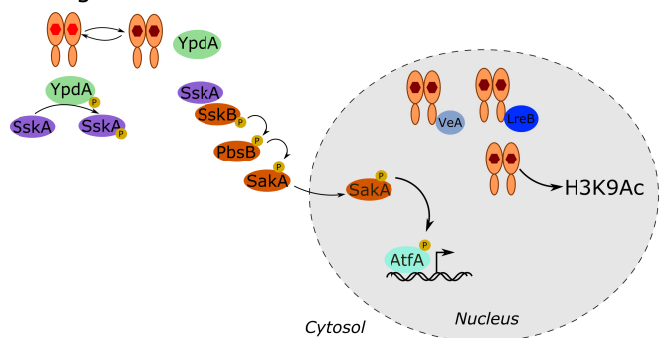


Figure 12. Signaling mechanisms in phytochromes. A. Phytochromes can bind PCB, PΦB or BV as chromophore. Illumination triggers isomerisation around the C15-C16 double bond (see molecular structure on the right). Phytochromes exist in two conformation, Pr and Pfr. As they work as dimers, 3 different dimers type are found in equilibrium under illumination. Eventual thermal reversion might revert phytochrome to this most stable form. B. Different phytochrome structure found in Eukaryotic group 1 (first column), Eukaryotic group2 and the most common architecture for Bacteriophytochrome (Second column). Dotted contour indicates that this domain may not be present. C. Signalisation cascade. A. Simplified phytochrome signalisation cascade in plants. Upon illumination, phytochromes will be shuttled to the nucleus, where they interact with PIF transcription factors and inhibit them. PIFs will be phosphorylated and degraded or sequestered in the cytosol. In addition, phytochromes can interact with E3 ubiquitin ligase complexes such as the COP1/SPA1 complex. This will prevent the E3 ligase complex to address transcription factors to the proteasome for degradation. *Legend continues on next page*

B. In bacteria, phytochromes such as RpBphP4 function as two-component systems, and will phosphorylate a response regulator domain (left). However, it has been shown recently that some phytochromes can work as phosphatase (*Deinococcus radiodurans* DraBphP, middle) (Multamäki et al., 2021). Other bacteriophytochromes possess a REC domain fused at their C-terminal, suggesting that the signalisation cascade thus requires an additional phosphorelay (Hpt protein, right). C. In Fungi, the signaling cascade seem to begin as a two component system. In the dark YpdA, an Hpt protein, is phosphorylated and keeps Sska (REC) inactive. Upon illumination, YpdA dephosphorylates (YpdA interacts with FphA, but the light-dependency is unknown). SskA will interact with SskB, the first component of a MAPKinase cascade, which ends with the phosphorylation of SakA. SakA will be shuttled to the nucleus, where it will activates transcription factors (AtfA). FphA has also been shown to interact with LreB and VeA which are involved in blue-light signalling, and to have an impact on histone acetylation, potentially regulating gene expression.

### **Bacterial phytochromes**

Bacterial phytochromes show variations both in the PCM and in the OM (Auldridge and Forest, 2011) (Fig. 11). Most bacterial phytochromes have the PAS-GAF-PHY architecture, but some lack the PAS, like Cph2 from the cyanobacteria *Synechocystis*, and others only have the GAF domain (Cyanobacteriochromes, CBCR) (Fig 12). Bacteria phytochromes use biliverdin as a chromophore, bound at the N-terminal extremity (Bhoo et al., 2001), while cyanobacteria phytochromes and CBCR use phycocyanobilin bound to the GAF domain (Yeh et al., 1997; Hughes et al., 1997). Additional variation of the PCM includes the presence of a PYP (Photoactive Yellow Protein) upstream of the PCM; this additional domain binds another chromophore (p-hydrocinnamic acid) and absorbs blue light, conferring new light-sensing properties to these chimeric photoreceptors. The Output module is a typical two-component system domain, with histidine kinase eventually followed by a response regulator domain. This system is a phosphorelay-signaling cascade, with ATP-dependent phosphorylation of a Histidine residue in the HK domain; the phosphate group is then transferred to an aspartate in a response regulator (RR) protein (Yeh et al., 1997; Giraud et al., 2005) (Fig. 12). Phosphorylation of the RR affects its activity (usually, DNA binding and transcription regulation). Bacteriophytochrome genes are often part of operon containing also heme oxygenase for BV production and response regulator genes, and sometimes also the phytochromes-regulated (Bhoo et al., 2001; Giraud et al., 2002; Lamparter et al., 2002). Other output architecture exists, but their signaling mechanisms have not been elucidated.



Different aspects of bacterial physiology can be regulated by phytochromes. In cyanobacteria, one major response is complementary chromatic adaptation (CCA) (reviewed in Wiltbank and Kehoe, 2019). CCA is the process by which cyanobacteria remodel their light harvesting machinery (phycobilisomes) to optimize their light harvesting capacity according to the incoming light quality (Grossman, 2003). Different types of CCA exist, regulated by CBCR (Kehoe and Grossman, 1996). CBCR can have very different absorption spectra, enabling the detection of green, teal, yellow and red light for example (Wiltbank and Kehoe, 2019). A special case of CCA, far-red light photoacclimation, is regulated by a phytochrome with red/far-red absorption spectra (Gan et al., 2014; Gan and Bryant, 2015).

Anoxygenic photosynthetic bacteria also use phytochrome to regulate their photosynthetic apparatus (Giraud et al., 2002; Giraud et al., 2005; Jaubert et al., 2007). Phytochromes control both photosystem and light harvesting complexes synthesis, in response to far-red light in *Rhodospseudomonas palustris* and *Bradyrhizobium sp.* In some *Rps. palustris* strains, some phytochromes have lost their light-sensing abilities, but mediate the response to redox conditions (Vuillet et al., 2007).

In non-photosynthetic bacteria, the physiological role of phytochromes is less known. In *Agrobacterium fabrum*, recent studies showed that 2 phytochromes (Agp1 and Agp2) are involved in the control of conjugation and plant infection (Bai et al., 2016; Xue et al., 2021). In the non-photosynthetic bacteria *Deinococcus radiodurans*, phytochrome regulates the synthesis of carotenoid pigment that protects the bacteria during growth under high light (Davis et al., 1999). Recently, bacteriophytochrome from a deep-sea bacterium *Croceicoccus marinus* was proposed to promote growth in infra-red (940nm) by modulating the cell's metabolism (Liu et al., 2021).

### **Fungal phytochromes**

Phytochromes seem to be widespread in Fungi, at least in Ascomycota (Schumacher, 2017). Their structure is close to Bacteriophytochromes (Fig. 12B), but with a long N-terminus extension, and they bind biliverdin as chromophore. In the model *Aspergillus nidulans*, the phytochrome FphA has a red/far-red absorption spectra and represses

sexual reproduction in red light (Blumenstein et al., 2005). It is localized in the cytoplasm and signals through interaction with the phosphotransfer protein YdpA, as a two-component phosphorylation relay, followed by the HOG (high osmolarity glycerol) MAP kinase cascade and activation of transcription factors (Yu et al., 2016) (Fig. 12C). FphA also interacts with blue-light photoreceptors and with chromatin modification complexes (Purschwitz et al., 2009; Hedtke et al., 2015). Recombinant FphA shows no dark reversion *in vitro* (Consiglieri et al., 2019), but it was suggested that FphA is also a temperature sensor in Fungi (Yu et al., 2019).

### **Algal phytochromes**

Phytochromes have been found in many marine algae, and their characterization is gaining interest in recent years (Duanmu et al., 2014; Rockwell et al., 2014). Duanmu et al. (2014) showed that phytochrome from the green algae *Micromonas pusilla* (marine prasinophyte) is shuttled to the nucleus during the light period, suggesting conserved mechanism with land plant phytochromes. However, this phytochrome exhibits light-induced auto-phosphorylation activity (probably histidine kinase). This phytochrome and other prasinophyte phytochromes (*Dolichomastix tenuilepis*, *Tetraselmis astigmatica*, *N pyriformis*, *Prasinoderma coloniale*) bind phycocyanobilin and show variation of their absorption spectra, from red/far-red (648/734nm, *T.astigmatica*) to absorption spectra shifted towards shorter wavelength, i.e., orange/far-red (586-614/690-718 nm) (Fig. 13) (Rockwell et al., 2014). This shift towards shorter wavelength is supposed to be an adaptation of phytochrome absorption spectra towards light abundant in the aquatic environment, i.e. spectral tuning (fig 13). Other algal phytochromes have been spectrally characterized: 2 phytochromes from glaucophytes, showing peculiar absorption properties and photocycle with blue/red absorption shifts; *Ectocarpus siliculosus* phytochrome 1 (brown algae), which binds phytochromobilin and has a red/green photocycle. These variations in absorption spectra are once again discussed in the scope of spectral tuning, considering that phytochromes in aquatic environments are tuned towards blue and green wavelengths (Fig. 13) (Rockwell et al., 2014))

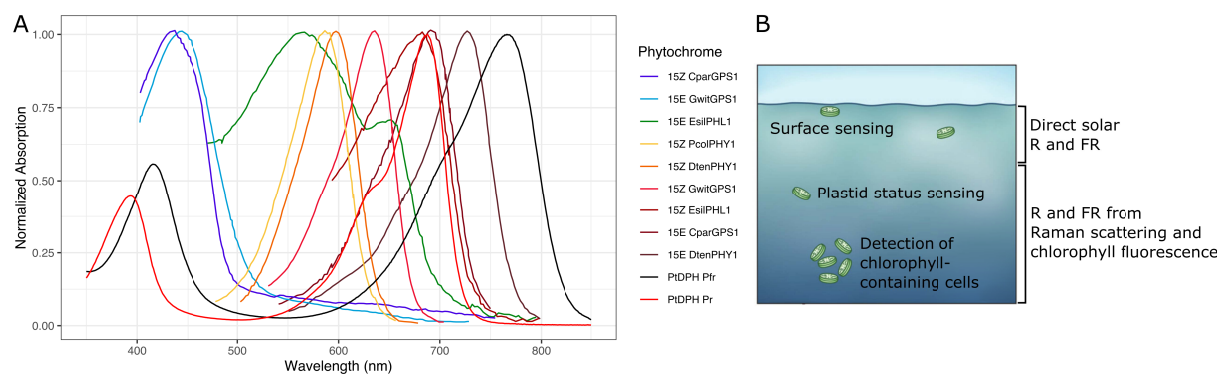


Figure 13. Algal phytochromes absorption spectra. PtDPH absorption spectra in compared to different algal phytochromes from (Rockwell et al., 2014). PtDPH Pfr is clearly the most FR-shifted phytochrome. Cpar and Gwit, *Cyanophora paradoxa* and *Guillardia theta* (Glaucophytes); Dten and Pcol, *Dolichomastix tenuilepis* and *Prasinoderma coloniale* (Prasinophytes); Esil, *Ectocarpus siliculosus* (Phaeophyceae) B. Hypotheses on DPH activation in the environment

Contrary to algal phytochromes presented above, phytochromes from the 2 marine diatom models *Thalassiosira pseudonana* (Tp) and *Phaeodactylum tricornutum* (Pt) bind biliverdin and have a red/far-red absorption spectra (686/764 nm for TpDPH, 700/750nm for PtDPH, Fig. 13), shifted towards longer wavelength compared to plant phytochromes (Fortunato et al., 2016). Other DPHs found from an initial exploration of available genome information are structurally similar to TpDPH and PtDPH and phylogenetically close to bacterial and fungal phytochromes. Surprisingly, pennate and centric phytochromes do not belong to the same clade, but pennate phytochromes are sister to phaeophyceae, while centric phytochromes are sister to phaeophyceae+pennate phytochromes. Interestingly, giant brown algal viruses carry phytochrome genes that branch close to these groups, suggesting that horizontal gene transfer could have contributed to the evolution of diatom phytochromes. In addition, we can note that not all diatoms possess phytochromes, but that centric diatoms do have it in one copy while pennate phytochromes can be present in several copies.

Before I started by PhD, functional studies in *P. tricornutum* using knockout (KO) lines have shown that diatom phytochrome (DPH) regulates gene expression in response to far-red light (Fortunato et al., 2016). The far-red light induced genes (80 genes) are of unknown function (55%) or involved in transcription and signaling (25%). DPH does not seem to be involved in red light gene regulation in *P. tricornutum*. DPH KOs do not show alteration of physiology so far, so the consequences of DPH activation by far-red light are

unknown. *In vitro* both Pt and Tp DPH show no dark reversion, but PtDPH has been shown to auto-phosphorylate in response to FR light, suggesting a phosphorylation cascade as in two-component systems. No homologs of PIFs have been found in diatom genomes, suggesting different phytochrome signalization pathway.

Hypotheses have been made concerning the light that could be perceived by DPH in the marine environment (Fortunato et al., 2016). DPH could sense direct solar R and FR light, but this would be limited to the first meters of the water column, which leads to the idea that DPH could be a surface sensor. Alternatively, DPH could sense R and FR photons from fluorescence of photosynthetic organisms; in that case, DPH could perceive changes in its own chloroplast photosynthetic activity, or light from surrounding photosynthetic organisms. This could bring information on cell density in an algal bloom for example. However, it was underlined that below the first meters the R/FR light ratio is rather constant in the marine environment, and thus DPH activity is probably not based on sensing this waveband ratio.

## MODEL SPECIES AND GENETIC TOOLS

### THE GROWING NUMBER OF DIATOM GENOMES, TRANSCRIPTOMES AND META-DATA

The first diatom genome to be sequenced was the one from the cosmopolitan centric diatom *T. pseudonana* (Armbrust et al., 2004). Whole-genome sequence of a second model diatom, *P. tricornutum* soon followed (Bowler et al., 2008). These species were also chosen because they were amenable to genetic transformation and show fast growth under laboratory conditions. Both their genomes are relatively small (32.1 and 27.4 Mb for *T. pseudonana* and *P. tricornutum*, respectively) with 11776 and 12177 coding genes, respectively. Some peculiarities of diatoms already arose from the analysis of their genomes: the presence of red and green algal genes, an important number of bacterial genes from horizontal gene transfer, in addition to species-specific genes. Even today,

with way more genomes available, the fraction of genes not shared between centric and pennate species is very high (40%), highlighting the diversity and rapid evolution of diatoms (Tirichine et al., 2017; Mock et al., 2022).

Table 1. Sequenced genomes of centric diatoms

Centric species	Reference and interest
<i>Thalassiosira pseudonana</i>	(Armbrust et al., 2004), first diatom sequenced
<i>Thalassiosira oceanica</i>	(Lommer et al., 2012), adaptation to iron-poor waters
<i>Cyclotella cryptica</i>	(Traller et al., 2016; Roberts et al., 2020), biofuel production
<i>Skeletonema costatum</i>	(Ogura et al., 2018; Sorokina et al., 2022), harmful algal blooms and abundant coastal diatom
<i>Sketletonema marinoi</i>	(Johansson et al., 2019), chain forming and abundant coastal diatom
<i>Chaetoceros tenuissimus</i>	(Hongo et al., 2021), widely distributed blooming diatom, very small
<i>Attheya</i> , <i>Chaetoceros calcitrans</i> , <i>C. decipiens</i> , <i>C.muelleri</i> , <i>C.neogracile</i> , <i>Skeletonema dorhrnii</i> , <i>S.marinoi</i> , <i>S.menzelii</i> , <i>Thalassiosira guillardii</i> , <i>T. nordenskioldii</i> , <i>T.weissflogii</i>	(Nelson et al., 2021), large sequencing project to get insights from diverse microalgae, with a focus on the role of viruses in algal evolution
<i>Chaetoceros gracilis</i>	(Kumazawa et al., 2022)

A lot of other diatom species were sequenced to study specific aspects of diatom life and explore diatom diversity, with different interests (evolutionary, ecological, biotechnologies), summarized in Table 1 and 2 below.

We can also mention ongoing projects such as the Chaetoceros sequencing project “Most Abundant Diatom Genus in the World’s Ocean” and the 100 Diatom genomes project with very different diatom species.

Databases and tools have started to emerge for the molecular study of diatoms. One of the first databases was the Diatom Portal, gathering gene expression microarray data from *T. pseudonana* and *P. tricornutum* (Ashworth et al., 2016). A new platform gathering expression data from RNASeq experiments is being developed (Diatomicbase, Villar et al,

unpublished but available at the end of this manuscript). In another project, PhaeoNet integrated Pt RNASeq data to provide co-expression analyses (Ait-Mohamed et al., 2020). Comparative genomics platform PLAZA Diatom is available to compare gene families across diatom genomes (Osuna-Cruz et al., 2020).

Table 2 Sequenced genomes of pennate diatoms

Pennate species	Reference and interest
<i>Phaeodactylum tricornutum</i>	(Bowler et al., 2008), first pennate genome; model species
<i>Pseudo-nitzschia multiseriis</i>	harmful algal blooms
<i>Pseudo-nitzschia multistriata</i>	(Basu et al., 2017), sexual reproduction
<i>Fragilariopsis cylindrus</i>	(Mock et al., 2017), adaptation of polar regions
<i>Fistulifera solaris</i>	(Tanaka et al., 2015), biofuel production
<i>Seminavis robusta</i>	(Osuna-Cruz et al., 2020), first benthic diatom
<i>Synedra acus (Fragilaria radians)</i>	(Galachyants et al., 2015), araphid pennate, freshwater
<i>Nitzschia inconspicua</i>	(Oliver et al., 2021), biofuels and bioproducts; suitable for large scale aquaculture
<i>Amphora coffeaeformis</i> , <i>Cylindrotheca fusiformis</i> , <i>Halamphora</i> , sp (2 strains), <i>Navicula incerta</i> , <i>Navicula pelliculosa</i>	(Nelson et al., 2021), large sequencing project to get insights from diverse microalgae, with a focus on the role of viruses in algal evolution
<i>Haslea ostrearia</i>	(Gabed et al., 2022) Unique marenine pigment production, epiphytic diatom

Genomic data are growing, but transcriptome resources are still the reference for many diatom species. The MMETSP (Marine Microbial Eukaryote Transcriptome Sequencing Project) gathered data from 411 microalgal strains including 92 diatoms (Keeling et al., 2014). This valuable resource was re-assembled three times (Johnson et al., 2019; Guita Niang et al., 2020; Van Vlierberghe et al., 2021) to produce robust and clean (from bacteria contamination notably) references with the latest bioinformatics tools.

Lately, a new type of genetic data arose: environmental genomics. The massive sequencing of complete communities allowed the assembly of expressed transcripts forming atlas of expressed genes, and metagenomes-assembled genomes and sequencing of single cells allowed the assembly of genomes of uncultured organisms. However, the annotation and use of these data depends on our understanding of lab strains (ie genomes and MMETSP transcriptomes). The main contributions to these types of data (for eukaryotes) are the Tara expeditions: *Tara Oceans* and *Tara Oceans Arctic circle*, for which data are already available (Carradec et al., 2018; Seeleuthner et al., 2018; Delmont et al., 2020); data from other expeditions (Pacific Ocean, Mediterranean Sea) are on their way. Sampling was performed with a repetitive protocol that allows sample comparison, and a lot of other data are available (HPLC, physico-chemical parameters of the water column (optics, nutrient, carbon chemistry) to link genetics data to environmental conditions. Another project recently made public is the Sea of Change project (Martin et al., 2021).

---

#### MOLECULAR MODEL SPECIES

As the first diatoms whose genome was sequenced, *T. pseudonana* and *P. tricornutum* are the most studied diatoms in molecular studies. Both species are small (10µm) marine diatoms, isolated from coastal waters, close to estuaries. Some studies suggest that *T. pseudonana* is in fact a freshwater species that re-colonized marine environments (Alverson et al., 2011). In the ocean, the *Thalassiosira* genera is one of the most abundant diatom genera (Malviya et al., 2016). *P. tricornutum* is not an abundant species in the open ocean, but several strains were isolated from coasts around the world (De Martino et al., 2007), suggesting a widespread distribution. This species has the unique feature of existing 3 distinct morphotypes, oval, elongated and triradiate, and is able to shift from one morphotype to the other depending on the strain and the environmental conditions. *P. tricornutum* is though be a benthic diatom in its oval form, and planktonic in its triradiate and elongated forms. Contrary to other diatoms, *P. tricornutum* does not have a strict requirement for silica and its cell wall is only poorly silicified. Sequencing of ecotypes in both *T. pseudonana* and *P. tricornutum* (Koester et al., 2007; Rastogi et al., 2019) unveiled their dominant reproduction mechanisms (sexual for *T. pseudonana*,

asexual for *P. tricornutum*) through population genetics, and allow the study of gene polymorphisms, eventually linked to adaptation to specific environmental conditions.

---

## GENETIC ENGINEERING

Genetic transformation has been set up for the model species *T. pseudonana* and *P. tricornutum* (Apt et al., 1996; Falciatore et al., 1999; Poulsen et al., 2006), and also for other sequenced diatom species (Moosburner et al., 2022). The first successful method to transform diatom was by biolistic transformation, with the delivery of a DNA vector on gold or tungsten beads. The insertion of the vector into diatom genomes is stable but random, and often in multiple copies, and this can lead to the random disruption of nuclear genes and variable levels of expression of the transgene (George et al., 2020). DNA can also be introduced by electroporation (Falciatore et al., 2020).

Lately, the use of a yeast derived sequence allowed the design of an episome for the delivery of genetic material by bacterial conjugation (Karas et al., 2015). In most of the cases, the plasmid is not integrated into the genome but maintained as an episome. Some variations of expression within a clone are observed, but expression is comparable between clones (George et al., 2020). Antibiotic selection has to be maintained to keep the episome, and without selection, the episome is lost. As consequence, this method can be used for transient expression of deleterious transgenes and a more homogenous expression between independent transgenic lines.

First modulations of gene in diatoms by reverse genetic approaches were done in diatoms by gene overexpression. Several plasmids were designed with different promoters to allow high expression of the transgene in the cell (reviewed in (Falciatore et al., 2020)). Down regulation of a gene was also achieved by gene silencing (De Riso et al., 2009; Shrestha and Hildebrand, 2015). However, gene silencing is not fully controlled in diatoms and does not always determine clear reduction of expression of the targeted genes (De Riso et al., 2009; Shrestha and Hildebrand, 2015). Genome editing techniques were recently developed, first with TALEN (Transcription activation Like Effector Nuclease)(Daboussi et al., 2014), and more recently with the CRISPR-Cas9 method



(Hopes et al., 2016; Nymark et al., 2016). Genome editing is achieved by introducing in diatom a vector containing the Cas9-coding gene and guide RNA genes, that have to be expressed in the cell, either by biolistic transformation and vector integration or as episome (Hopes et al., 2016; Nymark et al., 2016; Sharma et al., 2018). The Cas9 can also be directly delivered as ribonucleoproteins and introduced into the cell by biolistics, resulting in DNA-free edition of the genome (Serif et al., 2018).

Finally, plastid transformation has been reported in *P. tricornutum* (Materna et al., 2009; Xie et al., 2014), which should enable the study of chloroplast-encoded genes.

## OBJECTIVE OF THE THESIS

Light regulates different aspects of diatom life. Accordingly, diatoms possess different photoreceptors that are involved in light regulation, including a plethora of blue light photoreceptors, but also red/far red light photoreceptors, the diatom phytochromes (DPH).

At the time I joined this project, new information on some photoreceptors identified in the genomes of some diatom species began to emerge (Jaubert et al., 2017). Particularly, the discovery of a red and far-red light photoreceptor in diatom model species, using far-red light as an active signal, was particularly surprising, as aquatic environments are less transparent than the atmosphere to these radiations (Fortunato et al., 2016). However, a complete information on the full possible photoreceptors repertoire in diatoms was still lacking, which seemed particularly relevant given the enormous diversity of diatoms and their complex evolution. Therefore during my PhD training, I addressed several questions, and their results are described in different chapters of this thesis:

*What is the repertoire of diatom (and more largely Ochrophyta) photoreceptors? Can we link this repertoire to the algae environment? What is the evolutionary origin of these photoreceptors, considering the complex history of Ochrophyta? And more specifically, what is the origin of diatom phytochromes?*

In chapter 1, to draft an answer to these questions, I gathered data from Ochrophyta genomes and transcriptomes and started looking for known photoreceptor domains. The resulting data showed that some photoreceptor types are conserved in Ochrophyta (aureochromes, a group of stramenopile-specific LOV-based sensors, and cryptochromes), while others are restricted to some algal groups only. Diatoms in particular show a duplication of several cryptochrome-like proteins, some of which might be new photosensors. Phytochromes and rhodopsins probably evolved from horizontal gene transfer, possibly through viral infection. Although the evolutionary origin of diatom photoreceptors has not been resolved yet, this study has provided comprehensive information on the photoreceptor repertoire in Ochrophyta, which is the subject of a manuscript, in preparation.

Still very little is known about phytochrome action in diatoms. To gain insights into the physiological function of DPH and its relevance for diatom life in the oceans, I addressed in the second chapter the following question:

*What is the light phytochrome responds to in the marine environment known to lack red and far-red lights? Do different phytochromes of different diatom species have the same spectral properties? Where the diatom containing phytochrome lives? What is the contribution of phytochrome to the diatom life?*

In Chapter 2, I combined studies in the lab to establish the *in vivo* action spectra of DPH in *Phaeodactylum tricornutum* to projection of DPH activity in modeled and measured marine light fields. This showed that blue light is the prominent waveband favoring DPH activation in the marine environment, and red, as well as green lights, act mostly in reverting blue light activation. I confirmed the presence of DPH in the open ocean by taking advantage of the *Tara* oceans data (Carradec et al., 2018), and by characterizing novel DPH from different species and environmental sequences, I showed these DPH are mostly from centric diatoms and are shared very conserved light-sensing properties. Overall, these results provide completely novel insights both for DPH functioning and its putative role in marine diatoms. We hope to publish these results very soon after the defense.

As this study showed the quasi-absence of pennate diatom phytochromes in the open ocean, we considered the possibility that DPH presence in pennate diatom is associated to the distribution of some of these species in the benthic environment.

*Is there a link between DPH and the light field in benthic environments?*

As noticed in chapter 1, many pennate diatoms have several copies of DPH genes and are mostly benthic. I therefore also explored the photosensing properties of the DPH homologs in the benthic diatom *Amphora coffeaeformis*. I successfully cloned 3 of these DPHs and showed that they all possess a red-far red light spectra but with variations of the absorption maxima and of the photocycle. To get insights into the function of these DPH, we explored the response of *P.tricornutum* and other pennate diatoms to red-enriched light field like the ones occurring in sediments, and showed the involvement of DPH in acclimation to red light.

Finally, as underlined in this introduction, knowledge on the DPH transduction cascade is missing, and the integration of this bacteriophytochrome in a eukaryote is a puzzle. In the last chapter, I addressed the following question:

*Who are the actors of the different DPH signaling steps?*

I started the *in vitro* characterization of *P.tricornutum* chromophore-producing enzymes, i.e heme oxygenase and putative biliverdin reductases. I also addressed the question of DPH localization *in vivo*, and identify the first component of its signaling cascade by Yeast Two Hybrid. We noticed that other factors, such as cell density in the culture and agitation, also regulate DPH-regulated genes, suggesting that these genes are the targets of signaling networks integrating several environmental cues.

I will discuss the results obtained in these different chapters as a whole in a final conclusion.



# CHAPTER 1: PHOTORECEPTORS REPERTOIRE AND DISTRIBUTION IN OCHROPHYTA

Carole Duchêne, Jean-Pierre Bouly, Angela Falciatore, Marianne Jaubert

CNRS, Sorbonne Université, Institut de Biologie Physico-Chimique, Laboratoire de Biologie du chloroplaste et perception de la lumière chez les microalgues, UMR7141, F-75005 Paris, France

With the rapidly increasing genomic resources for diatoms (genomes and transcriptomes), it was timely to systematically mine for their photoreceptor content using the different characteristic domains of these proteins.

Some very basic questions can be asked, such as inventorying the different known photoreceptors types that are found in this eukaryotic branch. As this has only little been addressed, we expect to find new photoreceptors. Expanding more widely this analysis to the Ochrophytes allowed to determine what part of this repertoire is specific to diatoms, or specific to some other ochrophyte lineages, and what is shared among Ochrophyta. As light fields can be highly different in aquatic environments, it would be interesting, when possible, to link the photoreceptor repertoire of an algal group to its ecological niche (oceanic, freshwater, brackish), as well as their lifestyle (planktonic, benthic) or trophic characteristics (obligate phototroph, mixotroph, heterotroph).

My contribution to this study consisted in gathering the available genomes and transcriptomes sequences, performing HMM search to look for photoreceptors light-sensing domains, and Sequence Similarity Networks. I performed simple phylogenies, to approach the question of the evolutionary origins of the different photoreceptors. When needed, I ran additional analyses, except the homology modelling.

Manuscript in preparation

# PHOTORECEPTORS      REPERTOIRE      AND DISTRIBUTION IN OCHROPHYTA

---

Carole Duchêne, Jean-Pierre Bouly, Angela Falciatore, Marianne Jaubert

CNRS, Sorbonne Université, Institut de Biologie Physico-Chimique, Laboratoire de Biologie du chloroplaste et perception de la lumière chez les microalgues, UMR7141, F-75005 Paris, France

## ABSTRACT

Ochrophyta are a diverse class of photosynthetic eukaryotic algae resulting from secondary endosymbiosis and including ecologically important groups, such as the unicellular diatoms and the multicellular brown algae. Light regulates many aspects of the life of these organisms, likely through photoreceptor-mediated signalling mechanisms. Photoreceptor proteins have been identified in the genomes of model microalgal species, but the amount of genomic and transcriptomic data from marine phytoplankton organisms is rapidly increasing, allowing a more complete exploration of their occurrence and possible diversification. Using a large genome and transcriptome dataset, we provide here a detailed repertoire of the different photoreceptors classes of Ochrophyta: the light-oxygen-voltage (LOV)-based sensors, cryptochromes, rhodopsins and phytochromes. We show different distribution patterns, with some photoreceptors conserved across Ochrophyta (aureochromes, a new group of LOV photosensing protein or some cryptochrome/photolyase family members), but also families specific to some algal groups (some LOV-based photoreceptors, diatom-specific duplication of cryptochrome/photolyase family members), while other photoreceptors are sparsely distributed in Ochrophyta (phytochrome and rhodopsin), likely the result of various horizontal gene transfer events. Hence, it appears that Ochrophyta have evolved different sets of photoreceptors, most likely contributing to their adaptation to specific ecological niches.

Keywords : Ochrophyta, LOV, Cryptochrome, Phytochrome, Rhodopsin

## INTRODUCTION

Light is essential for life on Earth, both as a primary source of energy for phototrophs, but also as a source of information for most organisms, shaping different aspects of their physiology and life cycle. Accordingly, organisms have evolved biological mechanisms to perceive and respond to light (in colors and intensity) that suit different ecological niches (Björn, 2015). Light is absorbed by specific proteins called photoreceptors, which usually contain a chromophore molecule that gives photoreceptors their light-sensing properties (Möglich et al., 2010). Following light absorption, changes in the chromophore conformation are reflected in the structural changes of the protein, triggering signaling cascade(s) leading to cellular response.

Known types of photoreceptors can be classified based on their light-sensing abilities and the nature of the chromophore they bind (Möglich et al., 2010). Most UV-A-Blue light photoreceptors use flavin-based chromophores to perceive light. Flavin-based photoreceptors can be subdivided in three categories. Cryptochrome photoreceptors are known in all kingdoms of life (Bacteria, Archaea and Eukaryotes). They are part of a larger protein family called Cryptochrome/Photolyase Family (CPF) also including DNA photolyases (blue-light activated enzymes that repair UV-B DNA damage such as cyclobutane pyrimidine (CPD) dimer or 6-4 pyrimidine-pyrimidone photoproducts), Cry-DASH for which single and double strand DNA repair, and photoreceptors activities have been associated in a taxa-dependent manner (Kiontke et al., 2020), and light-insensitive transcription repressors (animal type II cryptochrome) (For review see (Chaves et al., 2011)). These proteins usually contain a non-covalently bound Flavin Adenine Dinucleotide (FAD) as cofactor, and eventually additional antenna chromophores (Essen et al., 2017). Other important flavin-blue light photoreceptors include LOV-based photoreceptors. LOV (Light-Oxygen-Voltage) domain, a subclass of Per ARNT Sim (PAS) super family domain, binds flavin mononucleotide (FMN), and can be found in



combination with different effector domains in all kingdoms of Life (Krauss et al., 2009; Losi and Gärtner, 2012; Glantz et al., 2016). The diversity of effector domains associated with a similar sensing domain allows modularity and customized biological response. To illustrate this aspect, one can consider the typical LOV-containing photoreceptors such as phototropins (green lineage specific photoreceptors), with 2 consecutive LOV domains associated with serine/threonine kinase (Christie, 2007), helminchrome proteins coupling two LOV domain and a Regulator of G protein Signaling (RGS) in tandem (Fu et al., 2016) or aureochromes (stramenopiles specific) that combine a LOV-domain with a DNA-binding bZIP domain, thus being a blue-light-driven transcription factor (Takahashi et al., 2007; Kroth et al., 2017). Last flavin-based group of blue-light photoreceptors are BLUF sensors (Blue Light Using Flavin), which are essentially known in Bacteria but also found in several Euglenozoa (Gomelsky and Klug, 2002; Iseki et al., 2002; Masuda and Bauer, 2002).

PYP (Photoactive Yellow Protein) also absorbs blue light but binds 4-hydroxycinnamic acid as chromophore. However, it has only been reported in bacteria so far (Meyer et al., 2012).

Plant UV-B light photoreceptors, UVR-8, do not bind an additional chromophore as their photosensing properties are based on light absorption by the amino acid tryptophan in the protein itself. Study of the phylogeny of these photoreceptors suggests that these are present in Viridiplantae only (Fernández et al., 2016).

Phytochromes are typically red/far-red light sensors that contain covalently attached bilin derivatives as chromophore to perceive light (Rockwell et al., 2006). They are known in bacteria and eukaryotes, and are typically composed of a photosensory module (PSM) containing the PAS, GAF and PHY domains, followed by highly variable output modules (for review see Rockwell et al, 2006; (Rockwell and Lagarias, 2020). Eukaryotic phytochromes form two separate branches on a phylogenetic tree, one containing Archaeplastida and Cryptophyta phytochromes while the other contains fungal and Stramenopile phytochromes (Duanmu et al., 2014; Fortunato et al., 2016). Bacterial phytochromes branch in between these two groups, and the origin of eukaryote phytochrome is a debated issue (Duanmu et al., 2014; Li et al., 2015; Kooß and Lamparter,

2017). Cyanobacteria possess an additional group of specific phytochrome-related photoreceptors (Cyanobacteriochromes, CBCR), which only require a bilin-binding GAF domain, and show a wide variation of absorption spectra reflecting potential light environments adaptation (Ikeuchi and Ishizuka, 2008).

All photoreceptors characterized and mentioned above are intracellular or plasma membrane-associated. Rhodopsins on the other hand are multi-pass membrane proteins. They contain 7 trans-membrane helices and bind a retinal chromophore giving them the ability to sense green and blue light. Rhodopsins are classified in 2 groups, microbial rhodopsins (type I) and animal rhodopsins (type II), that do not share strong sequence homology (Ernst et al., 2014). Animal rhodopsins (opsins) are part of a wider group of membrane proteins also containing chemical receptors (G protein coupled receptors). Microbial rhodopsins can mainly have two biological roles, a bioenergetic one with the use of light as source of energy (through the generation of an electrochemical gradient) or a photosensory function that uses light as source of signal. Several chemical activities can be associated with the light perception properties such as ion transport (H<sup>+</sup>, Na<sup>+</sup>, Cl<sup>-</sup>) associated with the two biological functions (bioenergetics roles or photoperception), light-induced enzymes (photoperception). The recently discovered heliorhodopsins have no ion transport activity and no effector domain, and their function remains unknown (Pushkarev et al., 2018) (For review see (Rozenberg et al., 2021)).

Finally, if different photosensory protein domains can be combined with diverse effector domains, combinations of photosensing domains have also been uncovered in recent years, giving rise to proteins with new light-sensing abilities. Examples of chimeric photoreceptors are present in bacteria, with Ppr, a chimeric photoreceptor from *R. centenum* in which a Photoactive Yellow Protein is fused at amino-terminal of a phytochrome (PYP+PHY) (Jiang et al., 1999), in ferns with neochrome, an association of LOV (amino-terminal) and phytochrome domains (LOV + PHY) (Kawai et al., 2003), or in green algae with the dualchrome a recently characterized photoreceptor composed of a cryptochrome and a phytochrome (CRY+PHY) (Makita et al., 2021).

In aquatic environments, photoperception raises different challenges. Underwater light field varies with depth, due to the wavelength-dependent absorption properties of

water, and the content in suspended particulate material (such phytoplankton and mineral) and coloured dissolved organic matter (Kirk, 2011)(Kirk, 1994; Mobley, 1994). This provides a full variety of photic niches from solar spectra at the surface, to blue/green light in clear oceanic waters or red-enriched light field in turbid waters. Photoreceptor-like sequences from marine microalgae have been recently identified by omics approaches but their physiological properties have been characterized on a very few of them (Jaubert et al., 2017; Petroutsos, 2017). Initial studies have revealed that some algal photoreceptors are similar to those known in plants but new variants with different spectral tuning and algal-specific light sensors have also been found, changing current views and perspectives on how photoreceptor structure and function have diversified in phototrophs experiencing different environmental conditions (for review (Jaubert et al., 2017). Stramenopiles are a class of eukaryotes containing a large number of photosynthetic algae (Ochrophyta) that are of ecological importance, especially in the marine environment (de Vargas et al., 2015). Ochrophyta are a monophyletic clade and include a diversity of organisms, from the unicellular diatoms which are important primary producers in the open ocean, to the brown algae kelp forest in the coastal regions (Dorrell et al., 2022). Moreover, beyond different potential adaptation to specific light environments, Ochrophyta can have different trophic modes: photoautotrophy, mixotrophy and heterotrophy (some genus have to potential for both auto- and heterotrophy, while a few of them have completely loss their photosynthetic abilities)(Beisser et al., 2017; Onyshchenko et al., 2019; Kamikawa et al., 2021; Onyshchenko et al., 2021). These algae result from secondary endosymbiosis of a red algae, and their genome (at least in diatoms) contains red algal genes transferred from the rhodophyte symbiont to its host, but also genes of green algal origins suggesting previous association with a green algae(Dorrell et al., 2022). Horizontal gene transfer from bacteria also occurred at a high rate in Ochrophyta, and their genome is often considered as an evolutionary mosaic (Dorrell et al., 2021). There are initial evidences that this branch of Eukaryotes harbors specific types of photoreceptors, such as aureochromes (bZIP-LOV photoreceptor) (Kroth et al., 2017), but there is currently no global view of the light sensing capabilities of these organisms. With the increasing number of omics data (genomics, transcriptomics) from marine phytoplankton, Ochrophyta thus appear especially interesting from an ecological

and evolutionary point of view to deeply explore marine photoreceptors and get novel insights on the relevance of light sensing for life in aquatic environments.

To this aim, in this work we gathered genetic data (transcriptome and genomes) from different Ochrophyta to explore diversity and distribution of photoreceptors in this branch of the eukaryotic Tree of Life. We focused on LOV-containing proteins, cryptochromes, phytochromes and microbial rhodopsins as no UVR8-, BLUF- or PYP-like photoreceptors have been identified. Even if the physiological roles of photosensing have not been assessed, we uncover new LOV-containing proteins with original domain association as well as a potential new cryptochrome sub-family. Moreover, while some photoreceptors are ubiquitous into Ochrophyta, some appear phyla-specific. Hypotheses on the acquisition of some photoreceptors in Ochrophyta are also discussed.

## RESULTS AND DISCUSSION

### DATASET AND CONSERVATION OF PHOTORECEPTORS

We gathered nearly 9 million genes from 69 Stramenopiles genomes and 260 transcriptomes available on a variety of platforms (NCBI, JGI, MetDB, private repository, see Supp.Data 1 for full description of the dataset). Among this dataset, we used 276 Ochrophyta datasets, to which we added 33 genomes from Oomycota (sister group of Ochrophyta) and 20 genomes and transcriptomes Bigyra (sister group of Gyrista, (i.e. Ochrophyta + Oomycota) to try to infer the origins of the different photoreceptors in this group of eukaryotes. As diatoms (Bacillariophyta) are one of the most studied groups, they represent the majority of the organisms in our dataset (44% (147/329) of the transcriptomes and genomes). Most of the Ochrophyta studied here are phototrophs, with the notable exception of some Chrysophyceae species (12 species, but for some the exact trophic mode is not known) and one diatom strain (*Nitzschia sp*) which are heterotrophs (photosynthesis loss). Moreover, one diatom strain (the diatom *Fragilaria accus*) and some Chrysophyta are known to live in freshwater, the others being marine or brackish.

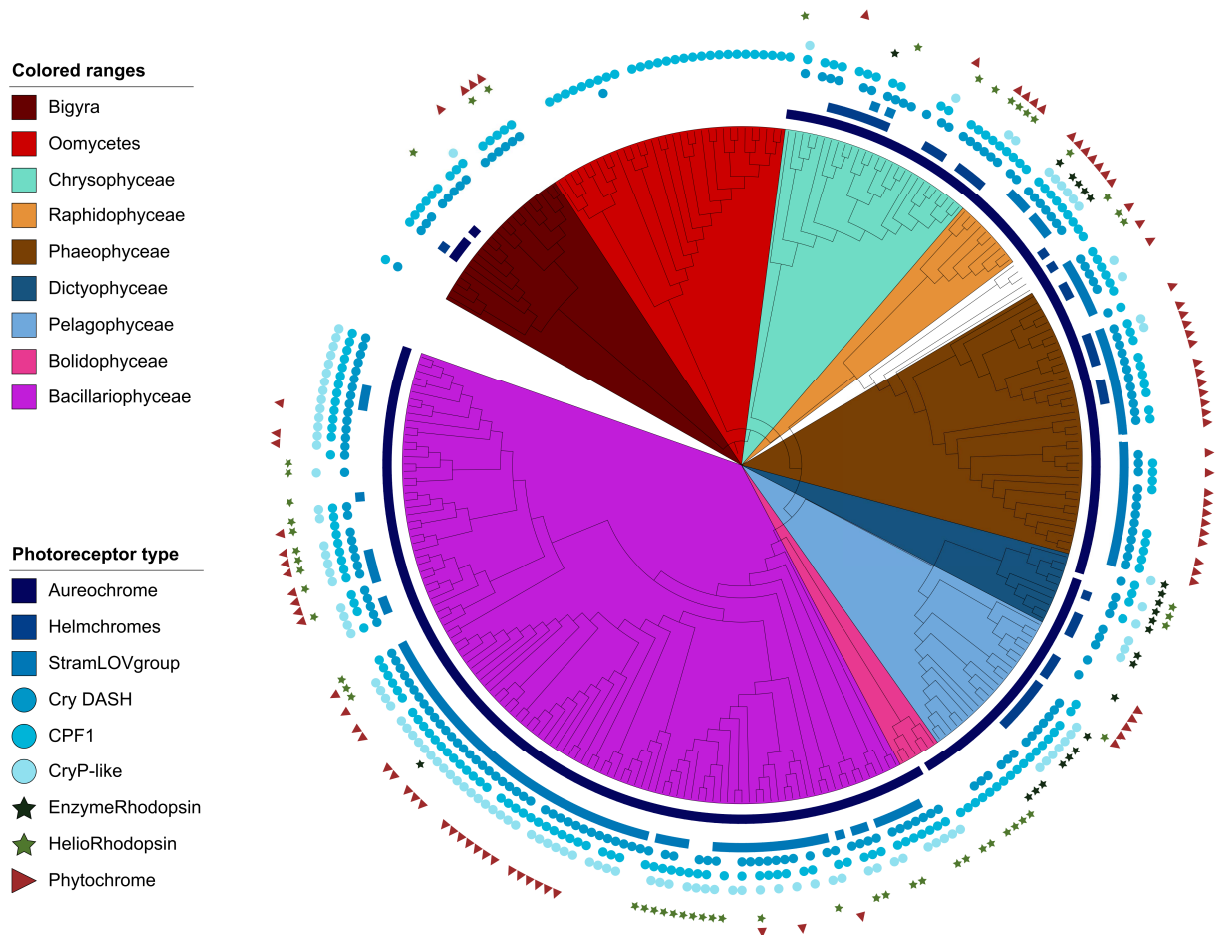


Figure 1. Photoreceptors repertoire in Stramenopiles. Presence of different photoreceptor types visualized on a multi-gene phylogenetic tree of Stramenopile species, colored according to the main algal groups. Outer circles indicate the presence of different photoreceptor types (squares: LOV-based, round: CPF, stars rhodopsins and triangles phytochrome). (White background in the Stramenopiles tree refers to two Eustigmatophyceae and one Xanthophyceae, See FigS1 for the tree with the species names)

We looked for known photoreceptor's light sensing domains in the dataset by using Pfam HMM models or custom HMM models (See Material and Methods and Supp.Data 2). A Sequence Similarity Network (SSN) (Atkinson et al., 2009) was computed for each photoreceptor type (including also reference sequences from other organisms) to classify proteins into sub-groups. Additional stringent filtering criteria specific for different photoreceptor types were further applied for the presence of conserved amino acids interacting with the chromophore, overall structure of the protein, and the presence of transmembrane helices. The protein domain architecture was also examined and eventually followed with phylogenetic tree reconstruction.

A very contrasted picture emerges regarding the photoreceptor domain content among the different Ochrophyta groups (Fig.1). Indeed, some photoreceptors appear ubiquitous in Ochrophyta such as aureochromes and Cry-DASH cryptochromes, while others (helmchromes, rhodopsins, phytochromes) are sparsely distributed on the Ochrophyta species tree. Interestingly, genomes of Oomycetes, sister group of Ochrophyta, exhibit no CPF of the 6-4 photolyase class and no aureochrome and other LOV-based sensors, while Bigyra also have some of the Cry-DASH class, and some LOV-based photoreceptors of the aureochrome family, suggesting a particularly rich content in photoreceptors amongst the Ochrophyta. It is to note that regarding the transcriptomes data, some photoreceptor's domain types may not have been detected in this dataset because they were not expressed. On the other hand, some species appear with high copy numbers of some genes, possibly due to assembly problems. For example, manual examination of phytochrome sequences showed that some truncated sequences, when put together, allow the reconstruction of a single photoreceptor. This manual examination was not done for all photoreceptor types, so the copy numbers are to consider with care. Details of each type of photoreceptors are presented in the following sections.

---

## LOV-BASED PHOTOSENSING: UBIQUITOUS AND SPECIFIC FAMILIES

LOV-domain model was designed with reference sequences from Glantz et al., 2016. Search for this domain in the stramenopile database resulted in 5800 sequences with LOV-domain hits.

Conserved amino acids defined to be necessary for the photochemical activity have been used as functional criteria (Glantz et al., 2016). More precisely the conservation of highly conserved FMN binding motif (GRNCRFLQ), especially the Cys involved in the adduct formation with FMN was verified. As shown by a representative alignment of the different sub-groups of LOV-domain (FigS2A), a strong conservation of the amino acids involved in FMN binding is observed, suggesting these proteins might be functional even if specific substitution affecting photochemicals properties cannot be excluded. Moreover, the presence of other associated protein domains was examined. To classify these genes, we first conducted a SSN using the full length sequences, including reference sequences from

plants, fungi, bacteria and other algae (sequences from Glantz, 2016), with an alignment score of 40 allowing separation of proteins with less than 45 % identity. This resulted in a separation of proteins, with 3 sub-groups containing no stramenopile sequences but only reference sequences from plants with the Adagio/Zeitlupe sub-group, and two sub-groups of fungal proteins including White Collar and Vivid proteins. These groups were removed for further analysis, allowing the identification of 6 sub-groups (alpha, gamma, zeta, theta, Stramenopile LOV sub-group and one larger group, FigS2B) containing Stramenopile sequences and carrying the LOV-FMN binding motif (FigS2A).

3 small sub-groups (of 10 to 40 sequences) contained only sequences of one alga type (alpha: Phaeophyceae, gamma: Pelagophytes, theta: pennate diatoms) (FigS2). Zeta sub-group contained a mix of different algae sequences, including reference sequences from Glaucophytes.

One sub-group contained 130 sequences from stramenopiles only and will be called Stramenopile-specific LOV sub-group (Fig1 and S2).

Last, one large sub-group contained more than 4000 proteins including reference sequences from Viridiplantae, aureochromes and heliochromes from Ochrophyta. This group was further subjected to a second, more stringent SSN step with an alignment score of 60 allowing separation of Viridiplantae and Ochrophyta photoreceptors (FigS2C; named SSN60\_group1 to 5 in FigS3).

Except for aureochromes and Stramenopile-specific LOV proteins, each of the resulting groups seem to be restricted to one branch on the Ochrophyta species tree (examples of the small sub-groups mentioned above, see also below and FigS3). The SSN approach used here does not enable us to decipher the relations between groups, and we would need to perform cautious phylogeny on the LOV-domain(s) to decipher the origin and evolutionary path of LOV-based photoreceptors. This analysis with 2 sequential SSN resulted in 13 LOV subgroups that will be discussed below.

## **Aureochromes: ubiquitous blue-light photoreceptors in Ochrophyta**

Proteins possessing bZIP domains and grouping with characterized aureochromes in SSN were annotated as aureochromes. These photoreceptors were first identified in the multicellular xanthophyte *Vaucheria frigida* as photomorphogenesis regulators, and were recently shown to be master regulators of blue-light responses in the model diatom *Phaeodactylum tricornutum* (Mann et al., 2020). These proteins have been identified in all Ochrophytes, some Bygira, but not in Oomycetes as previously described (Kroth et al., 2017).

Previous work distinguished different groups of aureochromes, named from *V. frigida* aureochromes: “group 1” and “group 2” (Takahashi et al., 2007; Schellenberger Costa et al., 2013). Based on functional characterization, both group 2 aureochromes from *V. frigida* and *P. tricornutum*, are not able to bind the flavin chromophore and thus not considered as bona fide blue light photoreceptors (Takahashi et al., 2007; Banerjee et al., 2016). To explore whether these two proteins might reflect all type 2 aureochrome, we looked at the conservation of a methionine in position M301, replacing a Valine in PtAureo1a (V253) and shown to impair FMN binding in PtAureo2 (Banerjee et al., 2016). In our alignment of type 2 sequences, this position was occupied by a cysteine in VfAureo2, but other proteins had a Valine or an Isoleucine, suggesting that some of the group 2 aureochromes from our dataset could be functional photoreceptors in several organisms (FigS2A). It has also to be noted that differences in the bZIP DNA binding site were also observed between group 1 and group 2 aureochromes suggesting different DNA binding abilities (FigS4B).

We made phylogenetic tree of the different aureochromes found in our data (Fig.2 A, see FigS4 for the unrooted tree). The sequences branched per “Aureotype” with three main branches: type 2, type 1a, and type 1b and c. Type 1b and c are close to each other, forming a clade named 1b/c. On each branch, sequences form clades by algal group, suggesting ancient multiplication and further evolution within the different species. Interestingly, in our data, all groups of algae have a copy of “type 2” aureochrome, one or more copies of “type 1a” and one or more copies of “type b/c” (Fig.2 A). As example, in the species *V. littorea*, in addition to type 2 and type 1a, we report a third aureochrome of type 1 b/c (see



red arrows in Fig2A) while in *V. frigida*, the species in which aureochromes were first discovered, only one of each type 1 and type 2 aureochrome has been characterized (Takahashi et al., 2007; Ishikawa et al., 2009). Interestingly, we also found aureochrome in heterotrophic species such as Chrysophyceae, and the diatom that has lost photosynthetic capacity (*Nitzschia sp* in our study) as already reported for *Nitzschia putrida*, another heterotroph diatom genome (Kamikawa et al., 2021).

Compared to diatoms, which all seem to have one type 2 and three type 1 (1a, 1b, 1c), Pelagophyte, Dictyophytes and Chrysophyte have only two copies of type 1 (1a-like and 1b/c like). Phaeophyceae show expansion of this family with 5 or 6 aureochromes, with two potential type 2, one being a small group containing phaeophyceae sequences close to *Ectocarpus siliculosus* (Esil)Aureo4 and branching at the base of the type 2 aureochrome; two type 1a (see sister Phaeophyceae branches in the 1a branch, also named EsilAureo1 and EsilAureo3 in *Ectocarpus siliculosus* annotation), one sister to diatom type 1b (also named EsilAureo 5) and one close to diatom type 1c (not found in *Ectocarpus siliculosus*, but found in the other *Ectocarpus* strain in our dataset). Whether all Phaeophyceae possess all 6 types of aureochromes has not been investigated in detail, but could be linked to the evolution of multicellularity.

Here, we show that these photoreceptors are ubiquitous within Ochrophyta, but not found in Oomycetes. Type 2, 1a and 1b/c thus seem common aureochrome types in Ochrophyta, suggesting that this multiplication was already present in their ancestor, while multiplication further occurred in brown algae and diatoms. Moreover, the presence of aureochromes in 4 species of Cafeteria/Halocafeteria genus, belonging to Bigyra clade, suggests that this photoreceptor might have originated in the common ancestor to Bigyra and Ochrophyta, with a loss in Oomycetes. Alternatively, this photoreceptor could have been transferred to Cafeteria after their divergence from Oomycetes and Ochrophyta. The exact evolutionary relations between these groups need to be further examined. Moreover, as the tree presented here is rooted with Cafeteria's Aureochrome, this might not be the appropriate outgroup to infer evolutionary history (See FigS4 for the unrooted tree, which shows the same topology).

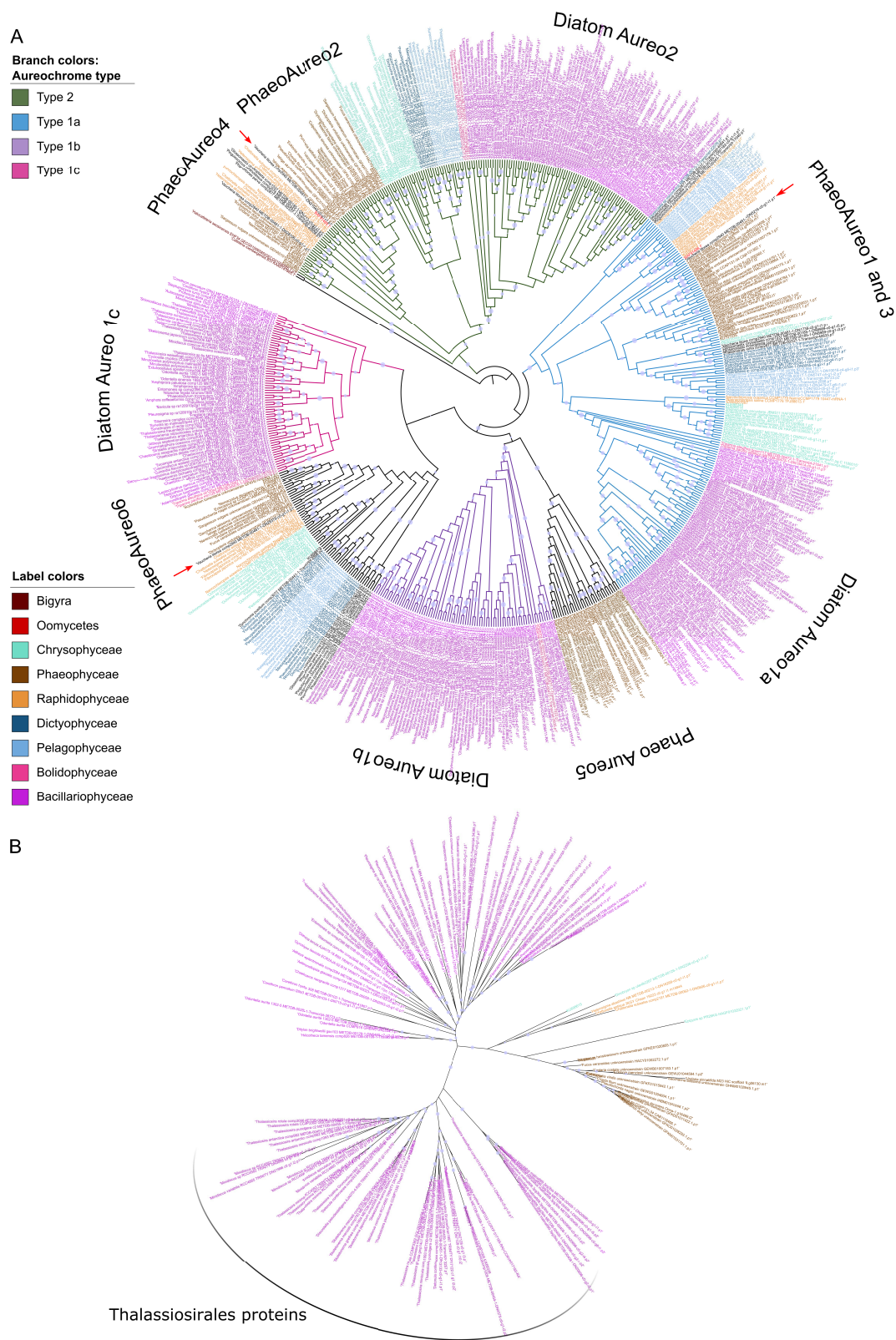


Figure 2. LOV-based photoreceptors in Ochrophyta. A, Aureochrome phylogenetic tree calculated from the alignment of bZIP and LOV domains. The tree was rooted with Cafeteria aureochrome sequences; see Fig S4 for unrooted tree. Branch colors correspond to aureochrome groups according to annotation in *P.*

*tricornutum*, Names are annotated according the *P. tricornutum* annotation (DiatomAureo 1a, b, c and 2) or *Ectocarpus* species (Phaeophyceae «Phaeo»Aureo1 to 6). Red arrows indicate *Vaucheria littorea* aureochromes. B. Unrooted phylogenetic tree of the new group of LOV-based photoreceptors that seems to be widespread in Ochrophyta (full length protein was used). In both trees, the colors of the taxa correspond to the algal groups to which the genes belong (*Vaucheria* and Eustigmatophytes on black), and purple circles on the branches indicate local-bootstraps support >0.7

### **A new group of LOV-based photoreceptors widespread among Ochrophyta**

A group of LOV domain-containing proteins was identified in Ochrophyta from SSN. These short proteins (on average 190 amino acids long) are all made of a single LOV domain (Fig S3B) with a conserved chromophore binding site (FigS2A). The LOV domain is present after 64 amino acids (on average), which might recall the structure of the Vivid proteins in Fungi (Yu and Fischer, 2018).

These proteins are present in a large number of species (3rd circle in Fig1, see also FigS3A): diatoms, brown algae and Raphidophyceae, but absent in Pelagophytes and Dictyophytes branches. This might suggest a common origin of this putative photoreceptor in the ancestor of diatoms, brown algae and Raphidophyceae, with a loss in the Pelagophyte/Dictyophyte branch. Phylogenetically (Fig2B), these proteins form separate branches according to species, with 2 groups of diatom genes. Some diatoms such as *Thalassiosira oceanica* possess genes in both branches, but one branch contains mostly genes from *Thalassiosirale* species.

### **Other LOV-based: Branch-specific specialization?**

As indicated above, SSN analysis allowed us to identify several interesting groups of LOV domain-containing proteins, found restricted to some Ochrophyta groups (Fig. S2 and FigS3).

The “Alpha” sub-group, identified in the first SSN (alignment score 40) contains sequences from Phaeophyceae and from *Vaucheria* (closely related to Phaeophyceae). Most of these sequences possess a PAS\_9 (PF13426) domain, followed by the LOV domain with FMN

binding site, an architecture common in all three domains of life (Glantz et al., 2016), and suggested to have photoperception roles in *Arabidopsis thaliana* (Ogura et al., 2008).

The "Gamma " sub-group, restricted to Pelagophytes, contains proteins with a long N-terminal region followed by the LOV domain. No known conserved domains could be identified in this N-terminal region.

The Zeta sub-group encompasses a mix of proteins from Labyrinthulomycetes (Bigyra), from one diatom species (*Odontella aurita*) and reference sequences from Glaucophytes and Chlorophytes (Glantz et al., 2016). From the Interpro database, this architecture is also known in two proteins from Rhizaria. These proteins possessed GTP cyclohydrolase II domain, which is involved in riboflavin biosynthesis (Ren et al., 2005).

An important group of LOV-based proteins for which experimental studies exist are the helmchromes (Fu et al., 2016), which contain PAS<sub>9</sub> and RGS domains (Regulator of G protein Signaling). These proteins are thought to be involved in phototaxis of flagellated cells, and are detected by immunofluorescence in the flagella of sperm cells in some Phaeophyceae and in the flagella of some Chrysophyceae. Here, these photoreceptors were not found ubiquitously in Ochrophyta: they are present in some branches of Chrysophyceae, Phaeophyceae, Pelagophytes and Dictyophytes, but they are absent from diatoms and from Raphidophytes. Our distribution of helmchrome in Ochrophyta agrees with experimental data detecting the presence of this protein in Phaeophyceae and Chrysophyceae (Fu et al., 2016), and also with the presence of helmchromes in Pelagophytes and Dictyophytes reported recently (Coesel et al., 2021). The domain architecture of these helmchromes is RGS-LOV-PAS<sub>9</sub>-RGS-LOV-PAS<sub>9</sub>, although sometimes sequences are too short to span all the domains. Fungi also possess RGS-LOV domain proteins of unknown function, which form a separate group in a phylogenetic tree with helmchromes (FigS4), suggesting independent evolution. Interestingly, these fungal proteins are recruited to the plasma membrane upon blue-light illumination (Glantz et al., 2016; Glantz et al., 2018), suggesting that helmchromes could perform their photosensing function through the same type of mechanisms.

Although concerning only a few diatom strains, some proteins showed new domain association such as LOV domains fused to CAP-GLY (Cytoskeleton Associated Proteins

with glycine-rich domain), which might bind the microtubules (Weisbrich et al., 2007) (FigS3). In the Interpro database, this structure was not reported for species other than diatoms, suggesting that this new LOV protein class is specific for diatoms. This is an interesting domain combination, as such proteins could regulate light-induced intracellular movements of organelles in the absence of phototropins, which are known to regulate plastid movement in Viridiplantae (Christie, 2007).

Finally, by looking for LOV domain-based photoreceptors, we identified a wide variety of proteins, from the ubiquitous photoreceptor aureochromes to novel putative photoreceptor families restricted to some species only. Eukaryotic LOV-containing proteins are thought to have evolved from bacterial gene transfer, either through endosymbiotic gene transfer from the mitochondria (plant phototropins and fungal LOV-photoreceptors) or from the chloroplast (plant *Zeitlupe*) (Krauss et al., 2009). This hypothesis has been challenged for the origin of *Zeitlupe*, but plant phototropins remain a sister group to alpha-proteobacteria LOV proteins in more recent analyses (Losi et al., 2015). Aureochromes group close to the plant phototropins (Ishikawa et al., 2009; Djouani-Tahri et al., 2011), suggesting common origin. However, independent horizontal gene transfer (HGT) is possible for the other LOV photoreceptors identified here. For example, the LOV-HK photoreceptors from the green algae *Ostreococcus* and *Micromonas* are closer to bacterial genes than to phototropins, suggesting more recent horizontal gene transfer from bacteria (Djouani-Tahri et al., 2011). The origin of the LOV domains in Ochrophyta (common origin with plant phototropins vs independent HGT), and the possible domain shuffling events (Di Roberto and Peisajovich, 2014), will need to be further analyzed with phylogenetic studies of the LOV domain only. In addition, we found here new eukaryotic representatives of LOV-based photoreceptors, particularly represented in Dictyophyceae and Pelagophyceae (FigS3), that might be included in broader studies on the origin of LOV domain in Eukaryotes.

## THE CRYPTOCHROME PHOTOLYASE FAMILY: CONSERVED BUT DIVERSE BIOCHEMICAL FUNCTION

CPFs constitute a large group of flavoproteins with an impressive diversification (Chaves et al., 2011; Fortunato et al., 2015). CPF proteins share a common domain called Photolyase Homology Region (PHR) but are associated with a multitude of functions. CPF can be phylogenetically separated into 5 different superclasses regrouping proteins with different functions : (i) the bacterial Fes-BCP or Cry-Pro class, (ii) the 6-4 photolyase class including animal light-dependent and light-independent cryptochromes, (iii) the class I/III CPD photolyases including the plants cryptochromes (photoreceptor) and plant-like cryptochromes, (iv) the class II CPD photolyases, and (v) Cry-DASH (*Drosophila Arabidopsis Synechococcus Homo*) class. Due to the large diversity among the different members of this family, CPF domain architectures alone did not lead to sufficient information for a functional prediction (i.e., photolyase vs blue light photoreceptors and light-independent transcriptional regulation). Therefore, several criteria such as SSN, phylogenetic tree, or biochemical information need to be used and combined to give insights into the possible function of these proteins. A new functional classification of this protein family based on the use of computational multiple profile models has been recently proposed (Vicedomini et al., 2022). The model diatom *Phaeodactylum tricornutum* possesses 7 CPF members: PtCPF1, a 6-4 photolyase with gene regulation activity (Coesel et al., 2009), PtCryP, a plant-like cryptochrome having a role in gene expression regulation (Juhas et al., 2014; König et al., 2017), PtCPF2 and 4 belonging to the Cry-DASH sub-family, two class II CPD, PtCPD3 and PtCPD4, and PtCPD1, a NCRY (New CRY) protein which might represent a new class of light photosensing protein (Emmerich et al., 2020; Vicedomini et al., 2022).

We tested several HMM domain models (FAD binding domain 7-PF03441, DNA photolyase-PF00875, and custom model using COG0415 sequences from the conserved protein domain family PHRB). Only the one using COG0415 sequences could recognize all 7 members of the CPF family in the model diatom *Phaeodactylum* and was further used for the search in our database. The resulting SSN of the dataset (1830 sequences including references) separated proteins into 13 groups (FigS5) when applying an alignment score of 90.

Interestingly, sub-classes containing only diatoms and Bolidophyceae sequences (group 1 and 2) and a mix of stramenopile ones (group 3 to 5) did not group with other known sequences, even with lower alignment scores (FigS5). Reconstruction of a phylogenetic tree agreeing with the SSN (Fig3), placed group 1 and 2 close to Plant-like Cryptochrome including PtCryP (Juhas et al., 2014), and these were further named “CryP2a” and “CryP2b”.

Group 3 to 5 form an independent clade. All sequences in this clade had DNA photolyase domain (PF00875) followed by alpha beta hydrolase 6 domain (PF12697), and most of them are possibly addressed to the chloroplast based on HECTAR prediction (Gschloessl et al., 2008) (FigS5A). Mining other proteins showing this structure on Interpro retrieved unreviewed proteins from land plants and green algae (Viridiplantae). This observation was already done in a work on photolyases by Emmerich et al, 2020. As these proteins did not possess FAD binding domain, they are unlikely to function as photoreceptors.

CPF from the 6-4 photolyases, Cry-DASH class and class II CPD photolyase sub-families are present in all Ochrophyta and Bigyra (see Fig1 and S5). However, Oomycetes lack Cry-DASH as already observed in some fungi such as *Aspergillus nidulans* (Schumacher, 2017; Corrochano, 2019). Plant-like cryptochromes group is more sparsely distributed but appears conserved in diatoms. Interestingly duplication of Plant-like cryptochromes, CRY-DASH and class II CPD photolyase seems to be uniquely found in diatoms and in their sister group, the bolidophyceae, suggesting duplication occurred in their common ancestor. Last, NCRY proteins are only found in diatoms, Bolidophyceae and Raphidophyceae.

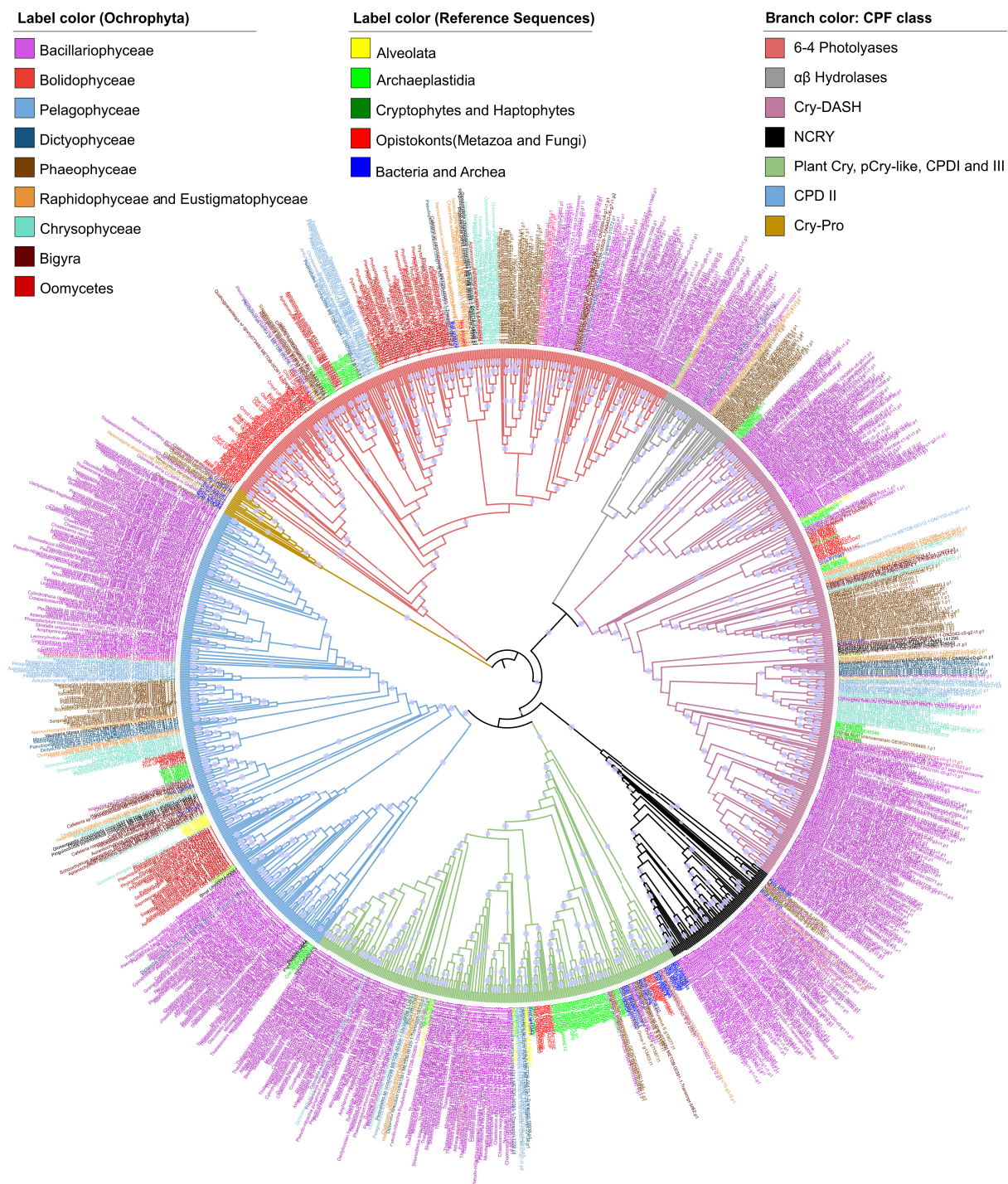


Figure 3. Ochrophyta CPF family protein tree. Tree was rooted at the base of the CryPro class. Branch colors corresponds to major CPF class, see also FigS5 to 8 for zoom on different branches. Taxa colors correspond to the algal groups to which the genes belong, and purple circles on the branches indicate local bootstraps support >0.7



Looking more in detail the Class II-like photolyase CPD, duplication seems to have occurred in Chlorophyta and Diatoms, forming a sister clade to the main CPD photolyase class II branch (Fig3, FigS6). Due to a strong divergence of the primary sequence, multi sequence alignment and homology modeling of the structure did not allow to identify active sites of photolyase activity. However, while 25% of the diatom class II CPD photolyases are predicted to be addressed to organelles (22/87 diatom proteins with mito/chloro/signalP prediction), the duplicated Class II CPD photolyase in diatoms do not seem to be addressed to the organelles (0 mito/chloro/signalP prediction with HECTAR) (FigS5). These results, in agreement with functional characterization of PHR1 and PHR2 in *Chlamydomonas* and PiPhr1 and PiPhr2, from the antarctic diatom *Phaeodactylum tricornutum* ICE-Hand suggest that only diatoms and Chlorophytes may have specific nuclear and chloroplast-targeted double strand CPD DNA photolyase activity (Small and Greimann, 1977; Petersen et al., 1999; An et al., 2021).

Similar results have been observed with CRY-DASH proteins whose function is still unclear (ssDNA photolyase or light sensors). We clearly distinguished 2 clades of Cry-DASH in Ochrophyta, one CRY-DASH-like clade containing only diatoms and another one containing sequences from all Ochrophyta (Fig3, FigS7). This includes a diatom clade, a brown algae clade, then a group with mixed Pelagophytes, Dictyophytes and Raphidophytes. Cry-DASH proteins have been found to be localized in chloroplasts or mitochondria in many organisms. Here, 44% of the diatom CRY-DASH proteins were predicted to be addressed to organelles (PtCPF2 group), while only 16% of the diatom CRY-DASH-like ones were (PtCPF4 group, FigS5). This might also suggest cellular compartmentalization and possible functional specialization. Interestingly, like for the class II CPD photolyases, duplication also occurred in Chlorophyta, and the duplicated Chlorophyta genes are on the same branch as CRY-DASH-like diatom genes.

Only one gene copy of the 6-4 photolyase class has been identified in Ochrophyta, a potential common feature to photosynthetic organisms, but this was also the case in Bigyra. The Ochrophyta 6-4 photolyases form a clade with a Chrysophyceae, a Phaeophyceae and 2 diatom clades (one sister to Phaeophyceae, the other basal to these), while Oomycete proteins group close to Viridiplantae (Fig3, FigS8). In *Phytophthora* species in particular, we could identify up to 3 copies (FigS8). As in insects, this could

mean sub-functionalization of the different copies with either 6-4 photolyase activity or cryptochrome functions. To support this idea, we calculated the isoelectric point (pI) of the 6-4 photolyase family members from oomycetes, considering that proteins with different pI might have different functions (Kiraga et al., 2007; Mohanta et al., 2019). Most photoreceptors of the CPF family (animal or plant) are characterized by an acidic pI. We found that the 3 groups of oomycota sequences have different pI: mean±sd of  $5.37\pm 0.44$  for group CPF1aa,  $7.3\pm 0.99$  (CPF1ab) and  $6.05\pm 0.8$  (CPF1b, FigS8B). Moreover, CPF with photoreceptor functions potentially exhibit a C-terminal extension, a feature found in at least two copies of the proteins present in Oomycetes. Indeed, the DNA photolyase domain starts at position 105 ( $\pm 11$ ) for CPF1aa, position 3 (except one sequence) for CPF1ab and position 96 ( $\pm 23$ ) for CPF1b. This suggests at least two different functions are carried by these proteins, and that CPF1aa probably are photoreceptors. Although speculative, it would also seem surprising that multiple proteins with a DNA repair of pyrimidine–pyrimidone (6–4) photoproducts which represent only ~20% of the DNA UV damage, would have been kept in Oomycetes while no CPD photolyase has been detected, suggesting that at least one of these oomycetes 6-4 photolyase family members might be a photoreceptor.

Plant-like Cryptochromes are the CPF that branch closest to the plant blue light photoreceptor cryptochromes (Fig. 3, S9 and (Juhas et al., 2014)). In the diatom *P. tricornutum*, both *CryP* knock-down and knock-out lines showed that this protein has a role in gene expression regulation (Juhas et al., 2014; König et al., 2017). Two groups of stramenopile sequences can be distinguished from the SSN and on the phylogenetic tree, forming 2 sister clades (named CryP1 and CryP2) themselves sisters to plant cryptochromes. CryP1 group contains a few sequences from Raphidophyta, Pelagophytes and Dictyophytes but mostly diatom ones, with the representative PtCryP. CryP2 group contains only diatom and Bolidophyte sequences, including the “CryP-like” sequence from the model diatom *Thalassiosira pseudonana*. Most diatoms have a gene from each group, such as *Thalassiosira oceanica* or *Navicula sp* (FigS9). Moreover, the CryP2 itself is separated into 2 groups (in both SSN and phylogenetic tree), called CryP2 a and b. By looking for different biochemical properties and different FAD fixation sites, CryP1 has a much higher average pI ( $8.2\pm 1.13$  for CryP1 compared to  $6.09\pm 0.7$  for CryP2, FigS9)

suggesting that CryP1 and CryP2 would bear different functions. Sequences of the FAD binding sites show also some differences between CryP1, CryP2a and b (FigS10A) and structure reconstruction on Swiss-model for one of each CryP sub-families members gave slightly different results suggesting also different photochemical properties (FigS10B). Overall, in Ochrophyta, we have seen duplication in diatoms into two sub-family (CryP1 and CryP2) probably associated with neofunctionalization of putative new light photosensors.

Finally, the last class is the NCRY, with PtCPD1 as diatom representative. Its position in the phylogenetic tree is not clear, either linked with the CRY-DASH or the CPD photolyase class I family in different studies (Lucas-Lledo and Lynch, 2009; Oliveri et al., 2014; Fortunato et al., 2016; Vicedomini et al., 2022). This gene was found mostly in diatoms, with a few sequences from Phaeophyceae and Raphidophyceae. This group was identified by SSN by Emmerich 2020, but only one functional study exists. Experimental approaches on NCRY from *Vibrio cholera* have shown that this protein was able to bind FAD as chromophore but was lacking photolyase activity. In addition, an *in silico* model of the NCRY sequences showed that it presents a unique FAD active site, the absence of few amino acids involved in CPD interactions, and the presence of a C-terminal extension (Vicedomini et al., 2022). All these characters support the existence of a novel functional class within the CPF, with new photochemical properties in diatoms.

Here we showed that Cry-DASH and (6-4) photolyases are well conserved amongst ochrophyta, while there is an expansion of the different CPF families in diatoms (NCRY, CRY-DASH-like, CPD Photolyase-like class II, and CRYP2) (FigS5). Given the complexity of the CPF, we cannot decipher whether these are photoreceptors but members of CryP2 group might be good candidates as new photosensible proteins.

---

## RHODOPSINS: DIVERSITY AND SPARSE DISTRIBUTION

We found 3 types of rhodopsins in Ochrophytes: proteorhodopsins, histidine kinase rhodopsins (HK rhodopsins) and heliorhodopsins (Fig. S11, S12 and 4). All proteins

reported here (86 proteins) have at least 5 trans-membrane helices as predicted by TMHMM, and were separated by SSN.

As already reported before, some diatoms possess Proteorhodopsins that may function as proton pumps using light to produce ATP. It has been proposed that these proteins could therefore have a bioenergetic role, particularly useful in iron-limited oceanic regions (Marchetti et al., 2012; Marchetti et al., 2015). We found in our dataset a similar distribution of Proteorhodopsins as reported in Marchetti et al, 2015, with the presence of this family in diatoms (with only a few additional new diatom sequences from species not included in previous analyses), in Cryptophytes and in Haptophytes (Fig S12). Within this family, some Proteorhodopsins are known to show spectral tuning, at least in bacteria: their absorption spectra is shifted towards blue if a glutamine is at position 106 (of the gamma-proteobacteria rhodopsin isolated at the Hawaii Ocean Time series station, Uniprot reference Q9AFF7) (Man, 2003). Here all diatom sequences had a leucine or a methionine at this position, suggesting green-absorbing spectra.

A group of Ochrophyta rhodopsins possessed additional signaling domains, with HATPase and Response Regulator domains (HK rhodopsins). HK rhodopsins are reported here in Raphidophytes and Pelagophyte, and have been previously described in green algae as *Chlamydomonas reinhardtii*, where it seems to work as a UV/BL photoreceptor (Luck et al., 2012). However, its biological function in green algae remains unknown. Phylogenetically, Stramenopile enzyme rhodopsins form a clade grouped with enzyme rhodopsins from other marine phytoplankton groups (*Emiliana huxleii* Coccolithophyceae, Cryptophytes and Mameliophyceae) (Fig4A). The sequences from green algae *Chlamydomonales*, including *C. reinhardtii* histidine kinase rhodopsin 1, are sister to this group.

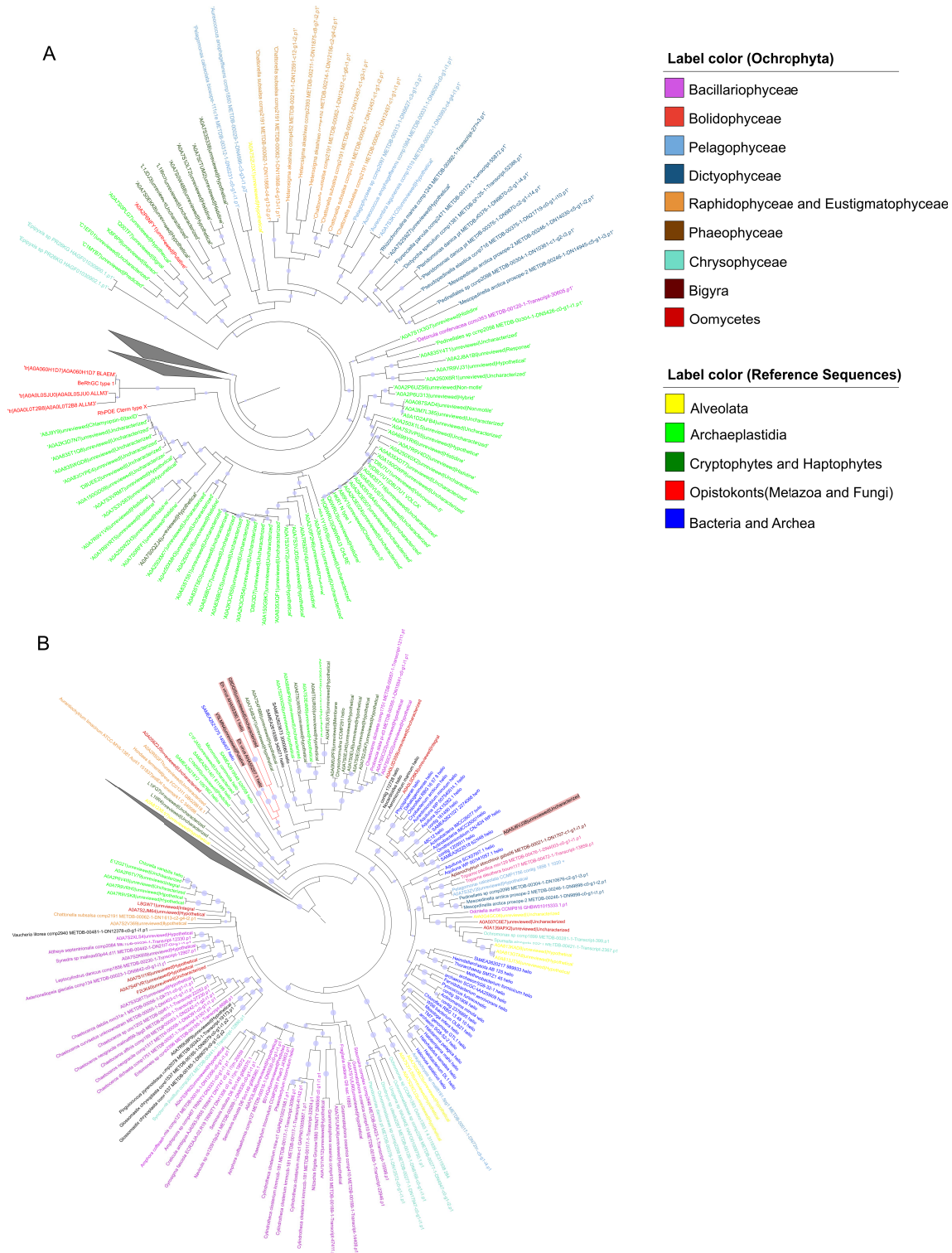


Figure 4 Phylogeny of the rhodopsins. A. HK rhodopsins and B heliorhodopsins, Taxa colors correspond to the algal groups the genes belong to, with a red background for viral sequences, and purple circles on the branches indicate local bootstraps support >0.7. Both trees were rooted with other type 1 rhodopsin from (Pushkarev et al., 2018), collapsed clades.

We also studied the presence of heliorhodopsins, which has been identified in a functional screen of environmental sequences and has already been found in *P. tricornutum* (Pushkarev et al, 2018). Heliorhodopsins have a slow photocycle, suggesting that it would function as photoreceptor (Pushkarev et al., 2018). The majority of stramenopile heliorhodopsins identified here groups with other eukaryotes rhodopsins (Fig4B), including Chlorophyta, Glaucophyta and Euglenozoa sequences. Within this branch, at least 3 groups of diatom heliorhodopsins are found. A few stramenopile heliorhodopsins group with Fungi or Haptophyte. We underline that at the base of the clade containing haptophyte sequences, 2 sequences come from Haptophyta viruses, suggesting horizontal transfer of heliorhodopsin gene between viruses and their host.

#### PHYTOCHROMES: CONSERVED BACTERIAL PHYTOCHROME (BPHP)-LIKE STRUCTURE BUT SPARSE CONSERVATION

The photosensing module of phytochrome is composed of three domains including the specific PHY domain which connects the chromophore binding domains PAS-GAF with the output module. Search for Pfam PHY domain (PF00360) resulted in proteins all showing the same domain architecture (except for one brown alga *Cladosiphon* sequence, and some sequences that are truncated): (PAS)-GAF-PHY-HisKA-HATPase-(REC), which represents the structure of bacteriophytochromes. Full length sequences were aligned and used to design stramenopiles- and phytochrome-specific domain models for each domain except for the PAS domain, which has been found not well-conserved in some diatoms (Fortunato et al., 2016). These Stramenopile- and phytochrome-specific models were used in combination with SSN to identify truncated sequences. The presence of the conserved cysteine residue involved in the linkage of the bilin chromophore was verified. Most of the proteins had the Cys residue at the N-terminal extremity of the protein, conserved with the chromophore-binding cysteine in bacteriophytochromes, but brown algae and some phytochromes from the diatom *Amphora coffeaeformis* also possessed an additional cysteine in the GAF domain, involved in chromophore binding in plant and cyanobacteria phytochromes (Rockwell et al., 2014; Fortunato et al., 2016). In addition, we noticed that Chrysophyte sequences (*Ochromonas* species) only possessed the

cysteine in the GAF domain, suggesting that these phytochromes bind a different chromophore and have different light-sensing abilities (FigS12).

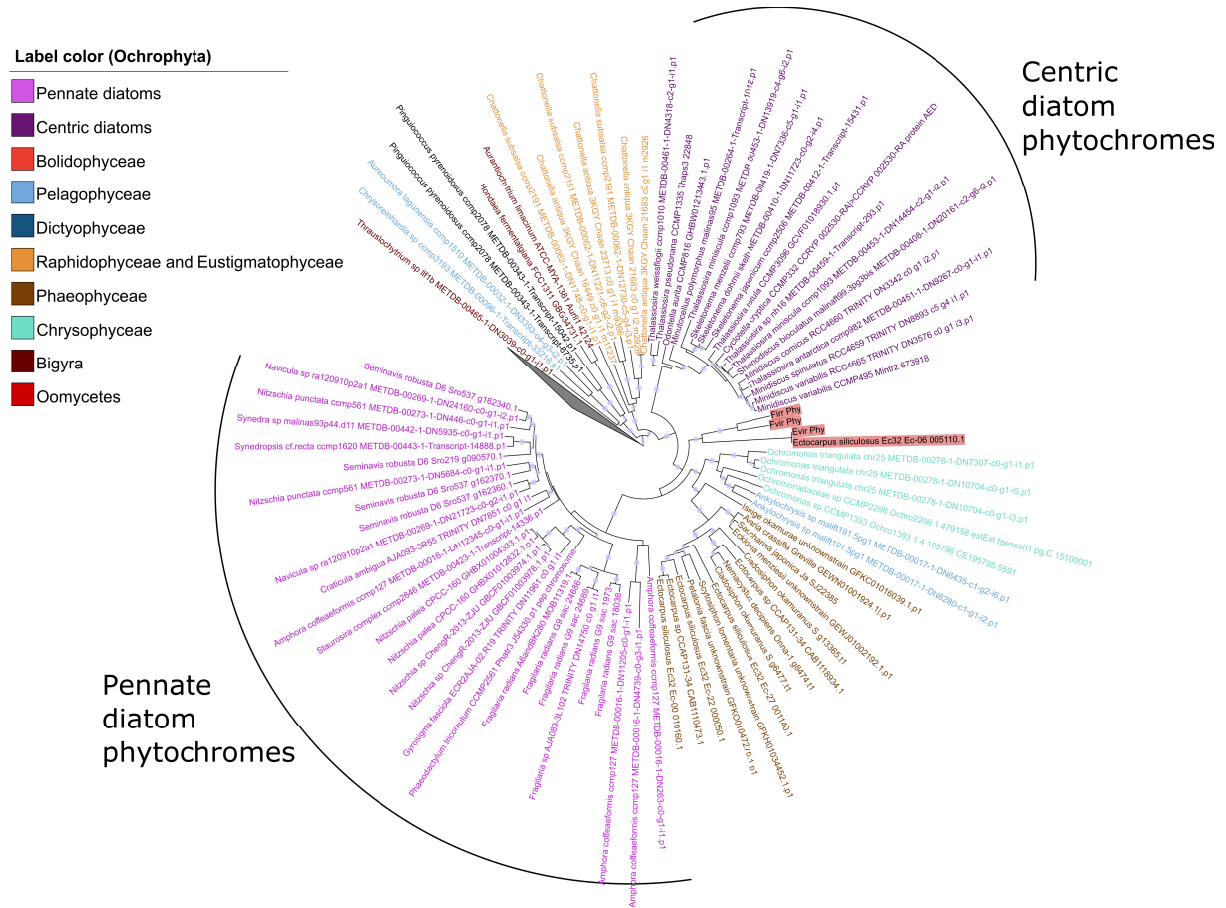


Figure 5 Phylogeny of the Stramenopile phytochromes. Full-length sequences were aligned and the tree was rooted with fungal and bacterial phytochromes (collapsed clade). Taxa colors correspond to the algal groups the genes belongs to, with a red background for viral sequences, and purple circles on the branches indicate local bootstraps support >0.7

Phylogenetic tree of phytochrome full-length proteins reveals a complex story (Fig 5). As reported before (Duanmu et al., 2014; Fortunato et al., 2015), all stramenopile phytochromes seem to share a common origin, branching as sister clade to fungal phytochromes (Fig5). However, the relative position of phytochrome branches does not reflect the species phylogenetic relationships. The most striking example is the relation between diatoms, brown algae and chrysophyceae. In the phylogenetic species tree, diatoms are separated into pennate species forming a monophyletic clade, and centric

species that are paraphyletic, branching as sister groups basal to pennate. Bolidophyceae form the sister group to all diatoms (Kooistra and Medlin, 1996; Kessenich et al., 2014), but in our search of phytochrome, Bolidophyceae do not possess any phytochromes. Brown algae diverged much earlier in the history of Ochrophyta. Pennate phytochrome sequences form a monophyletic clade, sister to chrysophyceae and multicellular brown algae phytochromes; the resulting group (pennate-brown algae-chrysophyceae) being sister to centric phytochromes. Interestingly, some sequences from brown algae viruses are branching just outside the Pennate+brown algae+chrysophyceae group (see red background in Fig5). This latter point could suggest that virus-mediated gene transfer may have also contributed to phytochrome evolution amongst ochrophytes. The ancestor of pennate and centric diatom could have acquired phytochrome, which then diverged into centric and pennate branches. Brown algae and chrysophyceae would have acquired phytochromes from a pennate species, possibly through viral horizontal gene transfer.

Another interesting point is the phytochrome copy number (FigS11). From first studies (Fortunato et al., 2016) and further investigation of this omic dataset, it is clear that not all Ochrophyta possess a phytochrome gene. Even in diatoms, within both the pennate and centric diatoms, there are species which do not possess this putative photoreceptor. At present, we have not been able to find a clear explanation linking the presence/absence of the phytochrome and the species that possess it. Multiplication seems also to have occurred several times in different algal lineages. In the centric diatoms, there seem to be only one phytochrome gene copy, while pennate species have up to 4 different phytochromes suggesting that gene duplication occurred after separation between centric and pennate diatoms. Moreover, brown algae and Raphidophyceae also have several phytochrome genes (Fig5, FigS13). As suggested already by previous analyses (Fortunato et al., 2016; Osuna-Cruz et al., 2020), it seems that gene duplication occurred in benthic species preferentially (*Seminavis* and *Amphora* especially). Because these environments are heterogeneous, characterized by sharp and sometimes dynamic gradients of light, oxygen, nutrient availability, and redox state (Stockdale et al., 2009; Cartaxana et al., 2016; Marques da Silva et al., 2017), it is possible to speculate that phytochrome gene duplication represented an adaptation to thrive in this complex environments. Considering the huge biodiversity of diatoms and the important amount of



novel information deriving from environmental genomics, it will be interesting to further investigate phytochrome distribution in different diatom species, and try to associate it to their environmental distribution and diversity of life style and strategies.

The origin of stramenopile phytochromes remains puzzling. When compared to other phytochromes, there is clearly a stramenopile specific phytochrome branch, sister to fungal phytochromes, but separated from other eukaryotic algal phytochromes (Archaeplastida and Cryptophyceae) by bacterial phytochromes. This would suggest a common origin for stramenopile phytochrome. However, within the stramenopile phytochromes, the phytochrome phylogenetic tree does not follow the phylogenetic species tree. One possibility could be acquisition of phytochrome in one ochrophyta species, then within-ochrophyta horizontal gene transfer through viral infection.

## CONCLUSIONS

In this study, we have addressed the presence of different photoreceptors types in Ochrophyta. We have shown that the LOV-based photosensing seems to be very important in this eukaryote branch, with Aureochrome photoreceptors present in all Ochrophyta. Moreover, we also discover a Stramenopile-specific group containing a single LOV domain as potential ubiquitous photoreceptors in diatoms, while phyla-specific proteins appeared in some algal groups. However, whether all these photoreceptors evolved from a unique ancestral protein has not been assessed here, but could be done in an other study focusing on the phylogeny of the LOV domain. We also uncovered new domain combinations, for example with cytoskeleton-interacting domains, with possibly new functions for LOV-based photoreceptors. Members of the cryptochrome-photolyase family are also present in all species studied here, and among Ochrophyta, diatoms show gene multiplication in several CPF classes (CryP, Cry-DASH and class II CPD photolyase). In particular, we have been able to show a multiplication of Plant-like cryptochrome in diatoms with the potential presence of new photoreceptors (CRYP2) and duplication associated with potential specific subcellular localization (Cry-DASH and class II CPD photolyase).

The conservation of phytochromes and rhodopsins is much more sparse on the Ochrophyta species tree. Interestingly, each of these proteins can be found in giant virus genomes (phycodnaviridae). There are several arguments in favor of a viral transfer for both rhodopsins and phytochromes. Rhodopsins are found in algal viruses, including Chlorophyta and coccolithophore viruses, and transfer of rhodopsin genes between virus and host have been demonstrated in other algae (Needham et al., 2019; Rozenberg et al., 2020). For phytochromes, the presence of a copy of a giant phycodnavirus genome carrying a phytochrome in the *Ectocarpus* genome is a good argument for virus-mediated horizontal gene transfer (Cock et al., 2010). Giant phycodnavirus are known to infect brown algae (McKeown et al., 2017), raphidophytes (Ogura et al., 2016) and pelagophytes (Moniruzzaman et al., 2014), and have been suggested to infect chrysophytes too (Endo, 2020), but not diatoms. These can be infected by small ssDNA and ssRNA viruses, but the giant phycodnaviridae are thought to be too big to get through the pores of the frustule (Tomaru and Nagasaki, 2011). Interestingly, if we focus on pennate diatom phytochromes, we can see that these genes are sparsely distributed on the diatom phylogenetic tree, but the genes form a clade on a phytochrome phylogenetic tree. This could mean that the ancestor of pennate diatoms had a phytochrome, but that it was repeatedly lost in the current lineages. However, we also noticed that some pennate diatoms that carry phytochrome are benthic species, living in dense sea floor communities, where viral pressure is high (Stal et al., 2019; McMinn et al., 2020). Viral-mediated gene transfer could therefore be very successful in this environment, and bring phytochrome and heliorhodopsin genes to these species. On this basis, one can infer that pennate diatom phytochromes form a clade not because they come from the same diatom ancestor, but because these diatoms share the same environment and are infected by the same viruses. Mining the available giant DNA virus databases (Moniruzzaman et al., 2020; Schulz et al., 2020) for rhodopsins and phytochromes genes and comparing them to Ochrophyta genes will be the next step in this analysis.

Our database is composed of a wide range of Ochrophyta species. Because our search is based on known characterized photoreceptors domains, it is nearly certain that other photoreceptors types exist and could be identified by enriching this search with additional genomic data now available from algae (Nelson et al., 2021), or environmental

sequences derived from meta-omic data and single cell genomes (Carradec et al., 2018; Delmont et al., 2020). It would be also interesting to add to this search dinotoms (dinoflagellates that possess a chloroplast of diatom origin) to identify conserved photoreceptors and eventual transfers from diatoms to the dinoflagellate host genome (Hehenberger et al., 2016). Finally, it would be informative to look more into species characteristics of the different ochrophyta groups possessing different photoreceptor classes, such as trophic mode and habitat (freshwater or marine, planktonic or benthic), and identify putative blind spots in our analysis.

Finally, representing the first comprehensive study of the photoreceptor repertoire in an important group of eukaryotes that have so far been largely ignored in photobiology research, this study provides important support for the discovery of new photosensitive proteins and for the study of their evolution and functional diversification in aquatic environments.

## MATERIAL AND METHODS:

### **Database:**

All transcriptome and genome resources used in this analysis are listed in Supplementary Data 1. We used only genomes with protein predictions and assembled transcriptomes. Protein prediction was performed with Transdecoder (version 5.5.0, default settings) on assembled transcriptomes if needed.

### **Photoreceptor search (summarized in Table1):**

We used the hmmsearch algorithm (HMMER 3.2.1) to retrieve candidate sequences. HMM model details are as follow:

LOV: the LOV domain is recognized by the PAS\_9 model in Pfam (PF13426), however this model also recognizes other domains (lacking the flavin binding site notably). To define a specific LOV domain signature, we aligned all LOV-containing sequences from Glantz et

al, 2016 to the PAS\_9 model with hmalign (--trim option) and constructed LOV hmm profile with hmmbuild.

CPF: We tested Pfam models PF and PF models , but a custom model recognized more sequences. This custom model was built by aligning reference sequences from COG0415 (ncbi conserved domain database) with mafft (v7.247) and built with hmmbuild.

Rhodopsins: Pfam models PF01036 (bac-rhodopsin) and PF18761 (heliorhodopsins) have been used.

Phytochromes: Pfam model PF00360 (PHY) has been used to retrieve full length sequences, as defined by sequences with GAF and REC domains, aligned with mafft). Alignment of the full-length sequence has been used to build stramenopile specific models for each domain (GAF, PHY, HisKA, HATPase, REC).

Table 1. Details of the parameters and reference sequences used in our photoreceptor search

Photoreceptor type	HMM domain	SSN alignment score	Reference sequences
LOV	custom (PF13426 and reference sequences from Glantz 2016)	40 and 60	(Glantz et al., 2016) and haptophyte RGS-LOV sequences reported by (Coesel et al., 2021)
CPF	custom (alignment of the COG0415 sequences from ncbi conserved domain database)	90	(Vicedomini et al., 2022)
Rhodopsins	PF01036 (bac-rhodopsin) and PF18761 (heliorhodopsins)	20	(Govorunova et al., 2017; Pushkarev et al., 2018) enriched in eukaryotic sequences from Interpro for the phylogenetic trees
Phytochromes	PF00360 and custom from Stramenopiles sequences	80	(Fortunato et al., 2016)

Candidate sequences were clustered at 90% identity with Cdhit (<http://weizhong-lab.ucsd.edu/cdhit-web-server/cgi-bin/index.cgi>) and submitted to Sequence similarity network (SSN) at EFI/EST (<https://efi.igb.illinois.edu/efi-est/>, fasta option) (Zallot et al., 2019). SSN alignment scores were chosen to separate proteins at about 40% identity and

the score has been increased if necessary. We proceeded with trial and error to find optimal alignment scores.

Reference sequences were added to the SSN (see Table1) to help to visualize known photoreceptor types.

Other domain examination and sequence visualization was done in CLCworkbench 8.0.1 (Pfam database version 33, built-in hmmer algorithm version 3.1b1).

### **Phylogenetic reconstruction:**

Alignment of selected proteins was done with muscle (v3.8.31), except for CPF, where sequences were aligned to a reference alignment of the DNA photolyase and FAD-binding domains with hmalign with the --trim option. Alignments were trimmed with trimAL (v1.2) (80% conservation), and tree was reconstructed with FastTree (Price et al., 2010) (version 2.1.11), default mode (JTT+CAT; local support values: Shimodaira-Hasegawa test with 1000 resampling).

For species phylogeny, we used the mutligene method from Dorell et al, 2021: 63 genes from *Phaeodactylum tricornutum* were used as blastp queries to identify homologs in each strain of our database (best blast hit for each strain was used). Genes were aligned individually with mafft, trimmed with trimAL (-gt 0.2) and alignments concatenated. Tree was reconstructed with FastTree.

Trees were visualized with iTOL v6 (<https://itol.embl.de>, (Letunic and Bork, 2021)).

### **Complementary analyses:**

Sequences logos (aureochromes binding site and CryP family) were done at <https://weblogo.berkeley.edu/logo.cgi>. Protein targeting predictions were done with HECTAR (Gschloessl et al., 2008), diatom sequences only, at <http://www.sb-roscoff.fr/hectar/>. Isoelectric points were calculated with the online tool IPC2 ((Kozlowski, 2021), <http://ipc2.mimuw.edu.pl>). Transmembrane prediction for the

rhodopsins were done with TMHMM 2.0 at <https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>.

Structure homology modelisation was performed with SwissProt ((Waterhouse et al., 2018), <https://swissmodel.expasy.org> ).

## BIBLIOGRAPHY

**An M, Qu C, Miao J, Sha Z** (2021) Two class II CPD photolyases, PiPhr1 and PiPhr2, with CPD repair activity from the Antarctic diatom *Phaeodactylum tricornutum* ICE-H. 3 *Biotech* **11**: 377

**Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC** (2009) Using Sequence Similarity Networks for Visualization of Relationships Across Diverse Protein Superfamilies. *PLoS ONE* **4**: e4345

**Banerjee A, Herman E, Kottke T, Essen L-O** (2016) Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**: 171–178

**Beisser D, Graupner N, Bock C, Wodniok S, Grossmann L, Vos M, Sures B, Rahmann S, Boenigk J** (2017) Comprehensive transcriptome analysis provides new insights into nutritional strategies and phylogenetic relationships of chrysophytes. *PeerJ* **5**: e2832

**Björn LO, ed** (2015) *Photobiology: The Science of Light and Life*. doi: 10.1007/978-1-4939-1468-5

**Carradec Q, Pelletier E, Da Silva C, Alberti A, Seeleuthner Y, Blanc-Mathieu R, Lima-Mendez G, Rocha F, Tirichine L, Labadie K, et al** (2018) A global oceans atlas of eukaryotic genes. *Nat Commun*. doi: 10.1038/s41467-017-02342-1

**Cartaxana P, Ribeiro L, Goessling J, Cruz S, Kühl M** (2016) Light and O<sub>2</sub> microenvironments in two contrasting diatom-dominated coastal sediments. *Mar Ecol*

Prog Ser **545**: 35–47

**Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T, Brettel K, Essen L-O, van der Horst GTJ, Batschauer A, Ahmad M** (2011) The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. *Annu Rev Plant Biol* **62**: 335–364

**Christie JM** (2007) Phototropin Blue-Light Receptors. *Annu Rev Plant Biol* **58**: 21–45

**Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury J-M, Badger JH, et al** (2010) The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature* **465**: 617–621

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

**Coesel SN, Durham BP, Groussman RD, Hu SK, Caron DA, Morales RL, Ribalet F, Armbrust EV** (2021) Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. *Proc Natl Acad Sci* **118**: e2011038118

**Corrochano LM** (2019) Light in the Fungal World: From Photoreception to Gene Transcription and Beyond. *Annu Rev Genet* **53**: 149–170

**Delmont TO, Gaia M, Hinsinger DD, Fremont P, Vanni C, Guerra AF, Eren AM, Kourlaiev A, d'Agata L, Clayssen Q, et al** (2020) Functional repertoire convergence of distantly related eukaryotic plankton lineages revealed by genome-resolved metagenomics. doi: 10.1101/2020.10.15.341214

**Di Roberto RB, Peisajovich SG** (2014) The role of domain shuffling in the evolution of signaling networks: SIGNALING NETWORKS EVOLUTION. *J Exp Zool B Mol Dev Evol* **322**: 65–72

**Djouani-Tahri EB, Christie JM, Sanchez-Ferandin S, Sanchez F, Bouget FY, Corellou F** (2011) A eukaryotic LOV-histidine kinase with circadian clock function in the picoalga *Ostreococcus*. *Plant J* **65**: 578–588

- Dorrell R, Lui F, Bowler C** (2022) Reconstructing Dynamic Evolutionary Events in Diatom Nuclear and Organelle Genomes. *Mol. Life Diatoms*, Springer. pp 147–177
- Dorrell RG, Villain A, Perez-Lamarque B, Audren de Kerdrel G, McCallum G, Watson AK, Ait-Mohamed O, Alberti A, Corre E, Frischkorn KR, et al** (2021) Phylogenomic fingerprinting of tempo and functions of horizontal gene transfer within ochrophytes. *Proc Natl Acad Sci* **118**: e2009974118
- Duanmu D, Bachy C, Sudek S, Wong C-H, Jimenez V, Rockwell NC, Martin SS, Ngan CY, Reistetter EN, van Baren MJ, et al** (2014) Marine algae and land plants share conserved phytochrome signaling systems. *Proc Natl Acad Sci* **111**: 15827–15832
- Emmerich H-J, Saft M, Schneider L, Kock D, Batschauer A, Essen L-O** (2020) A topologically distinct class of photolyases specific for UV lesions within single-stranded DNA. *Nucleic Acids Res* **48**: 12845–12857
- Endo H** (2020) Biogeography of marine giant viruses reveals their interplay with eukaryotes and ecological functions. **4**: 23
- Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H** (2014) Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chem Rev* **114**: 126–163
- Essen L, Franz S, Banerjee A** (2017) Structural and evolutionary aspects of algal blue light receptors of the cryptochrome and aureochrome type ✱. *J Plant Physiol* **217**: 27–37
- Fernández MB, Tossi V, Lamattina L, Cassia R** (2016) A Comprehensive Phylogeny Reveals Functional Conservation of the UV-B Photoreceptor UVR8 from Green Algae to Higher Plants. *Front Plant Sci* **7**: 1–6
- Fortunato AE, Annunziata R, Jaubert M, Bouly JP, Falciatore A** (2015) Dealing with light: The widespread and multitasking cryptochrome/photolyase family in photosynthetic organisms. *J Plant Physiol* **172**: 42–54
- Fortunato AE, Jaubert M, Enomoto G, Bouly J-P, Raniello R, Thaler M, Malviya S, Bernardes JS, Rappaport F, Gentili B, et al** (2016) Diatom Phytochromes Reveal the



Existence of Far-Red-Light-Based Sensing in the Ocean. *Plant Cell* **28**: 616–628

**Fu G, Nagasato C, Yamagishi T, Kawai H, Okuda K, Takao Y, Horiguchi T, Motomura T** (2016) Ubiquitous distribution of helmchrome in phototactic swimmers of the stramenopiles. *Protoplasma* **253**: 929–941

**Glantz ST, Berlew EE, Jaber Z, Schuster BS, Gardner KH, Chow BY** (2018) Directly light-regulated binding of RGS-LOV photoreceptors to anionic membrane phospholipids. *Proc Natl Acad Sci* **115**: E7720–E7727

**Glantz ST, Carpenter EJ, Melkonian M, Gardner KH, Boyden ES, Wong GKS, Chow BY** (2016) Functional and topological diversity of LOV domain photoreceptors. *Proc Natl Acad Sci U S A* **113**: E1442–E1451

**Gomelsky M, Klug G** (2002) BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem Sci* **27**: 497–500

**Govorunova EG, Sineshchekov OA, Li H, Spudich JL** (2017) Microbial Rhodopsins: Diversity, Mechanisms, and Optogenetic Applications. *Annu Rev Biochem* **86**: 845–872

**Gschloessl B, Guermeur Y, Cock JM** (2008) HECTAR: A method to predict subcellular targeting in heterokonts. *BMC Bioinformatics* **9**: 393

**Guajardo M, Jimenez V, Vaultot D, Trefault N** Transcriptomes from the diatoms *Thalassiosira* and *Minidiscus* from the English Channel and Antarctica. 13

**Hehenberger E, Burki F, Kolisko M, Keeling PJ** (2016) Functional Relationship between a Dinoflagellate Host and Its Diatom Endosymbiont. *Mol Biol Evol* **33**: 2376–2390

**Ikeuchi M, Ishizuka T** (2008) Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. *Photochem Photobiol Sci* **7**: 1159

**Iseki M, Matsunaga S, Murakami A, Ohno K, Shiga K, Yoshida K, Sugai M, Takahashi T, Hori T, Watanabe M** (2002) A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. *Nature* **415**: 1047–1051

**Ishikawa M, Takahashi F, Nozaki H, Nagasato C, Motomura T, Kataoka H** (2009)

Distribution and phylogeny of the blue light receptors aureochromes in eukaryotes. *Planta* **230**: 543–552

**Jackson C, Salomaki ED, Lane CE, Saunders GW** (2017) Kelp transcriptomes provide robust support for interfamilial relationships and revision of the little known *Arthrothamnaceae* (Laminariales). *J Phycol* **53**: 1–6

**Jaubert M, Bouly JP, Ribera d'Alcalà M, Falciatore A** (2017) Light sensing and responses in marine microalgae. *Curr Opin Plant Biol* **37**: 70–77

**Jiang ZY, Swem LR, Rushing BC, Devanathan S, Tollin G, Bauer CE** (1999) Bacterial photoreceptor with similarity to photoactive yellow protein and plant phytochromes. *Science* **285**: 406–409

**Juhas M, Von Zadow A, Spexard M, Schmidt M, Kottke T, Büchel C** (2014) A novel cryptochrome in the diatom *Phaeodactylum tricornutum* influences the regulation of light-harvesting protein levels. *FEBS J* **281**: 2299–2311

**Kamikawa R, Mochizuki T, Sakamoto M, Tanizawa Y, Nakayama T, Onuma R, Cenci U, Moog D, Speak S, Sarkozi K, et al** (2021) Genome evolution of a non-parasitic secondary heterotroph, the diatom *Nitzschia putrida*. doi: 10.1101/2021.01.24.427197

**Kawai H, Kanegae T, Christensen S, Kiyosue T, Sato Y, Imaizumi T, Kadota A, Wada M** (2003) Responses of ferns to red light are mediated by an unconventional photoreceptor. *Nature* **421**: 287–290

**Kessenich CR, Ruck EC, Schurko AM, Wickett NJ, Alverson AJ** (2014) Transcriptomic Insights into the Life History of Bolidophytes, the Sister Lineage to Diatoms. *J Phycol* **50**: 977–983

**Kiontke S, Göbel T, Brych A, Batschauer A** (2020) DASH-type cryptochromes – solved and open questions. *Biol Chem* **401**: 1487–1493

**Kiraga J, Mackiewicz P, Mackiewicz D, Kowalczyk M, Biecek P, Polak N, Smolarczyk K, Dudek MR, Cebrat S** (2007) The relationships between the isoelectric point and: length of proteins, taxonomy and ecology of organisms. *BMC Genomics* **8**: 163

**Kirk JTO** (2011) Light and photosynthesis in aquatic ecosystems, 3rd Editio. Camb Univ Press. doi: 10.1017/S0025315400044180

**König S, Eisenhut M, Brutigam A, Kurz S, Weber APM, Bchel C** (2017) The Influence of a Cryptochrome on the Gene Expression Profile in the Diatom *Phaeodactylum tricornutum* under Blue Light and in Darkness. *Plant Cell Physiol* **58**: 1914–1923

**Kooistra WHCF, Medlin LK** (1996) Evolution of the Diatoms (Bacillariophyta). *Mol Phylogenet Evol* **6**: 391–407

**Kooß S, Lamparter T** (2017) Cyanobacterial origin of plant phytochromes. *Protoplasma* **254**: 603–607

**Kozlowski LP** (2021) IPC 2.0: prediction of isoelectric point and p *K* a dissociation constants. *Nucleic Acids Res* **49**: W285–W292

**Krauss U, Minh BQ, Losi A, Gärtner W, Eggert T, Von Haeseler A, Jaeger KE** (2009) Distribution and phylogeny of light-oxygen-voltage-blue-light-signaling proteins in the three kingdoms of life. *J Bacteriol* **191**: 7234–7242

**Kroth PG, Wilhelm C, Kottke T** (2017) An update on aureochromes: Phylogeny – mechanism – function. *J Plant Physiol* **217**: 20–26

**Kvernvik AC, Rokitta SD, Leu E, Harms L, Gabrielsen TM, Rost B, Hoppe CJM** (2020) Higher sensitivity towards light stress and ocean acidification in an Arctic sea-ice-associated diatom compared to a pelagic diatom. *New Phytol* **226**: 1708–1724

**Letunic I, Bork P** (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**: W293–W296

**Li F, Melkonian M, Rothfels CJ, Villarreal JC, Stevenson DW, Graham SW, Wong GK, Pryer KM, Mathews S** (2015) Phytochrome diversity in green plants and the origin of canonical plant phytochromes. *Nat Commun* **6**: 1–12

**Losi A, Gärtner W** (2012) The Evolution of Flavin-Binding Photoreceptors: An Ancient Chromophore Serving Trendy Blue-Light Sensors. *Annu Rev Plant Biol* **63**: 49–72

**Losi A, Mandalari C, Gärtner W** (2015) The Evolution and Functional Role of Flavin-

based Prokaryotic Photoreceptors. *Photochem Photobiol* **91**: 1021–1031

**Lucas-Lledo JI, Lynch M** (2009) Evolution of Mutation Rates: Phylogenomic Analysis of the Photolyase/Cryptochrome Family. *Mol Biol Evol* **26**: 1143–1153

**Luck M, Mathes T, Bruun S, Fudim R, Hagedorn R, Tran Nguyen TM, Kateriya S, Kennis JTM, Hildebrandt P, Hegemann P** (2012) A Photochromic Histidine Kinase Rhodopsin (HKR1) That Is Bimodally Switched by Ultraviolet and Blue Light. *J Biol Chem* **287**: 40083–40090

**Ma L, Wang X, Guan Z, Wang L, Wang Y, Zheng L, Gong Z, Shen C, Wang J, Zhang D, et al** (2020) Structural insights into BIC-mediated inactivation of Arabidopsis cryptochrome 2. *Nat Struct Mol Biol* **27**: 472–479

**Makita Y, Suzuki S, Fushimi K, Shimada S, Suehisa A, Hirata M, Kuriyama T, Kurihara Y, Hamasaki H, Okubo-Kurihara E, et al** (2021) Identification of a dual orange/far-red and blue light photoreceptor from an oceanic green picoplankton. *Nat Commun* **12**: 3593

**Man D** (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**: 1725–1731

**Mann M, Serif M, Wrobel T, Eisenhut M, Madhuri S, Flachbart S, Weber APM, Lepetit B, Wilhelm C, Kroth PG** (2020) The Aureochrome Photoreceptor PtAUREO1a Is a Highly Effective Blue Light Switch in Diatoms. *iScience* **23**: 101730

**Marchetti A, Catlett D, Hopkinson BM, Ellis K, Cassar N** (2015) Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *ISME J* **9**: 2745–2748

**Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, Morales R, Allen AE, Armbrust EV** (2012) Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci* **109**: E317–E325

**Marques da Silva J, Cruz S, Cartaxana P** (2017) Inorganic carbon availability in benthic

diatom communities: photosynthesis and migration. *Philos Trans R Soc B Biol Sci* **372**: 20160398

**Masuda S, Bauer CE** (2002) AppA Is a Blue Light Photoreceptor that Antirepresses Photosynthesis Gene Expression in *Rhodobacter sphaeroides*. *Cell* **110**: 613–623

**McKeown DA, Stevens K, Peters AF, Bond P, Harper GM, Brownlee C, Brown MT, Schroeder DC** (2017) Phaeoviruses discovered in kelp (Laminariales). *ISME J* **11**: 2869–2873

**McMinn A, Liang Y, Wang M** (2020) Minireview: The role of viruses in marine photosynthetic biofilms. *Mar Life Sci Technol* **2**: 203–208

**Meyer TE, Kyndt JA, Memmi S, Moser T, Colón-Acevedo B, Devreese B, Van Beeumen JJ** (2012) The growing family of photoactive yellow proteins and their presumed functional roles. *Photochem Photobiol Sci* **11**: 1495

**Möglich A, Yang X, Ayers RA, Moffat K** (2010) Structure and Function of Plant Photoreceptors. *Annu Rev Plant Biol* **61**: 21–47

**Mohanta TK, Khan A, Hashem A, Abd\_Allah EF, Al-Harrasi A** (2019) The molecular mass and isoelectric point of plant proteomes. *BMC Genomics* **20**: 631

**Moniruzzaman M, LeClerc GR, Brown CM, Gobler CJ, Bidle KD, Wilson WH, Wilhelm SW** (2014) Genome of brown tide virus (AaV), the little giant of the Megaviridae, elucidates NCLDV genome expansion and host–virus coevolution. *Virology* **466–467**: 60–70

**Moniruzzaman M, Martinez-Gutierrez CA, Weinheimer AR, Aylward FO** (2020) Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nat Commun* **11**: 1710

**Needham DM, Yoshizawa S, Hosaka T, Poirier C, Choi CJ, Hehenberger E, Irwin NAT, Wilken S, Yung C-M, Bachy C, et al** (2019) A distinct lineage of giant viruses brings a rhodopsin photosystem to unicellular marine predators. *Proc Natl Acad Sci* **116**: 20574–20583

**Nelson DR, Hazzouri KM, Lauersen KJ, Jaiswal A, Chaiboonchoe A, Mystikou A, Fu W, Daakour S, Dohai B, Alzahmi A, et al** (2021) Large-scale genome sequencing reveals the driving forces of viruses in microalgal evolution. *Cell Host Microbe* **29**: 250-266.e8

**Nishitsuji K, Arimoto A, Higa Y, Mekar M, Kawamitsu M, Satoh N, Shoguchi E** (2019) Draft genome of the brown alga, *Nemacystus decipiens*, Onna-1 strain: Fusion of genes involved in the sulfated fucan biosynthesis pathway. *Sci Rep* **9**: 4607

**Ogura Y, Hayashi T, Ueki S** (2016) Complete Genome Sequence of a Phycodnavirus, *Heterosigma akashiwo* Virus Strain 53. *Genome Announc* **4**: e01279-16, /ga/4/6/e01279-16.atom

**Ogura Y, Tokutomi S, Wada M, Kiyosue T** (2008) PAS/LOV proteins: A proposed new class of plant blue light receptor. *Plant Signal Behav* **3**: 966–968

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

**Onyshchenko A, Roberts WR, Ruck EC, Lewis JA, Alverson AJ** (2021) The genome of a nonphotosynthetic diatom provides insights into the metabolic shift to heterotrophy and constraints on the loss of photosynthesis. *New Phytol* **232**: 1750–1764

**Onyshchenko A, Ruck EC, Nakov T, Alverson AJ** (2019) A single loss of photosynthesis in the diatom order Bacillariales (Bacillariophyta). *Am J Bot* **106**: 560–572

**Osuna-Cruz CM, Bilcke G, Vancaester E, De Decker S, Bones AM, Winge P, Poulsen N, Bulankova P, Verhelst B, Audoor S, et al** (2020) The *Seminais robusta* genome provides insights into the evolutionary adaptations of benthic diatoms. *Nat Commun* **11**: 3320

**Parks MB, Wickett NJ, Alverson AJ** (2018) Signal, Uncertainty, and Conflict in Phylogenomic Data for a Diverse Lineage of Microbial Eukaryotes (Diatoms, Bacillariophyta). *Mol Biol Evol* **35**: 80–93

**Petersen JL, Lang DW, Small GD** (1999) Cloning and characterization of a class II DNA photolyase from *Chlamydomonas*. *Plant Mol Biol* **40**: 1063–1071

**Petroutsos D** (2017) *Chlamydomonas* Photoreceptors: Cellular Functions and Impact on Physiology. In M Hippler, ed, *Chlamydomonas Biotechnol. Biomed.* Springer International Publishing, Cham, pp 1–19

**Price MN, Dehal PS, Arkin AP** (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **5**: e9490

**Pushkarev A, Inoue K, Larom S, Flores-Uribe J, Singh M, Konno M, Tomida S, Ito S, Nakamura R, Tsunoda SP, et al** (2018) A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* **558**: 595–599

**Ren J, Kotaka M, Lockyer M, Lamb HK, Hawkins AR, Stammers DK** (2005) GTP Cyclohydrolase II Structure and Mechanism. *J Biol Chem* **280**: 36912–36919

**Rockwell NC, Duanmu D, Martin SS, Bachy C, Price DC, Bhattacharya D, Worden AZ, Lagarias JC** (2014) Eukaryotic algal phytochromes span the visible spectrum. *Proc Natl Acad Sci* **111**: 3871–3876

**Rockwell NC, Lagarias JC** (2020) Phytochrome evolution in 3D: deletion, duplication, and diversification. *New Phytol* **225**: 2283–2300

**Rockwell NC, Su Y-S, Lagarias JC** (2006) Phytochrome Structure and Signaling Mechanisms. *NIH Public Access* **57**: 837–858

**Rozenberg A, Inoue K, Kandori H, Béjà O** (2021) Microbial Rhodopsins: The Last Two Decades. *Annu Rev Microbiol* **75**: 427–447

**Rozenberg A, Oppermann J, Wietek J, Fernandez Lahore RG, Sandaa R-A, Bratbak G, Hegemann P, Béjà O** (2020) Lateral Gene Transfer of Anion-Conducting Channelrhodopsins between Green Algae and Giant Viruses. *Curr Biol* **30**: 4910-4920.e5

**Scheerer P, Zhang F, Kalms J, von Stetten D, Krauß N, Oberpichler I, Lamparter T** (2015) The Class III Cyclobutane Pyrimidine Dimer Photolyase Structure Reveals a New Antenna Chromophore Binding Site and Alternative Photoreduction Pathways. *J Biol*

Chem **290**: 11504–11514

**Schellenberger Costa B, Sachse M, Jungandreas A, Bartulos CR, Gruber A, Jakob T, Kroth PG, Wilhelm C, Schellenberger Costa B, Sachse M, et al** (2013) Aureochrome 1a Is Involved in the Photoacclimation of the Diatom *Phaeodactylum tricornutum*. PLoS ONE. doi: 10.1371/journal.pone.0074451

**Schulz F, Roux S, Paez-Espino D, Jungbluth S, Walsh DA, Denev VJ, McMahon KD, Konstantinidis KT, Eloie-Fadrosch EA, Kyrpides NC, et al** (2020) Giant virus diversity and host interactions through global metagenomics. *Nature* **578**: 432–436

**Schumacher J** (2017) How light affects the life of *Botrytis*. *Fungal Genet Biol* **106**: 26–41

**Shan T, Yuan J, Su L, Li J, Leng X, Zhang Y, Gao H, Pang S** (2020) First Genome of the Brown Alga *Undaria pinnatifida*: Chromosome-Level Assembly Using PacBio and Hi-C Technologies. *Front Genet* **11**: 140

**Small GD, Greimann CS** (1977) Photoreactivation and dark repair of ultraviolet light-induced pyrimidine dimers in chloroplast DNA. *Nucleic Acids Res* **4**: 2893–2902

**Stal LJ, Bolhuis H, Cretoiu MS** (2019) Phototrophic marine benthic microbiomes: the ecophysiology of these biological entities. *Environ Microbiol* **21**: 1529–1551

**Stockdale A, Davison W, Zhang H** (2009) Micro-scale biogeochemical heterogeneity in sediments: A review of available technology and observed evidence. *Earth-Sci Rev* **92**: 81–97

**Takahashi F, Yamagata D, Ishikawa M, Fukamatsu Y, Ogura Y, Kasahara M, Kiyosue T, Kikuyama M, Wada M, Kataoka H** (2007) AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *Proc Natl Acad Sci* **104**: 19625–19630

**Tomaru Y, Nagasaki K** (2011) Diatom Viruses. *In* J Seckbach, P Kociolek, eds, *Diatom World*. Springer Netherlands, Dordrecht, pp 211–225

**de Vargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R, Lara E, Berney C, Le**



**Bescot N, Probert I, et al** (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605–1261605

**Vicedomini R, Bouly JP, Laine E, Falciatore A, Carbone A** (2022) Multiple Profile Models Extract Features from Protein Sequence Data and Resolve Functional Diversity of Very Different Protein Families. *Mol Biol Evol* **39**: msac070

**Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al** (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**: W296–W303

**Weisbrich A, Honnappa S, Jaussi R, Okhrimenko O, Frey D, Jelesarov I, Akhmanova A, Steinmetz MO** (2007) Structure-function relationship of CAP-Gly domains. *Nat Struct Mol Biol* **14**: 959–967

**Yu Z, Fischer R** (2018) Light sensing and responses in fungi. *Nat Rev Microbiol*. doi: 10.1038/s41579-018-0109-x

**Zallot R, Oberg N, Gerlt JA** (2019) The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **58**: 4169–4182



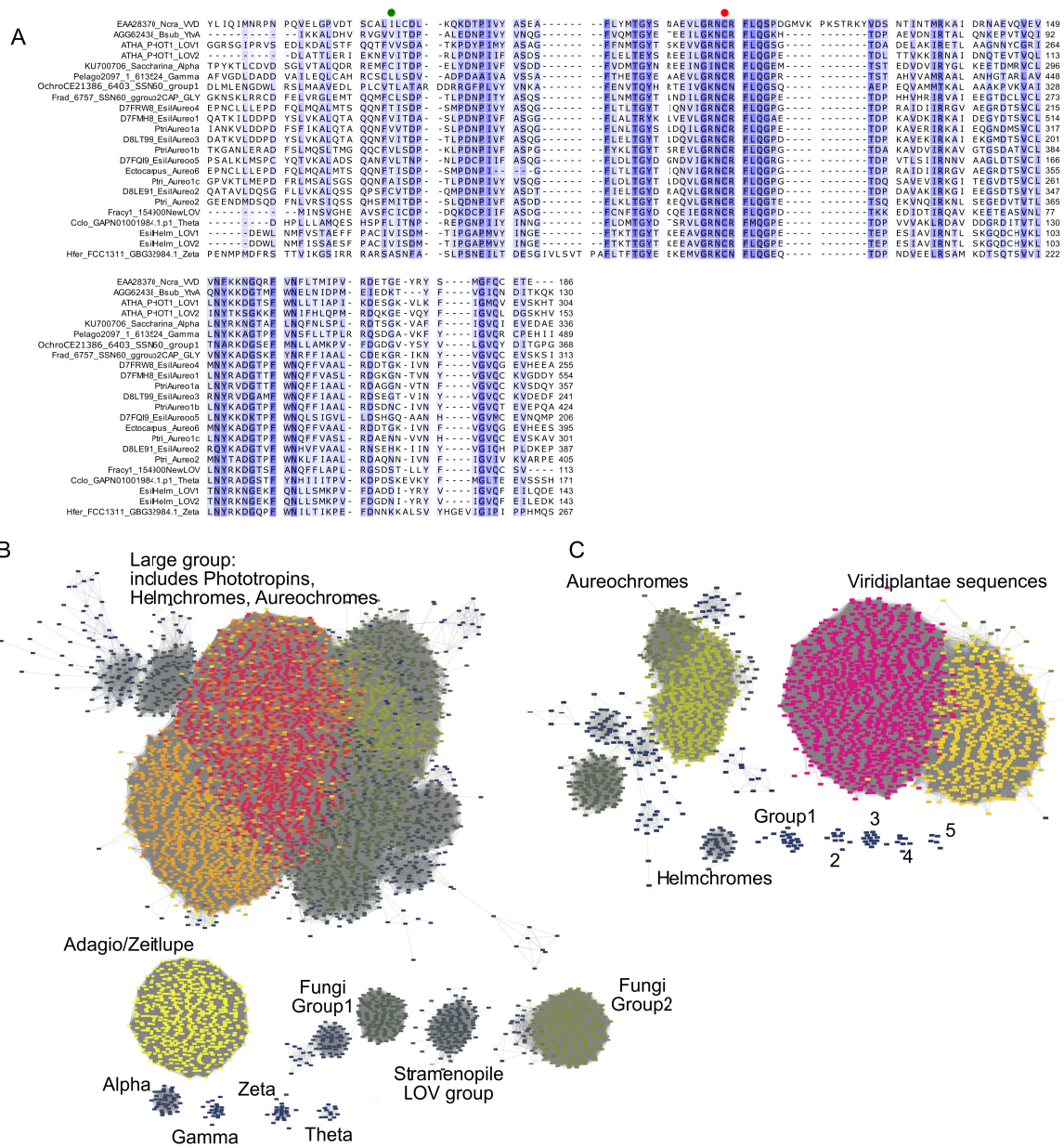


Figure S 2. Bioinformatic analysis of the putative LOV-containing photoreceptors. A Alignment of selected sequences from LOV-containing Ochrophyta proteins, with references LOV containing proteins (*Arabidopsis* phototropins and known aureochromes from *Ectocarpus siliculosus* and *Phaeodactylum tricorutum*). The cysteine required for chromophore interaction is marked with a red spot; the position suggested to prevent photoreceptor activities in Aureo2 proteins is marked with a green spot. Shades of blue indicate sequence conservation. B and C Sequence similarity network of the LOV candidate proteins. Colors indicate the network connectivity, i.e to how many sequences each sequences is linked. This helps to detect the different groups B, alignment score 40; C alignment score 60 performed on the sequences in the large group in B.

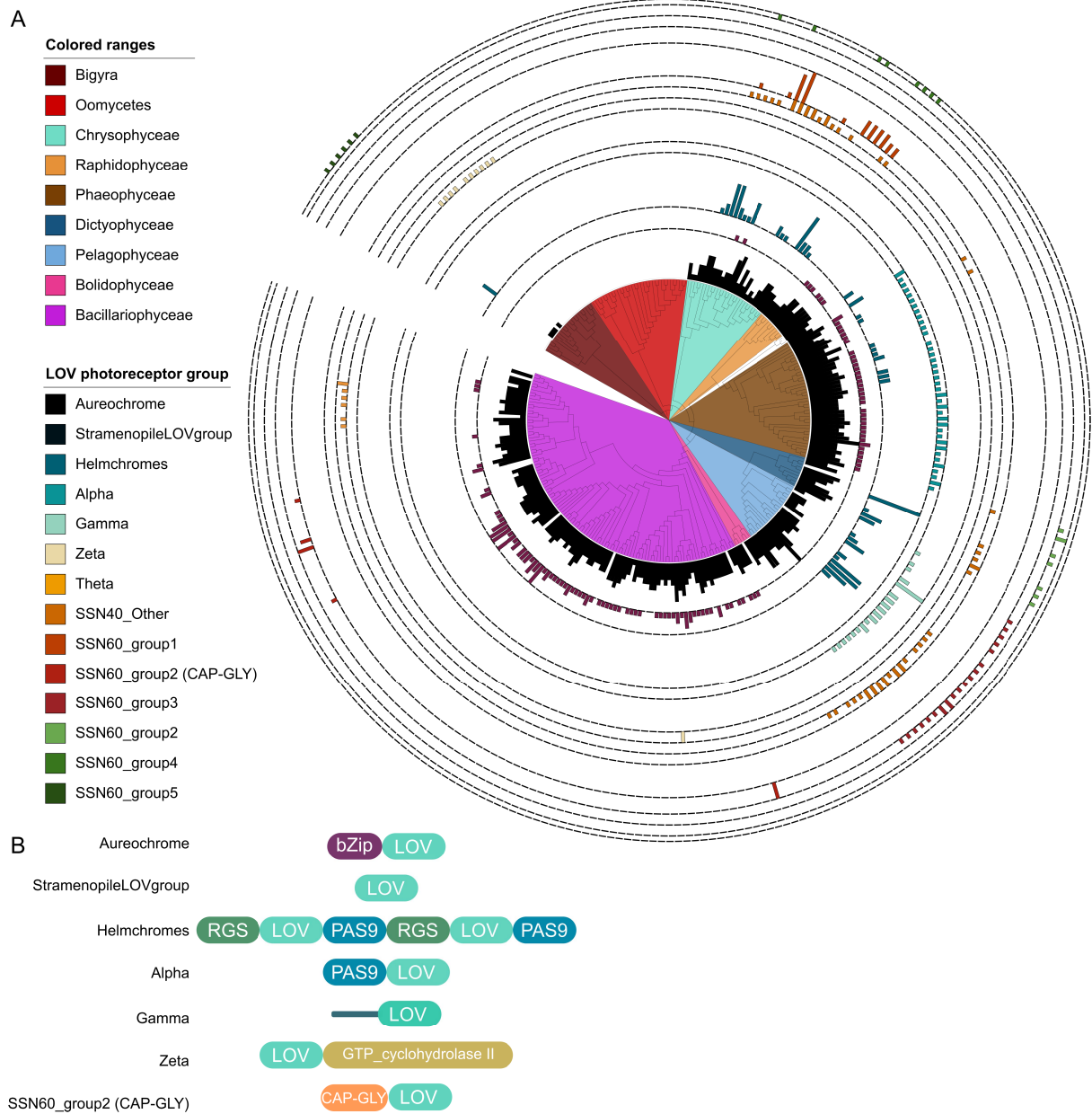
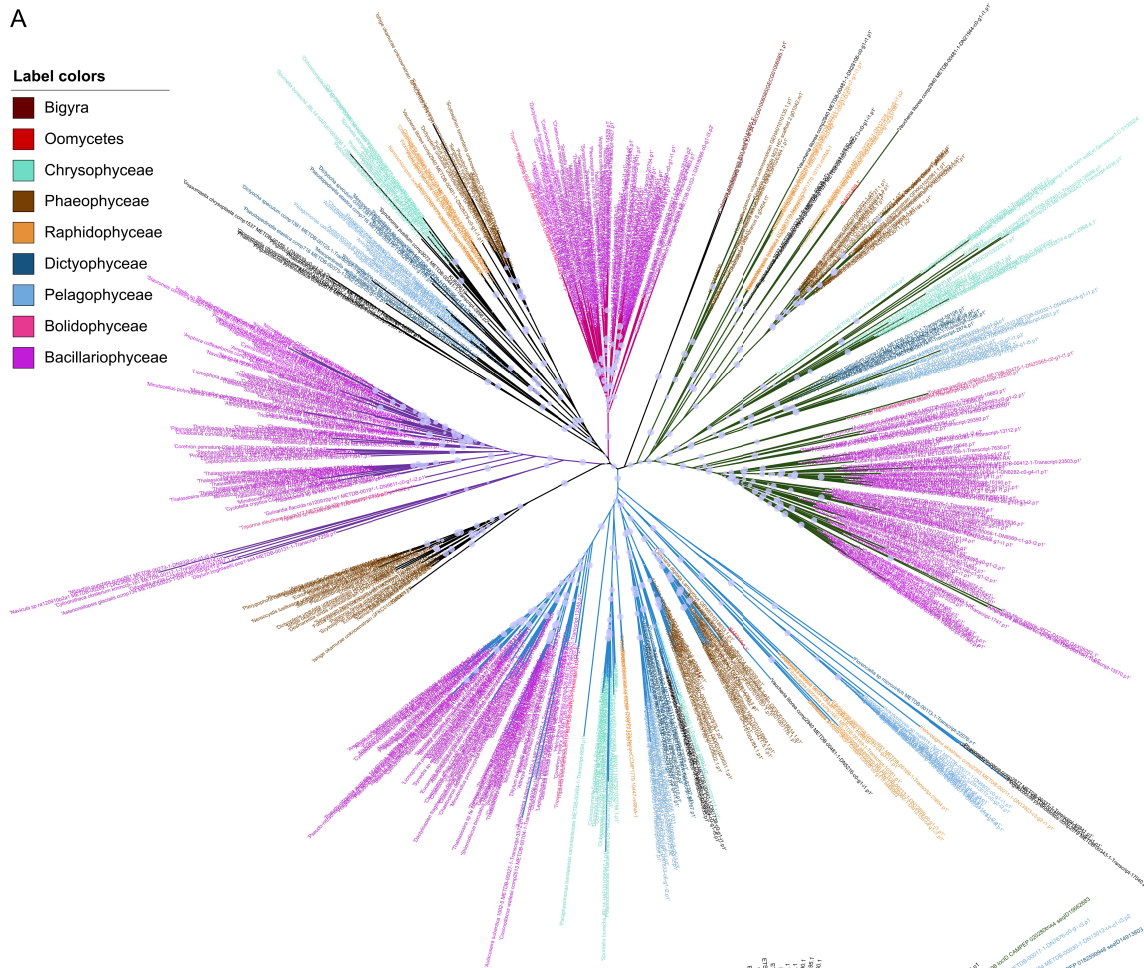


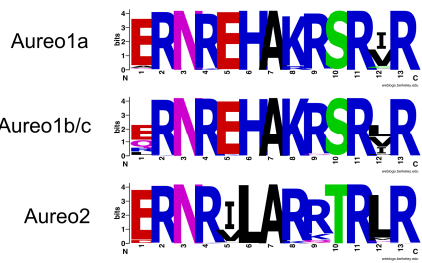
Figure S 3. LOV-based photoreceptors families in Ochrophyta. A. Repartition of the LOV families on the Stramenopile tree of species. Outer circles represent the presence of the different photoreceptor types and the bar height is proportional to the number of copies. Except for aureochromes and the Stramenopile LOV group, all other LOV families appear specific to one branch of algae. B, domain architecture of the LOV-containing proteins.

A

- Label colors**
- Bigyra
  - Oomycetes
  - Chrysophyceae
  - Phaeophyceae
  - Raphidophyceae
  - Dictyophyceae
  - Pelagophyceae
  - Bolidophyceae
  - Bacillariophyceae



B



C

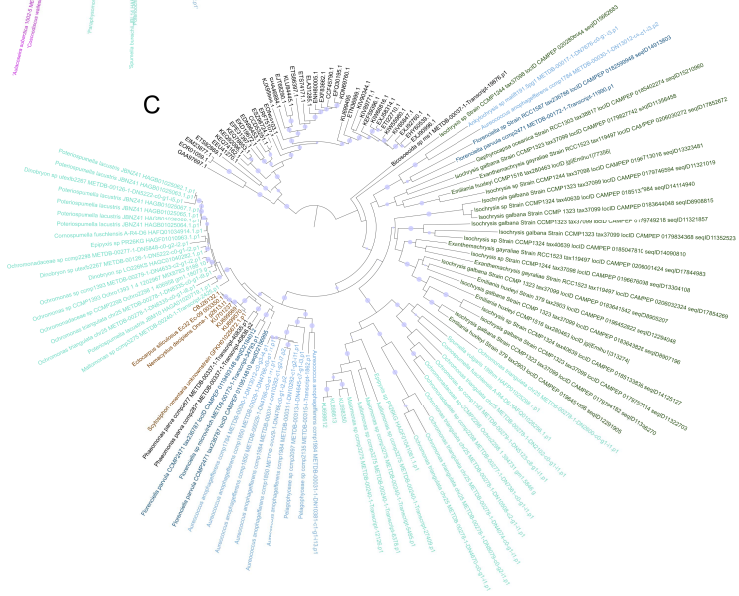


Figure S 4. Analysis of the aureochrome and helmchromes. A Unrooted tree of the aureochromes photoreceptors. B, DNA-binding motif of the bZIP domains of Aureochrome 1a, 1b/c and 2 clades. C. Tree of the helmchrome sequences (RGS-LOV alignment). In A and C, label colors correspond to the algal groups the genes belongs to, and purple circles on the branches indicate local bootstraps support >0.7

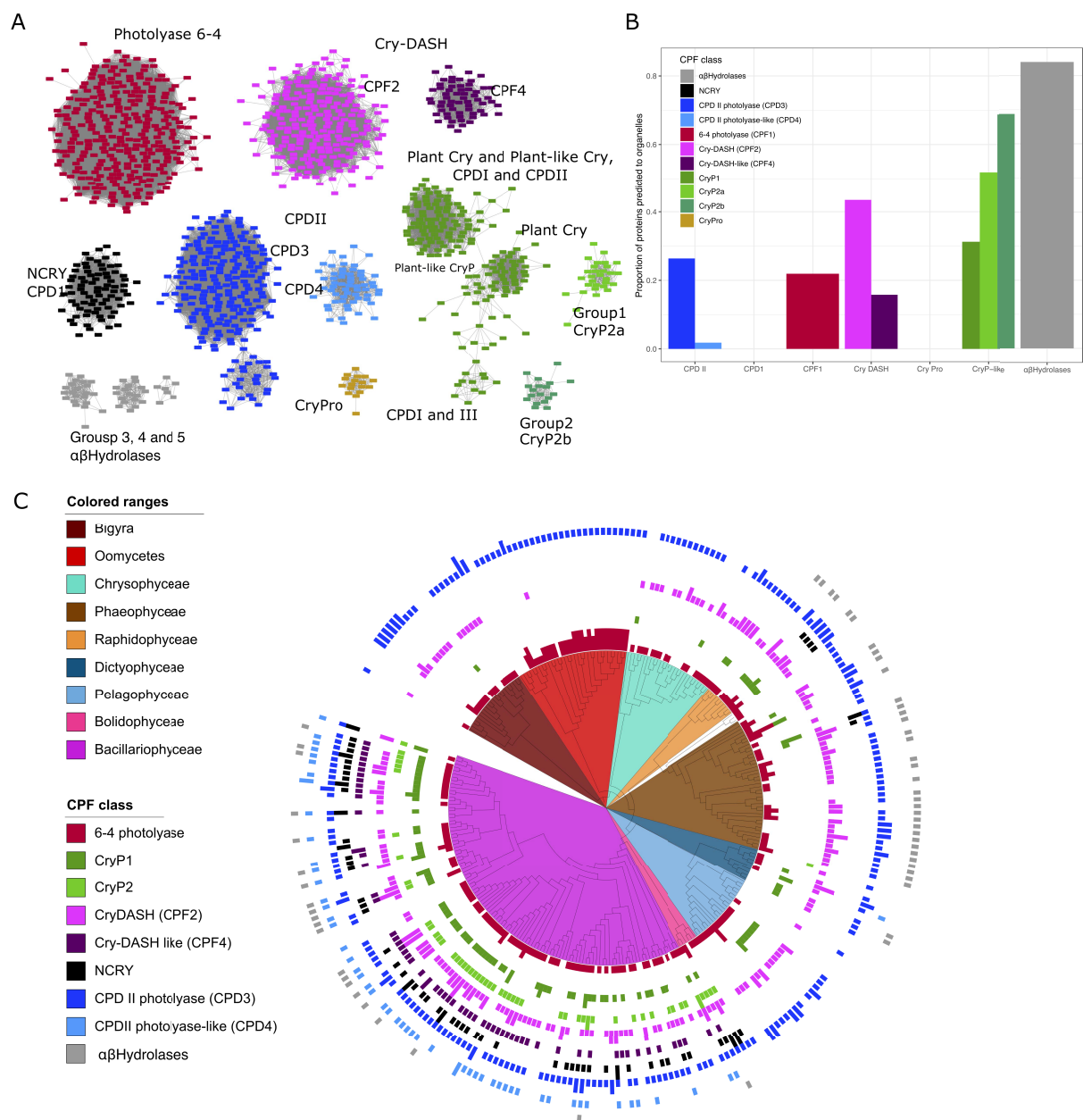


Figure S 5 Analysis of the CPF family in Ochromophyta. A Sequence similarity network of the CPF (alignment score 90). B Targeting prediction by HECTAR on diatom CPF proteins. Proportion of sequences addressed to the organelles are prediction to the chloroplast, mitochondria or presence of a signal peptide. C. CPF repertoire visualized on the Ochromophyta tree of species. Outer circle represent the presence of the different photoreceptor types and the bar height is proportional to the number of copies.



Figure S 6. Zoom on the CPD branch of the CPF tree presented in Fig3



Diatom CPF4

Diatom CPF2

Figure S 7. Zoom on the Cry-DASH branch of the CPF tree presented in Fig 3





Figure S 8. The 6-4 photolyases : A. Zoom on the 6-4 photolyase branch of the CPF tree presented in Fig3. B Isoelectric point calculated with IPC for oomycetes CPF1, separated per genus. Note that *Alphanomyces* CPF1 do not branch with other Oomycetes CPF.



Figure S 9 The plant-Cry like proteins : A. Zoom on the plan Cry branch of the CPF tree presented in Fig3. B. Isoelectric point calculated for the different type of Diatom CryP

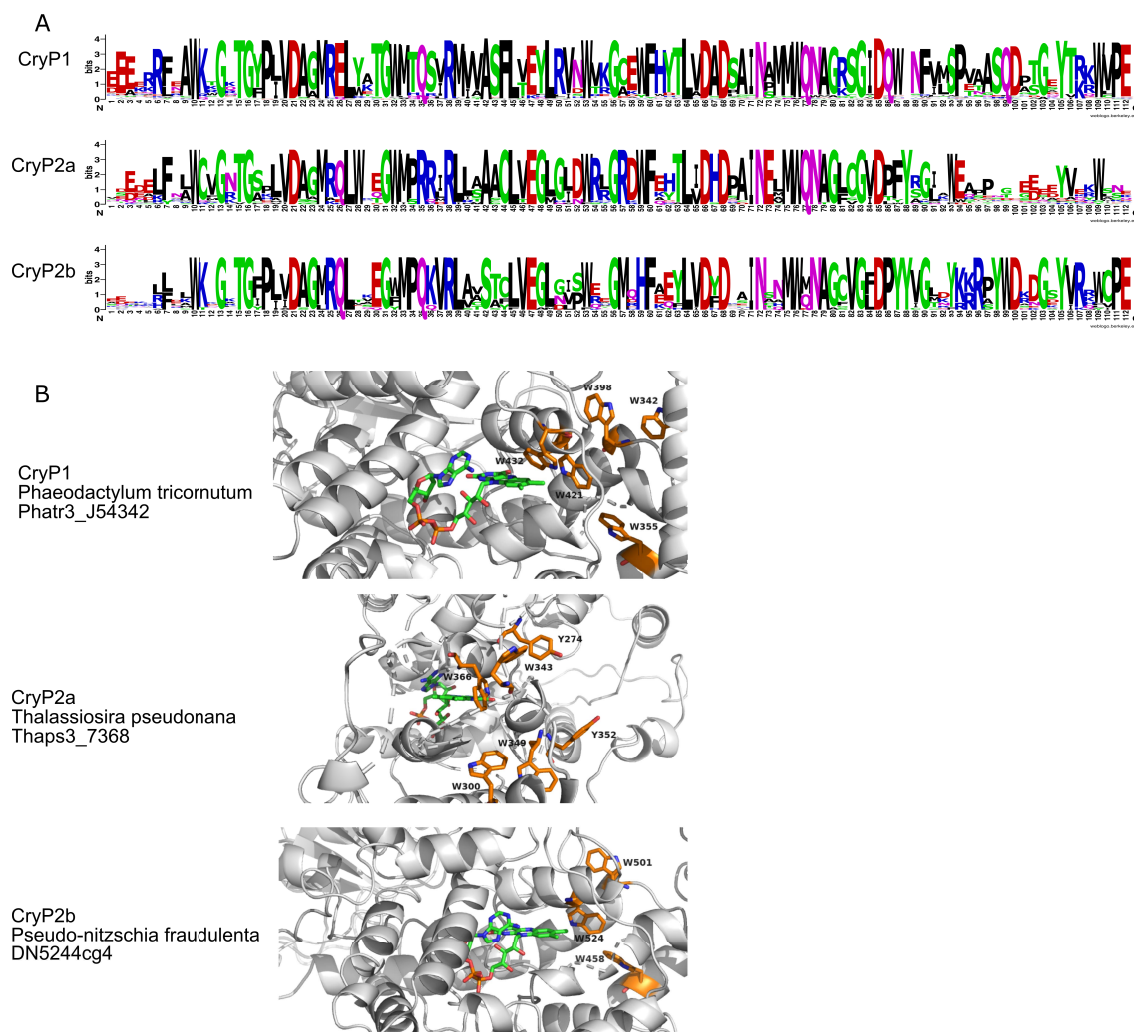


Figure S 10 Sequence analysis of the CryP sub-families. A Sequences logo of the FAD binding site of the different sub-families. PtCryP has been shown to bind FAD as chromophore and to be capable of a photoreduction of the FAD upon illumination with blue light (Juhas et al., 2014). Comparison of the representative motifs of CryP1, CryP2a and CryP2b shows that all three subclasses share amino acids involved in FAD binding such as R38, D66, D68 or N72, the latter one interacting with the N5 of the isoalloxazine ring of the FAD (numbers from the logo) (Zhong, 2015).. Moreover, while CryP1 has a methionine at position 39 that has been shown to be essential for CPD binding, Cryp2a/b have a leucine. The absence of this M39 with the absence of N71 and W78 suggest that Cryp2a/b proteins cannot interact with CPDs (Tan et al., 2015).

B. Spatial arrangement of tryptophan and tyrosine which may be involved in an electron transfer chain. CryP1 and CryP2b homology models based on the crystal structure of a bacterial class III photolyase from *Agrobacterium tumefaciens* (Scheerer et al., 2015) PDB 4u63.1. CryP2a homology models based on the crystal structure of *Arabidopsis thaliana* BIC2-CRY2 complexa (Ma et al., 2020) PDB 6k8k.1. In the CPF protein, FAD is embedded inside the protein, and transfer of electrons from the solution (photoreduction) occurs through an intramolecular electron transfer chain involving mostly tryptophan and tyrosine and is

an indispensable process for the CPF enzymatic or photoperception activity (Chaves et al., 2011). Notable differences are observed in the amino acids that may be involved in the electron transfer chain. While W10, W53, and W76 seem to be conserved between all CryP, alternative electron transfer pathways are observed in CryP1 and CryP2a with conserved W or Y such as W87, W59 allowing to reconstitute in silico a complete electron transfer chain.

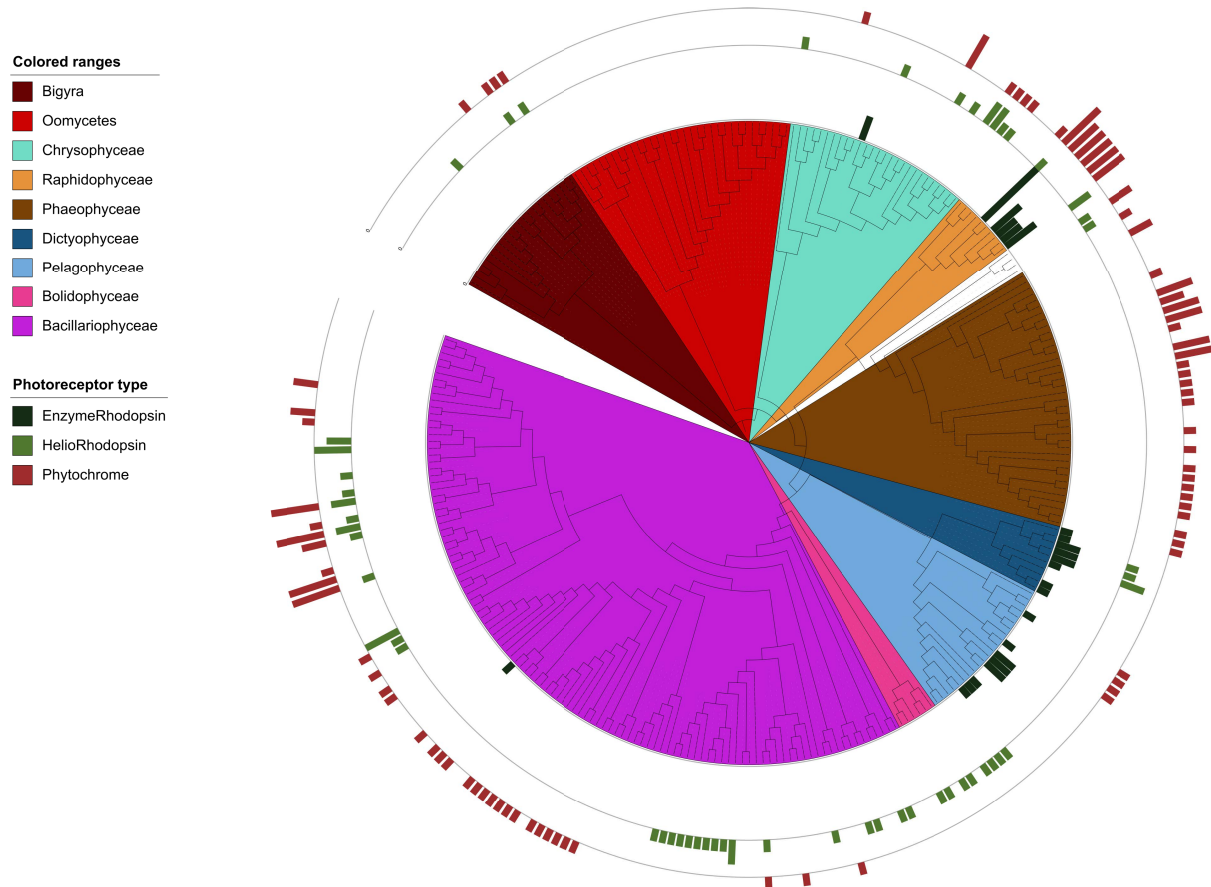


Figure S 11. Repartition of rhodopsins and phytochromes visualized on the Stramenopile species tree. Outer circles represent the presence of the different rhodopsins and phytochromes and the bar height is proportional to their number of copies.

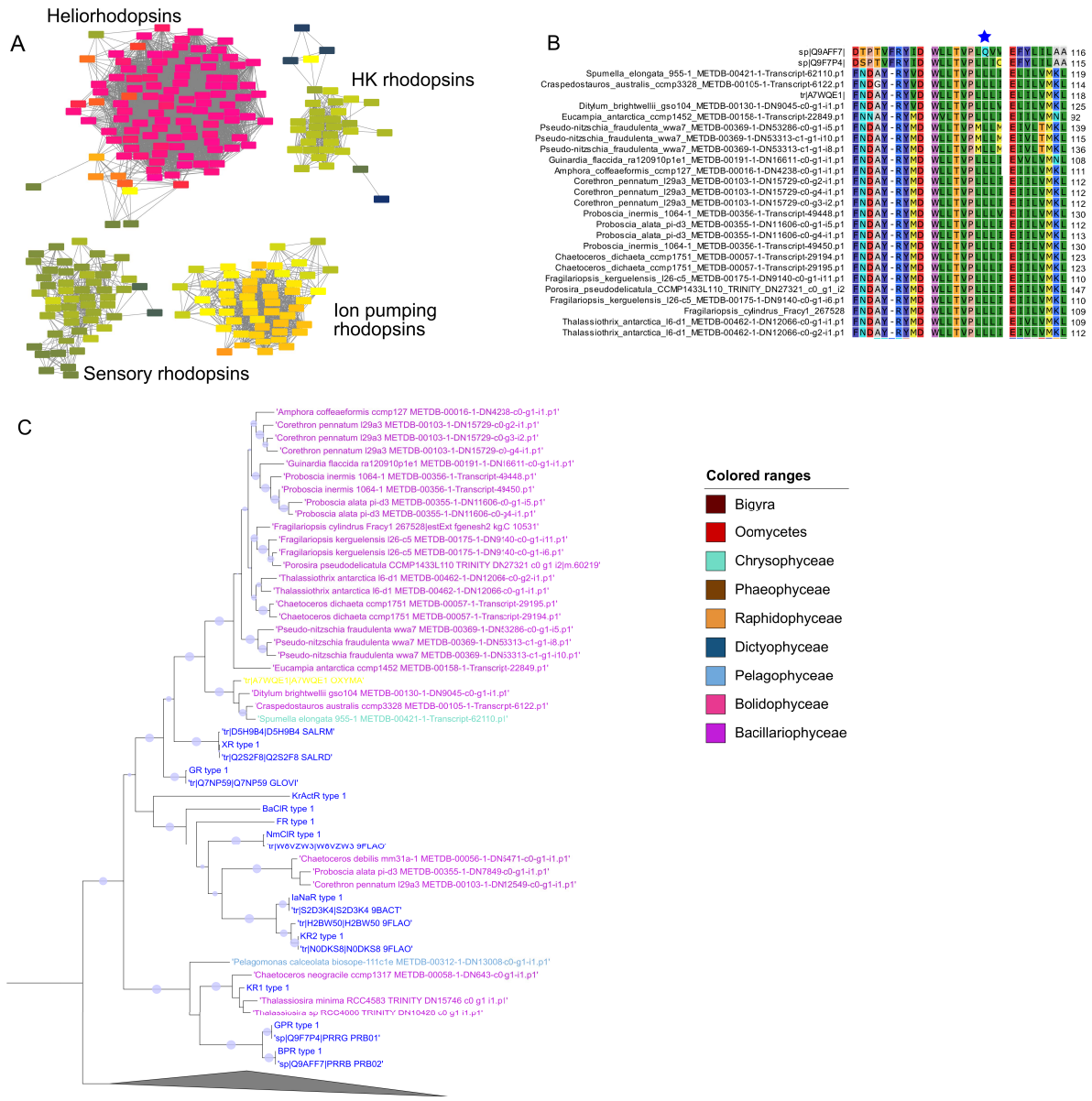


Figure S 12. Analysis of the rhodopsin photoreceptor types. A. Sequence similarity network of the rhodopsin family. B alignment of the porteorhodopsin sequences. The glutamate residue responsible for the spectral tuning is marked with a blue star. C. Phylogenetic tree of the proteorhodopsins identified in this study. Label colors correspond to the algal groups the genes belongs to, as in Fig4, and purple circles on the branches indicate pseudo-bootstrap support >0.7.









# CHAPTER 2 : MARINE DIATOM PHYTOCHROMES ARE SENSORS OF DEPTH AND PHYTOPLANKTON CONCENTRATION THROUGH SENSING OF UNDERWATER LIGHT FIELD VARIATIONS

Carole Duchêne<sup>1</sup>, Jean-Pierre Bouly<sup>1</sup>, Juan Pierella Karlusich<sup>2</sup>, Julien Sellès<sup>1</sup>, Chris Bowler<sup>2</sup>, Maurizio Ribera d'Alcalà<sup>3</sup>, Angela Falciatore<sup>1</sup>, Marianne Jaubert<sup>1</sup>

1. CNRS, Sorbonne Université, Institut de Biologie Physico-Chimique, Laboratoire de Biologie du chloroplaste et perception de la lumière chez les microalgues, UMR7141, F-75005 Paris, France

2. Ecole Normale Supérieure, PSL Research University, Institut de Biologie de l'Ecole Normale Supérieure, CNRS UMR 8197, INSERM U1024, F-75005 Paris, France

3. Stazione Zoologica Anton Dohrn, 80121 Naples, Italy

Phytochrome photoreceptors from *T.pseudonana* and *P.tricornutum* display a red/far-red absorption spectra, which led to complex hypotheses concerning the light source that could modulate DPH activity in the ocean, given the poor abundance of these wavelengths in this environment. To get insights into the function DPH could play, we wanted to determine the conditions in which DPH can be active in the environment, we aborded the questions by different approaches. On one side, we determined the action spectra of the DPH from the model diatom *Phaeodactylum tricornutum* thanks to the use of a reporter

system in which the fluorescent reporter eYFP is under the control of a DPH-regulated promoter in *P.tricornutum*, allowing to build a mathematical model describing DPH activity in a given light field. On the other side, we investigated in which environment are found DPH-containing diatoms, which could explain the adaptive value that could confer DPH activity. For this, we collaborated with Chris Bowler and Juan Pierella Karlusich to dig into the Tara Oceans resources. Both approaches brought complementary and convergent information toward the environmental impact on DPH activity.

In this study, I realized the wide majority of the work, from the setting of the experimental conditions to the realization of fluence response curves of the action spectra, expression analysis of different DPH-regulated genes, and the mathematical modeling of DPH activity based on already established equations, using different light field I modeled or environmental data provided by Maurizio Ribera d'Alcala. I analyzed the sequences retrieved from the Marine Atlas for Tara Oceans Unigenes (MATOU) and quantified their respective abundance, and characterization of novel DPH recombinant proteins.

Manuscript in preparation

# MARINE DIATOM PHYTOCHROMES ARE SENSORS OF DEPTH AND PHYTOPLANKTON CONCENTRATION THROUGH SENSING OF UNDERWATER LIGHT FIELD VARIATIONS

---

Carole Duchêne<sup>1</sup>, Jean-Pierre Bouly<sup>1</sup>, Juan Pierella Karlusich<sup>2</sup>, Julien Sellès<sup>1</sup>, Chris Bowler<sup>2</sup>, Maurizio Ribera d'Alcalà<sup>3</sup>, Angela Falciatore<sup>1</sup>, Marianne Jaubert<sup>1</sup>

1. CNRS, Sorbonne Université, Institut de Biologie Physico-Chimique, Laboratoire de Biologie du chloroplaste et perception de la lumière chez les microalgues, UMR7141, F-75005 Paris, France

2. Ecole Normale Supérieure, PSL Research University, Institut de Biologie de l'Ecole Normale Supérieure, CNRS UMR 8197, INSERM U1024, F-75005 Paris, France

3. Stazione Zoologica Anton Dohrn, 80121 Naples, Italy

## ABSTRACT

Aquatic environments are more penetrative to blue and green wavelengths while red and far-red ones are rapidly attenuated. So marine organisms are expected to have adapted their photosensing capabilities to the most abundant wavebands. However, marine diatoms possess photoreceptors of the phytochrome family (DPH) that respond to red/far-red light. To shed light on this puzzling evidence, we set up a reporter system in the model diatom *Phaeodactylum tricornutum* to determine the action spectra of PtDPH activity. We assessed that this photoreceptor responds, effectively, not only to red and far-red bands, but also to blue and green lights. Projecting PtDPH capability, quantified by the photochemical parameters determined *in vivo*, in different ocean light fields, we determined that, counterintuitively, PtDPH activity increases with depth, while being sensitive to the concentration of optically active components, as phytoplankton concentration. We then investigated the occurrence of DPH *in situ*, using the ocean wide data set from the *Tara* Oceans expedition, and found that DPH containing, and expressing, diatoms were almost exclusively present in temperate and polar regions, which experience high annual variability in phytoplankton abundance and a variable

photoperiod. We expanded the number of characterized DPH family members with sequences found to be abundant in the open ocean. We found that their photochemical properties are strongly conserved and similar to those of PtDPH, and that they show a similar activation pattern in the environment. Overall, these results set a milestone in marine photoreception, showing light sensitivity of DPH along the whole photic zone, and open new perspectives both for DPH functioning and putative role in marine diatoms.

Keywords: Phytochrome, diatoms, light sensing, oceans

## INTRODUCTION

For photosynthetic organisms, the perception of their light environment is crucial for multiple aspects of their life, from the regulation of photosynthesis to cellular and developmental processes such as acclimation, adjustment or synchronization of physiology and metabolism (Briggs and Spudich, 2005). Photosensing relies on photoreceptor proteins that perceive particular bands of the light spectrum (Möglich et al, 2010). Among them, the phytochromes are red/far-red light sensing proteins, present in terrestrial plants but also in photosynthetic and non-photosynthetic bacteria, fungi and diverse algae such as streptophytes, prasinophytes, glaucophytes, cryptophytes and heterokonts (Butler et al, 1959; Hughes et al, 1997; Yeh et al., 1997; Bhoo et al, 2001; Giraud et al, 2002; Blumenstein et al, 2005; Froehlich et al, 2005; Falciatore and Bowler, 2005; Rockwell et al, 2014). Phytochromes are characterized by their ability to photoreversibly convert between an active and an inactive state: upon the perception of red light, the red-absorbing form Pr switches to a far-red absorbing form, Pfr, which can subsequently revert back into the Pr form upon far-red light (Butler et al, 1959). Thus, the activity level of phytochrome is related to the photoequilibrium formed upon a given light exposure (Morgan and Smith, 1976). A well-known example of phytochrome photoequilibrium dependent response in plants is the perception of the low R/FR ratio under a canopy, allowing the detection of neighbors (i.e., competing phototrophs) and mediating shade avoidance responses (Casal, 2012). The photoequilibrium state as well

as the phototransformation rates between the two forms can be predicted mathematically at any wavelength, based on the phytochrome photoconversion cross sections ( $\sigma_\lambda$ ) (that depend on the quantum yield of photoconversion ( $\phi$ ) and the extinction coefficient of each form ( $\epsilon_\lambda$ )), and the spectral fluence rate of a light source (Butler, 1972). These photochemical properties can be estimated from measurements made *in vitro* on recombinant phytochromes (Giraud et al, 2010) but also *in vivo* (Schmidt et al., 1973). However, the determination of action spectra for specific biological responses that reflect phototransformation reactions (Hartmann 1966, Beggs, 1981, Shinomura et al, 1996) have shown some differences between *in vitro* and *in vivo* approaches due to distortions of the light transmitted within cells (Kazarinova-Fukshansky, 1985; Kusuma, 2021) or/and biological events acting downstream of phytochrome photochemical reactions (Klose, 2015; Olson, 2017).

Phytochromes are dimeric chromoproteins which sense light via heme-derived linear tetrapyrroles as chromophore (Quail, 1991; Rockwell et al, 2006). These chromophores covalently bind to a conserved cysteine residue in the N-terminal photosensory module (PSM), composed of PAS (Per-Arnt-Sim), GAF (cGMP phosphodiesterase/ adenylyl cyclase/FhlA) and PHY (phytochrome-specific) domains (Montgomery and Lagarias, 2002; Karniol et al, 2005; Lamparter et al, 2002, Blumenstein et al, 2005). Upon excitation by light, changes in the chromophore generate protein conformational changes which activate a more diversified C-terminal output module initiating a signaling cascade (Rockwell and Lagarias, 2006).

The presence of phytochromes in marine phytoplankton was quite unexpected because water differentially absorbs light bands in a manner inconsistent with phytochrome activation. For example, light intensity decreases exponentially with depth, and short (UV) and long (red and far-red) wavelengths of the spectra are attenuated more strongly than green and blue wavelengths (Mobley, 1994). Additionally, the presence of other components in the seawater, such as phytoplankton, and their pigments, inorganic particles and colored dissolved organic matter (CDOM) also alter the light field because they absorb and/or scatter specific wavelengths depending on their bioptical properties (Kirk, 2011). These properties give rise to a strong vertical structuring of light in the ocean

with a continuum of spectra, going from a high intensity solar spectrum near the surface to a dim blue-green light field at depth, therefore progressively deprived of the red-far red band. Absorption spectra variations have been shown for some algal phytochromes from prasinophytes, glaucophytes and brown algae, that are tuned to the shorter wavelengths that are more abundant in the water column, with orange-yellow/far-red, red/blue or far-red/green photocycles, respectively, which has been interpreted as an adaptive spectral tuning of the photoreceptors (Rockwell et al, 2014).

Among marine phytoplankton, diatoms are one of the most abundant and diverse groups, and are considered responsible for about 20% of the primary production on Earth (Field et al, 1998; Malviya et al, 2016; Pierella Karlusich et al, 2020). These secondary endosymbiosis-derived microalgae have been able to colonize very different aquatic niches. They particularly thrive in turbulent environments (Margalef, 1978; Esposito et al, 2009), that can imply drastic variations in light conditions, both in term of intensity and quality. Diatoms exhibit a palette of photoreceptors including blue light photoreceptors of the Cryptochrome/photolyase family (CPF), the stramenopile-specific Aureochrome family, and putative green light photoreceptors of the Heliorhodopsin family (Jaubert et al, 2022). Some diatoms also possess phytochromes (DPH), similar to the bacterial class, with a conserved photosensory module followed by an output domain composed of an histidine kinase/ATPase and REC (response receiver) domains (Bowler et al, 2008; Fortunato et al, 2016).

The DPH from the model diatom species *Thalassiosira pseudonana* (TpDPH), from the centric group, and *Phaeodactylum tricorutum* (PtDPH), from the pennate group, were both shown to be able to bind as chromophore biliverdin (BV), the bilin bound to bacterial and fungal phytochromes (Bhoo et al, 2001; Blumenstein et al, 2005). TpDPH and PtDPH both display absorption spectra of the Pr and Pfr maximally peaking at 686 and 700, and 764 et 750 nm, respectively, which are red-shifted in respect to canonical plant phytochromes and other algal phytochromes (Fortunato et al, 2016).

Attempts to understand the function of this photoreceptor in *P. tricorutum* showed that exposure to far-red light leads to expression changes of a set of genes. These patterns are lost in PtDPH knockout (KO) mutants. Most of these genes are of unknown function, or

encoding proteins involved in transcriptional regulation and signaling (Fortunato et al, 2016). Hypotheses have been proposed about a role of these phytochromes as depth sensors, given the light spectrum differences along the water column, or as sensors of phototrophic cells in the surroundings, during algal blooms in particular, through the sensing of chloroplast fluorescence (Fortunato et al, 2016). However, up to now, the physiological function algal phytochrome remains a mystery, and the presence of red/far-red photoreceptors in marine diatoms that are exposed to a mostly blue-green light environment is still puzzling.

In this work we went beyond the bias of DPH being activated only in the red-far red bands and explored over the whole spectrum the response of this unique photoreceptor.

We first characterized the *in vivo* sensing properties of PtDPH, using an ad hoc designed reporter system which allowed to monitor and quantify PtDPH activation in *P. tricornutum* over all the wavelength range available in the ocean surface layer. We then modeled the PtDPH activation state in various computed or real marine light fields, to infer its *in situ* response. We additionally investigated the occurrence and distribution of DPH in *Tara Oceans* expedition data, characterized photochemical properties of a suite of novel DPHs from different ecological niches and different taxa, and verified similar properties of those of the model systems previously analyzed. In brief, the interplay between the varying underwater light field and the highly specific spectral properties of DPH produces a unique vertical pattern in the photoreceptor response. Altogether, these results permit new perspectives both for DPH functioning and for its putative role in marine diatom biology.

## RESULTS

### **In vivo action spectra of DPH**

To monitor DPH activity *in vivo*, we introduced in wild-type (WT) and KO-PtDPH *P. tricornutum* strains the coding sequence of the enhanced Yellow Fluorescent Protein

(eYFP) under the control from a promoter of a PtDPH-regulated gene (*Hsf4.6a*) (ProHSF4.6a::YFP WT and ProHSF4.6a::YFP KO, respectively) (Fortunato et al. 2016)(Fig. 1A). Reporter strains were grown in continuous green light to avoid diurnal regulation and because green light is the least absorbed by PtDPH *in vitro*.

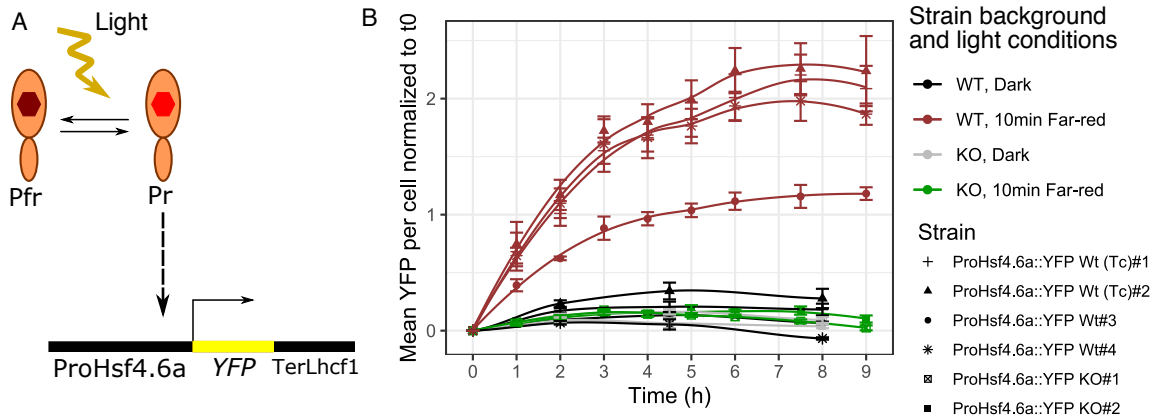


Figure 1. PtDPH activity reporter system and experimental conditions. A. Schematic illustration the PtDPH activity reporter system. B. Kinetics of the eYFP induction in different reporter strain backgrounds: *P. tricornutum* wild-type (WT), knockout PtDPH mutants (KO), or non mutated cells originated from the same transformed colony than the KO (WT (Tc)(Transformation control)). Cells were grown in continuous green light (22  $\mu\text{moles}$  of photon/ $\text{m}^2/\text{sec}$ ) and exposed for 10 min to 800-nm far-red light (15  $\mu\text{moles}$  of photon/ $\text{m}^2/\text{sec}$ ) at  $t_0$  then left in the dark («10min far-red») up to 9 h post-irradiation, or directly transferred from green to darkness («Dark»). eYFP signals were measured by flow cytometry and normalized to  $t_0$ .

To verify the PtDPH-dependent control of reporter gene expression, ProHSF4.6a::YFP WT, ProHSF4.6a::YFP WT (Tc) (transformation control, i.e., cells originated from the same transformed colony than its corresponding KO but not having undergone PtDPH gene editing as described in Fortunato et al, (2016)), and ProHSF4.6a::YFP KO lines were submitted or not to a far-red (800 nm) light exposure of 15  $\mu\text{moles}$  of photon/ $\text{m}^2/\text{sec}$  for 10 minutes and then left in the dark. The YFP signal was measured by flow cytometry at different time points. As shown in Fig. 1B, YFP signal increased over 6h in the dark in ProHSF4.6a::YFP WT and remained constant up to 9h post-irradiation. On the contrary, it remained at a basal level when the ProHSF4.6a::YFP WT lines were not exposed to far-red light over the same period. Moreover, eYFP induction was not observed in ProHSF4.6a::YFP KO lines submitted to the same far-red light exposure (Fig. 1B). These results indicate that our reporter system specifically reports DPH- and FR- light-dependent inductions, and is thus suitable for the characterization of *in vivo* DPH activity.



Using the same experimental setup as described above, fluence rate response curves were then generated from reporter lines submitted for 10 min to a gradient of different monochromatic lights (Fig. S1A) from the far-red to violet-blue regions (850, 810, 765, 740, 730, 470, 430 and 405 nm (Fig. S1B)). As shown in Fig. 2A, for all the colors tested, the curves exhibited the same sigmoidal shape, with no response measured at the lowest light intensities, followed by a log-linear phase as intensity increases, before reaching a plateau when light intensity is saturating (Fig. 2A). Notably, photoinduction of YFP was effective in the blue band and even up to 850 nm, a wavelength at which the recombinant PtDPH protein barely absorbs. These responses were all DPH-dependent, as the reporter ProHSF4.6a::YFP KO lines showed no induction of YFP at any wavelength (Fig. S2). Furthermore, the same fluence rate response curves performed with the light-regulated promoter *LHCF2* and the constitutive promoter *H4* (Siaut et al, 2007), showed no wavelength- or intensity-dependent induction (Fig. S2). The induction of expression in saturating blue as well as in far-red light was also verified by RT-qPCR for *HSF4.6a* and other PtDPH-regulated genes previously identified in Fortunato et al (2016) (Phatr3\_J15138, Phatr3\_J46431, Phatr3\_J18096, Phatr3\_J45662) (Fig. S3). On the contrary, no induction occurred in *PtDPH* KO lines. This loss of induction by blue light in the KO lines was shown to be specifically regulated by PtDPH because blue light-mediated induction of expression of the Phatr3\_J18180 gene (encoding LHCR7 fucoxanthin chlorophyll a/c light-harvesting protein and not found regulated by PtDPH), was still observed in KO as well as in WT lines (Fig. S3).

The log linear phase of each curve occurs at different intensities depending on the light used. It extends to the far-red (765nm) from 0.1 to 10  $\mu\text{mol photon/m}^2/\text{s}$ , with an intensity needed to reach half the maximum YFP ( $I_{1/2}$ ) of 1.47  $\mu\text{mol photon/m}^2/\text{s}$ . Surprisingly, the log linear phase in blue (430 nm) is shifted to lower intensities, from 0.01 to 1  $\mu\text{mol photon/m}^2/\text{s}$ , with an  $I_{1/2}$  of 0.3  $\mu\text{mol photon/m}^2/\text{s}$  while PtDPH absorbs less in the blue than in the far-red light.

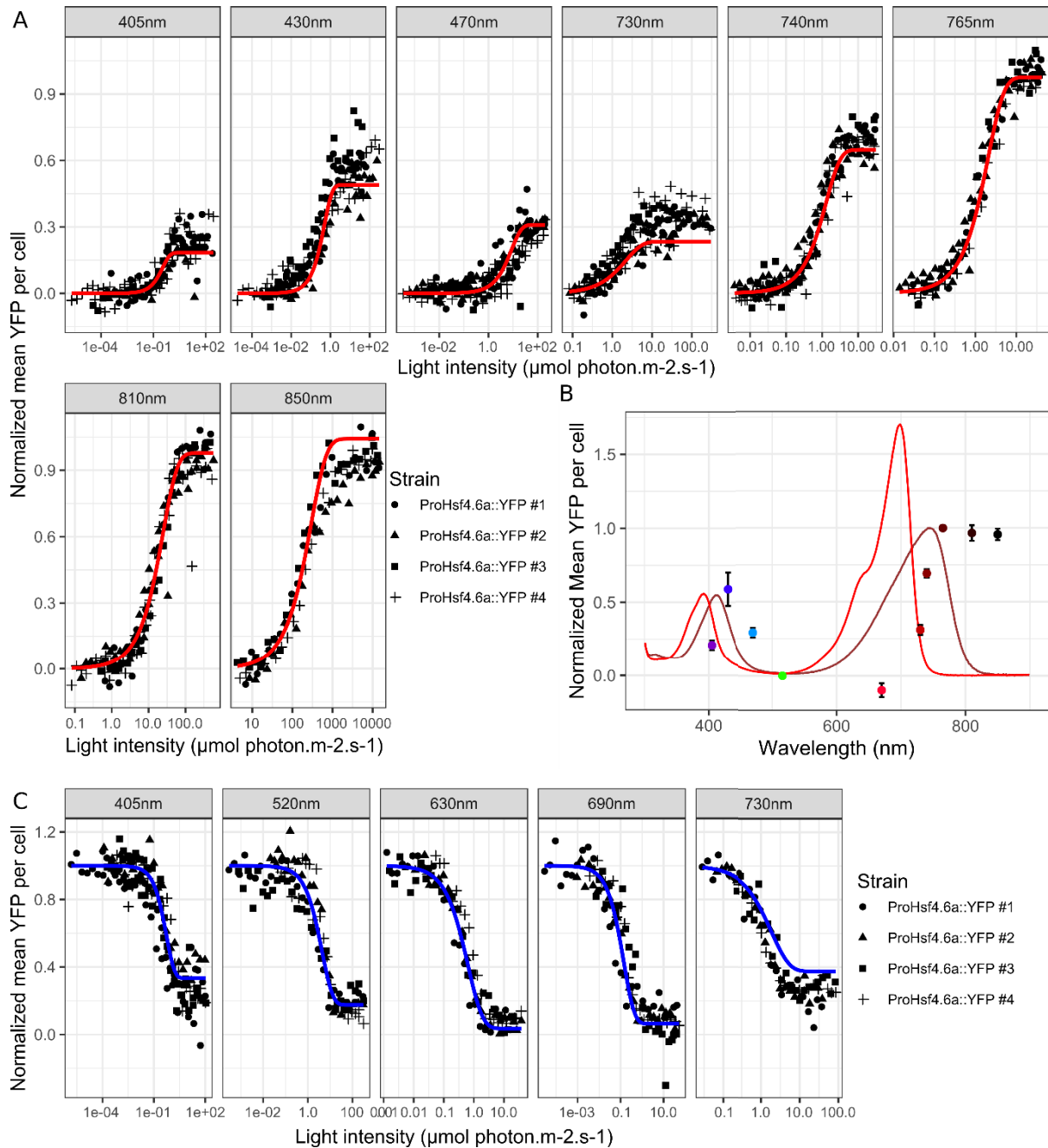


Figure 2. Action and inhibition spectra for PtDPH-mediated responses. A. Action spectra of PtDPH (via activation of the YFP reporter) in response to different light colors and intensities. Cells were grown in continuous green light and then exposed to different intensities of monochromatic light for 10 min and the YFP signal measured after 6h in the dark, as in Fig.1. B. YFP signal at saturation, determined from the 4 highest intensities in A, and superimposed to PtDPH absorption spectra (red lines, Pr, dark red line, Pfr (Fortunato et al, 2016)). Values correspond to normalized means  $\pm$ sd from 4 independent WT strains (WT and WT (Tc) (Transformation control)). C. Inhibition spectra of PtDPH, as photoreversibility of the far-red-induced response. Cells were grown in continuous green light, then treated to saturating 765-nm LED for 10 min, and then exposed to different intensities of monochromatic LED or dark. In A and C, the YFP signal was averaged from the measure of 29700 cells by flow cytometry, and normalized between YFP signal of cells unexposed to light (=0) and exposed to saturating far-red light (765 nm) (=1). Colored lines in A and C reflect the fitting of the normalized YFP for each wavelength as function of the light intensity (Supplemental method).

The levels of YFP reached by the ProHSF4.6a::YFP WT lines in saturating light are different for each wavelength (Fig. 2A). These levels plotted as a function of wavelengths fit with the ratio of absorbance of Pfr compared to Pr (Fig. 2B). This suggests that the observed responses are primarily regulated by PtDPH under our conditions, and that the level of YFP under a given light reflects the photoequilibrium between the active and inactive forms of PtDPH *in vivo*. Fluence rate response curves for the inhibition of YFP induction were also generated, by exposing reporter strains first to 10 min of saturating far-red light (800nm) followed by 10 min of an intensity gradient of 5 different monochromatic LEDs (730, 690, 630, 520, 405 nm (Fig. S1B)) (Fig. 2C). For all these wavelengths, the same sigmoidal-shaped fluence rate response curves, in an opposite orientation to the induction experiments, were obtained, clearly showing that PtDPH-activated responses can be photoreversed by red, but also by green and violet wavelengths (Fig. 2C).

Altogether, these results indicate that, *in vivo*, PtDPH is a photoreversible photoreceptor, not only sensitive to red and far-red light but also effectively responding to blue and green wavelengths. This implies that the integration of the full absorption spectra of PtDPH as well as its photoreversibility properties are necessary to fully understand the activation of DPH in an highly variable light field strongly dependent on depth.

### ***In situ* DPH activation model**

Because our reporter system appeared to be a specific and quantitative marker of the PtDPH activation state *in vivo*, we then sought to construct a general model describing DPH activity *in situ*. The rate of phytochrome photoconversion depends on phytochrome photoconversion yield, extinction coefficient, and light intensity. Assuming that PtDPH acts as a dimer and that the active form is PrPr (Rockwell et al, 2006; Fortunato et al, 2016 and this study), the percentage of PrPr in a given light field can thereby be expressed by the equations previously established (Mancinelli, 1994) (detailed in Suppl methods). The ratio of photoconversion yield,  $\eta$ , and the photoconversion rate of phytochrome at a

wavelength  $\lambda$ ,  $k_\lambda$  (corresponding to the sum of the two kinetic constants for activation and inhibition of PtDPH), were determined by exploiting the action and inhibition fluence rate response curves generated.

To estimate the *in vivo* ratio of photoconversion yield,  $\eta$ , we plotted the normalized YFP signal of action curves at saturating light intensity, reflecting the %PrPr, as a function of the ratio of absorbance ( $A_{Pr}/A_{Pfr}$ ) deduced from the PtDPH recombinant protein absorption spectra (Fig. 3A). We also enriched this dataset with normalized YFP levels of the reporter strains exposed to mixed wavelength light fields (mix of 2 LEDs) at intensities saturating the PtDPH response. By fitting the equation of photoconversion at equilibrium (Fig. 3A), we could estimate a  $\eta$  of  $0.98 \pm 0.092$ . It should be noted that, in this simple model, the photoconversion ratio integrates the photoconversion coefficient of both forms but also all the different downstream events.

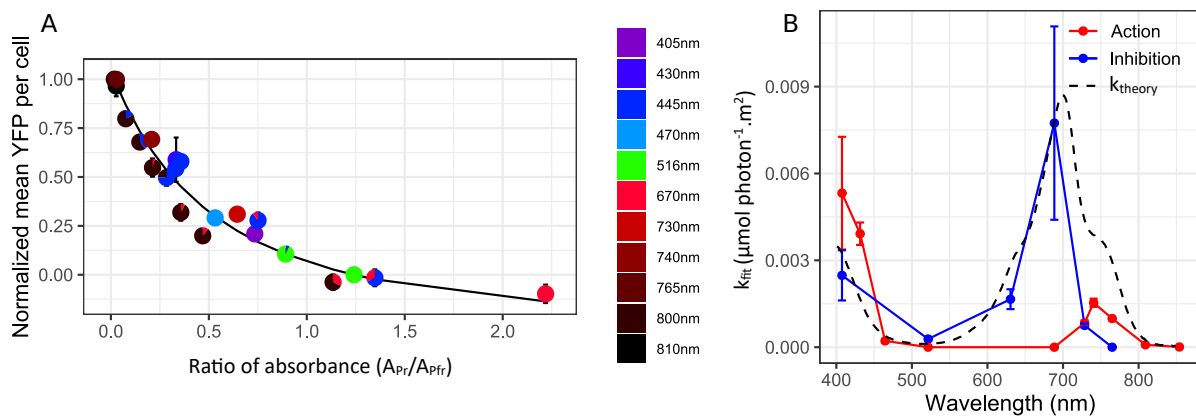


Figure 3. Modelling of PtDPH activity *in vivo*. A. Estimation of the ratio of photoconversion rates from YFP signal at saturating light (from Fig. 2A) and in response to mixed wavelengths. Colors indicate the wavelength of the LED, and pie charts, the proportion the relative intensity of each LED in mixed illuminations. Line represents the fitting of the PtDPH photoequilibrium as a function of the ratio of absorbance of PtDPH Pr and Pfr forms (Equation 5 in Supplementary method). Values are means  $\pm$  sd on 4 independent WT strains. B. Sensitivity of PtDPH to different wavelengths represented from the fitting done on the curves in Fig. 2A and 2C.  $k_{Fit}$  is the resulting exponential constant. Mean  $\pm$  sd on the 4 independent WT strains; dashed line represents the theoretical  $k_{Fit}$ , calculated with the absorption spectra of the recombinant protein and the determined ratio of photoconversion yield (Supplementary method).

With this estimation of  $\eta$ , we could fit exponential curves to each normalized eYFP-fluence rate response data, both for induction and inhibition experiments (Fig. 2A and 2C), and fit an exponential constant  $k_{Fit}$  for each wavelength. When plotting the  $k_{Fit}$  values against wavelengths (Fig. 3B), we found the inhibition and activation curves exhibit peaks in the blue and red or far-red regions, as the absorption spectra of the Pr and Pfr forms,

respectively (Fig. 3B). However, the relative height of these peaks does not fit the theoretical ones calculated with the absorption spectra and ratio of photoconversion yield (Fig. 3B, dashed line). In the action spectra,  $k$  at 430 nm,  $k_{430\text{nm}}$ , was expected to be about a third (0.3) of  $k_{730\text{nm}}$ , but we observed it to be 4.68 times higher. The same difference is observed in the inhibition spectra with  $k_{405\text{nm}}$  vs  $k_{730\text{nm}}$  (theoretical factor: 0.59, observed factor:3.3) or  $k_{690\text{nm}}$  vs  $k_{730\text{nm}}$  (theoretical factor:1.4, observed factor:10.3) (Fig. 3B). A linear relation between theoretical and observed values only reaches adjusted  $R^2=0.4311$ , but this rose to 0.6247 when ignoring the values in far-red bands (>700nm). Given the poor abundance of far-red light in the marine environment, the second relation was kept to calculate  $k$  in a complex light field and to model DPH activity in the environment.

### **DPH activity in the environment**

To scan DPH activation state in the ocean in a wide range of conditions, we projected the DPH activity model established above in a large suite of modeled marine light fields, using a relatively simplified bio-optical model. Briefly, we computed the sea surface light spectra using the Tropospheric ultraviolet and visible radiation model (Madronich 1987). The underwater spectrum at any depth depends on the sea surface light irradiation, depth and the spectral diffuse attenuation coefficient of the water environment  $K$ .  $K$  can be modeled as the sum of the absorption spectrum of the water ( $K_w$ ), the attenuation coefficient of the phytoplankton ( $K_{ph}$ ) and the attenuation coefficient of the background turbidity ( $K_{bg}$ ) due to non-algal particles and dissolved organic matter. The formula and tabulated values from Morel and Maritonema (2001) have been used to calculate light attenuation (350-700nm) of phytoplankton in the water column for different Chl *a* contents. We also added absorption by dissolved organic matter and inanimate particles ( $K_{bg}$ ) using  $K_{bg(\lambda)} = K_{bg(440\text{nm})} \exp(-S(\lambda-440\text{nm}))$ , with  $S=0.017$  (Babin et al, 2003, Holtrop et al, 2021). It is worth noting that this approach neglects transpectral processes, but allows a rapid exploration of a large suite of scenarios. We started considering that all parameters are homogeneous throughout the vertical column water. Very different spectra were calculated at the bottom of the photic zone (1% of surface light intensity) (Fig. S4A and D), with particularly varying ratios of blue to red and blue to green bands

(Fig. S4B, C, E, F) with depth, and the angle the sun light hits the water surface (Fig. S4G and H).

Projection of %PrPr at saturation - using the model developed with the fluence rate response results - showed that for Chl a concentration from 0 to 1mg/m<sup>3</sup>, the percentage of active PtDPH increases with depth, rapidly in the first approx. 20 m and more slowly going deeper (Fig. 4A). Interestingly, at the same time, red and far-red bands are strongly absorbed by water and blue light becomes dominant (Fig. S4A to F). For higher, though homogeneous, values of Chl a (1 to 5 mg/m<sup>3</sup>), % of PrPr increases compared to the surface down to approx. 20m, and then begins to decrease (Fig. 4A). The depth of this subsurface maxima decreases when Chl a increases (17m for 2mg/m<sup>3</sup>, 8m for 5mg/m<sup>3</sup>) (Fig. 4A). Because Chl a absorbs mainly blue and red bands, the increase in its concentration results in a reduced intensity predominantly of blue light (Fig. S4A to C). These results indicate that DPH activation state is thus depth- and Chl a- concentration/distribution dependent (Fig. 4A and S5C). *K<sub>bg</sub>* variations, mimicking coastal (*K<sub>bg</sub>* from 0.005 to 0.6 m<sup>-1</sup>) and clear oceanic (*K<sub>bg</sub>* around 0.003 m<sup>-1</sup>) environments resulted in the same effects as increasing Chl a concentration due to the absorption of blue light by organic matter (Fig. S4D, E and F), although at low but not null *K<sub>bg</sub>* values, DPH activity will slightly increases (Fig. S5A and D).

To take into account the light intensity and the sensitivity of PtDPH, we used an integration time of 10 min (as in experimental data), hypothesizing 100% of PfrPfr as initial state. This showed that PtDPH can be active even below 100 m in clear waters, while it would be active only at the surface in turbid environments (Fig. 4B and S5B).

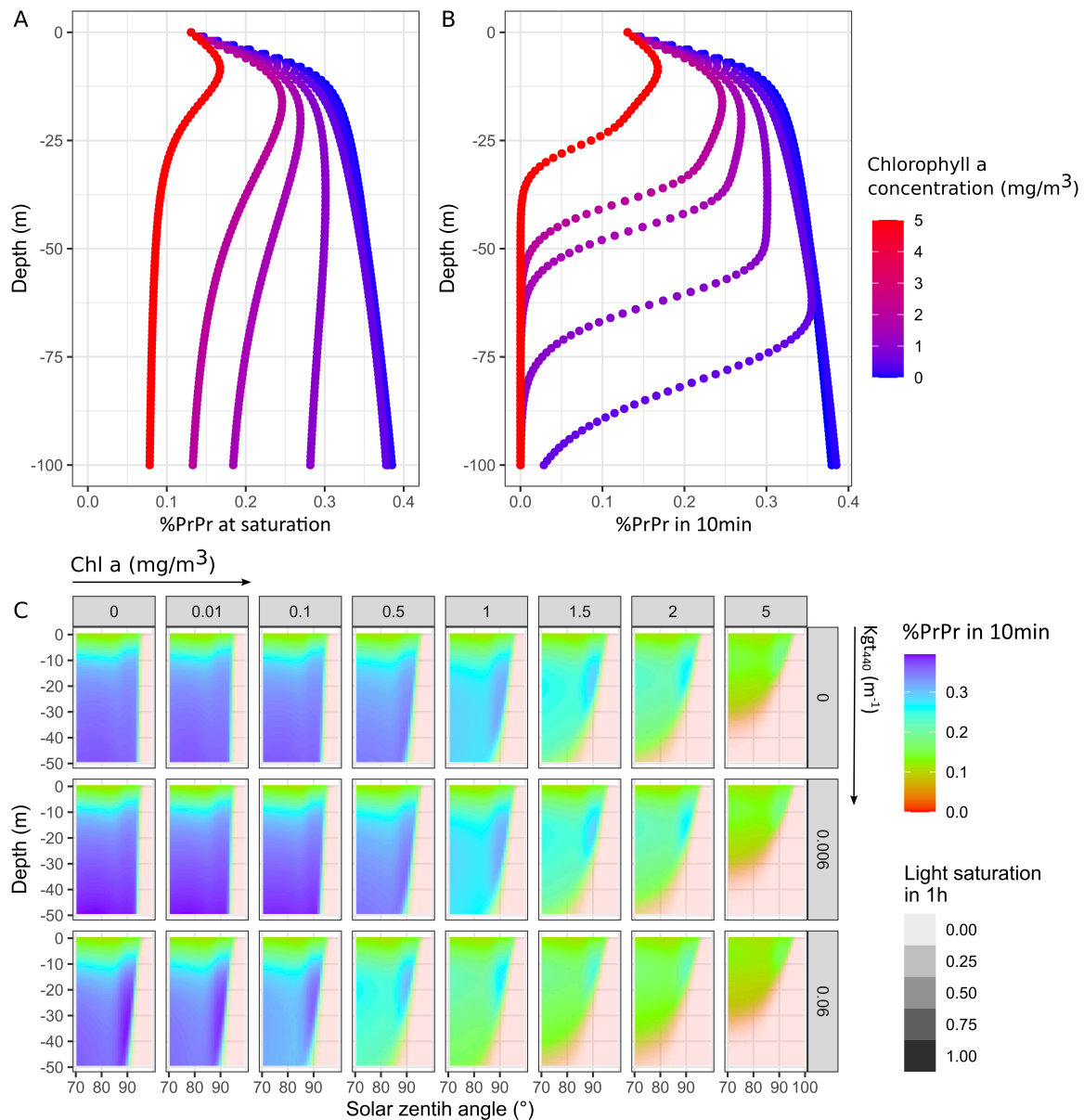


Figure 4. PtDPH activity as proportion of PrPr formed in modeled environmental fields varies with depth, water turbidity, and solar angle. Effect of chlorophyll a (Chl a) and depth on DPH photoequilibrium at saturation (A) or after 10 min of illumination (B). C. Combined effects depicted as color of depth (Chl a), background turbidity ( $K_{bg}$ ) and sun zenith angle on PtDPH activity, calculated for 1h of illumination. Transparence levels reflect the saturation of the photoreactions in 1h.

To include all the components of the underwater light field in a realistic scenario, we ran the photoresponse model using profiles measured *in situ* with light wavelengths up to 800 nm compared to the modeled ones that were only going to 700 nm. The shape of the DPH activity profiles and Chl a-dependence were clearly close to those calculated with

modeled light (Fig. 5) while reflecting the dishomogeneous distribution of bio-optically active components and transpectral processes. These *in situ* light profiles are also examples of non-homogeneous distributions of Chl a in the water column, with the presence of a deep chlorophyll maximum. Using the environmental light profiles, results matched those obtained with the modeled light (Fig. 4B and 5). Interestingly, PtDPH activity drops below the deep chlorophyll maximum (when present) (Fig. 5).

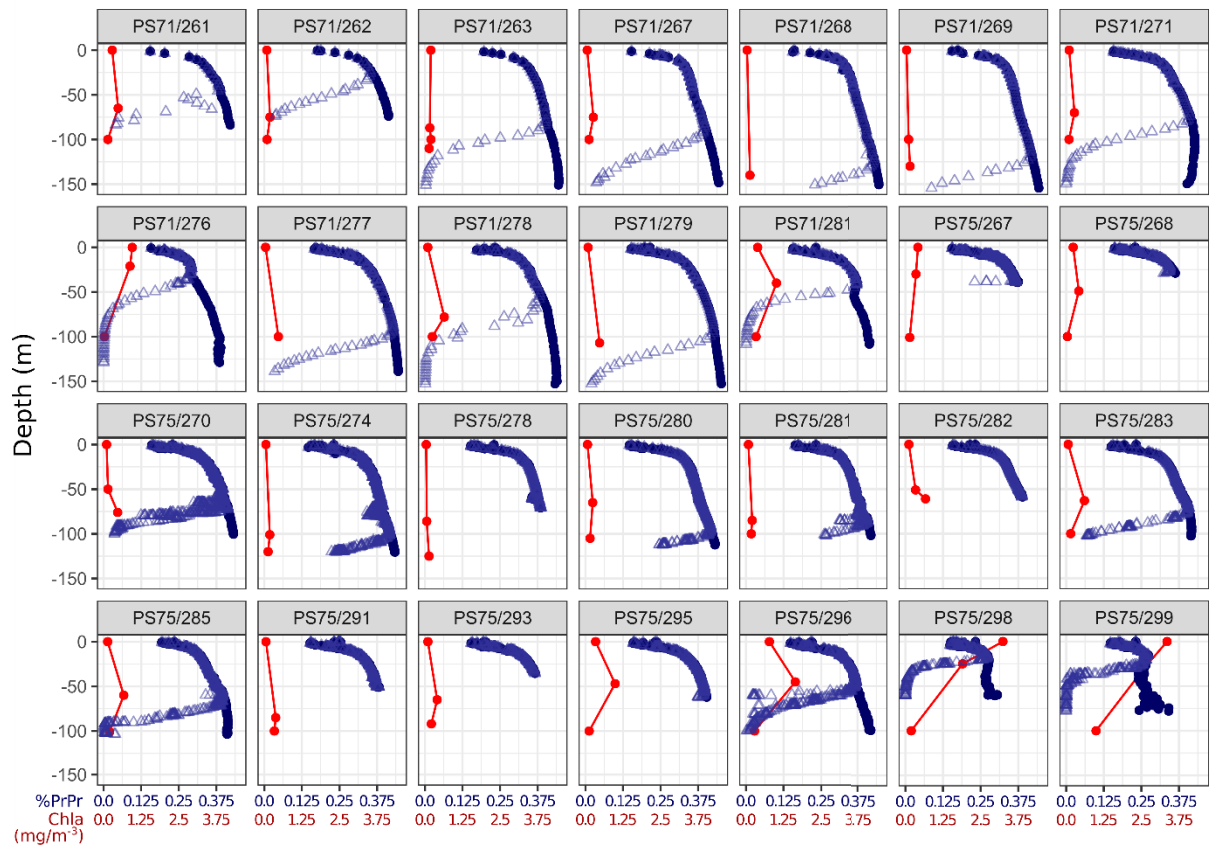


Figure 5. PtDPH activity in *in situ* measured field fields. A. DPH activity as proportion of PrPr formed (%PrPr), at saturation (dot) and after 10min illumination (triangles), in *in situ* environmental light fields measured at different stations in the Atlantic Ocean (data gathered in Bracher et al, 2015). In red, chlorophyll a (Chl a) concentrations measured by HPLC at different depths.

We also examined the effect of solar angle. Solar zenith angle changes with time of day, but maximal zenith angle in the day also changes with the season and the geographical position. The effect of solar angle on the underwater spectral field (Fig. S4G and H), and therefore PtDPH activity, appears to be only impacted for very low solar angles (from  $80^\circ$ ), i.e., sunrise and sunset, or at high latitudes in winter (Fig. 4C and S5E). The exposure time was critical to determine if PtDPH would sense the spectral variation



at low solar angle. If exposure time was kept to 10min, PtDPH would not perceive the increase of blue compared to red and green lights at the end of the day. However, increasing exposure time to 1h, hypothesizing a fluence dependent response, would allow PtDPH activity to increase at low solar angles (Fig. 4C, Fig. S5E). In that case, the % of PrPr can increase from 20% during the day to 27% at dawn/dusk (Fig. S5E, at 10 m). We also used time series light data from Veedin Rajan et al (2021), and calculated an increase of DPH activity at the beginning and end of day of about 5 % of %PrPr variation between middle of the day (28%+/-2.8 at noon) and dawn/dusk (33%+/-0.8 at 9pm)) (Fig. S6).

In conclusion, we show here that in environmental light spectra, PtDPH activity increases at depth in clear water, but decreases when turbidity increases; to a lesser extent, DPH can sense spectral variations at low solar angles. This behavior matches the relative enrichment of blue bands compared to red and even green bands in the light fields.

### **DPH biogeography and its global drivers**

As DPH is not conserved in all diatoms (Fortunato et al, 2016), we conducted biogeography studies to identify environmental factors that could influence DPH distribution. To do so, we looked directly for *DPH* genes in the Marine Atlas for *Tara* Oceans Unigenes (MATOU, version2) (Carradec et al, 2018).

We designed DPH-specific HMM models based on known DPH protein sequences. This search resulted in 95 *DPH* genes attributed to abundant diatom genera such as the centric Thalassiosirales or Cymatosirales, with 2 near-full-length proteins, while only 2 pennate DPH genes were found (Supplementary Data, online). This is consistent with DPH not being present in the most abundant oceanic pennate diatom genera, *Fragilariopsis* and *Pseudo-Nitzschia* (Malviya et al, 2016). Searches in the recently produced SMAGs data (Single cell and Metagenome Assembled Genomes) from *Tara* Oceans (Delmont et al, 2020), produced similar results with only MAGs corresponding to the Thalassiosirales or Cymatosirales phyla possessing DPH. No other diatom MAGs possessed DPH or other types of phytochrome-derived photoreceptors (i.e., neochromes or other combinations of domains).

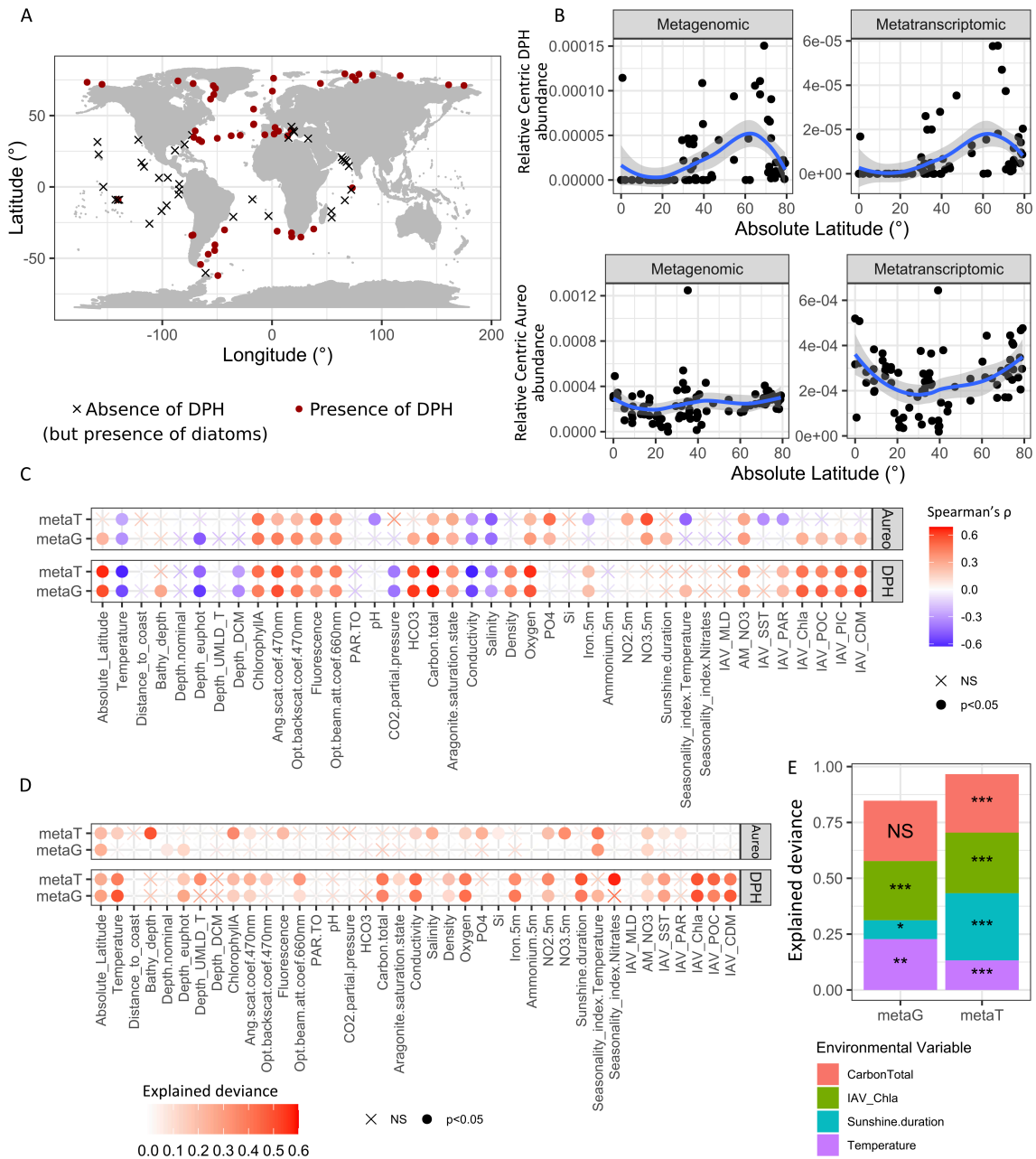


Figure 6: DPH distribution is linked to latitude, temperature and optical parameters. A. Map of the presence and absence of *DPH* genes in *Tara* Oceans sampling stations. B. Centric *DPH* (top panels) and Aureochrome (*Aureo*, bottom panels) gene abundance relative to total centric diatom genes as a function of absolute latitude, from metagenomic (left panels) and metatranscriptomic (right panels) data. C. Spearman's correlation of *Aureo* or *DPH* relative abundance with different environmental parameters. D. Generalized additive models (GAM) of *Aureo* and *DPH* relative abundances with different environmental parameters ("individual gams"). E. Complex GAMs explaining *DPH* relative abundance with a combination of environmental parameters.

Mapping the presence and absence of DPH in the different *Tara* Oceans stations (Fig. 6A) revealed that these genes are hardly found in tropical regions (only 2 stations with DPH) while they are present in nearly all stations in extratropical regions (absolute latitude > 30°) (Fig. 6A and Fig. S7A and B). Biogeography of the DPH-containing diatom MAGs reflected those found for DPH genes in the *Tara* Oceans metatranscriptomes, with nearly no presence detected in tropical regions (Fig. S8).

To confirm that this localization was specific to DPH, we designed HMM models specific for Aureochromes (Aureos), which are blue light photoreceptors ubiquitous in diatoms. 1305 genes were found as diatom Aureos in the *Tara* Oceans metatranscriptomes of which 817 belonging to centric species. When mapping the presence/absence of Aureo genes in *Tara* Oceans stations, Aureos were found to be ubiquitous (Fig. 6B and Fig. S7C and D).

Given the low number of pennate DPH sequences, we performed a quantitative analysis of DPH distribution only with centric genes. The abundance of centric DPH reads was normalized against total centric reads from the MATOU2 gene set. As shown in Fig. 6B (and Fig. S9), the relative abundance of DPH sequences increases with absolute latitude, while Aureo abundance did not exhibit any particular latitudinal pattern, suggesting a specific biogeography of DPH containing diatoms. To decipher the impact of environmental conditions on DPH biogeography, we used relative gene abundance to perform Spearman's correlations with a set of environmental variables (Fig. 6C and Fig. S10A). These were a mix of parameters measured *in situ* (available at PANGEA, Pesant et al., 2015), but also seasonality indexes derived from the World Ocean Atlas (nitrate concentrations and temperature annual variations) and from satellite measurements (annual variation of chlorophyll a concentration, downwelling attenuation coefficient Kd, dissolved organic matter (Ardyna et al, 2017)). In agreement with previous analyses (Fig. 6A and B), a strong positive correlation was detected between the relative centric DPH abundance and the absolute latitude, and a negative correlation with temperature, salinity and conductivity (Fig. 6C et Fig. S10A). Strong positive correlations were found with Chl a concentrations, total carbon and other related *in situ* optical parameters such as backscattering at 470 nm, fluorescence, particle backscattering coefficient or beam attenuation coefficient at 660 nm, and a negative correlation was found with depth of the

euphotic zone (Fig. 6C and Fig. S10A). DPH abundance also positively correlated with seasonality indexes related with Chl a and light. However, DPH relative abundances were not strongly linked to nutrient concentrations (no correlation with  $\text{NO}_3$ ,  $\text{PO}_4$ , Si, small with Fe). These tendencies are common to metatranscript reads and metagenomic reads, and are conserved when separating samples by filter size (Fig. S10A). Aureo relative abundance also correlated strongly with *in situ* optical parameters, but not with their annual variations. Interestingly, Aureo relative abundance in metatranscriptomes showed positive correlations with nutrients (especially  $\text{NO}_3$  and  $\text{PO}_4$ ) and  $\text{CO}_2$  as well as a negative correlation with pH (Fig. 6C and Fig. S10A). Such a pattern was largely absent in metagenomes. This could suggest regulation of Aureo transcription by nutrients, contrary to what is observed for DPH.

To go further with this analysis of DPH biogeography and capture non-monotonous relations between DPH abundance and environmental parameters, we performed “individual” generalized additive models (GAMs, DPH abundance explained by one environmental variable). Individual GAMs analysis summarized the above correlation analysis, but detected also additional relations, such as with iron concentration and sunshine duration (Fig. 6D and Fig. S10B). More complex GAMs were performed by selecting environmental variables which were found with high explanatory power in the individual GAM analysis and did not correlate with each other. We ended with 4 variables: *in situ* temperature, total carbon (which correlates with angular scattering coefficient at 470nm and Chl concentration in the water), sunshine duration, and annual variability of Chl a concentration. Together, these parameters could explain 84.7% to 96.6% of the deviance in metagenomes and metatranscriptomes, respectively (Fig. 6E, Fig. S11). The analysis done with the same environmental variables using Aureo genes explained only little deviance (14.2 and 25.1% in metatranscriptomes and metagenomes), with a many non-significant terms (only temperature was found to be significant).

These analyses indicate that DPH can be found in open ocean regions, especially in temperate and polar regions where high variations of primary productivity occur.

## Photochemical properties of DPH are strongly conserved

TpDPH and PtDPH absorption spectra are so far the only 2 examples of DPH absorption spectra, and both exhibit peaks in red and in far-red bands, characteristic of bacteriophytochromes (Fortunato et al, 2016). To determine whether these features were widely conserved among DPH or if there can be spectral tuning in some species, as seen in other algal groups (Rockwell et al, 2014), we produced recombinant versions of additional DPH photosensory modules (PSM), by co-expressing in *Escherichia coli*, PSM sequences together with the *Synechocystis* Heme Oxygenase gene, encoding the enzyme producing biliverdin, and purified the recombinant proteins. Four sequences were cloned from different distantly related diatom species, with distinct ecological niches, isolated at different locations around the world (Table S1). Moreover, DPH from two synthetic genes from sequences identified in *Tara* Oceans dataset were also produced (Table S1).

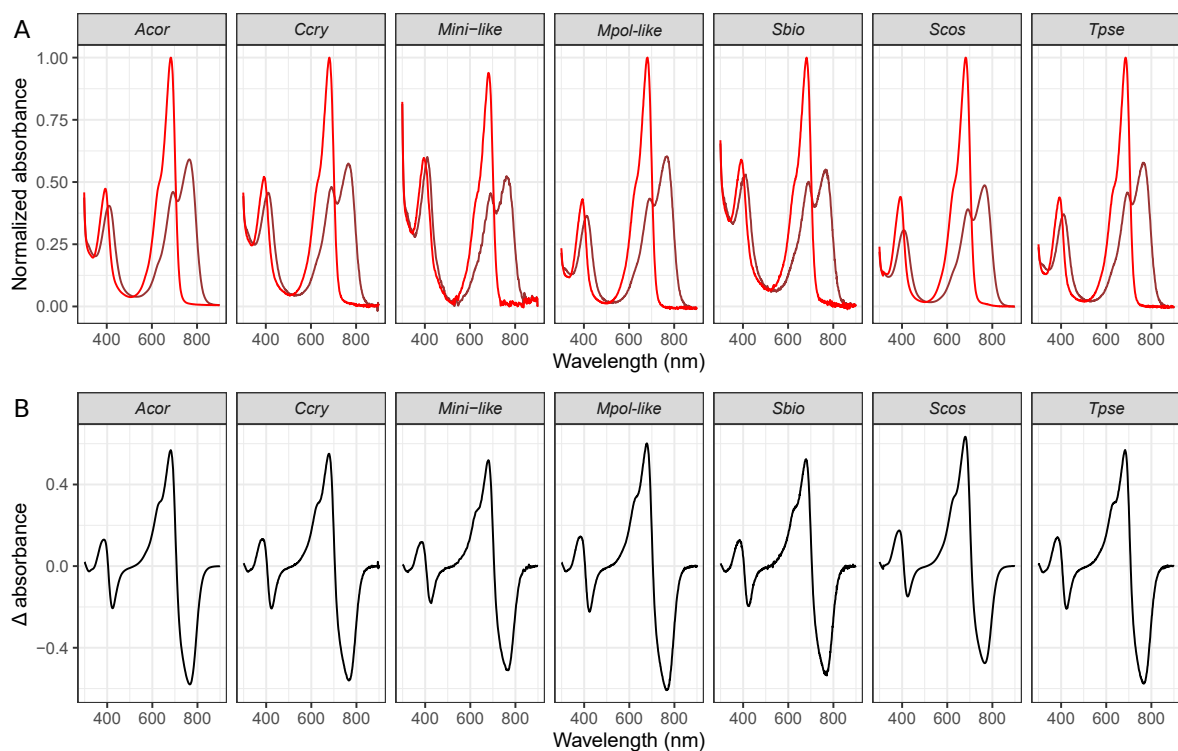


Figure 7. Spectral properties of DPH from various species are conserved. A. Absorption spectra of far-red light illuminated (red line) and red light illuminated (darker red line) recombinant photosensory domains of DPH expressed with biliverdin as chromophore. B. Differential absorption spectra between red- and far-red-illuminated DPH. *Acor*, *Arcocellulus cornucervus*; *Ccry*, *Cyclotella cryptica*; *Mini-like*, *Minidiscus-like* sequence; *Mpol*, *Minutocellus polymorphus-like* sequence; *Sbio*, *Shionodiscus bioculatus*; *Scos*, *Skeletonema costatum*; *Tpse*, *Thalassiosira pseudonana*.

Table 1: Spectral and photochemical properties of different recombinant DPH

Phytochrome	max $\Delta A$ in B	min $\Delta A$ in B	max $\Delta A$ in R	min $\Delta A$ in FR	$\eta$
Acor	385	424	682	766	0.65
Ccry	383	424	678	767	0.76
Mini-like	388	427	680	769	NA
Mpol-like	383	424	678	766	0.82
Scos	388	425	680	767	0.61
Sbio	385	423	678	767	NA
Tp*	384	425	684	767	0.68
Pt*	380	420	To 698	751	0.66

\*absorption spectra from Fortunato et al, 2016. Max and min  $\Delta A$  : maximum and minimum of the differential absorption spectra in the blue (B), red (R), and far-red (FR) bands.

$\eta$  : ratio of quantum yield. NA, indicates that the ratio of the quantum yield could not be determined from the absorption spectra obtained (recombinant proteins were not pure and abundant enough, which would lead to inaccurate values).

Acor, *Arcocellulus cornucernis* RCC2270; Ccry, *Cyclotella cryptica* CCMP332; Mini-like, synthetic gene related to *Minidiscus* species identified in MATOU; Mpol-like, synthetic gene related to *Minutocellus polymorphus* species identified in MATOU; Scos, *Skeletonema costatum* RCC1716; Tp, *Thalassiosira pseudonana* CCMP1335; Sbio, *Shionodiscus bioculatus* RCC1991; Pt, *Phaeodactylum tricornutum* CCMP2561.

All these sequences exhibit a cysteine residue in the N-terminal extremity domain, at a conserved position with the cysteine residue found to bind the biliverdin chromophore in bacterial and fungal phytochromes (Fig. S12). All purified proteins were chromophoric and exhibit similar absorption spectra, with absorption maxima in the red (682 to 686 nm) and far-red (762 to 768 nm) light, essentially the same as previously observed in Pt- and TpDPH (Fig. 7 and Fig. S13). As already observed for Pt- and TpDPH, no thermal reversion occurred over 24h with these recombinant proteins (data not shown). We used the method from Giraud et al (2010) to estimate the absorption spectra of the pure Pfr form and estimate the ratio of photoconversion reactions ( $Pr \rightarrow Pfr$  and  $Pfr \rightarrow Pr$ ) of these

recombinant proteins. These latter little vary, with values from 0.61 for *S. costatum* DPH to 0.82 for the synthetic *Minidiscus*-related DPH (Table 1). It is noteworthy that the value determined for the recombinant PtDPH is pretty close to the one determined *in vivo* with the reporter system (0.66 and 0.98 respectively).

The photochemical properties of phytochrome, both in terms of absorption spectra and ratio of quantum yield, thus appear to be strongly conserved in phylogenetically and ecologically distant diatom species.

Additionally, we introduced in a ProHSF4.6a::YFP KO reporter line, the *DPH* gene of *T. pseudonana*. We observed that YFP induction in response to far-red and blue light was restored in the transformed line (Fig. S14), showing that TpDPH can complement the PtDPH KO mutant. This result indicates that TpDPH under its Pr form can activate the same signaling cascade as PtDPH.

We then used the photochemical values determined for recombinant DPH from *T. pseudonana* and *S. costatum* (Table 1), to model their activity in the marine light fields. Overall, the effects of depth, turbidity and time of day were the same as we found for PtDPH, as could be expected from the very close biophysical properties of the recombinant proteins (Fig. S15). However, TpDPH in particular seems to be more responsive to Chl a variations (Fig. S15B). In summary, here we showed that the modeled DPH activity using PtDPH can be extended to DPH of ecologically relevant species, and reveals the ubiquitous property of DPH to effectively sense blue and green lights variations along the water column, in addition to red and far-red lights. This allows exquisite sensing of depth and chlorophyll concentrations.

## DISCUSSION

### **DPHs from distant diatom species display conserved spectral and photochemical properties**

Prior to the current study, absorption spectra from only two diatom phytochromes were known: from the pennate *P. tricornutum* and the centric *T. pseudonana*. Additionally, the

phytochrome from the closely related multicellular brown alga, *Ectocarpus siliculosus*, showed spectral variation with a far-red/green photocycle that was hypothesized to be linked to the intertidal and shallow sublittoral habitats of this alga (Rockwell et al, 2014). Moreover, the blue-shifted dark state of several phytochromes from marine Prasinophyte and Glaucophyte living in coastal and open waters was suggested to be an adaptation to marine blue-enriched light fields (Rockwell et al, 2014). This is clearly not the case in diatoms, as spectra of different DPHs show strong conservation of spectral properties. Even DPH from species that are evolutionary more distant (*Minutocellus* sp. and *Arcocellulus cornucervis* compared to the Thalassiosirales) or synthesized from environmental sequences, and therefore likely to be chimeric, share close photophysical properties in terms of absorption and ratio of photoconversion yield (Table1).

We also found that TpDPH can complement PtDPH function in *P. tricornutum*, at least for the phenotype studied here (far-red- or blue light-induced expression of the reporter protein). This suggests that at least the first signaling step downstream of DPH activation could be shared between *T. pseudonana* and *P. tricornutum* and that Pr could be the active form for TpDPH, and possibly also that of the other DPHs characterized in this study.

Overall, these results imply that the conserved spectral features of phytochrome among diatom species/genera must confer a specific adaptive value to the life of this group of organisms, different from those provided by phytochrome spectral tuning observed for algae belonging to other phyla.

### **Blue light-enriched fields favor DPH activation**

Analysis of diatom phytochrome activity and distribution has revealed several novel aspects about the adaptation of this photoreceptor to work in the marine environment. We found that *P. tricornutum* DPH-dependent responses exhibit photoreversibility, characteristic of phytochrome photoreceptors. Moreover, we also found that these responses could be elicited upon blue light exposure, with an even higher efficacy than the far-red ones, which was unexpected based on the PtDPH absorption spectra.



The PtDPH-dependent blue light induction was observed for the *HSF4.6a* gene, whose expression was monitored with the eYFP reporter, but also for other genes, previously identified as far-red light- and PtDPH- regulated (Fortunato et al, 2016) (Fig. S3). Phytochrome-dependent blue light responses have also been observed in fungi, terrestrial heterotrophic organisms possessing phytochromes similar to bacterial phytochrome ones and phylogenetically close to DPH. For instance, *Aspergillus nidulans* phytochrome FphA was shown to be a dominant photoreceptor controlling both red- and blue-light mediated gene expression, with 86% of the blue-light induced genes requiring the presence of this sensor (Yu et al, 2021). A milder role of phytochrome in blue light responses has also been observed in the fungus *Ustilago maydis* (Brych et al, 2021). On the contrary, in plants, while phytochromes also exhibit a minor absorption peak in the blue band (Butler et al, 1959; Beggs, 1981), a very limited effect of blue light has been reported for phytochrome-mediated responses (Shinomura et al, 1996). This weak efficacy of blue photons can be explained by a possible distortion of light in leaves due to the presence of absorbing photosynthetic pigments as well as scattering by cell walls (Kazarinova-Fukshansky, 1985; Kusuma, 2021).

In diatoms, despite the presence of blue light-absorbing photosynthetic pigments (i.e., chlorophyll c and fucoxanthin), a higher responsiveness to blue, compared to far-red light, has been observed for PtDPH. This behavior could suggest a local increment of blue photons actually reaching DPH within the cell. It is known from diatom species with a frustule that the light field can be modified when it goes through the silica cell wall, the valves trapping and scattering photons, blue ones in particular, resulting in an enhancement of photosynthesis (Goessling et al, 2018). However, the *P. tricornutum* cell wall is very thin, and no such light field alteration has been reported so far. Because the level of the measured PtDPH-mediated response at saturating light matches the absorption ratio of the Pfr/Pr forms, we hypothesize a reduction in the number of far-red photons reaching DPH, rather than an increase of blue photons. In addition, this also suggests that in our condition, PtDPH would be the sole photoreceptor responsible for this response, but we cannot exclude that the sensitivity to blue compared to far-red light derives from the synergistic action of other blue-light photoreceptors on the DPH-

signaling pathway, as shown for FphA (Yu et Fischer, 2018; Yu et al, 2021) and in plants (Casal et al, 1998; Usami et al, 2004; Hughes et al, 2012).

In any case, our modeling reveals that blue light is the prominent waveband favoring DPH activation in the marine environment. Indeed, both our characterization of the *in vivo* PtDPH features, and the rapid depletion of the long wavebands in the water column drive this effect. Consequently, red, as well as green light, act mostly in reverting blue light activation. This new scenario minimizes our previous efforts to find the far-red light sources eliciting DPH activation underwater (Fortunato et al, 2016). It also makes the contribution of the red/far-red fluorescence emitted by the chloroplast(s) negligible since it represents <5% of the exciting blue light, and is therefore largely overcome by the light applied.

### **DPH is an extremely sensitive photosensor in marine environments**

Our projection of DPH activity in different marine environments, from clear oceanic waters to turbid coastal environments, shows that DPHs sense both the intensity and the ratio of wavebands. The first parameter controlling DPH activity is depth. In our model, DPH will be less active at the surface due to a higher red-to-blue ratio, but can quickly change its photoequilibrium with the mild spectral changes occurring in the first ~ 20 meters of the water column, as the light intensity is high and not limiting. With depth, DPH becomes more active while the red- and green- to-blue ratio get lower, but then, its activity becomes fluence rate-dependent, as light is limiting. Because of the dynamic of DPH equilibrium and the effect of different wavebands on its sensitivity all along the photic zone, DPH is predicted to be an efficient depth sensor measuring green-red/blue ratio. At depth, in a blue-enriched environment, DPH could also mediate long-term low light acclimation responses. DPH could also have a role in perceiving changes in the light field, both in amplitude and spectrum, associated with vertical displacements, due to convective events or wind mixing, both being more intense at mid-high latitudes. Considering vertical velocities in the range of 0.01 to 0.1 m.s<sup>-1</sup>, which are typical of convection or Langmuir circulation (Moum and Smyth, 2019), in one hour cells can be

displaced by a few to tens of meters producing a change in the ratio between the two DPH forms by a factor of more than 3.

In addition to depth, our model also shows that DPH activity increases when solar zenith angles are low, such as at dawn and dusk, during winter, or at high latitudes. Parallel analysis of DPH biogeography in the *Tara Oceans* dataset showed that DPH-containing diatoms are distributed with a latitudinal pattern, in areas with strong seasonal variations. DPHs are quasi-absent from the tropics (metatranscriptomes from only 2 stations contained DPH sequences), which is consistent with a recent analysis of metatranscriptomic data in a tropical region (north of Hawaii), indicating an absence of *DPH* genes (Coesel et al, 2021). This is not due to diatoms being absent in these regions, as the *Thalassiosirales*, one of the most represented genera in *Tara Oceans*, are particularly abundant in the tropical Pacific Ocean (Malviya et al, 2016). In addition, we did find the blue light Aureo photoreceptors from diatoms in these regions. The strong latitudinal distribution pattern observed for DPH opens interesting novel hypotheses about the role of this photoreceptor in the detection of seasonal light variations, and photoperiodic regulation, processes that are still largely unexplored in phytoplankton. Interestingly, we also found that DPH activity in the environment is strongly modulated by light field changes induced by phytoplankton concentration (i.e. Chl a concentration), and tightly follows the displacement of the deep chlorophyll maximum. This result is evocative of the neighbor detection response triggered by plant phytochromes (Casal et al, 2012), though strikingly different, as in diatom it would be based on the reduction of blue/green ratio light absorbed by the photosynthetic pigments and not in a reduction in the ratio of red/far-red light, as in plants. DPH have been found at high latitudes where seasonal blooms are more common than in the oligotrophic ocean (Malviya., et al. 2016). In blooms, phytoplankton concentration increases drastically, which, according to our model, will reduce activity of DPH. Blooms of *Thalassiosirales* species in late winter/early spring, followed by *Chaetoceros* species (a non DPH-containing genus) have been recorded (Caracciolo et al, 2021; Arsenieff et al, 2020). It is thus appealing to speculate that a change in DPH activity because of increased phytoplankton concentrations could be an end-of-bloom signal, controlling species succession and bloom timing. This modulation of DPH activity could also play a role in the germination of diatom resting

spores (Eilertsen et al, 1995) or even sexual reproduction as a signal that cell concentration is high enough to allow gametes to meet. This last hypothesis is particularly appealing, as red light is known to induce sexual reproduction in the centric diatom *Stephanopyxis palmeriana*, in a far-red reversible manner (Ren et al, 2005). Also, phytochromes are known to be primary regulators of sexual reproduction in fungi (Blumenstein et al, 2005). Unfortunately, no genomic and genetic tools are available for *Stephanopyxis palmeriana*, and the transcriptome data from a related species, *Stephanopyxis turris*, does not show DPH expression. In the future, we could potentially test this phenotype in other diatom species expressing DPH and showing sexual reproduction.

Finally, although the function of this photoreceptor remains unestablished in diatoms, as in other microalgae, this study identifies critical parameters influencing the activity of DPH in an environmental context. By completely changing perspectives on phytochrome activity in the watercolumn, it identifies DPH as a highly sensitive photoreceptor, whose activity would be intimately tuned to the light distribution pattern and critical biological processes in the marine environment.

## MATERIAL & METHODS

### Culture conditions

Wild-type *P. tricornutum* (Pt1 8.6; CCMP2561) cells, PtDPH-KO mutants (Fortunato et al, 2016), and the reporter lines derived from them, *Thalassiosira pseudonana* CCMP1335, *Cyclotella cryptica* CCMP332, and *Skeletonema costatum* RCC1617 were maintained in Enriched Seawater, Artificial Water at 19 °C under a 12L/12D regime using 50  $\mu\text{mol photons.m}^{-2}.\text{s}^{-1}$  white light (Philips TL-D De Luxe Pro 950). *Arcocellulus cornucervis* RCC2270 and *Shionodiscus bioculatus* RCC1991 were grown in enhanced artificial seawater at 4°C in 12hL/12hD cycles with RGB LED panel. See also Supp Table1 for details of the strains.

## **Experimental conditions for action and inhibition spectra**

For action and inhibition spectra cells, reporter lines were maintained in continuous green light (LEDs, 520nm, 22  $\mu\text{mol}/\text{m}^2/\text{sec}$ ) with shaking (160 rpm) for at least two weeks before the experiments. On the day of the experiments, 2 or 4L cultures at about  $1.5 \times 10^6$  cells/mL were split into 20mL in plastic culture flasks (Corning). 25 flasks per LED were then irradiated one behind the other to generate a gradient of light. For the inhibition spectra flasks were exposed in an incubator to 10min of far-red light (800nm, 10  $\mu\text{mol m}^2/\text{sec}$ ) prior to illumination with the LED to test. The LED was turned on for 10min with no other light source then turned off, and the flasks remained in the dark for an additional 5h50min time needed to have a stable amount of eYFP/cells. Shaking was maintained during the illumination and the dark period. The same experiment was repeated for all the LEDs to test (405, 430, 470, 730, 740, 765, 810, 850nm) for the action fluence responses, and 405, 520, 630, 690 and 730 nm for the inhibition fluence responses; see figure S1 for LEDs spectra). Multi-wavelength expositions were performed in incubators equipped with Blue (450nm), Green (520nm), Red (680nm) and FR (800nm) LEDs. Light mixes were chosen based on the action and inhibition spectra results (ratio of wavebands and intensities) to cover ratio of DPH absorption not covered before, and to use light intensity supposed to saturate DPH reaction in 10min. Cultures grown in the same conditions as for the spectra and split to 20mL cultures, exposed to 10min of LED mix then transferred in the dark for 5h50min before eYFP measurements. These experiments were done twice for each strain, with two technical replicates each time.

At the end of the dark period, each flask was sampled and passed through a MAQSQant flow cytometer. eYFP signal per cell was recorded in the B1 channel (excitation 450nm, 500-550nm emission filters). Each sample was run until 29700 events from the target population were measured, and eYFP signal measures per cell were averaged on the whole diatom population. eYFP signal was normalized between the signal from cells that went directly from green (growth) light to darkness (normalized to 0) and to cells that were exposed to 10min of FR (765nm) and then darkness (normalized to 1). This allowed us to compare different strains with different basal expressions of the eYFP.

The different equations used for the modeling of DPH activity are issued from Mancinelli (1994), and are presented in detail in Supplementary method. Analysis of the normalized eYFP signal at “saturating” light intensities was done using the 4 flasks exposed to the highest intensities for monochromatic illumination and the data from the bichromatic illuminations. For each strain, equation corresponding to the normalized YFP for each wavelength ( $\lambda$ ) at light saturation (Equation 5 in Supplementary method) was fitted with the nls function in R, with only the ratio of quantum yield as parameter to estimate. For the analysis of the eYFP-light intensity curves, for each strain and for each LED, the eYFP signal was normalized between 0 and 1. The equation corresponding to the normalized YFP for each wavelength ( $\lambda$ ) as function of the light intensity and time (equation (7) in Supplementary methods) was fitted to the data, with only the exponential constant  $k$  as missing parameter (Supplementary method).

### **Modeled and real environmental light spectra.**

We used TUV v5.3 (Madronich 1987) to compute sea-levels light spectra (direct solar and diffuse) for different sun zenith angles (downwelling irradiance only). Underwater attenuation coefficients were calculated with the formula from Morel and Maritonema (2001) and for CDOM and particles (Babin 2003; Kirk, 2021; Holtrop 2021). Light spectra were then calculated down to 100m deep for different solar zenith angles, Chl a concentration and particles content (hypothesizing that Chl a, CDOM and particles concentration would be homogeneous in the water column). Measured environmental light spectra were retrieved from various studies (Bracher et al, 2015).

### **Construction of Pt-DPH activity reporter lines**

The 931-nt upstream region of the *Hsf4.6a* (Phatr3\_J49557) and the *eYFP:FcpAt* fragment were amplified by PCR using using the Phusion High-Fidelity DNA polymerase (Thermo Fisher, USA) from genomic DNA and the primer pairs : *Hsf4.6ap\_Fw0* and *Hsf4.6ap-YFP\_Rv*, and from the pDEST-C-HA vector (**Erreur ! Référence de lien hypertexte non valide.**)**Erreur ! Référence de lien hypertexte non valide.** and the primer pairs : *YFP-Hsf4.6ap\_Fw* and *FcpAT\_Rv0*, respectively (Table S2). The two fragments were assembled

by PCR using the *Hsf4.6ap\_Fw0* and *FcpAT\_Rv0*, and the final product cloned into pGEM-T (Promega). Transgenic lines were obtained by biolistic, by co-transformation with Phleomycin resistance plasmid as described in De Riso et al. (2009), in WT, WT (Tc) (Transformation control) and PtDPH-KO mutants obtained in Fortunato et al (2016).

### **RNA extraction and gene expression analysis**

WT and PtDPH KO mutant lines were grown as described for the action spectra experiments, and exposed for 30 min of saturating blue (450 nm, 15  $\mu\text{mol}$  of photons/ $\text{m}^2/\text{sec}$ ) or far-red (800 nm, 15  $\mu\text{mol}$  of photons/ $\text{m}^2/\text{sec}$ ) light, or placed in the dark. About  $1.5 \times 10^8$  cells were harvested by filtration and total RNA was extracted as in Huysman et al. (2013). For qRT-PCR, 500 ng of total RNA were reverse-transcribed using the QuantiTect Reverse Transcription Kit (Qiagen, USA) and the reaction accomplished with 12.5 ng cDNA as template, following the SsoAdvanced Universal SYBR Green Supermix instructions (Bio-Rad, USA), in a CFX 96 Real-Time Detection System (Bio-Rad). Primer sequences are indicated in Table S2. *H4* were used as reference gene as in Siaux et al. (2007) and normalization was performed to the values before specific exposure (i.e. cells grown in continuous green light at 15  $\mu\text{mol}$  of photons/ $\text{m}^2/\text{sec}$ ).

### **Pt-DPH KO mutant complementation :**

Tp-DPH coding sequence was amplified and domesticated from *T. pseudonana* cDNA by fusing with the primer pair UNS1FL.FW and UNSXFL.RV, the PCR products obtained with the primer pairs: TpPHY.D.Fw and TpPHY.Sap1.Rv, TpPHY.Sap1.Fw and TpPHY.BsaI.Rv, TpPHY.BsaI.Fw and TpPHY.Sap2.Rv, TpPHY.Sap2.Fw and TpPHY.E.Rv.

The 1041 nt upstream PtDPH start codon and the 368 nt downstream the PtDPH stop codon were amplified from *P. tricornutum* gDNA with the primer pairs : PrPtPHY.A.Fw and PrPtPHY.C.Rv, and TrPtPHY.E.Fw and TrPtPHY.F.Rv, respectively. The Tp-DPH domesticated coding sequence, Pt-DPH promoter region and terminator regions were each cloned into the pL0 vector from the modular cloning system uLoop and assembled with the Venus fluorescent protein sequence in pL1 as in Pollack et al, (2020). The assembled product was amplified with the primer pair UNS1FL.FW and UNSXFL.RV, and

inserted into the SmaI-linearized pUC19 vector containing the Blastocidin resistance cassette of the pPTbsr (Buck et al, 2018) in Eco53KI site.

### **Expression, purification and spectral analysis of Dph proteins**

Photosensory module (PSM) of DPH genes were obtained either from previous studies (Pt-DPH, Tp-DPH (Fortunato et al, 2016)), cloned from cDNAs generated as described above, from cultures of *S. costatum*, *C. cryptica*, *Minidiscus spinulatus*, *S. biocatulatus*, *A. cornucervis*, or synthesized by GenScript (USA) (for the 2 near full length sequences identified in MATOU, *Minidiscus*-like and *M. polymorphus*-like) with codons optimized for *E. coli* expression. Primer pairs used are provided in Suppl Table2). Sequences were amplified with the primers pairs indicated in the Table S2, and cloned into pET28a-HO vector generated by inserting the *Synechocystis* Heme oxygenase gene amplified from pKS270 vector (Mukougawa et al., 2006) with the primer pair HO1xpET.HindIII.Fw and HO1xpET.NotI.rv, in HindIII/NotI in the pET28a vector (Novagen). PSM sequences were expressed as N-terminal His6 tagged proteins in *E. coli* BL21 (DE3) strain, by auto-induction system (Studier, 2005). Recombinant proteins were purified as in Fortunato et al. (2016) and their absorption spectra of the recombinant proteins were measured immediately after purification on a Varian's Cary-50 spectrophotometer. Illumination with LEDs at 810, 630 and 405 nm were performed (approx. 1min illumination) to reach pure Pr spectra (after 810 illumination) or equilibriums between the Pr and Pfr forms. These spectra were used as in Giraud et al (2010) to calculate pure Pfr spectra and ratios of quantum yields.

### **Search for DPH and diatom Aureochrome genes in environmental sequence data**

Known DPH sequences were aligned with mafft v7.4 to generate separate Centric and Pennate HMM profiles for each DPH domain, covering nearly the full-length protein. HMM searches (HMMER version 3.2) were performed against different set of control sequences: protein sequences of known phytochromes from streptophytes, chlorophytes, glaucophytes, cryptophytes, bacteria, cyanobacteria, fungi and Stramenopiles, the MetDB dataset, diatom genomes (P.tri, T.pse, C.cry and F.cyl) and a set of proteins specific for each domain (either the proteins used in the corresponding Pfam domain alignment or if these were too few in numbers, 1000 randomly selected proteins from proteins known to



possess the specific domain in InterPro). All domains with a  $e\text{-value} < 1e\text{-}5$  were retrieved, clustered at 90% identity with cd-hit and submitted at EFI to generate Sequence Similarity Network (SSN). Alignment scores for the SSN were chosen so that domain sequences from diatom phytochromes clustered together, not linked to other diatom proteins nor to non-diatom phytochromes (as illustrated in Fig. S8; see Table S3 for the SSN alignment score values for each domain). The same search method was then applied to the different environmental sequence dataset (MATOU gene atlas (6-frame translated) and Tara SMAGs (Carradec et al, 2018; Delmont et al, 2020)

The same method was applied to the aureochromes except for a few changes: only one HMM profile was generated from diatom aureochromes protein sequences, covering both the bZIP and the LOV domain; SNN being insufficient to distinguish diatom aureochromes from other aureochromes, environmental sequences were placed on an Aureochrome phylogenetic tree. Briefly, sequences were aligned to the Aureochrome HMM model using hmalign and the phylogenetic tree was reconstructed with FastTree (v2.1, default settings). Only sequences branching in the diatom clades (as delimited by the sister Bolidophyceae clade) were annotated as diatom aureochromes.

Mapped reads to the MATOU sequences were retrieved from each Tara Oceans sample. Relative abundance was expressed as the sum of all reads mapping to centric DPH (or centric Aureos) divided by the sum of all the reads mapping to centric diatom genes. The analysis was performed on each filter size separately (Supplementary figures, 5-20 $\mu\text{m}$ , 20-180 $\mu\text{m}$ , 180-2000 $\mu\text{m}$  and 0.8-inf), or average values of the 4 filter sizes when all 4 were available. Environmental variables measured *in situ* during the Tara Oceans campaign are available at PANGEA (Pesant et al., 2015).

Influence of environmental parameters on DPH abundances were assessed with Spearman's correlations or GAMs (mgcv package in R). Only variables with high explanatory power in the individual GAM analysis (pvalue of the smooth term  $< 0.5$ , explained deviance  $> 0.35$ ) and that did not correlate to each other (spearman's rho  $< 0.7$ ) were used for complex GAMs. The "proportion of explained deviance" attributed to each environmental parameter was obtained by performing the same GAM without this parameter; we then used the difference of explained deviance between full and reduced

GAM as the “proportion of explained deviance”, i.e. the weight of this parameter in the model.

## REFERENCES

**Ardyna, Mathieu, d'Ovidio, Francesco, Speich, Sabrina, Leconte, Jade, Chaffron, Samuel, Audic, Stephane, Garczarek, Laurence, Pesant, Stephane, Tara Oceans Consortium, Coordinators, Tara Oceans Expedition, Participants** (2017) Environmental context of all samples from the *Tara Oceans* Expedition (2009-2013), about mesoscale features at the sampling location. 1847508 data points

**Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, et al** (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306: 79–86

**Arsenieff L, Le Gall F, Rigaut-Jalabert F, Mahé F, Sarno D, Gouhier L, Baudoux A-C, Simon N** (2020) Diversity and dynamics of relevant nanoplanktonic diatoms in the Western English Channel. *ISME J* 14: 1966–1981

**Babin M** (2003) Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *J Geophys Res* 108: 3211

**Beggs CJ, Geile W, Holmes MG, Jabben M, Jose AM, Schäfer E** (1981) High irradiance response promotion of a subsequent light induction response in *Sinapis alba* L. *Planta* 151: 135–140

**Bhoo S-H, Davis SJ, Walker J, Karniol B, Vierstra RD** (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature* 414: 776–779

**Blumenstein A, Vienken K, Tasler R, Purschwitz J, Veith D, Frankenberg-Dinkel N, Fischer R** (2005) The *Aspergillus nidulans* phytochrome FphA represses sexual development in red light. *Curr Biol* 15: 1833–1838

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

**Bracher A, Taylor MH, Taylor B, Dinter T, Röttgers R, Steinmetz F** (2015) Using empirical orthogonal functions derived from remote-sensing reflectance for the prediction of phytoplankton pigment concentrations. *Ocean Sci* 11: 139–158

**Briggs WR, Spudich JL** (2005) Handbook of photosensory receptors. Wiley, Weinheim

**Brych A, Haas FB, Parzefall K, Panzer S, Schermuly J, Altmüller J, Engelsdorf T, Terpitz U, Rensing SA, Kiontke S, et al** (2021) Coregulation of gene expression by White collar 1 and phytochrome in *Ustilago maydis*. *Fungal Genet Biol* 152: 103570

**Buck JM, Río Bártulos C, Gruber A, Kroth PG** (2018) Blastocidin-S deaminase, a new selection marker for genetic transformation of the diatom *Phaeodactylum tricornutum*. *PeerJ* 6: e5884

**Butler WL** (1972) Photochemical properties of phytochrome in vitro. *Phytochrome*, Mitrakos K, Shropshire W Jr (eds). Academic Press, London New York, pp 185–192

**Butler WL, Norris KH, Siegelman HW, Hendricks SB** (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc Natl Acad Sci* 45: 1703–1708

**Caracciolo M, Rigaut-Jalabert F, Romac S, Mahé F, Forsans S, Gac J-P, Arsenieff L, Manno M, Chaffron S, Cariou T, et al** (2021) Seasonal temporal dynamics of marine protists communities in tidally mixed coastal waters. *bioRxiv* 2021.09.15.460302

**Carradec Q, Pelletier E, Da Silva C, Alberti A, Seeleuthner Y, Blanc-Mathieu R, Lima-Mendez G, Rocha F, Tirichine L, Labadie K, et al** (2018) A global ocean atlas of eukaryotic genes. *Nat Commun* 9: 373

**Casal JJ** (2012) Shade Avoidance. *Arabidopsis Book*. 10:e0157. doi: 10.1199/tab.0157.

**Casal JJ, Mazzella MA** (1998) Conditional synergism between Cryptochrome 1 and Phytochrome B is shown by the analysis of *phyA*, *phyB*, and *hy4* simple, double, and triple mutants in *Arabidopsis*. *Plant Physiol* 118: 19–25

**Coesel SN, Durham BP, Groussman RD, Hu SK, Caron DA, Morales RL, Ribalet F, Armbrust EV** (2021) Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. *Proc Natl Acad Sci* 118: e2011038118

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

**Delmont TO, Gaia M, Hinsinger DD, Fremont P, Vanni C, Guerra AF, Eren AM, Kourlaiev A, d'Agata L, Clayssen Q, et al** (2020) Functional repertoire convergence of distantly related eukaryotic plankton lineages revealed by genome-resolved metagenomics. *bioRxiv* 2020.10.15.341214

**Eilertsen H, Sandberg S, Tøllefsen H** (1995) Photoperiodic control of diatom spore growth; a theory to explain the onset of phytoplankton blooms. *Mar Ecol Prog Ser* 116: 303–307

**Esposito S, Botte V, Iudicone D, Ribera d'Alcala M** (2009) Numerical analysis of cumulative impact of phytoplankton photoresponses to light variation on carbon assimilation. *J Theor Biol* 261: 361–371

**Field CB, Behrenfeld MJ, Randerson JT, Falkowski PG** (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281: 237–240

**Fortunato AE, Jaubert M, Enomoto G, Bouly J-P, Raniello R, Thaler M, Malviya S, Bernardes JS, Rappaport F, Gentili B** (2016) Diatom phytochromes reveal the existence of far-red-light-based sensing in the ocean. *Plant Cell* 28: 616–628

**Froehlich AC, Noh B, Vierstra RD, Loros J, Dunlap JC** (2005) Genetic and Molecular Analysis of phytochromes from the filamentous fungus *Neurospora crassa*. *Eukaryot Cell* 4: 2140–2152

**Giraud E, Fardoux J, Fourrier N, Hannibal L, Genty B, Bouyer P, Dreyfus B, Verméglio A** (2002) Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria. *Nature* 417: 202–205

**Giraud E, Lavergne J, Verméglio A** (2010) Characterization of Bbacteriophytochromes from photosynthetic bacteria. *Methods Enzymol.* Elsevier, pp 135–159

**Goessling JW, Su Y, Cartaxana P, Maibohm C, Rickelt LF, Trampe ECL, Walby SL, Wangpraseurt D, Wu X, Ellegaard M, et al** (2018) Structure-based optics of centric diatom frustules: modulation of the *in vivo* light field for efficient diatom photosynthesis. *New Phytol* 219: 122–134

**Hartmann KM** (1966) A general hypothesis to interpret ‘high energy phenomena’ of photomorphogenesis on the basis of phytochrome. *Photochem Photobiol* 5: 349–365

**Holtrop T, Huisman J, Stomp M, Biersteker L, Aerts J, Grébert T, Partensky F, Garczarek L, Woerd HJ van der** (2021) Vibrational modes of water predict spectral niches for photosynthesis in lakes and oceans. *Nat Ecol Evol* 5: 55–66

**Hughes J, Lamparter T, Mittmann F, Hartmann E, Gärtner W, Wilde A, Börner T** (1997) A prokaryotic phytochrome. *Nature* 386: 663–663

**Hughes RM, Vrana JD, Song J, Tucker CL** (2012) Light-dependent, dark-promoted interaction between *Arabidopsis* cryptochrome 1 and phytochrome B proteins. *J Biol Chem* 287: 22165–22172

**Jaubert M, Duchene C, Kroth PG, Rogato A, Bouly J-P, Falciatore A** (2022 in press) Sensing and signalling in diatom responses to abiotic cues. Chap 21, In *The Molecular Life of Diatoms*, Falciatore & Mock Eds. Springer. doi: 10.1007/978-3-030-92499-7\_21

**Karniol B, Wagner JR, Walker JM, Vierstra RD** (2005) Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors. *Biochem J* 392: 103–116

**Kazarinova-Fukshansky N, Seyfried M, Schäfer E** (1985) Distortion of action spectra in photomorphogenesis by light gradients within the plant tissue. *Photochem Photobiol* 41: 681–688

**Kirk JTO** (2011) *Light and photosynthesis in aquatic ecosystems*, 3rd Edition. Camb Univ Press. doi: 10.1017/S0025315400044180

**Klose C, Venezia F, Hussong A, Kircher S, Schäfer E, Fleck C** (2015) Systematic analysis of how phytochrome B dimerization determines its specificity. *Nat Plants* 1: 15090

**Kusuma P, Bugbee B** (2021) Improving the predictive value of phytochrome photoequilibrium: consideration of spectral distortion within a leaf. *Front Plant Sci* 12: 596943

**Lamparter T, Michael N, Mittmann F, Esteban B** (2002) Phytochrome from *Agrobacterium tumefaciens* has unusual spectral properties and reveals an N-terminal chromophore attachment site. *Proc Natl Acad Sci* 99: 11628–11633

**Madronich S** (1987) Photodissociation in the atmosphere: 1. Actinic flux and the effects of ground reflections and clouds. *J Geophys Res* 92: 9740

**Malviya S, Scalco E, Audic S, Vincent F, Veluchamy A, Poulain J, Wincker P, Iudicone D, de Vargas C, Bittner L, et al** (2016) Insights into global diatom distribution and diversity in the world's ocean. *Proc Natl Acad Sci* 113: E1516–E1525

**Mancinelli AL** (1994) The physiology of phytochrome action. *In* RE Kendrick, GHM Kronenberg, eds, *Photomorphogenesis Plants*. Springer Netherlands, Dordrecht, pp 211–269

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

**Mobley CD** (1994) Light and water: radiative transfer in natural waters. Academic Press, San Diego

**Möglich A, Yang X, Ayers RA, Moffat K** (2010) Structure and function of plant photoreceptors. *Annu Rev Plant Biol* 61: 21–47

**Montgomery BL, Lagarias JC** (2002) Phytochrome ancestry: sensors of bilins and light. *Trends Plant Sci* 7: 357–366

**Morel A, Maritorena S** (2001) Bio-optical properties of oceanic waters: A reappraisal. *J Geophys Res Oceans* 106: 7163–7180

**Morgan DC, Smith H** (1976) Linear relationship between phytochrome photoequilibrium and growth in plants under simulated natural radiation. *Nature* 262: 210–212

**Moum JN, Smyth WD** (2019) Upper Ocean Mixing. *Encycl. Ocean Sci.* Elsevier, pp 71–79

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

**Olson EJ, Tzouanas CN, Tabor JJ** (2017) A photoconversion model for full spectral programming and multiplexing of optogenetic systems. *Mol Syst Biol* 13: 926

**Pesant S, Not F, Picheral M, Kandels-Lewis S, Le Bescot N, Gorsky G, Iudicone D, Karsenti E, Speich S, Trouble R, et al** (2015) Open science resources for the discovery and analysis of Tara Oceans data. *Scientific Data* 2: 1–16

**Pierella Karlusich JJ, Ibarbalz FM, Bowler C** (2020) Phytoplankton in the *Tara* Ocean. *Annu Rev Mar Sci* 12: 233–265

**Pollak B, Matute T, Nuñez I, Cerda A, Lopez C, Vargas V, Kan A, Bielinski V, von Dassow P, Dupont CL, et al** (2020) Universal loop assembly: open, efficient and cross-kingdom DNA fabrication. *Synth Biol* 5: ysaa001

**Quail PH** (1991) PHYTOCHROME: A Light-activated molecular switch that regulates plant gene expression. *Annu Rev Genet* 25: 389–409

**Ren H, Junmin L, Qiuqi L, Boping H** (2005) Study of light signal receptor of *Stephanopyxis palmeriana* during sexual reproduction. *Chin J Oceanol Limnol* 23: 330–334

**Rockwell NC, Duanmu D, Martin SS, Bachy C, Price DC, Bhattacharya D, Worden AZ, Lagarias JC** (2014) Eukaryotic algal phytochromes span the visible spectrum. *Proc Natl Acad Sci* 111: 3871–3876

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

**Schmidt W, Marmé D, Quail P, Schäfer E** (1973) Phytochrome: First-order phototransformation kinetics *in vivo*. *Planta* 111: 329–336

**Shinomura T, Nagatani A, Hanzawa H, Kubota M, Watanabe M, Furuya M** (1996) Action spectra for phytochrome A- and B-specific photoinduction of seed germination in *Arabidopsis thaliana*. *Proc Natl Acad Sci* 93: 8129–8133

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

**Studier FW** (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* 41: 207–234

**Usami T, Mochizuki N, Kondo M, Nishimura M, Nagatani A** (2004) Cryptochromes and phytochromes synergistically regulate Arabidopsis root greening under blue light. *Plant Cell Physiol* 45: 1798–1808

**Veedin Rajan VB, Häfker NS, Arboleda E, Poehn B, Gossenreiter T, Gerrard E, Hofbauer M, Mühlestein C, Bileck A, Gerner C, et al** (2021) Seasonal variation in UVA light drives hormonal and behavioural changes in a marine annelid via a ciliary opsin. *Nat Ecol Evol* 5: 204–218

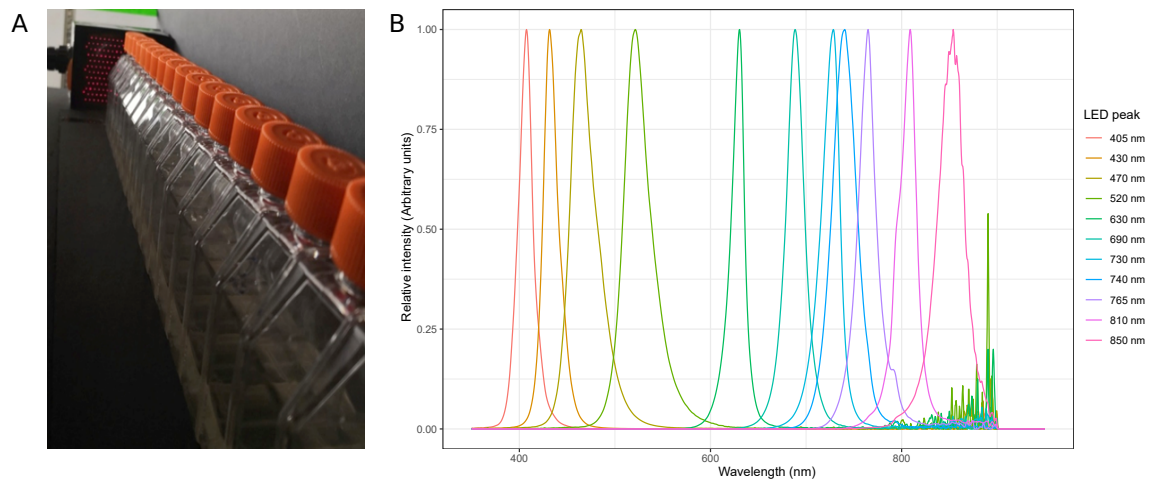


**Yeh K-C, Wu S-H, Murphy JT, Lagarias JC (1997)** A cyanobacterial phytochrome two-component light sensory system. *Science* 277: 1505–1508

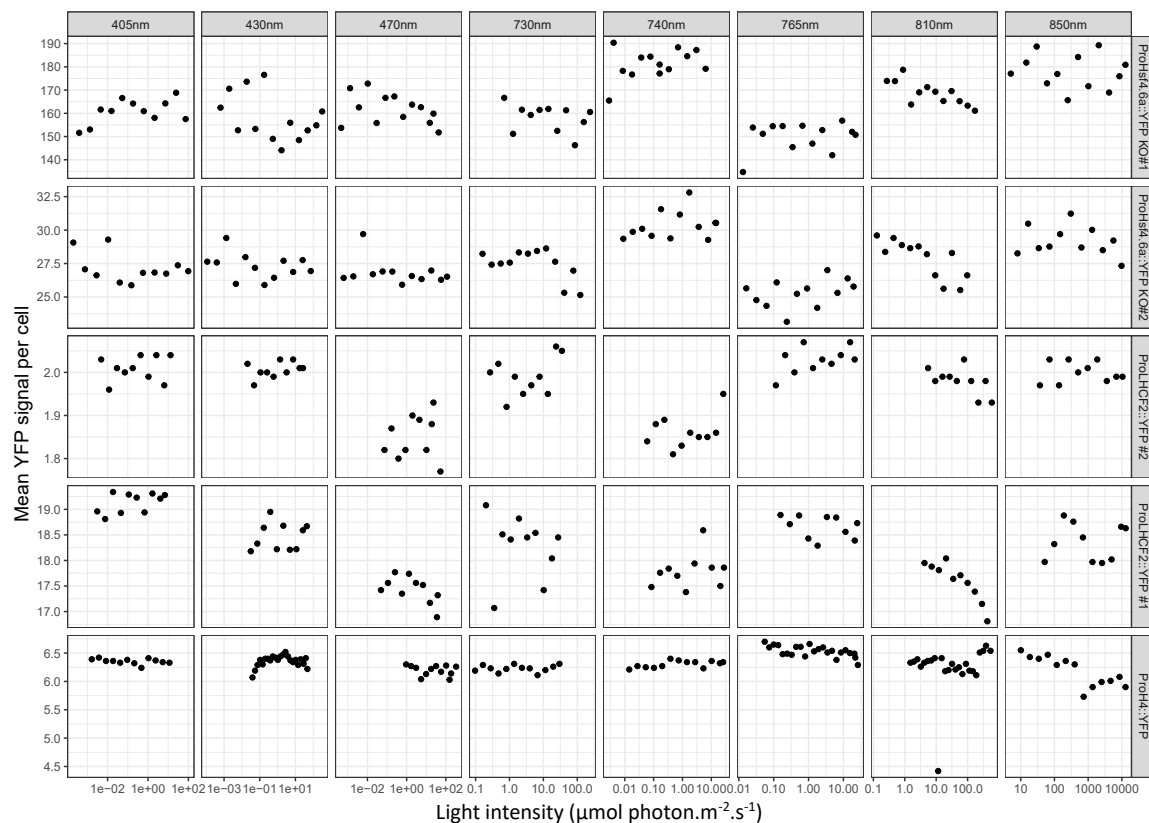
**Yu Z, Fischer R (2019)** Light sensing and responses in fungi. *Nat Rev Microbiol* 17: 25–36

**Yu Z, Streng C, Seibeld RF, Igbalajobi OA, Leister K, Ingelfinger J, Fischer R (2021)** Genome-wide analyses of light-regulated genes in *Aspergillus nidulans* reveal a complex interplay between different photoreceptors and novel photoreceptor functions. *PLOS Genet* 17: e1009845

## SUPPLEMENTARY FIGURES

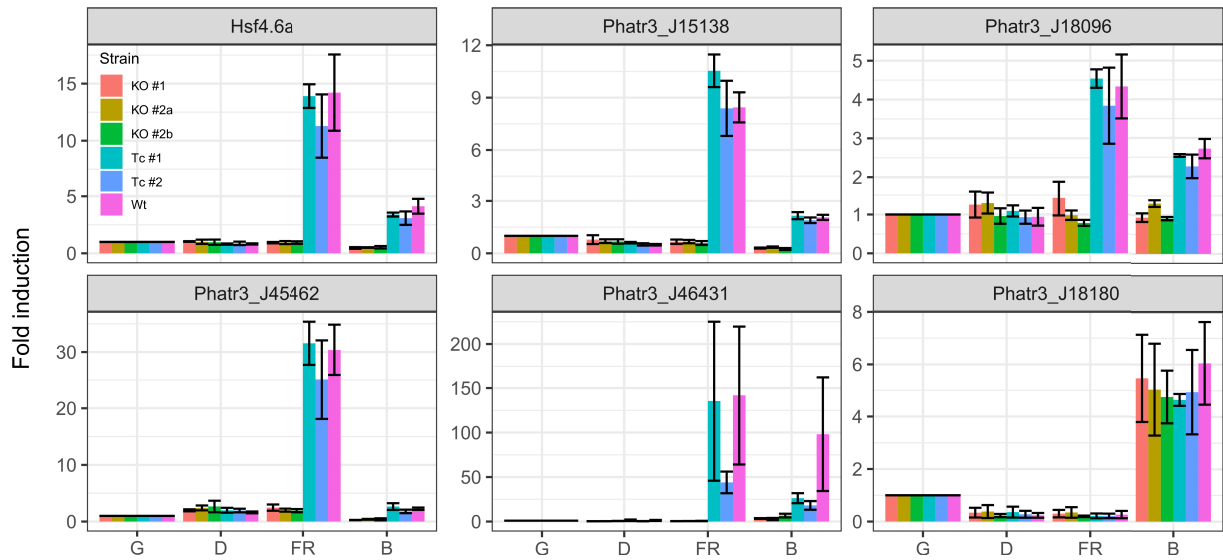


Supplementary Figure 1 : Experimental setup for the action and inhibition spectra. A. Illustration of an illumination of reporter lines to a gradient of a monochromatic LED. B. Spectra of the different LEDs used for the action and inhibition spectra.

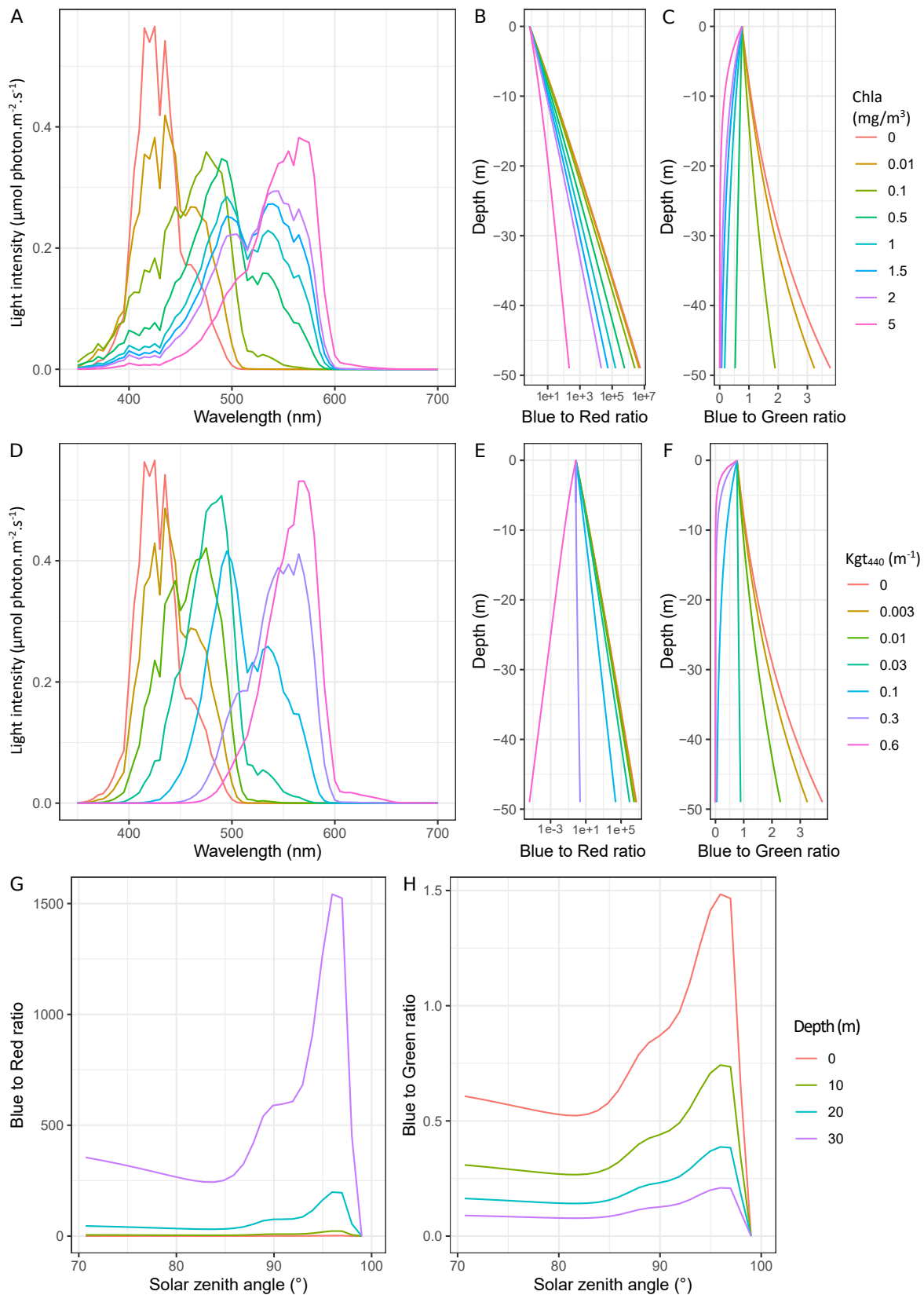


Supplementary Figure 2 : Control lines show no induction of YFP in the action spectra. Control lines correspond to: PtDPH KO background transformed with the ProHsf4.6a:: YFP construct, or WT background transformed with a DPH-independent promoter controlling eYFP expression, ProLHCF2 or ProH4 (Siaut et

al, 2007)). Cells were exposed to a gradient of monochromatic LED. YFP signal was measured by flow cytometry.

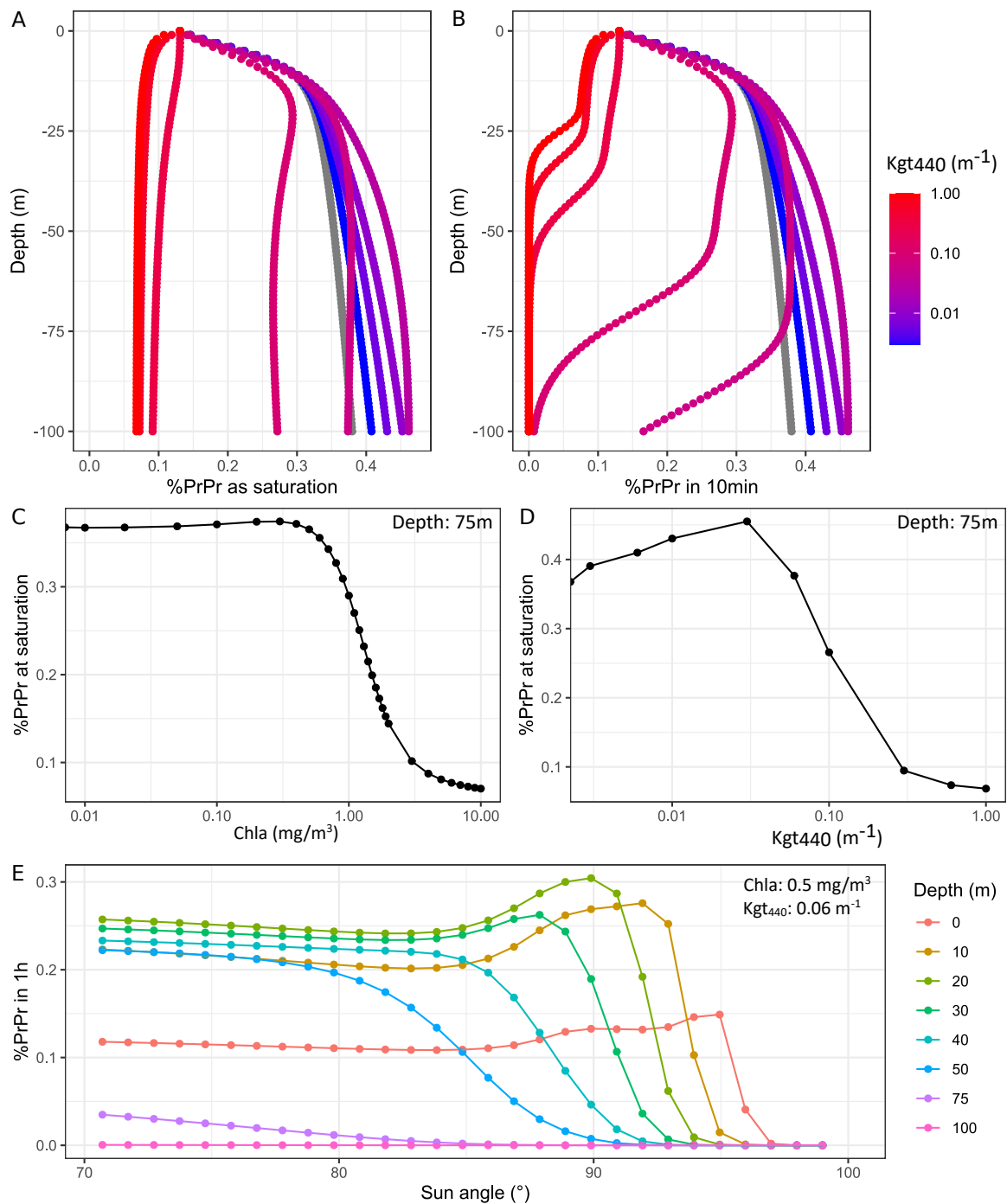


Supplementary Figure 3 : Expression analysis of PtDPH-regulated (Hsf4.6a, Phatr3\_J15138, Phatr3\_J18096, Phatr3\_J45662, Phatr3\_J46431) and PtDPH-non regulated (Phatr3\_J18180) genes upon far-red and blue irradiation. *P. tricornutum* WT, DPH KO mutant and their corresponding non mutated transformant (Tc) cells were collected in their continuous green light growth condition (G) and following a 30 min of irradiation with 470-nm (B) or 765-nm (FR) LED ( $15 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) or kept in the dark for the same time (D). Gene expression quantifications were performed by RT-qPCR, with *H4* used as normalizer gene and relativized to the G conditions. Values are the mean +/- SD on 3 independent biological replicates.

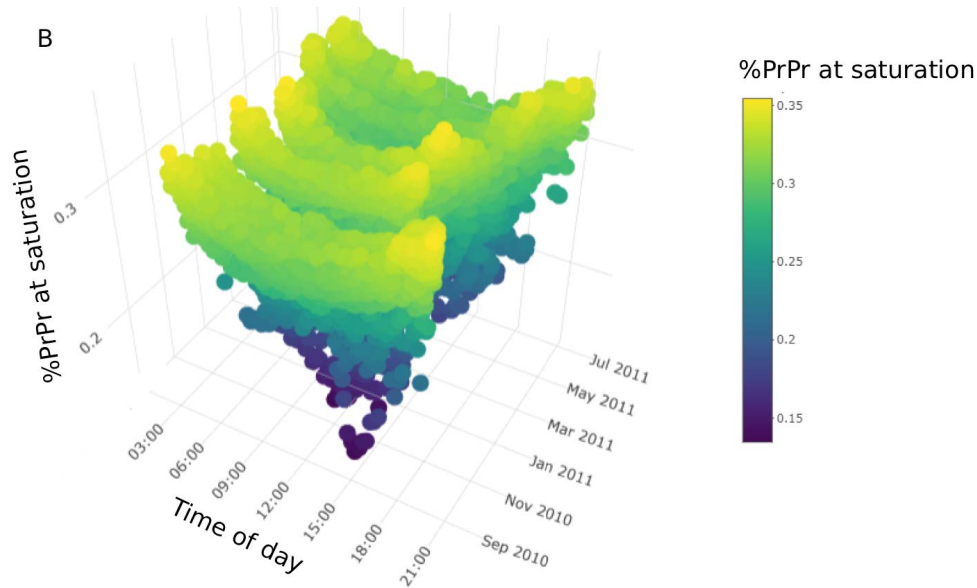


Supplementary Figure 4. Light modeling examples. A and D, Modelled light fields at the bottom of the photic zone (1% of surface irradiance) for sun at Zenith, for different Chl a concentrations (K<sub>bg</sub>=0) (A), and different K<sub>gt440</sub> (Chl a=0) (D). B,C,E,F,G and H, variation of the ratio of blue (400-450nm) to green (500-

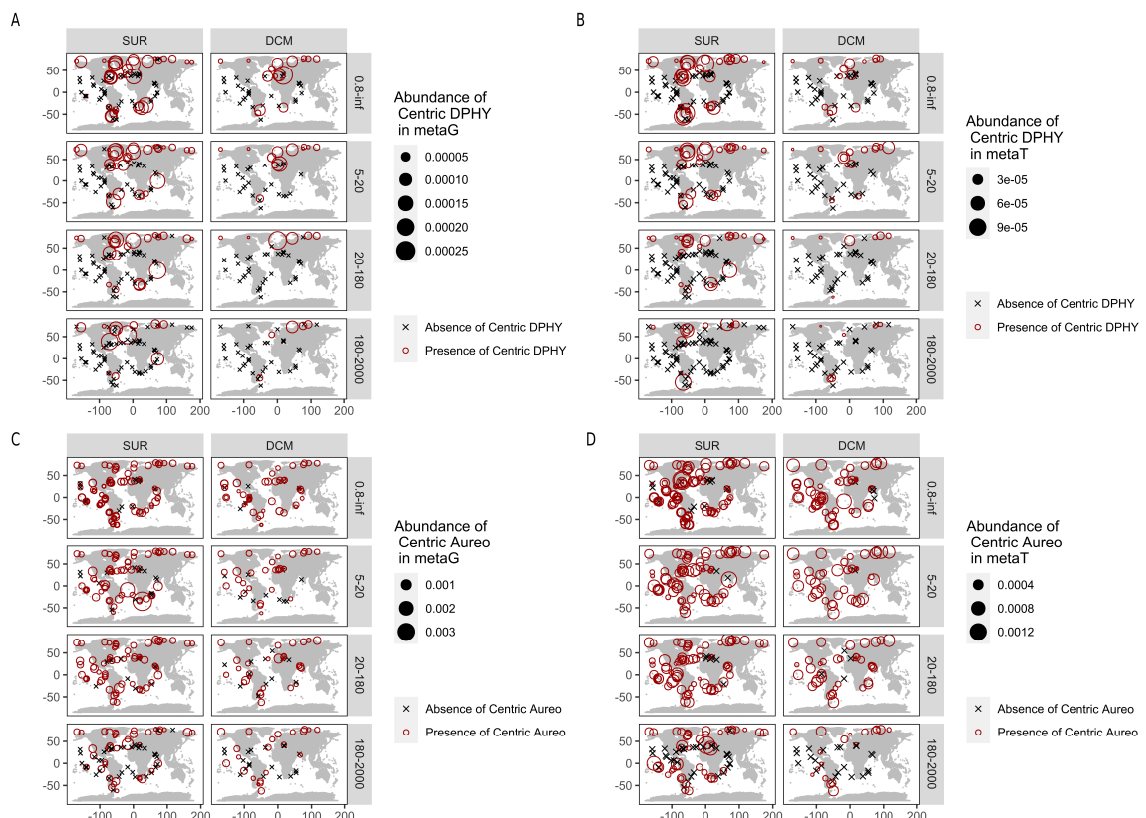
550nm)(C, F, H) or red (630-680nm) 'B, E, G) bands in different modelled light fields: B and C, ratios from light fields showed in A; E and F, ratios from light fields showed in D; G and H ratios from the light fields at different solar zenith angles as selected depths, for Chl a=0.5 mg/m<sup>3</sup> and Kgt<sub>440</sub>=0.06m<sup>-1</sup>



Supplementary Figure 5 : DPH activity as proportion of PrPr formed (%PrPr) in different modelled environmental conditions: at saturation (A) and upon 10 min of exposure (B), depending on the depth and different Kgt<sub>440</sub>; at saturation at 75-m depth depending on the Chl a concentration (C) and Kgt<sub>440</sub> (D); upon 1h of exposure of light spectrum at different depths and depending on the sun Zenith angle (E).

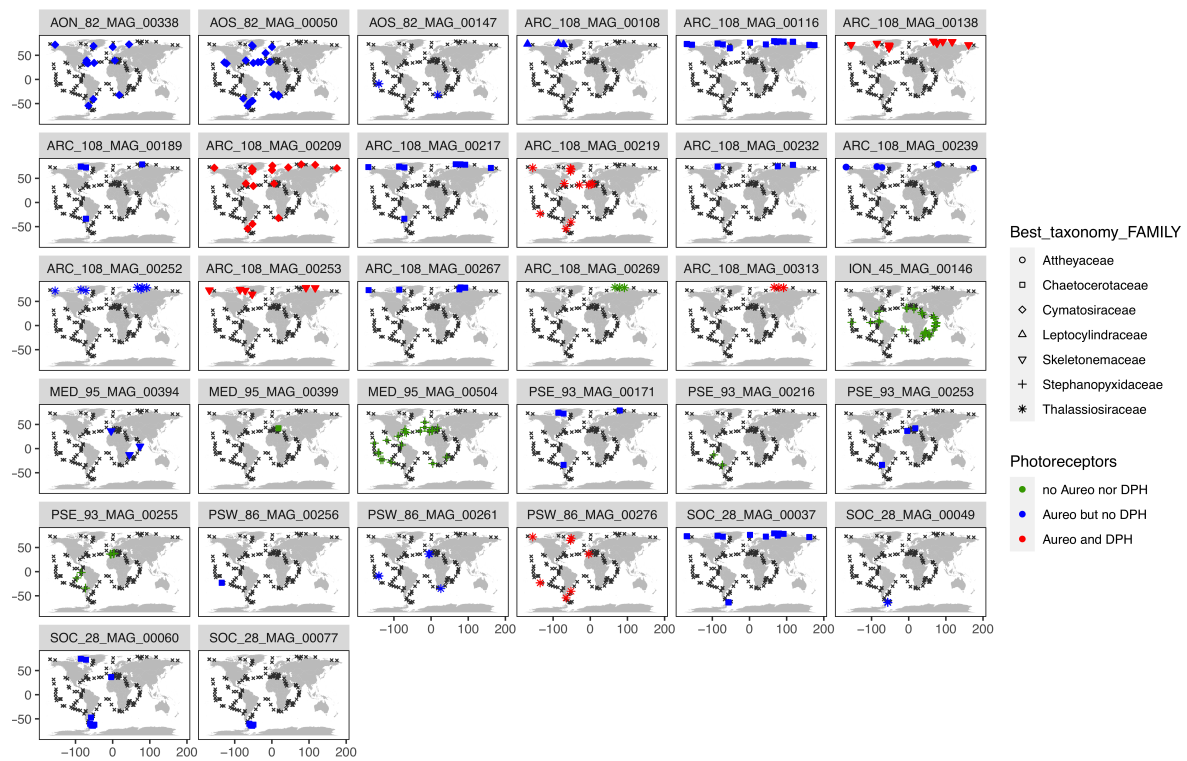


Supplementary Figure 6: DPH activity as proportion of PrPr formed (%PrPr) in different *in situ* measured environmental light fields as a function of the time of the day over a year time series. Data from Veedin Rajan, 2021, with a downwelling irradiance spectra taken every 15 min at 10 m-depth near Ischia (Gulf of Naples, Italy) in 2010-2011. DPH activity at saturation was calculated.

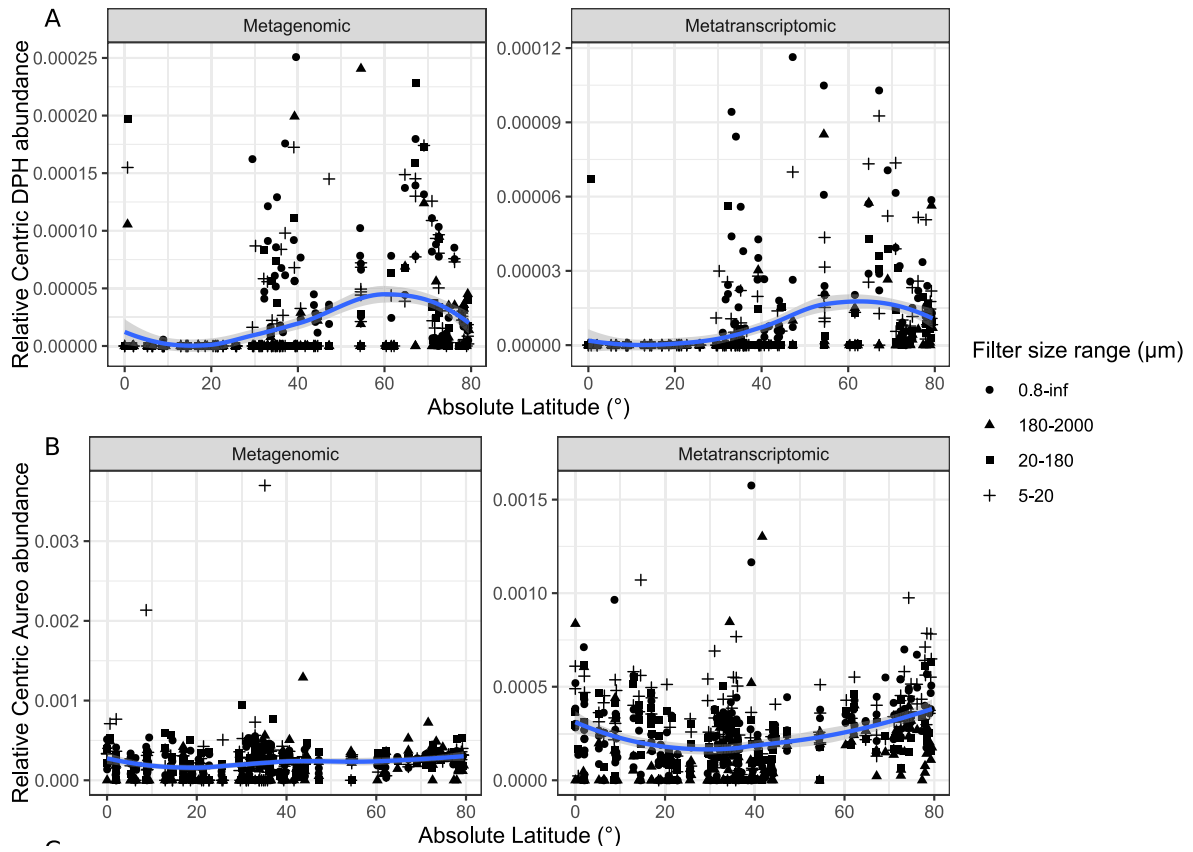


Supplementary Figure 7: Biogeography of DPH and Aureochromes (Aureo) from centric diatoms, per filter size, in metagenomic and metatranscriptomic data from *Tara* Oceans. Crosses indicate that no gene was

found at the station, while circles indicate presence. Circles size indicate the abundance of the DPH or Aureo genes relative to all centric genes.

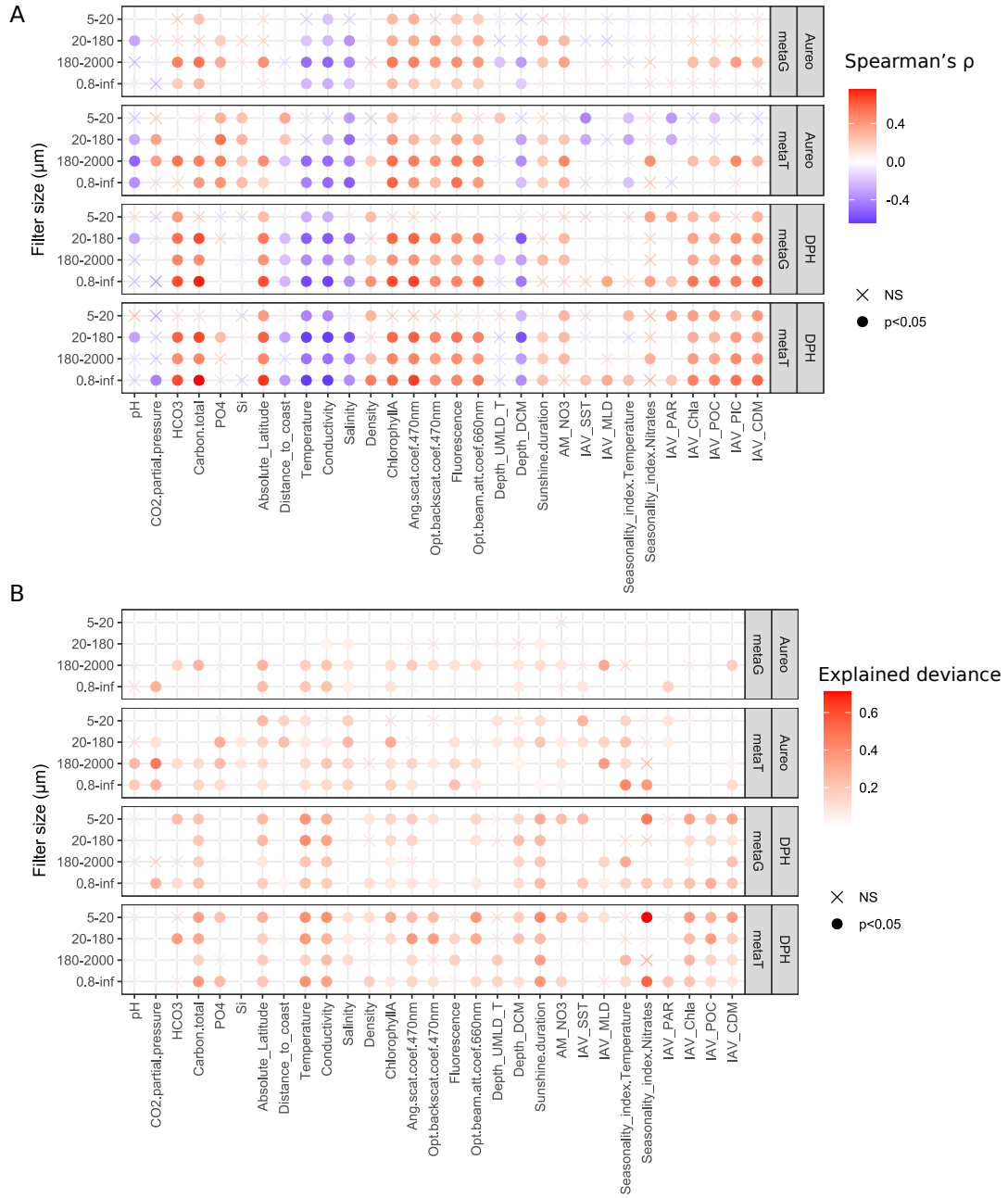


Supplementary Figure 8: Metagenome Assembled Genomes (MAGs) geographical distribution of centric diatoms. Black crosses indicate *Tara* Oceans station where the MAG was not found present; colored points indicate the photoreceptor content of the MAG, while the shape indicate the centric diatom family the MAG belongs too.

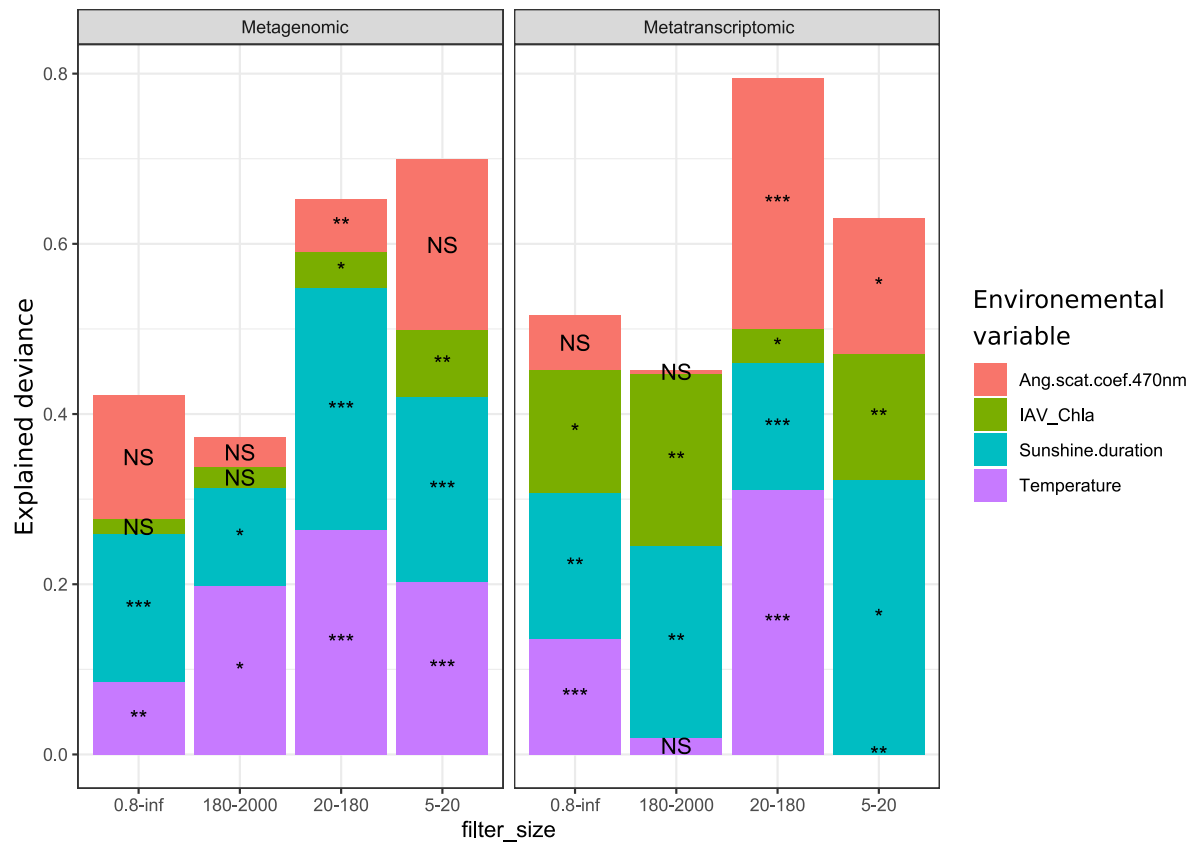


Supplementary Figure 9: DPH relative abundance (A) compared to Aureochrome (B) relative abundance as a function of absolute latitude. Point shapes indicate the filter size range.





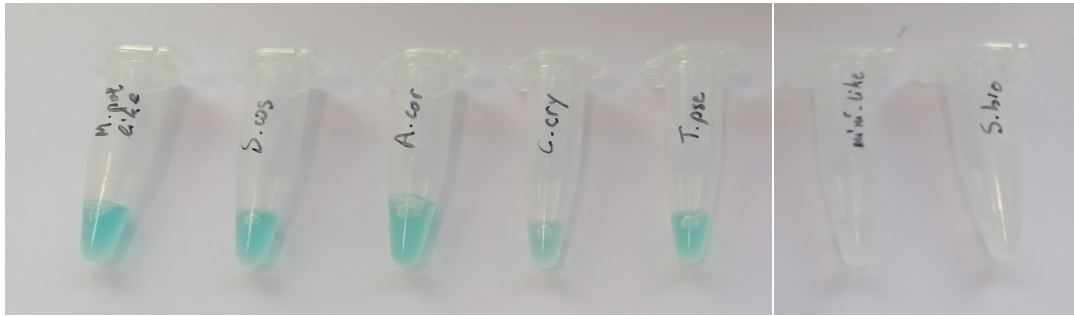
Supplementary Figure 10 : Correlation (A) and individual generalized additive models (B) analysis for DPH and Aureochromes (Aureo) relative abundance in metagenomic and metatranscriptomic data from *Tara* Oceans, separated by filter sizes.



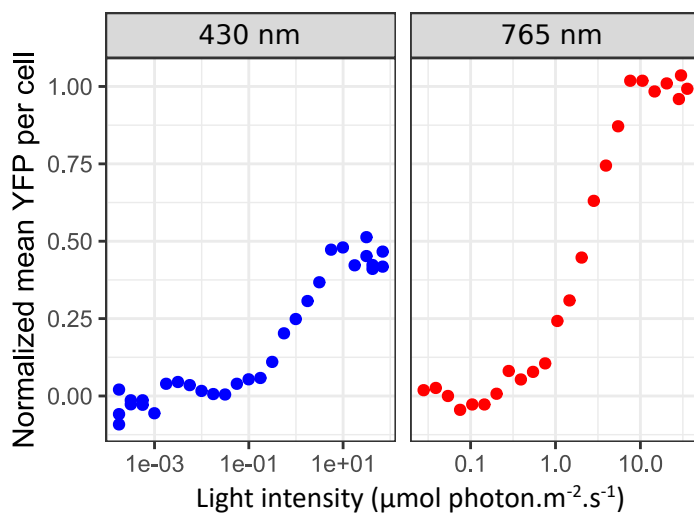
Supplementary Figure 11 : Complex generalized additive models explaining DPH relative abundance with a combination of environmental parameters for each filter size.

PAS					GAF				
Atha_PhyA	PVVENQPPRS	DKVTTTTLHH	IQKGKLIQPF	GCL	KLSFDLTLCG	STLRAPHSC	LQYMANMDSI	ASLVM	
Syne_Cph1	TTVQLSDQSL	RQLETLAIHT	AHL---IQPH	GLV	STNRAVDLTE	SILRSAYHC	LYLKNMGV	ASLTI	
tr Q7S5Q3	GVIQKARRAF	TTCDDEEPIHI	PGA---IQSY	GML	DMDVPLDLTH	AYLRAMSPVH	LKYLNSMGVR	SSMSM	
Drad_Bph	LYLGGPEITT	ENCEREPIHI	PGS---IQPH	GAL	QTNAPTPLGG	AVLRATSPMH	MQYLRNMGV	SSLSV	
Tp	TSPSEKRDVP	TECEKEPIHI	INR---IQPC	GYL	KELGYTDLSM	CRLRGSSFVH	LKYLKNMGVT	STMVI	
Mpol-like_(+1)	SNDYSKQNV	LQCEKEPIHI	VNR---IQPC	GTL	GGEHYTDLSM	CRLRGSSVYH	LQYLRNMGVT	STLVI	
Acor_(+1)	SNDYSKQNV	LQCEKEPIHI	VNR---IQPC	GTL	GGEHYTDLSM	CRLRGSSVYH	LQYLRNMGVT	STLVI	
Mini-like_(+1)	ETKAADMTPA	LSCEREPIHI	VNR---IHPC	GNL	ETTKYTDLSM	CRLRGSSH IH	LEYLRNMV	STLVI	
Scos_(+1)	NEDYARQVTP	LSCEREPIHI	VNR---IQPC	GTL	ESKKYTDLSM	CRLRGSSYIH	LDYLRNMGVT	SSLVI	
Sbio_(+1)	GADYAPQVTP	LSCEREPIHI	VNR---IQPC	GTL	EPEKYTDLSM	CRLRGSSYIH	LEYLRNMGVT	STLAI	
Ccry_(+1)	SADYAVQVTP	LSCAREPIHI	VNR---IQPC	GTL	GEEKYTDLSM	CRLRGSSYIH	LEYLRNMGVT	SSLVI	
Ptri_(+1)	HNNSITTKEL	TECDREPVHL	IAN---VQGG	TGH	LDNEKMDLSQ	IRMRAVAKPH	IYVLRNMGV	SSLSL	

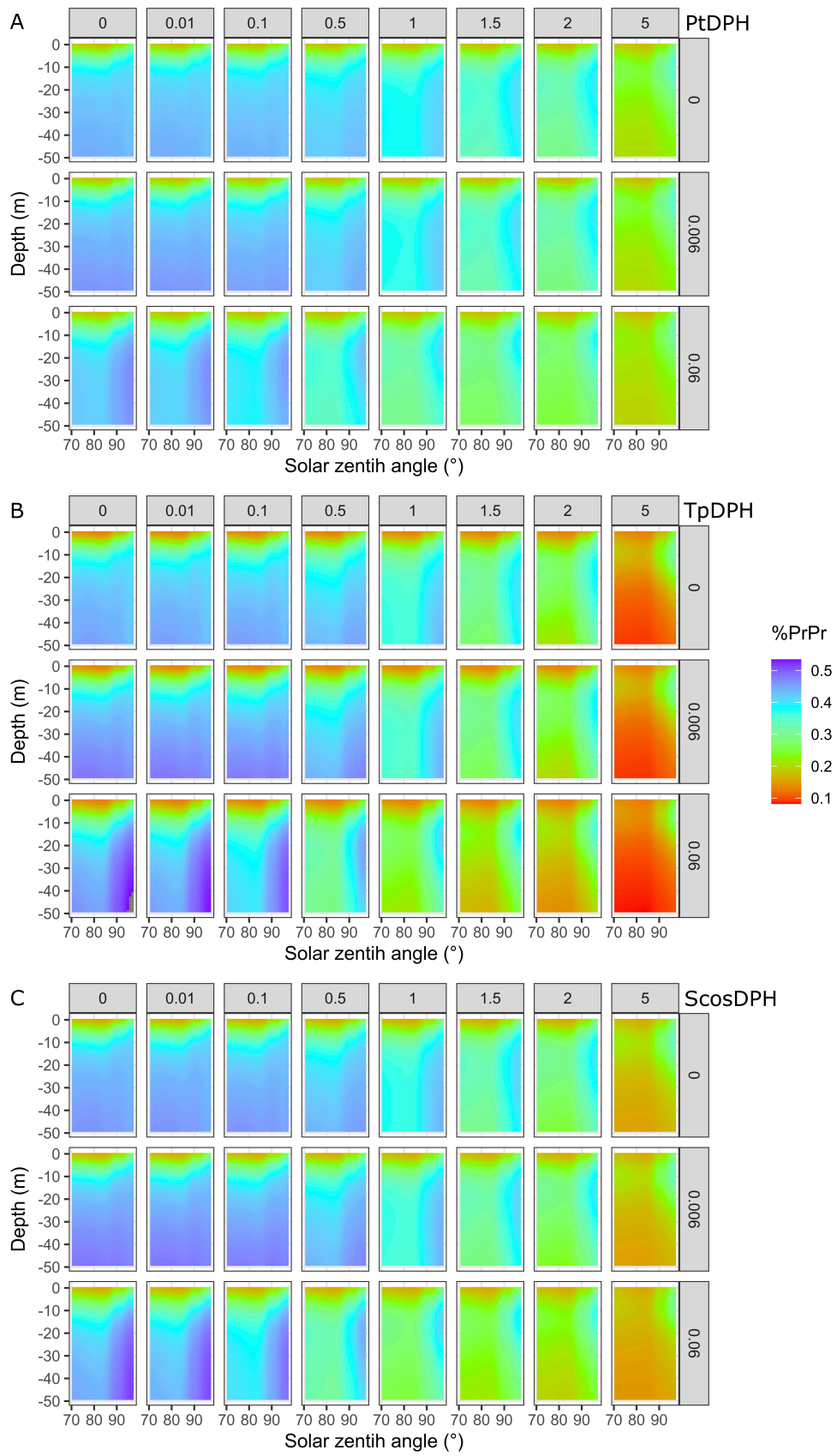
Supplementary Figure 12 : Alignment of a portion of the PAS and GAF domains of selected DPH and reference phytochrome sequences (*Arabidopsis thaliana*, Atha\_PhyA; *Synechocystis*, Syne\_Cph1; *Neurospora crassa* PHY2, tr-Q7S5Q3, and *Deinococcus radiodurans*, Drad\_Bph). In red is highlighted the cysteine residue conserved with the one used for chromophore binding in the PAS domain in bacterial and fungal phytochromes, and in the GAF domain, in plant and cyanobacteria phytochromes. Tp, *Thalassiosira pseudonana*; Mpol-like, *Minitocellus polymorphus*-like sequence from MATOU; Acor, *Arcocellulus cornucervis*, Mini-like, *Minidiscus*-like sequence from MATOU; Scos, *Skeletonema costatum*; Sbio, *Shionodiscus bioculatus*; Ccry, *Cyclotella cryptica*; Ptri, *Phaeodactylum tricornutum*, Acof, *Amphora coffeaeformis*.



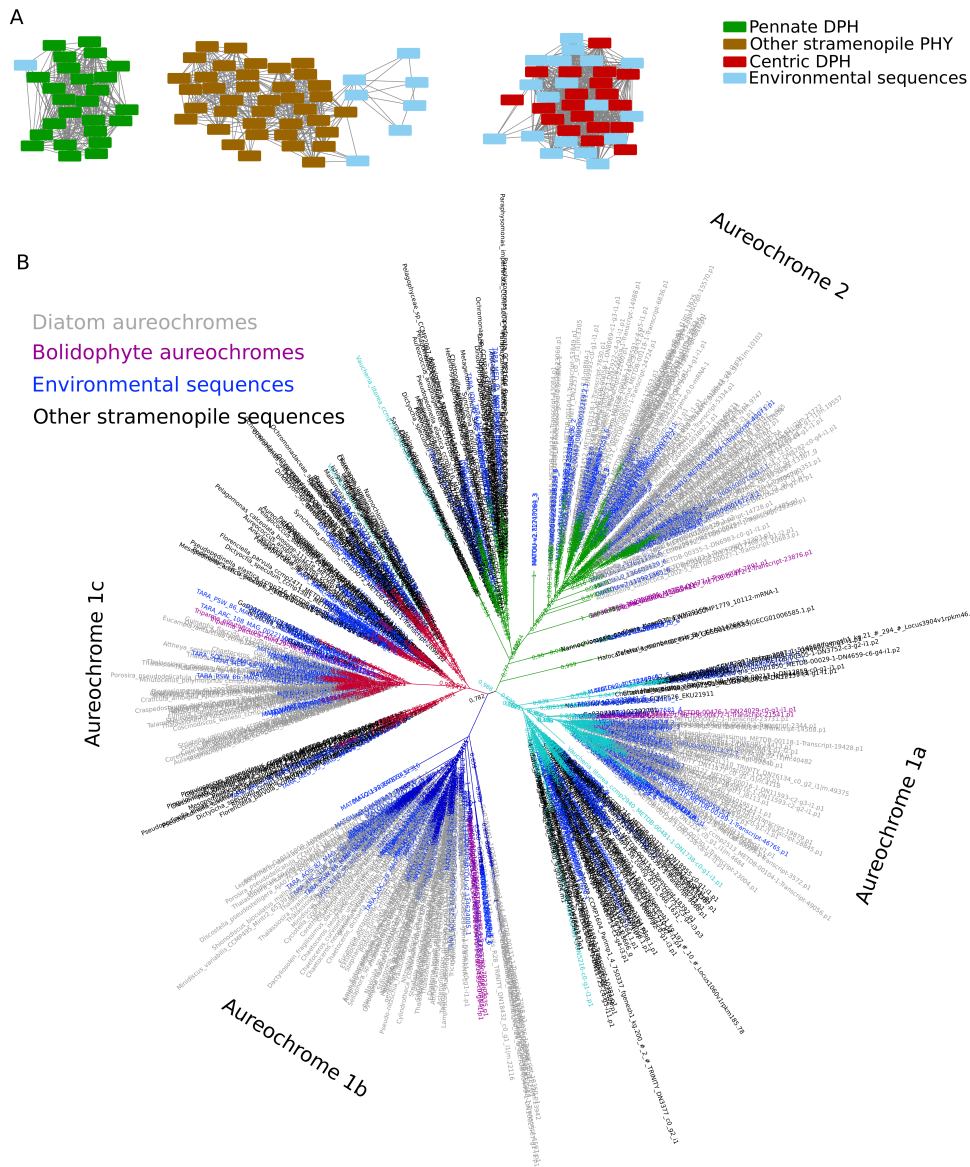
Supplementary Figure 13 : Purified recombinant DPH. Photosensor module of DPH genes expressed as a N-terminal 6 Histidine-tagged recombinant proteins along with *Synechocystis* Heme oxygenase gene encoding the enzyme allowing the synthesis of the biliverdin chromophore, were purified on Ni-NTA affinity chromatography. From left to right: *Minutocellus polymorphus*-like sequence; *Skeletonema costatum*; *Arcocellulus cornucervus*; *Cyclotella cryptica*; *Thalassiosira pseudonana*; *Minidiscus*-like sequence; *Shionodiscus bioculatus*.



Supplementary Figure14: TpDPH restores blue and far-red-dependent YFP induction in PtDPH KO line. PtDPH KO reporter strain (line KO #1 in Supplementary Figure 2) was transformed with TpDPH gene under the control of PtDPH promoter and terminator, and subjected to a gradient of blue (430 nm) or far-red (765) lights, according to the same experimental setup used for the action spectra.



Supplementary Figure 15 : Projection of DPH activity as proportion of PrPr formed (%PrPr), in modeled light fields, with variations of chlorophylla (Chl a), Kbg<sub>440</sub> and solar zenith angle, using properties of PtDPH (A), TpDPH (B) and ScosDPH (C), determined *in vitro*, from the recombinant protein absorption spectra. The color legend is common to all 3 figures.



Supplementary Figure 16 : Method for DPH and diatom Aureochrome gene search. A, Example of Sequence Similarity Network of the DPH REC domain, searched in Tara Oceans data showing a nice separation of the different phytochromes (centric diatom, pennate diatom, other stramenopiles) for an alignment score of 20. Environmental sequences grouping with the centric or pennate DPH will be annotated as such. B, Phylogenetic tree of diatom Aureochrome gene search. Diatom clades are identifiable (grey label), delimited by the Bolidophyceae (a sister picoplanktonic group of diatoms) Aureochromes (pink labels). Branches are colored by the different Aureochrome types : Aureochrome 1a, 1b, 1c and 2.

## SUPPLEMENTARY TABLES

Table S1. Species from which DPH have been characterized in this study and previous one\* (\*Fortunato et al,2016) and information about their isolation site.

Species	Strain	
<i>Phaeodactylum tricornutum</i> *	CCMP2561	North Atlantic, off Blackpool, England
<i>Thalassiosira pseudonana</i> *	CCMP1335	North Atlantic, Moriches Bay, Forge River, Long Island, New York USA
<i>Cyclotella cryptica</i>	CCMP332	North Atlantic, Martha's Vineyard, Massachusetts USA
<i>Minidiscus spinulatus</i>	RCC4659	Atlantic Ocean, English Channel, Brittany coast, pelagic
<i>Skeletonema costatum</i>	RCC1617	Atlantic Ocean, English Channel, Brittany coast, pelagic
<i>Arcocellulus cornucervis</i>	RCC2270	Arctic Ocean, Beaufort Sea, at 0m depth pelagic
<i>Shionodiscus bioculatus</i>	RCC1991	Arctic Ocean, Beaufort Sea, at 65m depth, pelagic,

Table S2 : Sequences of primers used in this study

qPCR-Phatr3_45662FW	CGAGGGAGCTCGGTTTATGG
qPCR-Phatr3_45662RV	TGATGGGAAGCTGTTCTGCCC
qPCR-Phatr3_46431FW	GGTTTGCGAGTGCATTTGGT
qPCR-Phatr3_45662RV	TGTCAGCAACCTCATTCCCC
qPCR-Phatr3_18180FW	CCGGAACGTAGGTTTGAT
qPCR-Phatr3_18180RV	CCGCGCCAACATAGCAAG
qPCR-H4FW	AGGTCCTTCGCGACAATATC
qPCR-H4RV	ACGGAATCACGAATGACGTT
<i>Hsf4.6ap_Fw0</i>	TCTAGAGAGCTCGGATTTTCAATCTGTTTTGGGCA
<i>Hsf4.6ap-YFP_RV</i>	GATATCGGATCCTGTCAAAGGTTTAAGAGAATCGGC
<i>YFP-Hsf4.6ap_Fw</i>	AAACCTTTGACAGGATCCGATATCATGGTGAGCAAGGGCGAG
<i>FcpAT_Rv0</i>	TCTAGATGAAGACGAGCTAGTGTATTCC
HO1xpET.HindIII.Fw	GCAAGCTTAGAAGGAGATATACATATGAG
HO1xpET.NotI.Rv	GCGCGCCCGCTAGCCTTCGGAGGTGGCGAG
TpPHY.D.Fw	AGGCTGTCTCGTCTCGTCTCAGGTCTCAAGGTATGAGTGTCAAAAAGAGCAC
TpPHY.Sap1.Rv	CAATGTGTTTTCGCATGAACAGCTCTCTCGCTTG
TpPHY.Sap1.Fw	CAAGCGAGAGAGCTGTTTCATGCGAAACACATTG
TpPHY.BsaI.Rv	CAAGCTGTTTATCAGGGTCACCACCCCACGTGACAC
TpPHY.BsaI.Fw	GTGTCACGTGGGGTGGTGACCCTGATAAACAGCTTG
TpPHY.Sap2.Rv	GAGATCGAACAAGTGCTCCTCAATCTGAAGGTCTGATG
TpPHY.Sap2.Fw	CATACGACCTTCAGATTGAGGAGCACTTGTTCGATCTC
TpPHY.E.Rv	TGGTAATCTATGTATCCTGTTGGTCTCTAAGCTCATCGTTTCATTTTTGTGAT

PrPtPHY.A.Fw	GGCTGTCTCGTCTCGTCTCAGGTCTCAGGAGCCCGGGGATATCGAAGATCC
PrPtPHY.C.Rv	TGGTAATCTATGTATCCTGGTGGTCTCGCATTTTTAAAGGCGTGGTTCCTTG
TrPtPHY.E.Fw	TCGTCTCGTCTCAGGTCTCAGTTTCATGGTCGTTTCATTCATAGAAG
TrPtPHY.F.Rv	TGGTAATCTATGTATCCTGGTGGTCTCAAGCGCGCTCTTTCCACCTCATCTC
UNS1FL.FW	CATTACTCGCATCCATTCTCAGGCTGTCTCGTCTCGTCTC
UNSXFL.RV	GGTGGAAGGGCTCGGAGTTGTGGTAATCTATGTATCCTGG
SbioPHY.NheI.Fw	GCGGCTAGCATGTCTGCCAGTTCACCAC
SbioPHY_PCD.SalI.Rv	GCGGTCGACCTAAAGATTTTCTTTTGGATCTTTG
AcorPHY.NheI.Fw	GCGGCTAGCATGTCCGCACCTGCGGCAGC
AcorPHY_PCD.SacI.Rv	GCGGAGCTCCTAGTAGCTTGTGTGTTCTCGTC
MspiPHY.NheI.Fw	GCGGCTAGCATGACCTCCTCCTCAACCAAC
MspiPHY_PCD.SacI.Rv	GCGGAGCTCCTACTTCTGATCGGCAATCAATTC
ScosPHY.SpeI.Fw	GCGACTAGTATGTCGTCCACCAATAGCAC
ScosPHY_PCD.SacI.Rv	GCGGAGCTCCTAGAGGTTTTCTTTTGGATCTTG
CcryPHY.NheI.Fw	GCGGCTAGCATGGCAGCACCCACAAAAC
CcryPHY_PCD.SacI.Rv	GCGGAGCTCCTACTTCTGATCTTTAATCAAATC

Table S3. SSN parameters for DPH and diatom Aureochrome search in *Tara* Oceans data

Domain model	SSN alignment score threshold
Centric GAF	30
Centric PHY	10
Centric HisKA	20
Centric HATPase	20
Centric REC	20
Pennate GAF	30
Pennate PHY	10
Pennate HisKA	20
Pennate HATPase	20
Pennate REC	20
Diatom Aureochrome	70

## SUPPLEMENTARY DATA

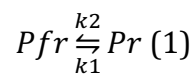
Supplementary Data : List of MATOU2 genes identified as diatom aureochromes and diatom phytochromes. Available at

<https://mycore.core-cloud.net/index.php/s/hD5zBsVlyhGkvfH>

## SUPPLEMENTAL METHOD

### Detailed description of DPH activity model construction based on equations described in (Mancinelli, 1994)

Phytochrome equilibrium can be theorized as the equilibrium between the Pfr→Pr transition (kinetic constant  $k_1$ ) and the Pr→Pfr reverse reaction (kinetic constant  $k_2$ ) with the following equation:



The rate of phytochrome photoconversion depends on phytochrome extinction coefficient, photoconversion yield and on light intensity, so  $k_1$  and  $k_2$  can be expressed as:

$$k_1 = N * \sigma_{Pfr} = N * 2.3 * \epsilon_{Pfr} * \phi_{Pfr} \text{ and } k_2 = N * 2.3 * \epsilon_{Pr} * \phi_{Pr}$$

where  $N$  is the fluence rate, or light intensity (mol photon/m<sup>2</sup>/s), and  $\sigma_{Pfr}$  is the photoconversion cross-section of the Pfr→Pr reaction,  $\epsilon_{Pfr}$  is the extinction coefficient of Pfr and  $\phi_{Pfr}$  the quantum yield of the Pfr→Pr reaction;  $\sigma_{Pr}$  is the photoconversion cross-section of the Pr→Pfr reaction,  $\epsilon_{Pr}$  is the extinction coefficient of Pr and  $\phi_{Pr}$  the quantum yield of the Pr→Pfr reaction.

The photoconversion rate of phytochrome can be expressed as the sum of  $k_1$  and  $k_2$  and expressed as  $k$ .

Considering that the total amount of phytochrome  $P$  stays constant, i.e. synthesis and degradation are negligible, then  $P = P_r + P_{fr}$

From equation 1, the rate of Pr production can be expressed as :

$$\frac{dPr}{dt} = k_1 * Pfr - k_2 * Pr \quad (2)$$



At equilibrium,  $\frac{dPr}{dt} = 0$  and  $\%Pr = \frac{Pr}{P} = \frac{k_1}{k_1+k_2} = \frac{1}{1+\frac{\epsilon_{Pr}*\phi_{Pr}}{\epsilon_{Pfr}*\phi_{Pfr}}} = \frac{1}{1+\frac{\epsilon_{Pr}*\eta}{\epsilon_{Pfr}}}$

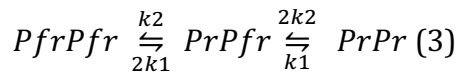
With  $\eta = \frac{\phi_{Pr}}{\phi_{Pfr}}$  the ratio of quantum yields.

Integration of equation (2) also gives:

$$\%Pr(t) = \left( \%Pr_0 - \frac{k_1}{k_1+k_2} \right) e^{-(k_1+k_2)t} + \frac{k_1}{k_1+k_2}$$

with  $\%Pr_0$  the percent of Pr at time  $t=0$ .

If one considers that phytochromes is in a dimeric form (Klose et al., 2015; Brockmann et al.), they exist in equilibrium between homodimer PrPr, homodimer PfrPfr and the intermediate heterodimer PfrPr. The rate of conversion of each monomer is supposed independent of the dimer state, and each monomer photoconverts as in the monomer model explained previously, and can be expressed as:



We always consider that synthesis and degradation are negligible, with

$$P = PrPr + PrPfr + PfrPfr$$

Considering PrPr as the active form and total amount of phytochrome constant, the PrPr photoconversion rate can be expressed as:

$$\frac{dPrPr}{dt} = k_1 * PrPfr - 2k_2 * PrPr$$

At equilibrium,  $\frac{dPrPr}{dt} = 0$  and :

$$\%PrPr_{eq} = \left( \frac{k_1}{k_1+k_2} \right)^2 = \frac{1}{\left(1+\frac{\epsilon_{Pr}*\phi_{Pr}}{\epsilon_{Pfr}*\phi_{Pfr}}\right)^2} = \frac{1}{\left(1+\frac{\epsilon_{Pr}*\eta}{\epsilon_{Pfr}}\right)^2} \quad (4)$$

In this latter equation,  $\varepsilon_{Pr}/\varepsilon_{Pfr}$  can be known from the recombinant PtDPH absorption spectra.

The  $\varepsilon_{Pr}/\varepsilon_{Pfr}$  ratio, in both monochromatic and bichromatic illumination, as been determined with the LED spectra used (and not only the  $\varepsilon_{Pr}$  value at the peak of intensity of the LED). More exactly, we express  $\varepsilon_{Pr}/\varepsilon_{Pfr}$  as:

$$\frac{\varepsilon_{Pr,LED}}{\varepsilon_{Pfr,LED}} = \frac{\sum_{\lambda=300}^{\lambda=900} A_{Pr,\lambda} * N_{\lambda}}{\sum_{\lambda=300}^{\lambda=900} A_{Pfr,\lambda} * N_{\lambda}}$$

where  $A_{Pr,\lambda}$  is the absorbance of Pr at wavelength  $\lambda$  and  $N_{\lambda}$  the intensity measured at this wavelength.

%PrPr from equation (4) can be estimated from the eYFP signal at saturating light at each wavelength. Our output measure is the eYFP signal for a LED at wavelength  $\lambda$  at saturating intensities (the 4 points at highest light intensity were used), normalized between the signal from cells that went from constant green light (520nm, growth condition) to darkness, and the signal of cells exposed 10 min to saturating far-red light (765nm). This is expressed as following:

$$NormalizedYFP(LED\lambda, sat) = \frac{eYFP(LED\lambda, sat) - eYFP(520nm, growth)}{eYFP(765nm, sat) - eYFP(520nm, growth)}$$

If we consider that the eYFP is linearly linked to %PrPr, then :

$$NormalizedYFP(LED\lambda, sat) = \frac{\%PrPr_{eq,LED} - \%PrPr_{eq,520nm}}{\%PrPr_{eq,765nm} - \%PrPr_{eq,520nm}}$$

And therefore :

$$NormalizedYFP(LED\lambda, sat) = \frac{\frac{1}{(1 + \frac{\varepsilon_{Pr,LED}}{\varepsilon_{Pfr,LED}} * \eta)^2} - \frac{1}{(1 + \frac{\varepsilon_{Pr,520}}{\varepsilon_{Pfr,520}} * \eta)^2}}{\frac{1}{(1 + \frac{\varepsilon_{Pr,765}}{\varepsilon_{Pfr,765}} * \eta)^2} - \frac{1}{(1 + \frac{\varepsilon_{Pr,520}}{\varepsilon_{Pfr,520}} * \eta)^2}} \quad (5)$$

In this equation,  $\eta$  is the only unknown.

To determine it, the equation was fitted with the nls function in R with  $\eta = 1$  as starting value, on the experimental values obtained from monochromatic and bichromatic illuminations.

Integration of Equation 4 gives the following expression for %PrPr as a function of time:

$$\begin{aligned} \%PrPr(t) = & \left( \frac{k_1}{k_1+k_2} * \%PfrPfr0 + \frac{k_2}{k_1+k_2} * \%PrPr0 - \frac{k_1*k_2}{(k_1+k_2)^2} \right) e^{-2*(k_1+k_2)t} + \\ & \frac{k_1}{k_1+k_2} (\%PrPr0 - \%PfrPfr0 - \frac{k_2-k_1}{k_1+k_2}) e^{-(k_1+k_2)t} + \left( \frac{k_1}{k_1+k_2} \right)^2 \end{aligned} \quad (6)$$

To determine  $k$ , the following equation (7) was fitted to the eYFP levels further normalized for each LED and each strain (resulting in value between 0 and 1, and the fitting of the exponential will be independent on the fitting of the values at equilibrium as done above. However, we used the ratio of photoconversion  $\eta$  estimated above):

$$\begin{aligned} & \text{NormalizedYFP}(\lambda, N) \\ & = \frac{\left( \frac{k_1}{k_1+k_2} * \%PfrPfr0 + \frac{k_2}{k_1+k_2} * \%PrPr0 - \frac{k_1*k_2}{(k_1+k_2)^2} \right) e^{-2*k_{Fit}*N*t} + \frac{k_1}{k_1+k_2} (\%PrPr0 - \%PfrPfr0 - \frac{k_2-k_1}{k_1+k_2}) e^{-k_{Fit}*N*t} + \left( \frac{k_1}{k_1+k_2} \right)^2 - \%PrPr0}{\left( \frac{k_1}{k_1+k_2} \right)^2 - \%PrPr0} \end{aligned} \quad (7)$$

where %PrPr at time 0 is either the equilibrium in green (for the action spectra) or in far-red (for the inhibition spectra),  $t$ , the time of illumination, i.e. 10min, and  $k_{Fit}$  is the exponential constant to be estimated.

Further analysis of  $k$  and  $k_{Fit}$  in Equations 6 and 7 gives:

$$k * t = (k_1 + k_2) * t = (2.3 * \varepsilon_{Pr} * \phi_{Pr} + 2.3 * \varepsilon_{Pfr} * \phi_{Pfr}) * N * t$$

$$\text{so } k = \alpha * (A_{Pr,LED} * \eta + A_{Pfr,LED}) * N = k_{Fit} * N$$

$$\text{where } A_{Pr,LED} = \frac{\sum_{\lambda=300}^{\lambda=900} A_{Pr,\lambda} * N_{\lambda}}{\sum_{\lambda=300}^{\lambda=900} N_{\lambda}}, N = \sum_{\lambda=300}^{\lambda=900} N_{\lambda} \text{ and } \alpha \text{ is a constant}$$

In theory,  $k_{Fit}$  should be a linear combination of  $(A_{Pr} * \eta + A_{Pfr})$  (dotted lined on Figure 3D).

Fitting linear relationship between  $k_{Fit}$  and  $(A_{Pr} * \eta + A_{Pfr})$  with the lm function in R gave  $\alpha=0.012033+/-0.001969$  (adjusted  $R^2=0.4311$ ) when considering all the data points (Figure 3C), but  $\alpha=0.02005+/-0.002906$  (adjusted  $R^2=0.6247$ ) when ignoring the data points above 700nm.

From the previous equations, we can calculate %PrPr(t) in a given environment.

The light environment is described with  $I_{tot} = l * \sum_{\lambda=300}^{\lambda=900} N_{\lambda}$  with l the bandwidth and  $N_{\lambda}$  the intensity at wavelength  $\lambda$ .

We can calculate Pr and Pfr absorption in this environment with  $A_{Pr} = \frac{\sum_{\lambda=300}^{\lambda=900} A_{Pr,\lambda} * N_{\lambda}}{\sum_{\lambda=300}^{\lambda=900} N_{\lambda}}$ , from which we calculate  $k_{Fit} = 0.02005 * (A_{Pr} * \eta + A_{Pfr})$

The ratio of absorption is calculated with  $\frac{\epsilon_{Pr}}{\epsilon_{Pfr}} = \frac{\sum_{\lambda=300}^{\lambda=900} A_{Pr,\lambda} * N_{\lambda}}{\sum_{\lambda=300}^{\lambda=900} A_{Pfr,\lambda} * N_{\lambda}}$ , from which we can calculate the ratio of photoconversion:

$$\frac{k_1}{k_1 + k_2} = \frac{1}{1 + \frac{\epsilon_{Pr}}{\epsilon_{Pfr}} * \eta}$$

( $\eta$  is either the value determined in vivo or in vitro)

We have all the elements to calculate %PrPr(t)

$$\begin{aligned} \%PrPr(t) = & \left( \frac{k_1}{k_1 + k_2} * \%PfrPfr0 + \frac{k_2}{k_1 + k_2} * \%PrPr0 - \frac{k_1 * k_2}{(k_1 + k_2)^2} \right) e^{-2 * k_{Fit} * I_{tot} * t} \\ & + \frac{k_1}{k_1 + k_2} (\%PrPr0 - \%PfrPfr0 - \frac{k_2 - k_1}{k_1 + k_2}) e^{-k_{Fit} * I_{tot} * t} + \left( \frac{k_1}{k_1 + k_2} \right)^2 \end{aligned}$$

with %PrPr0=0 and %PfrPfr0=1 and t=10min or t=1h.





# CHAPTER 3: DPH IN BENTHIC SPECIES

Following the results from chapters 1 and 2, it was shown that amongst pennate diatom possessing DPH, most are found in benthic environments, and some have multiple DPH copies. Therefore, we wanted to explore the links between DPH in pennate species and the benthic environment. For this, we characterized the 4 DPH from a benthic diatom, *Amphora coffeaeformis*, as recombinant proteins. We also explored the physiological response of pennate species to red light, which was shown to induce chromatic acclimation in some of them (appearance of fluorescence emission at 710nm). We observed this fluorescence in several benthic species, and investigated the role of DPH in this phenomenon.

I participated in amplifying AcDPH CDS from cDNA and I purified recombinant DPH expressed in *E.coli* and characterized their absorption spectra. I was responsible of the characterization of fluorescence emission at 710nm under some conditions.

## DPH IN BENTHIC SPECIES

---

In the previous chapter, we have seen that DPH is sensitive to the whole light spectra. As a consequence, in the open ocean, DPH will sense mostly variation of the blue to green wavebands due to the absence of red and far-red lights after few meters. In addition, we showed that in the open ocean, essentially centric species possess DPH with a strong latitudinal repartition gradient. In the chapter one, we saw (as previously shown by (Fortunato et al., 2016), that centric and pennate DPH form separate clades, suggesting independent acquisition, maybe from viruses. Moreover, most of the pennate species known to possess DPH have a benthic/tychoplanktonic lifestyle, and some show multiplication of DPH copy number while centric diatoms only possess one DPH. This suggests that DPH conservation and multiplication is linked to the adaptation to benthic environments. In these environments, such as sediments and biofilms, blue light is strongly absorbed, resulting in red and far-red light enriched light field, where DPH could be sensing either sediment depth or the density of other photosynthetic organisms through the red/far-red ratio.

This led us to consider that DPH could have a different role in centric and pennate species, or at least mediate the response to different light cues. Centric species living in the open ocean sense blue to green ratio through phytochrome, which could help them to adapt to low blue light, highly mixed environments, or trigger sexual reproduction for example. Pennate species living in the sediments would use DPH to sense red/far-red light (or absence of blue) to adapt to life inside the sediments, i.e. low red light, metabolism (heterotrophy in the dark), phototaxis or sexual reproduction.

Recent analyses of the *Seminavis robusta* genome (Osuna-Cruz et al., 2020) and omics approaches (Chapter 1) suggest that an increase in photoreceptor numbers in some pennate diatoms may be related to benthic lifestyle. Indeed, DPH phytochromes are present in several copies in pennate diatoms (See Table1). This expansion appears essentially for species known to live in sediments. We therefore wanted to explore the significance of DPH in sediments by addressing several questions. Do the different



phytochromes present in the same species have the same photochemical properties? Can we associate the phytochrome to a chromatic acclimation linked to this particular light environment? Is this adaptation benthic and/or DPH specific?

Table 1. Pennate diatom species with phytochrome(s). In bold, species found in a tidal flat in (An et al., 2020)

Diatom species	Number of phytochromes
<i>Phaeodactylum tricornutum</i>	1
<b><i>Amphora coffeaeformis</i></b>	<b>4</b>
<b><i>Seminavis robusta</i></b>	<b>4</b>
<i>Fragilaria radians</i>	4
<i>Fragilaria sp</i>	2
<b><i>Nitzschia punctata</i></b>	<b>2</b>
<b><i>Nitzschia palea</i></b>	<b>2</b>
<b><i>Nitzschia sp</i></b>	<b>2</b>
<b><i>Nitzschia inconspicua</i></b>	<b>2</b>
<b><i>Navicula sp</i></b>	<b>2</b>
<b><i>Gyrosigma sp</i></b>	<b>1</b>
<i>Craticula ambigua</i>	2
<i>Diatoma tenuis</i>	1
<i>Synedra sp</i>	1
<i>Synedropsis recta</i>	1
<i>Staurosira complex</i>	1

## RESULTS

### Different DPH copies show different light-sensing abilities

Exploring the genomes and transcriptomes of diatom in the search for DPH genes unveiled that not all diatom possesses DPH, but that a lot of pennate species possess several copies of DPH genes, suggesting that these DPH can perform different functions or have some specific properties such as observed in plants (Franklin and Quail, 2010; Burgie et al., 2021) (See Table 1). Among these diatoms, the benthic diatom *Amphora*

*coffeaeformis*, possesses 4 DPH genes, and 3 of them possessing not only a putative chromophore-binding cysteine in the N-terminal region, conserved with bacterial phytochromes; but also a second one in the GAF region. The second cysteine is conserved with cyanobacteria and plant phytochromes, but also with *Ectocarpus siliculosus* phytochrome 1, which binds phycocyanobilin and show a far-red/green photocycle (Rockwell et al., 2014). This suggested that different DPH copies could bind different chromophore and have different light-sensing abilities in *A. coffeaeformis*.

To address this question, we cloned the 4 *A. coffeaeformis* phytochromes photosensory domains and expressed them in *E. coli* with the enzymes producing either biliverdin (BV), phytochromobilin (PΦB) or phycocyanobilin (PCB).

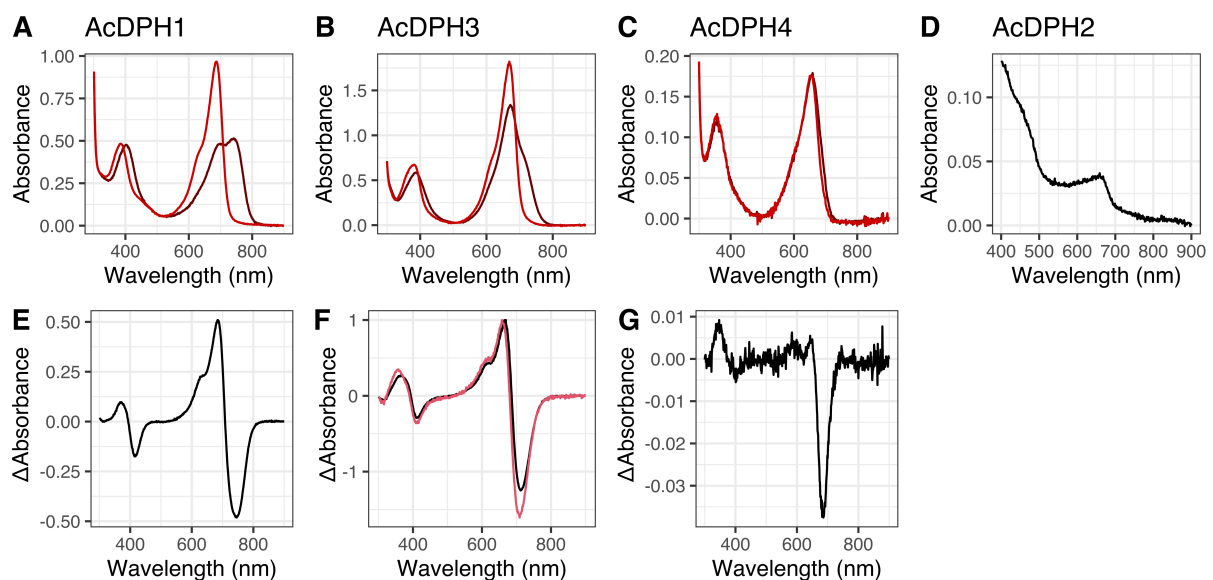


Figure 1. Absorption spectra and difference spectra between the Pr and Pfr forms for 4 *A. coffeaeformis* phytochromes. A. AcDPH1 with BV; red lines is the absorption spectra after illumination at 810nm and dark red lines correspond to illumination at 630nm. B AcDPH3 expressed with BV, red line is illumination at 740nm and dark line illumination at 66nm. C. AcDPH4 expressed with PCB, red line illumination at 740nm and dark red line is illumination 666nm. In D, AcDPH2 production (with BV) was very low, and the spectrum was only taken under white illumination. E to G, Difference spectra between far-red illuminated and red-illuminated AcDPH (E-AcDPH1, F-AcDPH3 and G-AcDPH4). In F, the comparison of the difference spectra between AcDPH3 expressed with BV (black) and PCB (red) is also presented, normalized to 1 for the maxima.

AcDPH1 seems to bind BV, and has an absorption spectra close to the one of PtDPH and other known DPH (see Chapter2), with peaks at 687nm for Pr form (810nm illumination)

and 742nm for Pfr-enriched mix (630nm illumination) (Fig1). For reference, PtDPH Pr/Pfr and TpDPH Pr/Pfr have peaks at 700/750nm and 686/764nm respectively.

AcDPH3 also seems to bind BV, as the absorption spectra of AcDPH3 expressed with PCB shows the same absorption peaks (Fig1F). However, AcDPH3 absorption spectrum is shifted towards shorter wavelength compared to AcDPH1 and PtDPH: the maxima and the minima of the absorbance difference are at 669 and 712nm, respectively.

AcDPH4 was binding PCB, but did not exhibit photocycle as other DPH. Absorption shows a maximum at 653nm, with only a slight bleaching at 686nm upon illumination at 740nm. This effect is photoreversible by illuminating at 666nm after irradiation at 740nm. AcDPH2 could not be produced in high amounts, but seems to show the same absorption peak as AcDPH4 (Fig1D).

With the analysis of these spectra, we can conclude that *A. coffeaeformis* DPH seem to be all red (R)/far-red (FR) phytochromes with light reversible photoconversion but have different spectral properties: AcDPH1 possesses conserved absorption spectra with known PtDPH, while AcDPH3 is shifted towards shorter wavelengths. AcDPH4 seems to have peculiar photocycle, is binding another chromophore (phycocyanobilin) and is expressed only in some conditions, suggesting that it might regulate different aspects of diatom life (Keeling et al., 2014). Indeed, AcDPH4 is more expressed in the presence of acetate in the diatom growth media, suggesting a link with the heterotrophic metabolism of *Amphora* species (Lewin and Lewin, 1960). We can notice that the wavelength difference between the absorption peak maxima of the Pr and Pfr forms seems smaller than for the other diatom phytochrome (43 nm for AcDPH3 (669/712nm) vs 78 nm for TpDPH (686/764nm) for example) and leading to a Pr/Pfr photoequilibrium rich in Pr form whatever the light used. Globally, this data suggests some spectral tuning of the different AcDPH, which might be linked to the benthic environment. This was not observed in the different oceanic and coastal centric diatoms DPH (chapter 2), and this is also different from the absorption spectra observed in *Ectocarpus siliculosus* (Rockwell et al., 2014).

## **DPH regulates red-light acclimation response in *P. tricornutum***

Sediments but also mats or biofilms are characterized by a strong depletion in green and blue light, resulting in red and far-red enrichment due to either wavelength-dependent light scattering or strong filtering of light by absorption by chlorophyll (Kühl and Jorgensen, 1992; Kühl et al., 1994; Kühl and Jorgensen, 1994). Some cyanobacteria have evolved a specific chromatic adaptation to enhance light harvesting in this red/far-red light environments called Far-Red Light Photoacclimation (FaRLiP)(Gan et al., 2014). Chromatic adaptation is a general mechanism observed in cyanobacteria allowing an adaptation of the photosynthesis pigment to the environmental light (Grossman, 2003). In the case of FaRLiP, this phenomenon is under the control of a phytochrome (RfpA) and two response regulators (RfpB and RfpC) to activate the expression of genes involved in this acclimation (Zhao et al., 2015). Little is known in diatoms about chromatic adaptation, but one of the responses of *P. tricornutum* to red light is the appearance of a fluorescence peak around 710nm at room temperature and slightly shifted at 714 nm at 77K (hereafter F710) (Fujita and Ohki, 2004; Herbstová et al., 2015; Oka et al., 2020). We thus tested whether F710 and this acclimation could be under DPH control. As shown in Fig2, fluorescence spectra at 77K of *P. tricornutum* showed two peaks one at 680nm characteristic of PSII, but also another one at wavelength around 710nm (F710). The fluorescence spectrum of *P. tricornutum* has been further investigated under different light conditions (wavelength) and all fluorescence spectra curves have been normalized to the peak at 680 nm for qualitative analysis. F710 was induced by red light or in high cell concentration cultures grown without agitation in accordance with previous reports (Brown, 1967; French, 1967; Shimural and Fujita, 1973; Fujita and Ohki, 2004; Herbstová et al., 2015). With agitation, in green or green plus far-red light, F710 did not increase even at high cell concentration, but in red light with agitation, the far-red fluorescence was clearly induced suggesting a specific red light induction of F710 (Fig2). Therefore, light color and “agitation” (also linked to culture age/cell concentration) are parameters controlling the red-shifted forms of chlorophyll allowing F710.

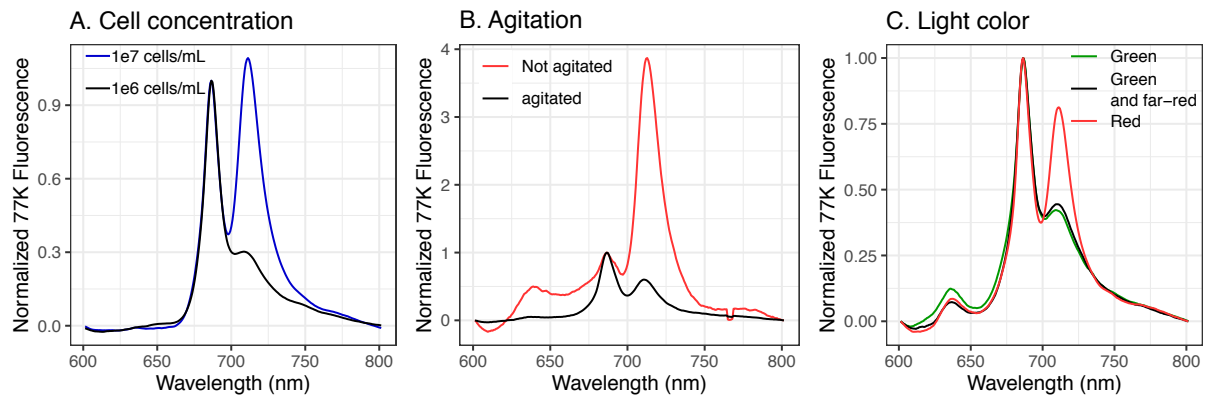


Figure 2. Fluorescence emission spectra at 77K of *P. tricornutum* grown in different conditions. A, Cells grown in white light without agitation and sampled at different two cell concentrations; B, cell grown in red light with and without agitation, sampled at  $\sim 2 \times 10^6$  cells/mL, C cells at  $\sim 2 \times 10^6$  cells/mL grown with agitation but under different lights.

In order to test a possible role of DPH in this red light acclimation, wild-type strain, transformation control lines (Wt (Tc), cells originated from the same transformed colony than its corresponding KO but not having undergone Pt-DPH gene editing as described in Fortunato et al, (2016)) and DPH-KO lines have been grown in the presence of red light but also with an increasing gradient of FR in order to vary the FR/R ratio, FR light being a light which activate PtDPH. Interestingly, we observed a FR and DPH-dependent regulation of F710 in red+far-red light (Fig 3 and FigS1). Indeed in Wt strains, F710 increases when FR light (and FR/R ratio) increases. In the DPH KO, F710 is present, but does not change with increasing ratio of FR/R lights. This experiment was repeated once for another pair of Wt Tc and DPHKO, and not in the exact same conditions, but showed the same type of phenotype (FigS1). F710 is thus induced by red light, but further regulated by DPH in a FR-dependent manner.

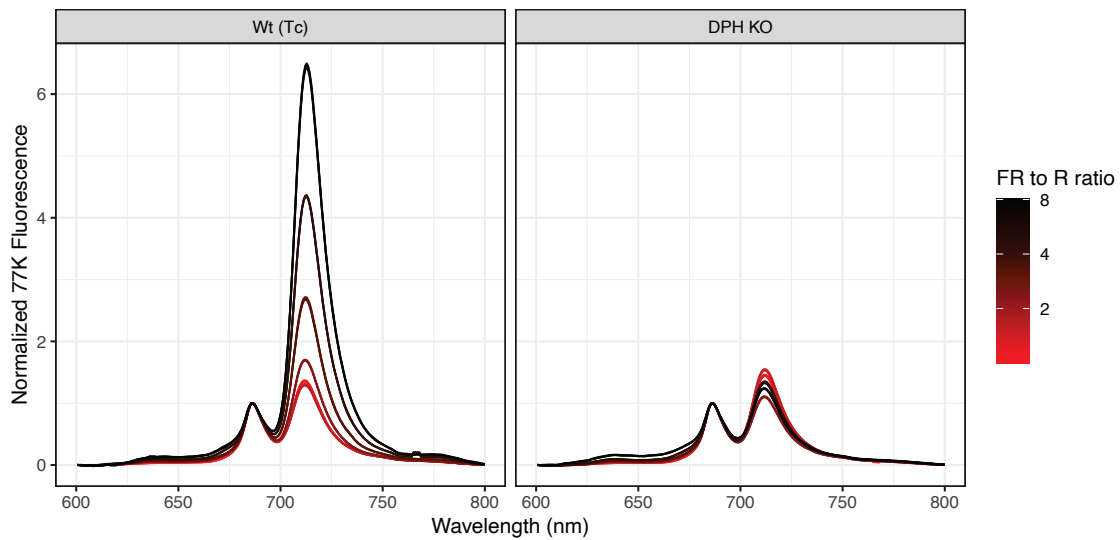


Figure 3. DPH controls F710 amplitude in response to increased FR/R ratio. Fluorescence emission spectra at 77K of *P.tricornutum* grown with agitation in red light with different far-red intensities (Tc tranformation control)

A kinetic follow-up of the appearance of F710 on cells grown at 735nm showed that this phenomenon took place after a few hours and that it was much faster and important on the wild type lines than on the KO lines (Fig4), confirming the potential role of DPH as a sensor of this light condition inducing the development of F710.

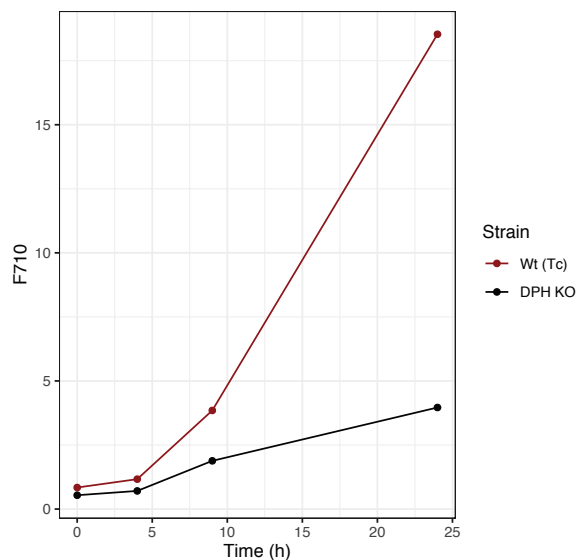


Figure 4. Kinetics of induction of F710. Cells were grown in blue light with agitation and transferred to FR light (735nm) at T0. F710 is the fluorescence emission at 77K at  $\lambda = 711\text{nm}$  (normalized to fluorescence at  $\lambda = 687\text{nm}$ ).

## F710 response seems restricted to benthic diatom species

The presence of F710 is known in Pt and in *Nitzschia closterium* (Fujita 2004), but it was not seen in centric species (*Chaetoceros*, *Skeletonema*, *Thalassiosira*) except for *Detonula sp* (Jupin and Giraud, 1971; Fujita and Ohki, 2004; Lavaud and Lepetit, 2013), and not in all pennate species (not observed for *Nitzschia punktata*, (Fujita and Ohki, 2004)). Therefore, we wanted to explore diatoms ability to develop the F710.

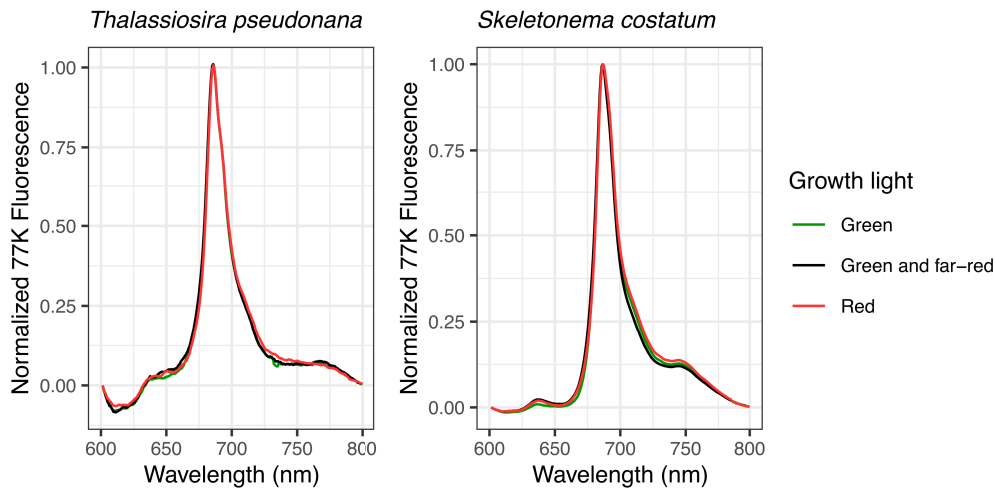


Figure 5. F710 in centric species : Fluorescence emission spectra at 77K of *Thalassiosira pseudonana* and *Skeletonema costatum* grown with agitation under different light colors.

As expected, the centric species tested did not show F710 induction (Fig5). Among the pennate diatoms tested, *Fragilariopsis cylindrus* is a planktonic and/or ice associated diatom. In the conditions we tested, which are known to trigger F710 in *P. tricornutum*, i.e white or red light without agitation at high cell concentrations, no F710 could be observed.

The benthic diatom *A. coffeaeformis* showed a high induction of F710 in red light (peak of fluorescence at 707nm) even with agitation, and in white light without agitation, but the F710 in green light with agitation remained low. It thus seems to be regulated in the same way as F710 in *P. tricornutum*.

The other three diatom species were grown only in white light without agitation, and both *Navicula salinicola* and *Nitzschia inconspicua* show a fluorescence peak at 708nm and

712nm respectively, while *Nitzschia agnita* did not show any peak around 710nm under this conditions.

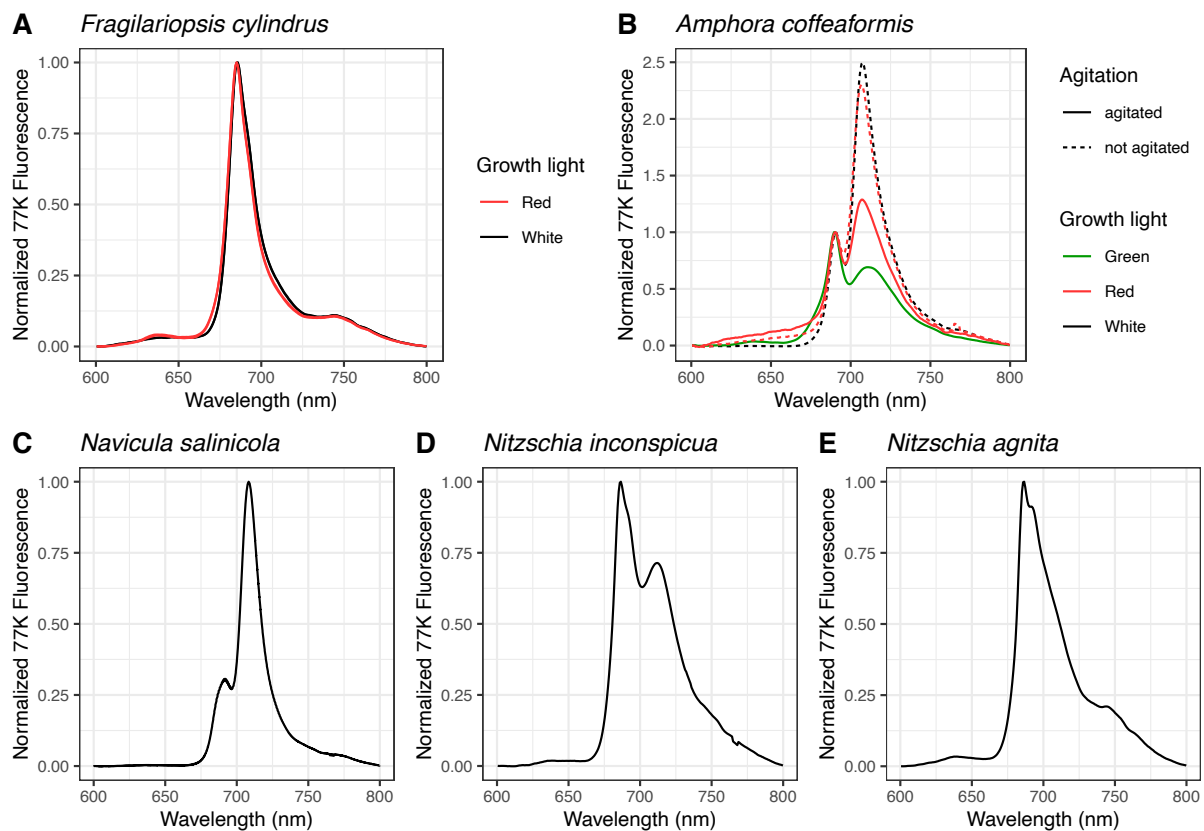


Figure 6. 77K fluorescence spectra of different pennate diatom species. A. *Fragilariopsis cylindrus* grown in red or white light without agitation, B. *A. coffeiformis* grown in green, red or white light, with and without agitation, C to E, *Navicula salinicola*, *Nitzschia inconspicua* and *Nitzschia agnita* grown in white light without agitation.

Therefore, different pennate species possess or not the ability to develop F710, and this might be linked to their planktonic vs benthic lifestyle.

## DISCUSSION

In this chapter, we focused on the role of phytochrome in benthic diatoms. We had previously seen that most of the pennate diatoms living in open water lack phytochrome and that on the contrary, pennate diatoms living in a benthic environment presented an



increase of their phytochrome number which could go up to 4 copies. In benthic environment, the quality and quantity of light changes significantly from the open water, which is dominated by blue/green light. One known red-induced phenotype in *P. tricornutum* is the adaptation of photosynthesis to red light. This is characterized by the appearance of a peak in far-red light (around 710nm) in the cell's fluorescence emission spectra. If the implication of DPH in such a phenotype has been suggested, it has not been tested so far. We therefore explore the possible links between phytochromes, adaptation to red light and life in the sediments.

### **F710 and life in the sediments**

In our study, we reproduced known effects of light, cell concentration and agitation on the induction of F710. Interestingly, the effect of the absence of agitation and high cell concentration could mimic the conditions experienced in a biofilm. Indeed, cell concentration in a biofilm is very high, with cells stacked on each other (Cartaxana et al., 2011). This results locally in high O<sub>2</sub> and low dissolved inorganic carbon due to photosynthetic activity and self-shading by other cells (Kühl and Jorgensen, 1992; Cartaxana et al., 2016; Marques da Silva et al., 2017). In a lab culture without agitation, cells sediment at the bottom of the flask; cell growth results locally in very dense cell layer, probably causing self-shading and high O<sub>2</sub>/low DIC and low nutrient availability. In the agitated cultures, we considered the gas concentration to be homogenous in the culture and equilibrated with ambient atmosphere, independent of cell concentration. At high cell concentration, one could suppose that self-shading is also occurring in agitated cultures, but this did not result in high F710, suggesting that gas equilibrium is a signal triggering F710.

Changes in light spectra due to high cell concentrations were not directly addressed in this preliminary work where we used mainly monochromatic light but may be addressed later. However, light is clearly an important signal for F710 development. Light intensity, quality and fluctuation are known to control F710 in *P. tricornutum*: induction in low white light (Brown, 1967) and intermittent light (Giovagnetti and Ruban, 2021), strong induction in red and yellow light, and small increase in green compared to white and blue

(Oka et al., 2020). In our work, F710 is clearly induced by red light, independently of DPH. This could be an effect of low photosynthesis efficiency in red, which is also supported by the appearance of F710 in low light. Alternatively, another red-light photoreceptor could be involved, but this photoreceptor has not been identified by genomic studies (Chapter 1), and would be of a new, unknown type, or a known type with atypical light sensing properties. In sediment, the light field is enriched in red and far-red light at depth, especially in the presence of photosynthetic organisms. Developing F710, and therefore red-shifted FCP, might be a response to the red-enriched light field in sediments.

The exact nature of this long-wavelength emission F710 is still unknown. F710 is clearly linked to photosystem II (PSII) activity but is debated whether this is due to an extra Light harvesting complex (LHC) binding to PSII and dissipating energy as FR fluorescence, or if this is spill-over of PSII over PSI (Fujita and Ohki, 2004; Herbstová et al., 2015; Bína et al., 2016; Giovagnetti and Ruban, 2021). However, the first hypothesis seems more convincing. In (Herbstová et al., 2015), the author performed sucrose gradient followed by gel filtration fractionation of thylakoid membranes of cells grown in red or white light. PSI fraction shows limited FR fluorescence and only at 77K but not at room temperature, while another fraction showed both 77K and room temperature FR fluorescence. Analysis of this second fraction showed that it contains LHCs, and one of them, Lhcf15 is specific of red light growth conditions. Lhcf15 seems to be a good candidate, as it was also found enriched in red-light grown cells compared to blue-acclimated cells in proteomic studies (Schellenberger Costa et al., 2013), and is transcriptionally induced upon 6 and 24 h of red light illumination (Valle et al., 2014) which is in accordance with our kinetic of induction of F710. We are currently trying to generate Lhcf15 knock out to validate (or not) the Lhcf15 hypothesis. If validated, we could explore the Lhcf family in diatoms to try to identify Lhcf15 homologs in other diatom species (database from Chapter 1) to assess if this antenna is restricted to benthic species. Lhcf phylogenetic trees exist (Herbstová et al., 2015; Bilcke et al., 2021; Kumazawa et al., 2022), but unfortunately without enough sequences from different diatom species to explore the link between Lhcf and the different environmental niches. One last comment is that the fitness advantage of developing F710 in red light is not established yet but to extend light-absorption abilities to longer wavelengths might play important roles in light adaptation of *P. tricornutum* and

benthic diatoms species. If we manage to control F710 induction, we could try to link F710 to other photosynthetic parameters and explore this question.

### **DPH and the fine-tuning of F710**

DPH clearly controls F710 levels in response to R and FR in *P. tricornutum*, although only two experiments were done. However, as many parameters control F710 (light, cell concentration, agitation), the experiments are difficult to reproduce with exactly the same F710 induction.

Interestingly, the pennate diatom *Fragilariopsis cylindrus*, lacking DPH, did not produce F710 while, *A. coffeaeformis* showed F710 induction in the same conditions as *P. tricornutum*, suggesting that in this diatom, F710 could be regulated by the same mechanisms including by DPH. Interestingly, the two other pennate diatoms showing F710 are likely to possess phytochromes too. The only *Navicula* species studied so far possess two DPH, so it might also be the case in the *Navicula salinicola* strain used here. However, *Nitzschia* species do not always possess DPH: some do (*Nitzschia palea*, sp, *N. punctata* see Table1), while some don't (see Chapter 1, although it is only transcriptome data). Here, *Nitzschia agnita*, such as *Nitzschia punktata* in (Fujita and Ohki, 2004) did not show F710, while *Nitzschia closterium* (Fujita and Ohki, 2004) and *inconspicua* (this study) have this red shift antenna phenotype. The *Nitzschia inconspicua* strain we used is not the one for which the genome is available (Oliver et al., 2021), but the genome sequenced from a close strain contains 2 DPH genes, what make us think that it might also be the case for the strain we used. However, we cannot exclude that we did not find the conditions inducing F710 in *F.cylindrus* and *N.agnita*, suggesting different regulations in these strains.

As mentioned above, light in the sediment is enriched in red and far-red light. In addition, the FR/R ration increases with depth in the presence of photosynthetic organisms. For example, in microbial mats, light below 700nm is quickly attenuated (Kühl and Jorgensen, 1992; Cartaxana et al., 2016), while far-red light is less affected. DPH could thus be activated when going under a layer of photosynthetic cells due to high FR/R ratio. Under

this condition, the photoequilibrium of DPH or the co-regulation of several DPH in species such as *A. coffeaeformis* could be determinant for this chromatic adaptation. It is therefore tempting to propose that DPH and F710 co-evolved in benthic diatoms, with DPH enabling the fine-tuning of F710 in sediments enriched in far-red light.

If the link between F710 and Lhcf15 mentioned above is confirmed, this would enable the exploration of diatom genomic database for co-occurrence of Lhcf15 and DPH.

### **DPH multiplication in benthic species**

In *P. tricornutum*, only one phytochrome is present and was shown here to control F710. However, several DPH are present in other benthic diatom species and could control F710. In *A. coffeaeformis*, we have shown here that the different DPH have different light-sensing abilities and are induced under different growth conditions, suggesting that they might have different functions. The strong conservation of AcPH1 spectra with PtDPH and centric DPH from Chapter 2, led us to propose that all benthic diatoms have such a DPH with light sensing in the red and far-red, that could control F710. AcDPH knock out are being generated in the lab by CRISPR-Cas9 and will allow us to address this question.

DPH could control other aspects of diatom life in the sediments. Interestingly, *A. coffeaeformis* is able to grow in the dark in the presence of glucose, and to respire nitrates in the dark (Lewin and Lewin, 1960; Kamp et al., 2011). Other benthic species also show metabolic adaptation to life in the sediments (Lewin and Lewin, 1960). In our study and (Keeling et al., 2014), we saw high expression of AcPH4 in the presence of acetate, which hints towards a link between DPH and carbon metabolism in *A. coffeaeformis*. However, growth did not seem to be higher in the presence of acetate, suggesting that acetate might be a signal rather than a carbon source for heterotrophic growth. Testing growth and photosynthesis parameters with and without acetate and red light will provide interesting answer to this question.

Phototaxis is another behavior important for diatom life in the sediments (Barnett et al., 2020). A phototaxis assay for *A. coffeaeformis* is being developed in the lab and could be used to test the effect of different wavelengths and the implication of DPH(s).

As a conclusion, some photosynthetic adaptations seem to be specific for the benthic lifestyle of pennate species, and are likely to be regulated by DPH. In addition, the multiplication of DPH in these diatoms suggests the importance of light-regulation of diatom life in this environment enriched in red and far-red lights.

## MATERIAL AND METHODS

### Diatom culture conditions

Wild-type *P. tricornutum* (Pt1 8.6; CCMP2561) cells, PtDPH-KO mutants and transformation controls (Fortunato et al., 2016) were maintained in Enriched Artificial Seawater without silica; *Amphora coffeaeformis*, *Fragilariopsis cylindrus*, *Navicula salinicola*, *Nitzschia inconspicua* IRTA\_CC1 and *Nitzschia agnita* were cultured with Si. All except *Fragilariopsis cylindrus* were grown at 19 °C under a 12L/12D regime using 50  $\mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  white light (Philips TL-D De Luxe Pro 950) ("White condition) or in continuous light in incubators equipped with green (520nm), red (670nm) or far-red (800nm) LEDs, with agitation at 160rpm. For the FR gradient, flasks were homogeneously illuminated from above with LED at 660nm at 25  $\mu\text{mol photons}/\text{m}^2/\text{s}$ , and from the side with FR LED at 765nm or at 810 nm, resulting in a gradient of FR light (from 30 to 250 $\mu\text{mol photons}/\text{m}^2/\text{s}$  for the LED at 760nm and 30 to 170  $\mu\text{mol photons}/\text{m}^2/\text{s}$  for 810nm). *Fragilariopsis cylindrus* was grown at 4°C with a white spot under a 12L/12D regime, and covered with a red filter for red-light growth. For F710 induction kinetics, cells were first grown in blue light (450nm, 10 $\mu\text{mol photons}/\text{m}^2/\text{s}$ ) until similar fluorescence spectra and a cell concentration around  $2\cdot 10^6$  cells/mL were obtained. The cultures were then placed at 735nm (40 $\mu\text{mol photons}/\text{m}^2/\text{s}$ ) to follow the induction of F710.

### AcDPH cloning and expression

Photosensory module (PSM) of DPH genes were cloned from *A.coffeaeformis* cDNAs grown in different conditions. Sequences were amplified with the primers pairs indicated in the Table S2, and cloned into pET28a-HO vector already containing *Synechocystis* Heme

oxygenase gene or pET28a-*pcyA* containing SynH01 and *pcyA* from *Synechocystis* PCC6803. PSM sequences were expressed as N-terminal His6 tagged proteins in *E. coli* BL21 (DE3) strain, by auto-induction system (Studier, 2005). Recombinant proteins were purified as in Fortunato et al. (2016) and absorption spectra of the recombinant proteins were measured immediately after purification on a Varian's Cary-50 spectrophotometer. Illumination with LEDs at 810, 740, 666 and 630 nm were performed (approx. 1min illumination) to reach pure Pr spectra (after far-red illumination) or equilibriums between the Pr and Pfr forms.

### 77K Fluorescence

Low temperature (77 K) fluorescence emission spectra were measured with an in housebuilt setup. Samples were immersed in liquid nitrogen and excited with a LED source (LS-450, Ocean Optics - blue LED, 450 nm). The emission spectra were recorded using a CCD spectrophotometer (QE6500, Ocean Optics).

#### SUPPLEMENTARY FIGURE:

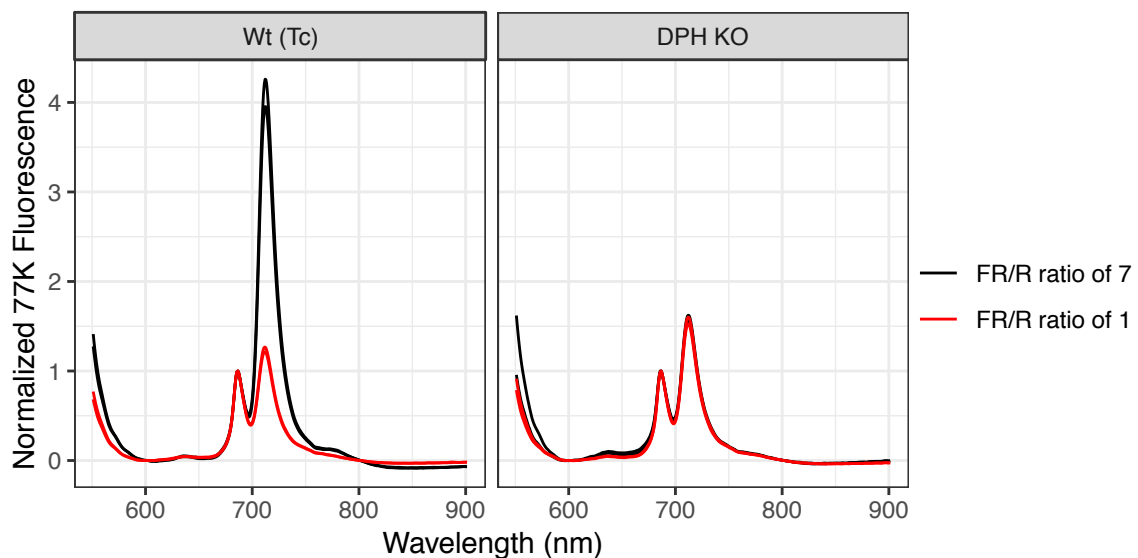


Figure S1. DPH controls F710 amplitude in response to increased FR/R ratio. Fluorescence emission spectra at 77K of *P.tricornutum* grown with agitation in red light with different far-red intensities (Tc transformation control). Compared to Fig3, the strains are an other pair of DPHKO and its transformation control, and the FR light used is a LED at 810nm.

Table S1. Primers used to clone *A. coffeaeformis* DPHs

Names	Sequence
AcDPHY1_Fw_SpeI	gcgACTAGTATGACCATGCCTCCTGACGAAG
AcDPHY1_PCD_Rv_SacI	gcgGAGTCCTATCCAACAAGGAAAAGTATC
AcDPHY2_Fw_SpeI	gcgACTAGTATGACAATCTTTCAGAAAATTGG
AcDPHY2_PCD_Rv_SacI	gcgGAGTCCTAATCAATGGCCGAGTAATAC
AcDPHY3_Fw_SpeI	gcgACTAGTATGGTCCTCCTGACTCCATGGAAG
AcDPHY3_PCD_Rv_SacI	gcgGAGTCCTATTCGAGTGCGTCAAAGCTTCGGATATTTG
AcDPHY4_Fw_SpeI	gcgACTAGTATGGTCCACCTCCTAAGGATATC
AcDPHY4_PCD_Rv_SacI	gcgGAGTCCTACATCGCCCGCATTGAAGCTCTT





# CHAPTER 4: DPH SIGNALIZATION

In this last part, we explored different levels of the DPH light-sensing mechanisms, from chromophore binding to the integration of several environmental cues into the DPH signaling cascade. For this part, I searched for putative chromophore synthesizing enzymes in *P.tricornutum* genome, and tested some by cloning and co-expressing them with DPH in *E. coli*. I used *P.tricornutum* transgenic lines expressing tagged version of the PtDPH to look at DPH localization. I also addressed the question of the signalization cascade by Yeast two hybrid. I cloned DPH (full length or fragments) into Y2H vectors to perform an interaction screen with an available Pt Y2H library. I finally turned to a candidate approach with the only histidine phosphotransferase found in *P. tricornutum* genome PtHpt, to test its interaction with PtDPH. Finally, I characterized the effect cell concentration on DH-mediated response in *P. tricornutum*.

# DPH SIGNALIZATION

---

In the previous chapter, I have explored DPH evolutionary history, modeled its activity *in vivo* and projected it *in situ* and started to explore potential DPH-regulated phenotypes in model species.

In this last chapter, I will explore DPH signaling-related mechanisms, i.e. the assembly of the apoprotein with its chromophore, DPH subcellular localization, and components of the signaling cascade triggered by DPH. I will also present evidence of the co-regulation of PtDPH-controlled genes by other factors. Each of these aspects aims at broadening our knowledge about DPH light-sensing in diatoms.

## RESULTS

### DPH chromophore

Phytochrome chromophores are all linear tetrapyrroles, also called bilins (Rockwell and Lagarias, 2017). Bilins derive from the oxidative cleavage of heme by heme oxygenase to form biliverdin (IX $\alpha$ )(BV), which can be further reduced by ferredoxin-dependent bilin reductases (BVRs), into phycocyanobilin (PCB), phytochromobilin (PB), phycourobilin (PUB) or phycoerythrobilin (PEB) (Fig. 1) (Dammeyer and Frankenberg-Dinkel, 2008).

Bacteriophytochromes (Bph) bind BV (Bhoo et al, 2001; Karniol et al, 2005). In bacteria, the heme oxygenase gene is often found in the same operon as the *Bph* gene (Bhoo et al, 2001, Giraud et al, 2002). BV was found to bind a conserved cysteine in the N-terminal extremity of the apoprotein Bph and fungal phytochromes (Lamparter et al, 2002; Karniol et al, 2005; Blumenstein et al, 2005)

Land plant phytochromes bind phytochromobilin (P $\Phi$ B), which is produced by the enzyme HY2 in a one-step process (Lagarias and Rapoport, 1980), while cyanobacteria bind phycocyanobilin produced by the enzyme PcyA (Hughes et al, 1997; Frankenberg et

al, 2001). In both cases, the chromophore is covalently bound to phytochrome apoprotein to a conserved cysteine in the GAF domain (Wagner et al, 2005).

PUB has been identified as chromophore of some cyanobacteriochromes (Sun et al, 2014), and as pigments of phycobilisome light-harvesting complexes, as the phycoerythrobilin produced by the enzymes *pebA* and *pebB*, in most red microalgae and in Cyanophyceae organisms (Dammeyer and Frankenberg-Dinkel, 2008).

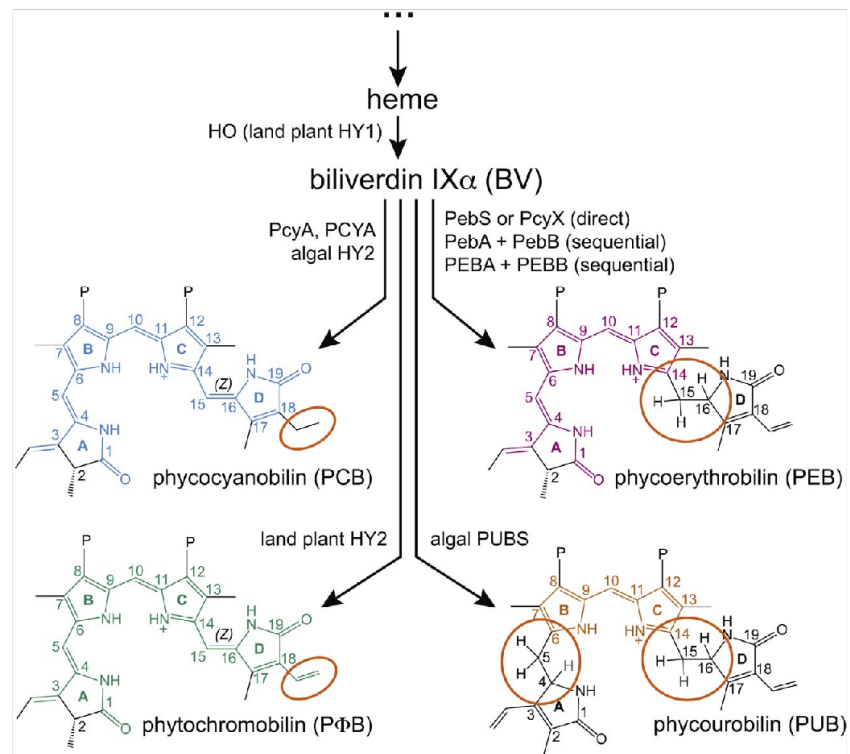


Figure 1. Phytochrome chromophores biosynthesis pathways (from Rochkwell and Lagarias 2017)

In plants, the phytochrome chromophore biosynthesis enzymes are located in the chloroplast, while in fungi it has been shown that heme oxygenases were attached to the mitochondrial outer membrane (Muramoto et al, 1999; Streng et al, 2021). Diatom phytochromes display a structure similar to the ones of bacteriophytochromes, with a conserved cysteine in the N-terminal part of the protein, and Fortunato et al (2016) showed that recombinant *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* DPH covalently bind BV as chromophore. PtDPH expressed with the enzymes for production of PΦB or PEB had the same absorption spectra as with BV, suggesting that BV is the preferentially bound chromophore, and therefore possibly the native one. However, this

has not been demonstrated, and it has also been reported that some DPH from the diatom *Amphora coffeaeformis* possess an additional cysteine residue in the GAF domain conserved with the one used in plant and cyanobacterial phytochromes to bind their chromophore, suggesting that different chromophores could be synthesized in diatoms (Fortunato et al, 2016)(see Chapter3).

In this first part, we will explore the possible DPH chromophore biosynthesis pathways by mining genome for HO and BVR homologous genes, and testing their activity, as well as the reconstitution of chromophoric recombinant DPH.

### Heme oxygenase

Candidate heme oxygenase genes have been previously identified in *P. tricornutum* and *T. pseudonana* genomes (Fortunato et al., 2016). Here, we confirmed the presence of the Heme oxygenase domain (PF01126.2) (Table 1), and that no other protein in Pt genome possesses this domain. We also explored the predicted localization of these enzymes with ASAFind (Gruber et al, 2015), HECTAR (Gschloessl et al, 2008) and SignalP5 (Almagro Armenteros et al, 2019) software. Two of these proteins seem to be close to plant and cyanobacterial HO1 and are predicted to be addressed to the chloroplast. One enzyme is closer to mammalian HO, while the other did not have any good hit in Swissprot database (best blast hit had an e-value of 4.2). Both the plant/cyanobacteria-like and the mammalian-like HO had homologs in *T. pseudonana* and in *Amphora coffeaeformis* (Ac) (Table 1).

Table 1 Characteristic of the candidate Heme oxygenase in *Phaeodactylum tricornutum*

Pt V3 Id Phatr3_	Best blast hit in swissprot	Blast e-value	Pfam (PF01126.2) evalue	SignalP5	HECTAR	ASAFind	Tp Thaps3_	Ac CAMPEP_
HO1 J12588	P09601.1 Heme oxygenase 1 [Homo sapiens]	1.13E-52	4.80E-55	other	other localisation	not plastid	17865	0186525012 0186524814
HO2 J5851	Q69XJ4.1 Heme oxygenase 1, [Oryza sativa]	1.09E-50	1.20E-07	signalP	chloroplast	not plastid	17854	0186534672
HO3 J35647	No good hit		4.20E-10	other	other localisation	not plastid		
HO4 J5902	Q8YVS7.1 Heme oxygenase 1 [Nostoc sp. PCC 7120]	9.11E-58	2.00E-56	signalP	chloroplast	plastid, high confidence		

We successfully cloned *PtHO2*, *PtHO3* and *PtHO4* from cDNA. As PtDPH is difficult to produce, we cloned these enzymes for recombinant expression in *E. coli*, along with *T. weissflogii* DPH Photosensory module (PSM) (TweiDPH), which is easily produced in high amounts and thus more suitable for screening (Fig3).

Only the strain expressing *PtHO2* produced high amounts of functional TweiDPH. Pellets of bacteria expressing *PtHO3* or *PtHO4* with TweiDPH were uncolored, suggesting that they did not produce the suitable chromophore (or that the enzymes were not functional in *E. coli*). TweiDPH produced with *PtHO2* could reversibly photoconvert to Pr upon far-red (765nm) light illumination, and had the same absorption spectra as TweiDPH when produced with *Synechocystis PCC6803* heme oxygenase 1 (SynHO1) (Figure 2). This suggests that *PtHO2* produced the same chromophore as SynHO1, which is efficiently incorporated in DPH and enables photoconversion.

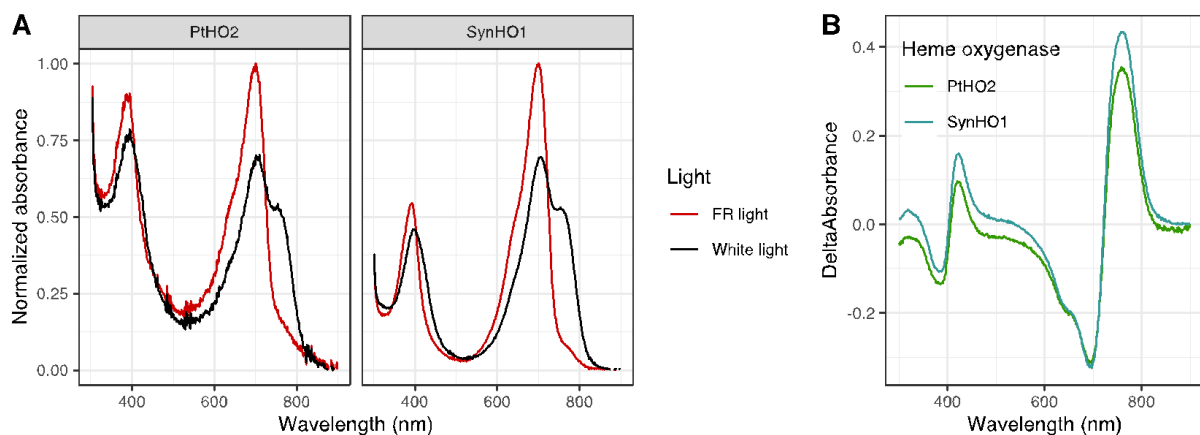


Figure 2. Spectral properties of TweiDPH expressed in *E.coli* with *Synechocystis* HO1 (SynHO1), producing biliverdin, or PtHO2. A Absorption spectra of TweiDPH at room (white) light (black line) or after illumination by a far-red LED (765nm) (red line), normalized to Pr peak, in red. B. Difference in absorption spectra for recombinant TweiDPH produced with SynHO1 or PtHO2.

Therefore, *PtHO2* is able to produce BV for phytochrome activity *in vivo*; *PtHO2* knock-out mutants are planned to be generated by CRISPR-Cas9 in *P. tricornutum*, to test whether or not this enzyme is required for phytochrome-mediated gene expression.

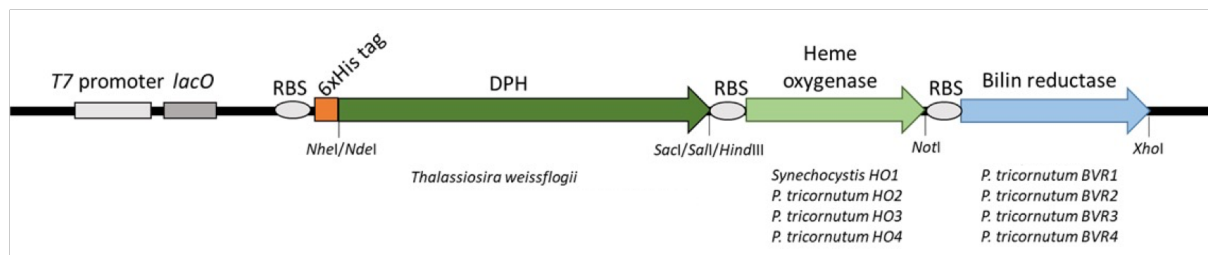


Figure 3 Scheme of the vector structure allowing to co-express *T. weissflogii* DPH photosensory module along with the putative enzymes of chromophore biosynthesis pathway identified in *P. tricornutum* genome. The backbone vector used was the pET28a, and the different coding sequences were differentially combined via sequential insertions according to the restriction sites indicated.

### Biliverdin Reductases

Putative bilin-reductase genes have been identified in *P. tricornutum* genome, raising the possibility of other chromophores synthesized in diatoms. Among these 4 candidates, 2 were predicted to harbor a peptide signal by the 3 software and possess as best hit counterparts in cyanobacteria and plant (Table 2).

These different BVR were amplified from cDNA and recombinantly expressed with *Synechocystis ho1* and *TweiDPH* as polycistronic sequence (Fig3).

When *TweiDPH* was expressed with *PtBVR3* and 4, the absorption spectra were the same as with *SynHO1* (Fig3A), suggesting that these phytochromes were not binding a new chromophore, but bound BV. However, when *TweiDPH* with *PtBVR1* or *PtBVR2*, a new absorption peak appeared around 620nm, suggesting that the DPH was binding another chromophore in addition or in place of BV. However in both cases, this peak did not show any photochemistry (Fig3B and C). As this DPH is produced in high amounts, it is possible that it binds non-optimal chromophores. Zinc induced fluorescence gel will tell if this new chromophore is covalently bound to the protein, and absorption spectra of the denatured protein will give indication on the nature of this chromophore.

Overall, alternative chromophores produced by *BVR1* and *BVR2* could be possible alternative chromophore but it seems that BV is the most convincing as DPH native chromophore.

Table 2 Characteristic of the candidate BVR in *P.tricornutum*

Accession	best blast hit in swissprot	evalue	Pfam domains	SignalP	HECTAR	ASAFind	Tp Thaps3
BVR1 J10830	Q7NL66.2, 15,16-dihydrobiliverdin:ferredoxin oxidoreductase [Gloeobacter violaceus PCC 7421]	4.00E-51	Ferredoxin-dependent bilin reductase (PF05996.15) 1.6e-53	signalP	signal peptide	Plastid, low confidence	32431
BVR2 J33770	Q93TL6.2, 15,16-dihydrobiliverdin:ferredoxin oxidoreductase [Nostoc punctiforme PCC 73102]	4.00E-39	Ferredoxin-dependent bilin reductase (PF05996.15) 9.2e-47	other	other localisation	Not plastid, SignalP negative	32431
BVR3 J42659	Q8H124.1, Uncharacterized protein At2g34460, chloroplastic [Arabidopsis thaliana]	8.00E-16	NAD(P)H-binding (PF13460.9) 6.7e-39	signalP	signal peptide	Plastid, high confidence	264007
BVR4 J45446	Q10UW6.1, Phycoerythrobilin:ferredoxin oxidoreductase [Trichodesmium erythraeum IMS101]	1.00E-30	Ferredoxin-dependent bilin reductase (PF05996.15) 1.5e-35	other	signal peptide	Not plastid, SignalP positive	262396

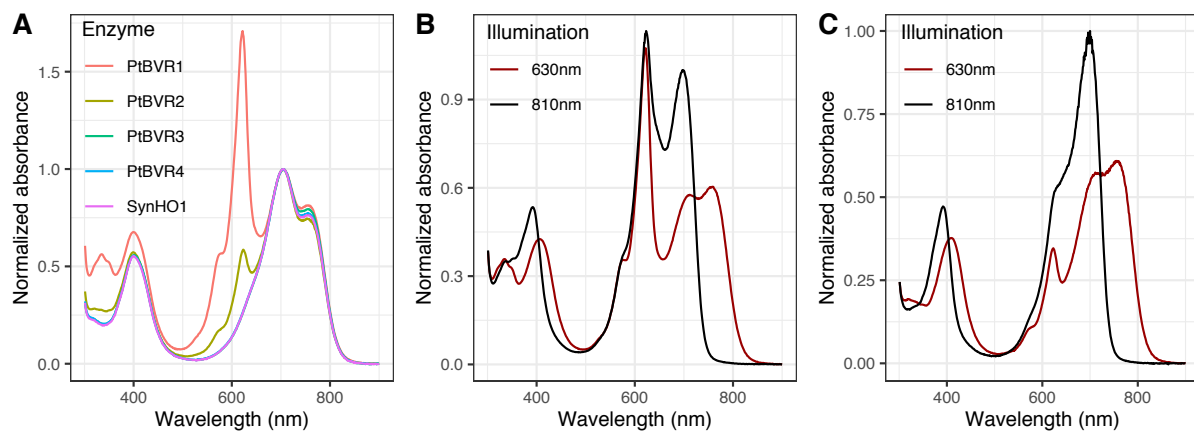


Figure 4. Spectral properties of TweiDPH expressed in *E.coli* with SynHO1 and different PtBVR, in white light; B TweiDPH in R of FR with PtBVR1, C TweiDPH in R and FR with PtBVR2

We have shown that *P. tricornutum* possesses at least one enzyme able to synthesize biliverdin as chromophore of phytochrome, and at least two enzymes are able to further

process BV. The nature and the role of this pigment in phytochrome activity need to be further investigated.

## DPH LOCALIZATION

Phytochrome localization is an important parameter for DPH signaling. In plants, phytochromes translocate to the nucleus upon light illumination, where they interact with transcription factors (Legris et al., 2019). This seems to be the case in green algae too, as *Micromonas pusilla* phytochrome is shuttled to the nucleus during the light phase of the day (Duanmu et al., 2014). In Fungi, FphA of *Aspergillus nidulans* interacts with the first component of its signaling cascade in the cytoplasm, and it is one of the downstream signaling component that is shuttled to the nucleus upon phytochrome light activation (Yu et al., 2016). Other studies also showed that FphA is interacting with blue light signalization components in the nucleus (Purschwitz et al., 2008).

We thus tested DPH subcellular localization in *P. tricornutum*. PtDPH was fused with Venus or NeonGreen fluorescent proteins either in N-terminal or C-terminal and expressed under the strong promoter of the *LHCF2* antenna gene in the episome system. Cells were grown in green light (DPH-inhibitory) and exposed to FR light to investigate eventual light-induced re-localization.

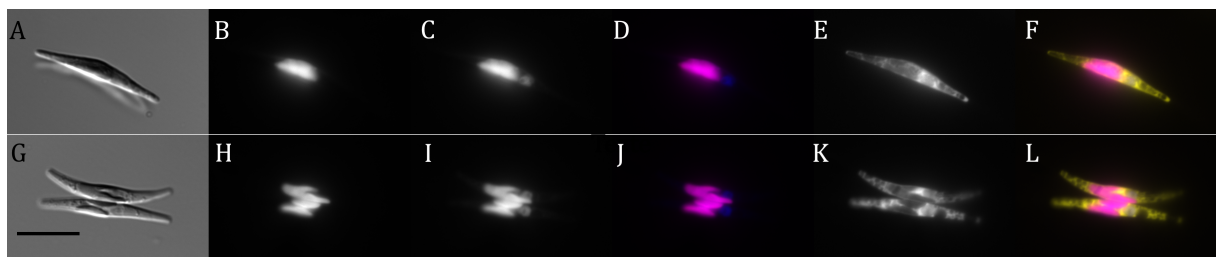


Figure 5. DPH localization. Venus-PtDPH-expressing *P. tricornutum* cells grown in green light (A to F), or far-red light (G to L). Bar scale : 5 $\mu$ m. A and G, Bright field ; B and H, Chlorophyll autofluorescence; C and I, Hoescht fluorescence (nucleus staining, also showing chlorophyll fluorescence) ; D and J, Hoeschst and chlorophyll overlay, nucleus appears blue when chloroplast appears magenta ; E and K, Venus fluorescence (PtDPH localization) ; F and L, overlay.



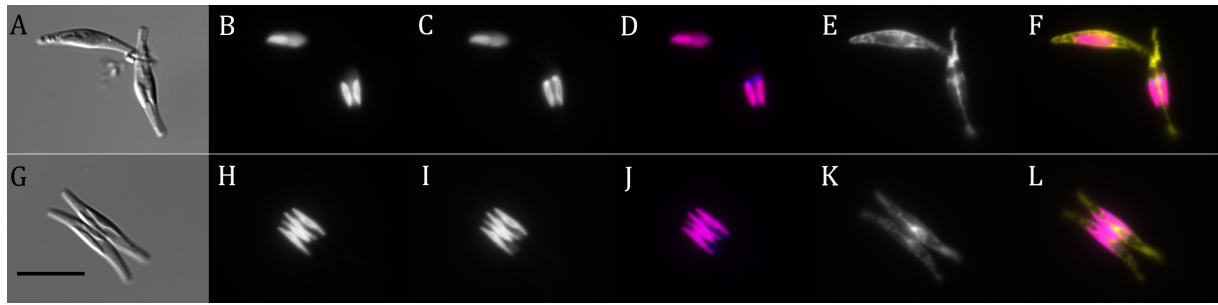


Figure 6. DPH localization. PtDPH-NeonGreen-expressing *P. tricornutum* cells grown in green light (A to F), or far-red light (G to L). Bar scale: 5 $\mu$ m. A and G, Bright field; B and H, Chlorophyll autofluorescence; C and I, Hoescht fluorescence (nucleus staining, also showing chlorophyll fluorescence) ; D and J, Hoeschst and chlorophyll overlay, nucleus appears blue when chloroplast appears magenta ; E and K, Venus fluorescence (PtDPH localization) ; F and L, overlay.

PtDPH seems to be cytoplasmic (Fig4 and 5), although it is difficult to affirm with certainty that PtDPH is absent from the nucleus. No major changes were observed upon FR illumination. We thus concluded that PtDPH is present in the cytoplasm.

As control, lines expressing only the Venus fluorescent protein (Fig7), or the fusion Venus-PtCPF1 (Fig8, previously shown to be addressed to the nucleus (Coesel et al, 2009)) under the same regulatory sequences and episomal vector, were observed and display cytoplasmic and nuclear localization, respectively.

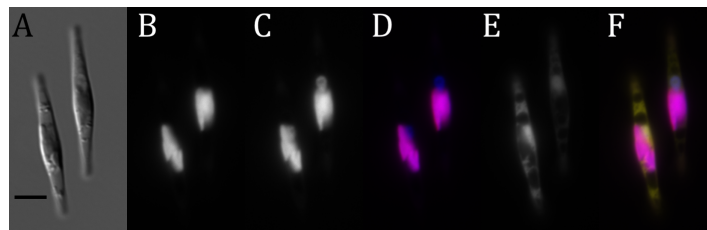


Figure 7. Venus localization. Venus-expressing *P. tricornutum* cells grown in white light (A to F). Bar scale: 2 $\mu$ m. A, Bright field; B, Chlorophyll autofluorescence; C, Hoescht fluorescence (nucleus staining, also showing chlorophyll fluorescence); D, Hoeschst and chlorophyll overlay, nucleus appears blue when chloroplast appears magenta ; E, Venus fluorescence ; F, overlay.

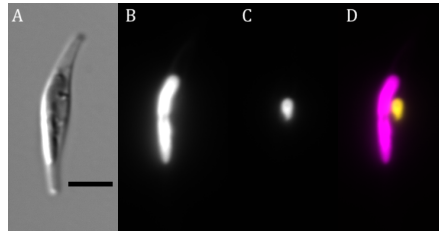
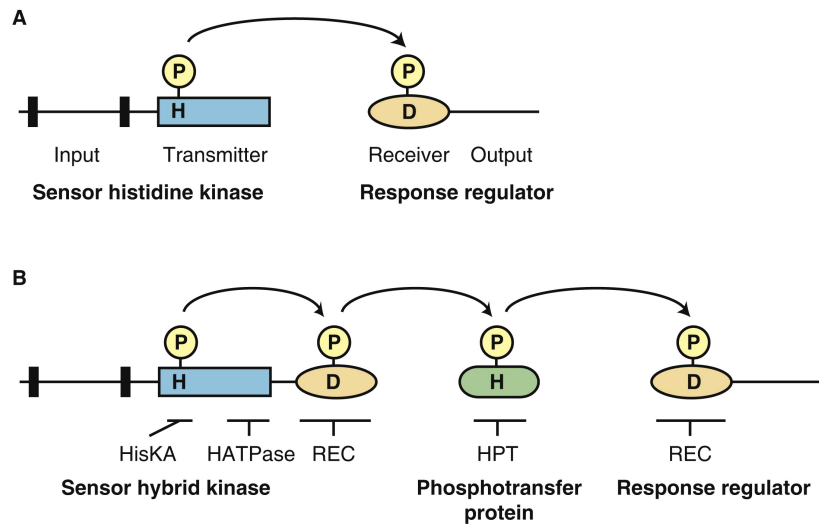


Figure 8. PtCPF1 localization. Venus-PtCPF1-expressing *P. tricornutum* cells grown in white light (Bar scale: 2  $\mu\text{m}$ . A, Bright field; B, Chlorophyll autofluorescence; C, Venus fluorescence (PtCPF1 localization in the nucleus); D, overlay.

---

## DPH SIGNALISATION PATHWAY

Phytochrome signalization is very different in plants compared to bacteria and fungi. In plants, phytochromes are shuttled to the nucleus upon light activation, where they interact with Phytochromes Interacting Factors, which are bHLH transcription factors. It has been described for several fungi and bacteria phytochromes that signaling cascade consists in a two-component phosphorelay system (Yeh et al, 1997; Bhoo et al, 2001; Giraud et al, 2005; Brandt et al, 2008). These systems start with the activation of a receptor histidine kinase (here, phytochrome), which phosphorylates a response regulator (REC) domain. In a simple two-components system, the response regulator directly regulates the functional response (for example gene expression). In multi-step system, the receptor histidine kinase is hybrid, as it also contains a REC domain. Upon activation, there is phosphorylation of this REC domain, then the phosphate group is transmitted to a phosphotransfer protein (characterized by a Hpt domain), which will bring the phosphate group to the effector response regulator (Fig9).



Current Biology

Figure 9. Elements from two-component systems, from (Schaller et al., 2011) A, simple two-component system, B, Multi-step phosphorelay where the hybrid histidine kinase contains both a histidine kinase domain and a response regulator. Phosphorylated residues Histidine (H) and Aspartate (D) are indicated.

Bacteriophytochromes show both simple and multi-steps two-component systems, and in most of the cases reported, the signalization components are encoded in the same operon as the *Bph*, which facilitates the determination of the downstream components (Fig9) (Yeh et al, 1997; Bhoo et al, 2001; Giraud et al, 2005; Vuillet et al, 2007; Jaubert et al, 2008). In fungi, FphA from *Aspergillus nidulans* interacts with the Hpt-containing protein YpdA, which itself controls the REC-containing SskA protein (Yu et al., 2016). Downstream of SskA, a MAPKinase cascade results in the phosphorylation of Saka, which is shuttled to the nucleus where it regulates the transcription factor AtfA (See also in the Introduction).

In diatoms, we know PtDPH autophosphorylates upon far-red light irradiation *in vitro* (Fortunato et al., 2016), which would be compatible with a two-component system phosphorelay as in Bacteria and Fungi.

I performed Yeast-two hybrid (Y2H) screens to find candidate interactors for PtDPH. I used a Y2H PtcDNA library from (Huysman et al., 2013), and performed the screen with PtDPH full length protein (one time) or PtDPH output module only (PtDPH\_OPM, screen was performed 3 times) as baits, fused to GAL4 binding domain. Despite the identification of some candidates during the screens, we were unable to reproduce these interactions

when re-transforming yeast strains with the candidates vectors, suggesting that these were not true interactants.

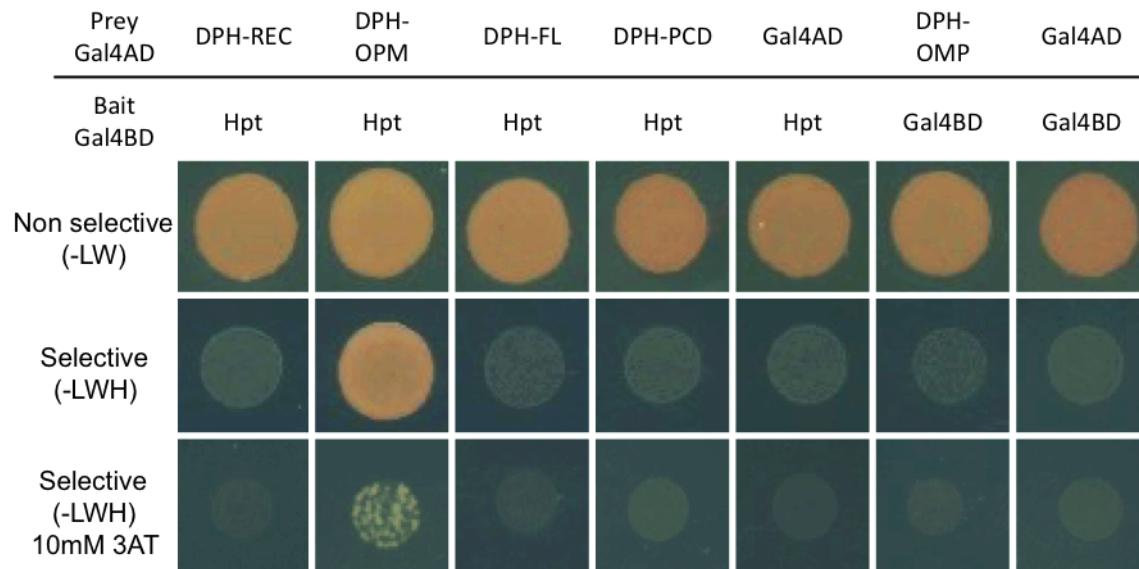


Figure 10. DPH-OPM interacts with PtHpt in Yeast Two Hybrid. Yeast spot growth on media selecting for the presence of 2 vectors (-LW) and on the protein interaction (-LWH). 3AT, which inhibits HIS3, was added to test the strength of the Hpt-DPH-OPM interaction. Last 3 columns are negative controls to test for prey or bait autoactivation.

I then turned towards a candidate gene approach. By looking for Histidine phosphotransferase protein in diatom genomes, we found only one protein with the Hpt domain (PF01627) in *P. tricornutum*. This protein is present in one copy in the diatom *Thalassiosira pseudonana*, but also in diatoms with no phytochrome (*Thalassiosira oceanica*, *Fragilariopsis cylindrus*). I tested the interaction of PtDPH with this protein by Y2H, using PtHpt as bait, fused to Gal4 binding domain, and PtDPH FL (full length), PCD (Photosensory domain), OPM (Output module) or just PtDPH REC (Response regulator) domain as preys (fused to Gal4 Activating domain, AD). I showed that PtHpt interacts with PtDPH OPM, but not with the other PtDPH fragments nor with the FL protein (Fig10). This is not surprising, as *Arabidopsis thaliana* PhyA and B FL protein interaction with SAP1 is light-dependent in Y2H, while *AthaPhyA* and B OPM recapitulates all interactions (Lu et al., 2015). We will test light-dependent interaction by adding BV chromophore to the media and grow yeast strains under different lights. We did not detect the interaction with

the reverse combination (Hpt-AD with PtDPH FL, PCD, OPM or REC-BD), maybe because of the absence of a nuclear localization signal in the GAL4BD fusion.

I also performed Y2H screen with PtHpt as bait, but once again we were unsuccessful in re-transforming with the candidate plasmids. However, Pt genome contains only a limited number of proteins with a response regulator (RR) (16 proteins, Table 3) that could be the target of Hpt, and which could be tested by direct interaction tests as done here for Hpt. Two of these proteins are especially interesting, as they combine RR and the LuxR transcriptional regulator domain, which binds DNA (Table 3). However, in fungi the response regulator element is not the transcription factor and the end of the signaling cascade, which involves also a MAPKinase cascade. The signaling cascade for PtDPH signaling could therefore be longer than 2 steps. We are currently trying to confirm the involvement of PtHpt in DPH signaling by generating PtHpt KO by CRISPR-Cas9 in *P. tricornutum*. In addition, we will test the phosphorelay between PtDPH and PtHpt *in vitro*. Alternatively, we could also start with the promoters of DPH-regulated genes, to analyze eventual binding motives, test the minimal promoter size *in vivo* with the reporter system used in Chapter 2, and perform a yeast one hybrid screen to identify transcription factors binding this promoter.

We have uncovered here a first good candidate as component of the signaling cascade initiated by DPH, even if work is needed to confirm it. Interestingly, Hpt is conserved across diatoms that do not possess DPH, suggesting that it could work as a signaling hub integrating different signals.

Table 3. Proteins with Response regulators domains in *P.tricornutum* genome (DPH not included)

Phatr3 ID	PFAM Domains
Phatr3_EG00123	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase
Phatr3_EG01983	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase;CHASE domain
Phatr3_EG02384	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain
Phatr3_EG02387	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain
Phatr3_J11994	Response regulator receiver domain;Bacterial regulatory proteins, luxR family
Phatr3_J12051	Response regulator receiver domain;Bacterial regulatory proteins, luxR family
Phatr3_J13255	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase
Phatr3_J16861	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase;GAF domain
Phatr3_J45485	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase;PAS domain
Phatr3_J46628	Response regulator receiver domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase
Phatr3_J47689	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;PAS fold;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase;PAS fold
Phatr3_J47930	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain
Phatr3_J55037	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;PAS fold;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase;PAS domain;PAS domain
Phatr3_J55187	Response regulator receiver domain;His Kinase A (phospho-acceptor) domain;Histidine kinase-, DNA gyrase B-, and HSP9NA-like ATPase
Phatr3_J8970	Response regulator receiver domain;Leucine Rich repeat

Note: I identified these proteins as interesting candidates to look at in the new database for diatom gene expression, *Diatomicbase*, for which a manuscript is in preparation

## CO-REGULATION OF DPH-REGULATED GENES

We used the reporter strains from Chapter 2 to explore the regulation of DPH targeted genes by other signals. While setting up the conditions for the action spectra, we found that the YFP signal varies depending on the cell concentration (Chapter2). YFP signal per cell increased when cell concentration increased in ProHsf6.4a ::YFP reporter lines, when grown in green light without agitation (Fig11A). This effect appeared specific for *Hsf4.6a* promoter, as lines containing the YFP under the control of other promoters (ProLHCF2 (light harvesting antenna protein gene) and the ProH4 (histone H4)) did not show this kind of regulation. To understand how this effect affects DPH-dependent gene regulation,

we grew cells without agitation, and under green light supplemented or not with far-red light (10  $\mu\text{mol photons/m}^2/\text{s}$ ). At very low cell concentration, the YFP signal was very low in both green and green plus far-red. From  $1\text{e}^5\text{Cells/mL}$ , YFP signal in green-light grown cells increased. Adding FR light increases the YFP signal, and the FR-induction itself seems to be cell-concentration dependent (Fig11B). The effect of cell concentration on cells grown in green light was also seen in PtDPH KO strains (Fig11C), implying that this effect is not DPH dependent. However, when cells were agitated (under green light), the increase in YFP with cell concentration was abolished (Fig11C; also true for the Wt background, data not shown).

Therefore, it seems that a signal other than light also regulates *Hsf4.6a* promoter. It could be oxygen or  $\text{CO}_2$  concentration in the media, as this effect appears when cell concentration is high in non-agitated cultures. Further investigations need to be done to decipher the different signals co-acting with PtDPH.

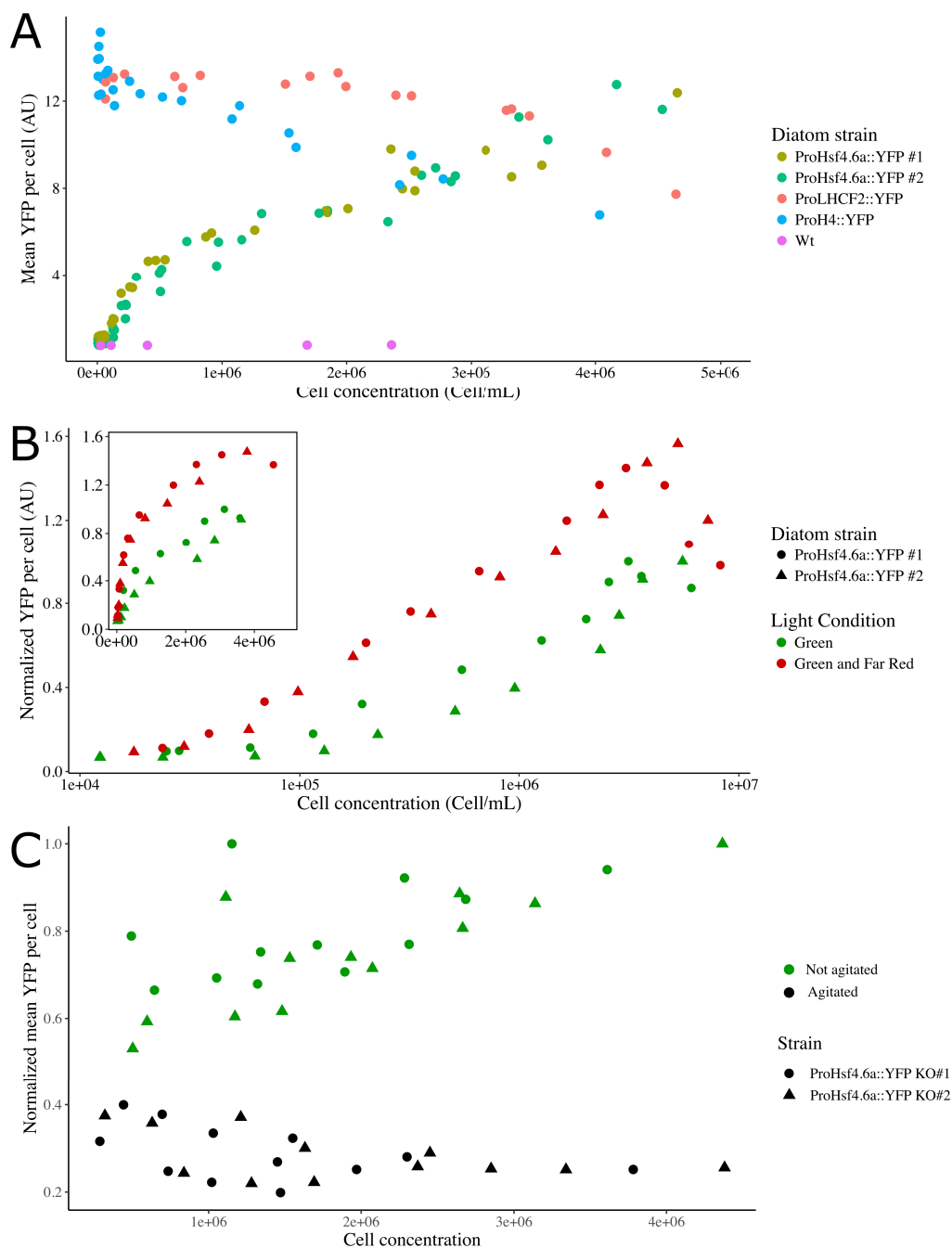


Figure 11 : Regulation of ProHsf4.6a by light, cell concentration and agitation. A. YFP expression under different promoters at different cell concentrations. Mean YFP per cell in *P. tricornutum* lines expressing the YFP under the control of different promoters during growth in constant green light. LHCF2: Light harvesting protein, H4: Histone H4. B. Effect of cell concentration of the FR induction of ProHsf4.6a ::YFP (inset, same data on a non-logarithmic scale). YFP was normalized to the maximum levels observed under green light. No agitation was applied to the cultures C. Effect of agitation on ProHsf4.6a activity. DPH KO strains were grown in green light with or without agitation, and YFP levels are normalized to the maximum observed in green light without agitation.



## CONCLUSION

In this chapter, we have gained insights into phytochrome signaling mechanisms. We propose that DPH binds BV produced by HO2, but that the chromophore is not further reduced by a BVR. DPH seems to be localized in the cytoplasm whatever its activation state, and interacts with the histidine phosphotransferase PtHpt, probably triggering a two-component phosphorelay. As PtHpt is conserved among diatoms, it could be a signaling hub integrating other signals such as the “cell concentration” signal, probably a gas pressure in the media, that regulates DPH-regulated genes. Overall, this resembles the situation in Fungi, where BV-binding phytochromes trigger a two-component signalization system that is also activated by osmolarity changes. As DPH and Fph are evolutionary close, these organisms might have found similar ways to integrate phytochromes in their cellular signalization network.

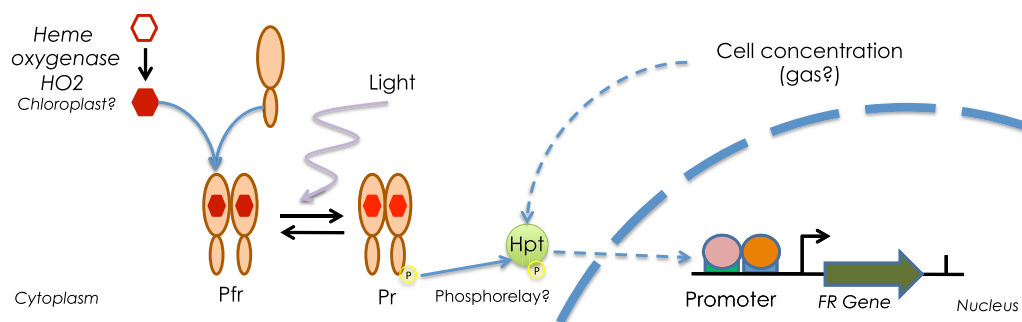


Figure 12 Proposition of DPH signaling mechanisms in *P. tricornutum* (cell concentration could act on Hpt, but also on independent signalisation pathways, and directly on DPH too ; Hpt can shuttle to the nucleus, but is could also be one of the downstream components.

## MATERIAL AND METHODS

### Diatom culture conditions

Reporter lines were the same as in Chapter 2. Cells were grown at 19°C under green light (530nm, 25µmol photon/m<sup>2</sup>/s) with or without agitation and far-red light (740nm at 10

$\mu\text{mol photon/m}^2/\text{s}$ ) was added for 24h. YFP fluorescence was measured with a MACSQuant®Analyser (Myltenyi Biotech) flow cytometer with excitation by a 488nm laser and detection at 500-550nm emission filters.

### **Phytochrome expression, purification and spectral analysis**

Photosensory modules (PSM) of DPH genes were obtained from previous studies (PtDPH (Fortunato et al, 2016), TweiDPH from chapter 2). Sequences were cloned as indicated in the Table 3, into pET28a-HO vector generated by inserting the *Synechocystis* Heme oxygenase gene amplified from pKS270 vector (Mukougawa et al., 2006) with the primer pair SynHO1xpET.HindIII.Fw and SynHO1xpET.NotI.rv, in HindIII/NotI in the pET28a vector. PtHO were cloned in Sall/NotI in the pET28a vector (see fig x and table 3 for the primers used, PtHO2, PtHO3 and PtHO4 were cloned without their peptide signal) ; PtBVR were cloned in NotI/XhoI in the pET28a vector (see Table S1)

PSM sequences were expressed as N-terminal His6 tagged proteins in *E. coli* BL21 (DE3) strain, by auto-induction system (Studier, 2005). Recombinant proteins were purified as in Fortunato et al. (2016). Absorption spectra of the recombinant proteins were measured after purification on a Varian's Cary-50 spectrophotometer. Illumination with LEDs at 810 and 630 nm were performed (approx. 1min illumination) to reach pure Pr spectra (after 810 illumination) or equilibriums between the Pr and Pfr forms.

### **Phytochrome localization**

Episomal vectors carrying fusion of PtDPH with fluorescent proteins were generated with the modular cloning system uLoop kit containing *P. tricornutum* regulatory sequences level 0 cloning block along with fluorescent reporter proteins and conjugative element (Pollak et al, 2019; 2020). PtDPH and PtCPF1 coding sequences were amplified by PCR from cDNA and using the primer pairs : PtDPH.D.Fw an PtDPH.E.Rv, PtDPH.C.Fw an PtDPH.D.Rv, and PtCPF1.D.Fw and PtCPF1.E.Rv, respectively, and cloned into pL0 receiver vector (Pollack et al, 2019). Each of these pL0 blocks were assembled with the FcpB promoter and terminator sequences, and the Venus or NeonGreen fluorescent proteins as 5' or 3' fusion, to generate level 1 expressing constructs. Finally, these expressing units were assembled with conjugative elements to express them from an episomal vector and the Bleomycin resistance cassette. these final constructs were transformed into *E. coli*

DH5alpha containing the pTA-MOB (plasmid encoding the machinery required for conjugal transfer), and delivered into *P. tricornutum* cells by conjugation according to the protocol described in Karas et al, 2015). Zeomycin 100µg/ml was used to select and grow conjugants.

Epifluorescence images were taken on an Axio Observer.Z1 inverted microscope (Zeiss) equipped with an ORCA-flash4.0 digital camera (Hamamatsu) and a Colibri.2 LED system (Zeiss) for excitation at 505 nm for Venus (520–550 nm emission) and 470 nm for chlorophyll autofluorescence (665–715 nm emission). Hoechst staining was performed by adding Hoeschst at a final concentration of 20µg/mL. To study eventual DPH re-localization, we illuminated the cells with FR light (765nm) while under the microscope.

### **Yeast two hybrid**

Phytochrome full length or domain (Photosensory Core Domain PCD, Output Domain (HK-ATPase-REC) or REC domain) sequences were amplified by PCR from plasmid containing the Pt-DPHPro:Pt-DPH-HA produced previously (Fortunato et al., 2016) and PtHpt was cloned from PtcDNAs with primers containing attB1 and attB2 GATEWAY recombination sites (see Table2). PCR products were introduced in the pDONR221 entry vector by BP GATEWAY recombination (Invitrogen). ENTRY clones were recombined in the pDEST32 vector (bait, for GAL4 DNA-binding domain N-terminal fusion, Invitrogen) or pDEST22 vector (prey, for GAL4 Activating domain N-terminal fusion, Invitrogen) by LR recombination. *Saccharomyces cerevisiae* PJ69-4a cells were transformed using the LiAc method (Gietz and Schiestl, 2007). Co-transformed yeast were selected on YNB media (recipe ?) lacking tryptophane and leucine, and interaction between the proteins was assessed by spotting on plates lacking tryptophane, leucine and histidine. 3-amino-1,2,4-triazole, which inhibits the HIS3 enzyme (that enables growth without histidine) was added at 10mM to test the strength of the interaction. Screens were performed by transforming yeast bait strain (ie transformed with just the bait vector) with the prey library obtained from (Huysman et al., 2013) and plated on selective -LWH media.

Table S1 Primers data used in this study

Name	Sequence
PtHpt_J33969.Gateway.Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGAATCCACCCGATGTGATCAATTG
PtHpt_J33969.Gateway.Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAACCGTAACAAAAAATTGTTTTTC
PtDPH_FL.Gateway.Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGAGCGGGGCAAATTATAGAG
PtDPH_PCD.Gateway.Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATGCCGCCAAATATTTCCGGTTT
PtDPH_OMP.Gateway.Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGGATTACTGAGAGGTGATA
PtDPH_REC.Gateway.Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGTCGGGTCCGATTTTGTATCG
PtDPH.Gateway.Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATTCGGGAACCCGCAGCCTC
PtHO2_J5851.Sall.fw	gcgGTCGACAGAAGGAGATATACATATGTTTCGACCCGGCCTTTTCGGTG
PtHO2_J5851.NotI.Rv	gcgGCGGCCGCTTAATTGTGGGGATGCTGAC
PtHO3_J35647.Sall.fw	gcgGTCGACAGAAGGAGATATACATATGCACTCGAACGACCCCCATG
PtHO3_J35647..NotI.Rv	gcgGCGGCCGCTCAAGATTTAATCATTCGTC
PtHO4_J5902.Sall.fw	gcgGTCGACAGAAGGAGATATACATATGTTTGCTCCTTCCGTCTCGAG
PtHO4_J5902.NotI.Rv	gcgGCGGCCGCTTACAATGCAACTCCAAGTTTTAG
J45446.NotI.fw:	cgcgccgcgagaaggagatataccatggcccgaaccaacacagcttttg
J45446.XhoI.rv:	ggcgctcgagctatggcgatgggaatagcac
J10830.NotI.Fw:	cgcgccgcgagaaggagatataccatgtacactaccggcaacagcggag
J10830.XhoI.rv	ggcgctcgagctaccgttgcgaagctgctac
J33770.NotI.Fw:	cgcgccgcgagaaggagatataccatggctccgcgctaccag
J33770.XhoI.Rv:	ggcgctcgagctacttttgcgagagtggaaag
J42659.NotI.Fw:	cgcgccgcgagaaggagatataccatgttccagcagctcctctttttgg
J42659.XhoI.Rv:	ggcgctcgagttaagcttctgttggccac
PtDPH.D.Fw	aggctgtctcgtctcgtctcaggtctcaaggtatgagcggggcaaattatagag
PtDPH.E.Rv	tggtaatctatgtgtcctgggtggtcttaagcctattccggaacccgcagcctc
PtDPH.C.Fw	aggctgtctcgtctcgtctcaggtctcaaatgagcggggcaaattatagag
PtDPH.D.Rv	tggtaatctatgtgtcctgggtggtctctaccttccggaacccgcagcctctg
PtCPF1.D.Fw	aggctgtctcgtctcgtctcaggtctcaaggtatggctaaatcggaagag
PtCPF1.E.Rv	tggtaatctatgtatcctgggtggtcttaagcttagttgcgacgttgcgc





# DISCUSSION

During this work, I have tried to get insights into diatom photoperception, and more specifically on diatom phytochromes.

Despite the availability of genomic data, there was no complete view of photoreceptors in diatoms. The genomes of the two diatom models, *T. pseudonana* and *P. tricornutum*, had been exploited to find photoreceptors, but the other diatom genomes had not been systematically mined for photoreceptors.

In the first chapter, I started to fill this gap by mining diatom genomic data for photoreceptors. I extended the search to all Ochrophyta, and included also other Stramenopiles to try to get insights into the conservation of these photoreceptors. We looked for known light-sensing domains like the LOV domain, cryptochrome photolyase-related domain, rhodopsins and phytochromes. Overall, it seems that diatoms, (but also more widely, Ochrophyta), possess a common set of blue light photoreceptors (aureochromes, cryptochromes) while some possess additional photoreceptors such as rhodopsins and phytochromes. The photoreceptor repertoire appears specific for an algal group or even down to a species, suggesting an adaptive value of the additional photoreceptors in specific environmental niches.

Some of the results confirmed already known photoreceptors distribution, such as the fact that all diatoms have 4 aureochromes. We also uncovered a new group of putative LOV-based photoreceptors that is quite conserved amongst Ochrophyta, including in diatoms, but is not found in *P. tricornutum*, showing the importance of looking in all available data. It was found here in *T. pseudonana*, but was not reported before.

Another example of new insights from the study described in chapter 1 is the CryP family of diatoms. CryP was reported recently in *P. tricornutum* as a photo-sensitive

protein with roles in both DNA repair and gene regulation. We highlighted here a new family, CryP2, which was already described in *T. pseudonana* (Oliveri et al., 2014; Fortunato et al., 2015), but it was not known if CryP2 was a species-specific light sensor, or something more conserved across diatoms. Based on the analysis in Chapter 1, it appears now as a broadly conserved CryP subgroup, probably representing a new type of photoreceptor with distinct properties.

We extended here the catalog for diatom photoreceptors, but we could also benefit from the environmental data to complete this view. Of course, the identified proteins are only candidate photoreceptors that would need to be biologically characterized.

Concerning phytochromes, we confirmed previous work (Fortunato et al., 2016), and better define the distribution and occurrence of DPH within centric and pennate diatoms. We did not resolve the origin of DPH, but mining other databases such as viral or bacterial sequences in addition to Ochrophyta and other eukaryotes might give new insights. Regarding centric group, investigations in genomic/transcriptomic resources (mostly from coastal species) indicate that DPH is found across nearly all Thalassiosirales species, suggesting that this is an ancient acquisition in all Thalassiosirales, with a few losses (for example in *Thalassiosira oceanica*). Thalassiosirales DPH was also found in the oceanic regions in the *Tara Oceans* dataset. Thalassiosirales are sister to a group of diatom containing *Minutocellus* species, which possess DPH, and *Chaetoceros* species, which do not. The absence of DPH in *Chaetoceros* species is well supported by the *Tara Oceans* data: centric DPH genes found in the environmental dataset were very close to *Thalassiosirales* or *Minutocellus* genes, and even in MAGs belonging to *Chaetoceros*, we did not find phytochrome genes. Of course MAGs are incomplete, and the other genomic data from Chapter 1 are mostly transcriptomes, but given the amount of transcriptomes in different species and different growth conditions, we could have expected DPH to be expressed in at least one of them. Looking for phytochrome in the *Chaetoceros* genomes recently published (Hongo et al., 2021; Kumazawa et al., 2022), and the new *Chaetoceros* genomes to come will definitely resolve the question of the absence of phytochromes in *Chaetoceros* species.



The more basal centric diatom species *Coscinodiscus* also possesses DPH, however this gene is incomplete and it is difficult to assess its exact position compared to other DPH. No DPH from this species was found in the environmental data.

Pennate DPH form a clade separated from centric DPH in phylogenetic studies, suggesting different origins, or an ancient origin in the common diatom ancestor followed by extensive gene loss, but also diversification and gene duplication. DPH was not found in the planktonic pennate genera *Pseudo-nitzschia* and *Fragilariospsis*. This is consistent with the quasi absence of pennate DPH in *Tara* oceans data, while pennate species are abundant in this dataset. Interestingly, the only 2 pennate genes found in the *Tara* oceans data are close to the two DPH copies of *Navicula sp.* (each *Tara* DPH is close to one copy of *Navicula* DPH, suggesting that these are the 2 copies from the same diatom species) and are found only in a few arctic stations. Some *Navicula* are known to thrive in sea ice, it could thus be considered that the DPH found in the *Tara* Oceans are from a sea-ice species released into the water column. The Dark Edge sampling cruise in October 2021 (coordinated by Prof. M. Babin), during which we sampled sea-ice and surrounding waters in the Arctic, might bring answer to this question.

Among pennate species with DPH there are well-known benthic species, such as the *Seminavis*, *Gyrosigma*, *Nitzschia*, *Navicula*, and *Amphora* genera. The araphid pennates *Fragilaria radians*, *Synedropsis recta* and *Synedra sp.* also possessed DPH, but we did not investigate their environmental niche yet. We also noticed that the benthic species often had several DPH copies (4 for *A. coffeaeformis* and *S. robsuta*, 2 for *Nitzschia* and *Navicula* species). *A. coffeaeformis* was reported before to have 2 putative chromophore-binding cysteines, one in the PAS and one in the GAF domain. However, we saw in chapter 1 that this seems to be a unique feature found in this species.

In the Ocean regions, the distribution of DPH showed a very strong latitudinal gradient, with absence in the tropics and presence in temperate and polar region of both Northern and Southern Hemispheres. In Tropical regions, diatoms, including Thalassiosirales, are found but no DPH could be detected.

It would be interesting to know if this is a feature specific for diatom phytochrome or if other marine organisms with phytochromes show the same trends. For example, the

cyanobacteria *Prochlorococcus* are abundant in Tropical regions and adapted to stratified waters, but do not possess phytochromes, while the *Synechococcus* cyanobacteria, which are more abundant in temperate regions, possess phytochromes. This could suggest that in cyanobacteria too, there is a latitudinal trend of phytochromes, maybe linked to adaptation to mixed waters such as the ones encountered in polar and temperate regions. An impressive set of data (*Tara* Oceans expedition, Global Ocean Sampling) are available for prokaryotes that could help to address this question. In addition, other Eukaryotic phytochromes found in the *Tara* Oceans data could extend this question to other eukaryotic microalgae, such as Cryptochytes or Glaucophytes.

This strong latitudinal gradient suggests that DPH function is linked to the high seasonal variability in these regions. This was also supported by the statistical analysis of the centric DPH abundance, which is linked to the variation of chlorophyll concentration in temperate regions. The variations of the light field and photoperiod in these regions are important, and in winter and early spring the phytoplankton experiences strong variations of light, from the surface to the bottom of the mixed layer, which can extend below the photic zone. In these conditions, the cells will spend most of their time in low blue light. The *Thalassiosirales* species seem to be early bloomers in the temperate species succession, which could be linked to an efficient adaptation to generally low but fluctuating light thanks to DPH. One important question raised here is the interplay with other photoreceptors. Indeed, we showed that in the marine environment, DPH activity is sensitive to blue, green and red light. Diatoms possess a plethora of blue-light photoreceptors, with 4 aureochromes and several light-sensing cryptochromes (Chapter 1), which makes it surprising that some diatoms have an additional blue-activated photoreceptor. The subtlety of DPH activity is that it is not only sensing the light intensity, but also measuring the ratio between blue and green to red wavebands. Some blue-light photoreceptors such as some aureochromes are thought to be involved in high light adaptation in diatoms, which diatoms experience only close to the surface. Thus, we can imagine reverse gradients of activity for blue light photoreceptors compared to phytochrome in the water column. The first ones would be active at the surface to shallow depth where (blue) light intensity is high, and their activity would decrease when going deeper due to decreasing light intensity. On the contrary, DPH is inactive at the surface,

and its activity increases when going deeper. The integration of DPH signal with blue light sensors signal could bring very precise information on the position in the water column in terms of light quality and intensity. This could also cross-talk with photosynthetic signals from the plastid and together mediate photosynthetic acclimation, such as low light acclimation at depth and photoprotection at surface. DPH KO strains in *Thalassiosira pseudonana* are available, and we could test if DPH mediates acclimation to light mimicking different environmental conditions, such as low blue light or fluctuating light.

We modeled that DPH activity was affected by phytoplankton concentration, such as the ones occurring during a bloom. DPH activity variation could therefore be a signal that a bloom is occurring, and mediate different physiological outputs such as nutrient uptake, resting spore formation or sexual reproduction (discussed in Chapter 2). The model diatom *Thalassiosira pseudonana* is sometimes used to study sexual reproduction, and we could also use another species which possesses DPH, such as *Skeletonema costatum*, to test the induction of sexual reproduction or the formation of resting spores under different light regimes, cell concentrations and nutrient status.

Finally, we showed that DPH activity increases when solar zenith angles are low, such as at dawn and dusk. DPH could thus be a sensor of the end (or beginning) of the day, but we miss critical information on the DPH state in the dark, i.e. its dark reversion rate. *In vitro*, no dark reversion could be observed for the DPH produced here nor in Fortunato et al., 2016. *In vivo*, our reporter system does not allow detection of inactivation, as the eYFP reporter is very stable in the cells. We made the hypothesis that PfrPfr is the dark state (and the form in which DPH is synthesized) and that dark reversion is negligible in the light. However, under photoperiodic conditions (i.e during the night or at sunrise/sunset), we do not know dynamics of the DPH equilibrium. On time scales such as 12h of darkness, dark reversion, synthesis and degradation of DPH are not negligible anymore. We know from tracking of tagged DPH that DPH synthesis is not stable during the day, with degradation of the protein during the light phase and re-accumulation during the night, and a maximum at the dark-light transition. This could point towards a role of DPH at dawn, where very sensitive DPH could detect the first photons and trigger acclimation to the incoming light. However, the tagged DPH might not completely reflect the native protein behaviour, and mRNA profiles have the reverse tendency: the transcription is at

its highest at the light to dark transition. If the DPH dark reversion is low, DPH could be activated at the end of the day and stay activated for a long time during the night, pointing towards a role of DPH in the dark. As DPH shows a strong latitudinal distribution, we could wonder if DPH could be linked to the high variability of daylength with the seasons. We could investigate growth of DPH KO under different photoperiods, or the DPH expression pattern in Wt strains under variable photoperiods. The implication of DPH in the regulation of rhythmic processes could also be tested.

We have raised several hypotheses on DPH role in planktonic diatoms, and also have the tools to test them, as we have in our hand DPH knockout for centric *T. pseudonana*. We mentioned above some phenotypes we could test. In addition, we will characterize the transcriptomic response of *T. pseudonana* to red, far-red and blue to determine the DPH-specific response and give hints on the role of this photoreceptor. The comparison with *P. tricornutum* genes regulated by DPH, established in Fortunato et al., 2016, will indicate the conservation of DPH regulation among diatoms.

We have described here the possible DPH-regulated phenomenon in the water column, where all DPHs characterized share a very conserved absorption spectra. However, we saw some differences in the DPH from benthic species, suggesting a specific adaptation to the benthic environment.

The 4 *Amphora coffeaeformis* DPH showed some different light sensing properties, suggesting spectral tuning and neo-functionalization. The different DPH may respond to subtle differences in the quality of light in their benthic environment rich in red and far-red light. We did not project pennate DPH activity in the benthic environment yet, but we can predict that DPH activity will increase under the sediments especially under phototrophic communities absorbing in blue and red light. Indeed, the light field under sediments is enriched in red and far-red light, and in the presence of photosynthetic environments the red is also quickly depleted, resembling the situation under a plant canopy in terrestrial environments.

These new information lead us to investigate the possible DPH-related phenotypes in pennate species in a light field mimicking the benthic environment, i.e. in red and far-red lights. We started with *P. tricornutum*, which is proposed to be a benthic species that can

be recruited in the plankton (See Introduction). We saw that DPH could regulate the amount of F710, a marker for photosynthetic red light acclimation (Herbstová et al., 2015), in red and far-red light. In the sediments, both light quality and low intensity, in addition to high cell density and special O<sub>2</sub>/dissolved inorganic carbon conditions could lead to the appearance of F710. The fitness advantage of F710 is not established yet, but extending the light absorption capacities to longer wavelength that are more abundant under the sediments might provide a growth advantage.

We also showed that this extends beyond *P. tricornutum*, as other benthic diatoms showed the same response, but not the planktonic species. Interestingly, all these species are suspected to possess DPH, suggesting that DPH could regulate F710 in these species too. In addition, we have mentioned earlier that benthic diatoms possess several DPH, with possibly different functions, probably fine-tuning adaptation to red and far-red light in sediments.

These exciting new results need to be further confirmed, both in characterization of F710, its regulation by DPH in *P. tricornutum* and other species, and possible advantage in an environmental context. We also raised hypotheses concerning the role of different DPH copies in benthic diatoms, and could test the effect of different lights on phototaxy and carbon metabolism. Environmental data, such as species inventory (or even metatranscriptomic data, but to our knowledge no such data currently exist for Eukaryotes) and physico-chemical conditions in sediments would be very useful to consolidate our hypotheses.

We also explored the DPH light-sensing and signaling pathways in *P. tricornutum*. One of the most persistent questions about the absorption spectra of DPH is whether the chromophore bound *in vivo* is BV or something else. In chapter 4, we found enzymes from *P. tricornutum* that were functional in bacteria for production of BV and an other chromophore. The implications for DPH chromophore *in vivo* remain to be addressed precisely.

I also addressed the question of DPH signaling cascade. I identified the first component of the signaling cascade, Hpt, which suggests a two-component phosphorelay system, as in bacteria and fungi. We also saw that TpDPH complements PtDPH KO, which

point towards the same signalization mechanisms, and the signalization cascade involving Hpt is probably conserved. Interestingly, in Fungi the same cascade is used for the response to osmolarity and to light through phytochromes. Here we identified cell concentration and agitation, so probably gas pressure, as other parameters regulating DPH-dependent genes. Hpt is present in other diatoms lacking DPH, suggesting that Hpt signalization is a conserved pathway in diatoms and integrates signals other than DPH activation. The *in vivo* reporter strains used to determine DPH action spectra are also good tools to study the cross-talk of DPH-mediated light inputs with other signals.

In this work, I studied DPH spanning different scales: from the protein *in silico*, to the protein *in vitro*, to the protein *in vivo* in diatoms, to the diatom life in lab conditions and ultimately to the diatom in its environment. I took advantage from the diatom incredible diversity to get new insights on their light sensing abilities and their adaptation to different light fields in specific environmental niches, hopefully bringing new knowledge for photobiology research.







# BIBLIOGRAPHY

- Ait-Mohamed O, Novák Vanclová AMG, Joli N, Liang Y, Zhao X, Genovesio A, Tirichine L, Bowler C, Dorrell RG** (2020) PhaeoNet: A Holistic RNAseq-Based Portrait of Transcriptional Coordination in the Model Diatom *Phaeodactylum tricornutum*. *Front Plant Sci* **11**: 590949
- Akiyama Y, Nakasone Y, Nakatani Y, Hisatomi O, Terazima M** (2016) Time-Resolved Detection of Light-Induced Dimerization of Monomeric Aureochrome-1 and Change in Affinity for DNA. *J Phys Chem B* **120**: 7360–7370
- Allen AE, Dupont CL, Oborník M, Horák A, Nunes-Nesi A, McCrow JP, Zheng H, Johnson DA, Hu H, Fernie AR, et al** (2011) Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature* **473**: 203–207
- Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H** (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**: 420–423
- Alverson AJ, Beszteri B, Julius ML, Theriot EC** (2011) The model marine diatom *Thalassiosira pseudonana* likely descended from a freshwater ancestor in the genus *Cyclotella*. *BMC Evol Biol* **11**: 125
- Amato A** (2010) Diatom Reproductive Biology: Living in a Crystal Cage. *The International Journal of Plant Reproductive Biology* **10**
- An SM, Choi DH, Noh JH** (2020) High-throughput sequencing analysis reveals dynamic seasonal succession of diatom assemblages in a temperate tidal flat. *Estuarine, Coastal and Shelf Science* **237**: 106686
- Annunziata R, Ritter A, Fortunato AE, Manzotti A, Cheminant-Navarro S, Agier N, Huysman MJJ, Winge P, Bones AM, Bouget F-Y, et al** (2019) bHLH-PAS protein RITMO1 regulates diel biological rhythms in the marine diatom *Phaeodactylum tricornutum*. *Proc Natl Acad Sci U S A* **116**: 13137–13142
- Apt KE, Grossman AR, Kroth-Pancic PG** (1996) Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Molec Gen Genet* **252**: 572–579
- Archibald JM** (2015) Endosymbiosis and eukaryotic cell evolution. *Curr Biol* **25**: R911–921
- Armbrust C** (1990) Role of light and the cell cycle on the induction of spermatogenesis in a centric diatom.
- Armbrust EV** (2009) The life of diatoms in the world's oceans. *Nature* **459**: 185–192
- Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, et al** (2004) The Genome of the Diatom *Thalassiosira Pseudonana*: Ecology, Evolution, and Metabolism. *Science (New York, NY)* **306**: 79–85
- Arrigo KR** (2016) Sea ice as a habitat for primary producers. *In* DN Thomas, ed, *Sea Ice*. John Wiley & Sons, Ltd, Chichester, UK, pp 352–369
- Ashworth J, Turkarslan S, Harris M, Orellana MV, Baliga NS** (2016) Pan-transcriptomic analysis identifies coordinated and orthologous functional modules in the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*. *Marine Genomics* **26**: 21–28

- Auldrige ME, Forest KT** (2011) Bacterial phytochromes: More than meets the light. *Critical Reviews in Biochemistry and Molecular Biology* **46**: 67–88
- Aumack CF, Juhl AR, Krembs C** (2014) Diatom vertical migration within land-fast Arctic sea ice. *Journal of Marine Systems* **139**: 496–504
- Azuma T, Pánek T, Tice AK, Kayama M, Kobayashi M, Miyashita H, Suzaki T, Yabuki A, Brown MW, Kamikawa R** (2022) An Enigmatic Stramenopile Sheds Light on Early Evolution in Ochrophyta Plastid Organellenogenesis. *Molecular Biology and Evolution* **39**: msac065
- Babenko I, Friedrich BM, Kröger N** (2022) Structure and Morphogenesis of the Frustule. *In* A Falciatore, T Mock, eds, *The Molecular Life of Diatoms*. Springer International Publishing, Cham, pp 287–312
- Bai Y, Rottwinkel G, Feng J, Liu Y, Lamparter T** (2016) Bacteriophytochromes control conjugation in *Agrobacterium fabrum*. *Journal of Photochemistry and Photobiology B: Biology* **161**: 192–199
- Bailleul B, Rogato A, de Martino A, Coesel S, Cardol P, Bowler C, Falciatore A, Finazzi G** (2010) An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proceedings of the National Academy of Sciences* **107**: 18214–18219
- Banerjee A, Herman E, Kottke T, Essen L-O** (2016a) Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**: 171–178
- Banerjee A, Herman E, Serif M, Maestre-Reyna M, Hepp S, Pokorny R, Kroth PG, Essen L-O, Kottke T** (2016b) Allosteric communication between DNA-binding and light-responsive domains of diatom class I aureochromes. *Nucleic Acids Res* **44**: 5957–5970
- Barnett A, Méléder V, Blommaert L, Lepetit B, Gaudin P, Vyverman W, Sabbe K, Dupuy C, Lavaud J** (2015) Growth form defines physiological photoprotective capacity in intertidal benthic diatoms. *ISME J* **9**: 32–45
- Barnett A, Méléder V, Dupuy C, Lavaud J** (2020) The Vertical Migratory Rhythm of Intertidal Microphytobenthos in Sediment Depends on the Light Photoperiod, Intensity, and Spectrum: Evidence for a Positive Effect of Blue Wavelengths. *Front Mar Sci* **7**: 212
- Bar-On YM, Phillips R, Milo R** (2018) The biomass distribution on Earth. *Proc Natl Acad Sci USA* **115**: 6506–6511
- Basu S, Patil S, Mapleson D, Russo MT, Vitale L, Fevola C, Maumus F, Casotti R, Mock T, Caccamo M, et al** (2017) Finding a partner in the ocean: molecular and evolutionary bases of the response to sexual cues in a planktonic diatom. *New Phytologist* **215**: 140–156
- Behrenfeld MJ, Halsey KH, Boss E, Karp-Boss L, Milligan AJ, Peers G** (2021) Thoughts on the evolution and ecological niche of diatoms. *Ecol Monogr*. doi: [10.1002/ecm.1457](https://doi.org/10.1002/ecm.1457)
- Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, et al** (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–1906
- Bhoo S-H, Davis SJ, Walker J, Karniol B, Vierstra RD** (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature* **414**: 776–779
- Bilcke G, Ferrante MI, Montresor M, De Decker S, De Veylder L, Vyverman W** (2022) Life Cycle Regulation. *Molecular Life of Diatoms*, Springer. pp 205–228

- Bilcke G, Osuna-Cruz CM, Santana Silva M, Poulsen N, D'hondt S, Bulankova P, Vyverman W, De Veylder L, Vandepoele K** (2021a) Diurnal transcript profiling of the diatom *Seminavis robusta* reveals adaptations to a benthic lifestyle. *Plant J* **107**: 315–336
- Bilcke G, Van Craenenbroeck L, Castagna A, Osuna-Cruz CM, Vandepoele K, Sabbe K, De Veylder L, Vyverman W** (2021b) Light intensity and spectral composition drive reproductive success in the marine benthic diatom *Seminavis robusta*. *Sci Rep* **11**: 17560
- Bína D, Herbstová M, Gardian Z, Vácha F, Litvín R** (2016) Novel structural aspect of the diatom thylakoid membrane: lateral segregation of photosystem I under red-enhanced illumination. *Sci Rep* **6**: 25583
- Blumenstein A, Vienken K, Tasler R, Purschwitz J, Veith D, Frankenberg-Dinkel N, Fischer R** (2005) The *Aspergillus nidulans* phytochrome FphA represses sexual development in red light. *Current Biology* **15**: 1833–1838
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otilar RP, et al** (2008) The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* **456**: 239–244
- Boylan MT, Quail PH** (1996) Are the phytochromes protein kinases? *Protoplasma* **195**: 12–17
- Brandt S, Von Stetten D, Günther M, Hildebrandt P, Frankenberg-Dinkel N** (2008) The fungal phytochrome FphA from *Aspergillus nidulans*. *Journal of Biological Chemistry* **283**: 34605–34614
- Bringloe TT, Starko S, Wade RM, Vieira C, Kawai H, De Clerck O, Cock JM, Coelho SM, Destombe C, Valero M, et al** (2020) Phylogeny and Evolution of the Brown Algae. *Critical Reviews in Plant Sciences* **39**: 281–321
- Brockmann J, Rieble S, Kazarinova-Fukshansky N, Seyfried M** (1987) Phytochrome behaves as a dimer in vivo. **8**
- Brown JS** (1967) Fluorometric evidence for the participation of chlorophyll a-695 in System 2 of photosynthesis. *BBA - Bioenergetics* **143**: 391–398
- Büchel C, Goss R, Bailleul B, Campbell DA, Lavaud J, Lepetit B** (2022) Photosynthetic Light Reactions in Diatoms. I. The Lipids and Light-Harvesting Complexes of the Thylakoid Membrane. *Molecular Life of Diatoms*, Springer. pp 397–422
- Burgie ES, Gannam ZTK, McLoughlin KE, Sherman CD, Holehouse AS, Stankey RJ, Vierstra RD** (2021) Differing biophysical properties underpin the unique signaling potentials within the plant phytochrome photoreceptor families. *Proc Natl Acad Sci USA* **118**: e2105649118
- Burki F, Kaplan M, Tikhonenkov DV, Zlatogursky V, Minh BQ, Radaykina LV, Smirnov A, Mylnikov AP, Keeling PJ** (2016) Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc R Soc B* **283**: 20152802
- Burki F, Roger AJ, Brown MW, Simpson AGB** (2020) The New Tree of Eukaryotes. *Trends in Ecology & Evolution* **35**: 43–55
- Caracciolo M, Rigaut-Jalabert F, Romac S, Mahe F, Forsans S, Gac J-P, Arsenieff L, Manno M, Chaffron S, Cariou T, et al** (2021) Seasonal temporal dynamics of marine protists communities in tidally mixed coastal waters. 2021.09.15.460302

- Carradec Q, Pelletier E, Da Silva C, Alberti A, Seeleuthner Y, Blanc-Mathieu R, Lima-Mendez G, Rocha F, Tirichine L, Labadie K, et al** (2018) A global oceans atlas of eukaryotic genes. *Nature Communications*. doi: [10.1038/s41467-017-02342-1](https://doi.org/10.1038/s41467-017-02342-1)
- Cartaxana P, Cruz S, Gameiro C, Kühl M** (2016a) Regulation of Intertidal Microphytobenthos Photosynthesis Over a Diel Emersion Period Is Strongly Affected by Diatom Migration Patterns. *Front Microbiol*. doi: [10.3389/fmicb.2016.00872](https://doi.org/10.3389/fmicb.2016.00872)
- Cartaxana P, Ribeiro L, Goessling J, Cruz S, Kühl M** (2016b) Light and O<sub>2</sub> microenvironments in two contrasting diatom-dominated coastal sediments. *Mar Ecol Prog Ser* **545**: 35–47
- Cartaxana P, Ruivo M, Hubas C, Davidson I, Serôdio J, Jesus B** (2011) Physiological versus behavioral photoprotection in intertidal epipelagic and epipsammic benthic diatom communities. *Journal of Experimental Marine Biology and Ecology* **405**: 120–127
- Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T, Brettel K, Essen L-O, van der Horst GTJ, Batschauer A, Ahmad M** (2011) The cryptochromes: blue light photoreceptors in plants and animals. *Annu Rev Plant Biol* **62**: 335–364
- Chen Y-R, Su Y -s., Tu S-L** (2012) Distinct phytochrome actions in nonvascular plants revealed by targeted inactivation of phytyl biosynthesis. *Proceedings of the National Academy of Sciences* **109**: 8310–8315
- Christie JM** (2007a) Phototropin Blue-Light Receptors. *Annual Review of Plant Biology* **58**: 21–45
- Christie JM** (2007b) Phototropin blue-light receptors. *Annual Review of Plant Biology* **58**: 21–45
- Coesel S, Mangogna M, Ishikawa T, Heijde M, Rogato A, Finazzi G, Todo T, Bowler C, Falcione A** (2009) Diatom PtCPF1 is a new cryptochrome/photolyase family member with DNA repair and transcription regulation activity. *EMBO Rep* **10**: 655–661
- Coesel SN, Durham BP, Groussman RD, Hu SK, Caron DA, Morales RL, Ribalet F, Armbrust EV** (2021) Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. *Proc Natl Acad Sci USA* **118**: e2011038118
- Cohn SA, Spurck TP, Pickett-Heaps JD** (1999) HIGH ENERGY IRRADIATION AT THE LEADING TIP OF MOVING DIATOMS CAUSES A RAPID CHANGE OF CELL DIRECTION. *Diatom Research* **14**: 193–206
- Cohn SA, Weitzell RE** (1996) ECOLOGICAL CONSIDERATIONS OF DIATOM CELL MOTILITY. I. CHARACTERIZATION OF MOTILITY AND ADHESION IN FOUR DIATOM SPECIES I. *J Phycol* **32**: 928–939
- Consalvey M, Paterson DM, Underwood GJC** (2004) THE UPS AND DOWNS OF LIFE IN A BENTHIC BIOFILM: MIGRATION OF BENTHIC DIATOMS. *Diatom Research* **19**: 181–202
- Consiglieri E, Gutt A, Gärtner W, Schubert L, Viappiani C, Abbruzzetti S, Losi A** (2019) Dynamics and efficiency of photoswitching in biliverdin-binding phytochromes. *Photochem Photobiol Sci* **18**: 2484–2496
- Costanzi S, Siegel J, Tikhonova IG, Jacobson KA** (2009) Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr Pharm Des* **15**: 3994–4002
- Crosson S, Rajagopal S, Moffat K** (2003) The LOV Domain Family: Photoresponsive Signaling Modules Coupled to Diverse Output Domains. *Biochemistry* **42**: 2–10

- Croteau D, Guérin S, Bruyant F, Ferland J, Campbell DA, Babin M, Lavaud J** (2021) Contrasting nonphotochemical quenching patterns under high light and darkness aligns with light niche occupancy in Arctic diatoms. *Limnol Oceanogr.* doi: [10.1002/lno.11587](https://doi.org/10.1002/lno.11587)
- Daase M, Berge J, Søreide JE, Falk-Petersen S** (2021) Ecology of Arctic Pelagic Communities. In DN Thomas, ed, *Arctic Ecology*, 1st ed. Wiley, pp 219–259
- Daboussi F, Leduc S, Maréchal A, Dubois G, Guyot V, Perez-Michaut C, Amato A, Falciaiore A, Juillerat A, Beurdeley M, et al** (2014) Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nature Communications* **5**: 1–7
- Dammeyer T, Frankenberg-Dinkel N** (2008) Function and distribution of bilin biosynthesis enzymes in photosynthetic organisms. *Photochem Photobiol Sci* **7**: 1121
- Davis SJ, Vener AV, Vierstra RD** (1999) Bacteriophytochromes: Phytochrome-Like Photoreceptors from Nonphotosynthetic Eubacteria. *Science* **286**: 2517–2520
- De Martino A, Meichenin A, Shi J, Pan K, Bowler C** (2007) Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions. *Journal of Phycology* **43**: 992–1009
- De Riso V, Raniello R, Maumus F, Rogato A, Bowler C, Falciaiore A** (2009) Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res* **37**: e96
- Delmont TO, Gaia M, Hinsinger DD, Fremont P, Vanni C, Guerra AF, Eren AM, Kourlaiev A, d'Agata L, Clayssen Q, et al** (2020) Functional repertoire convergence of distantly related eukaryotic plankton lineages revealed by genome-resolved metagenomics. doi: [10.1101/2020.10.15.341214](https://doi.org/10.1101/2020.10.15.341214)
- Di Roberto RB, Peisajovich SG** (2014) The role of domain shuffling in the evolution of signaling networks: SIGNALING NETWORKS EVOLUTION. *J Exp Zool (Mol Dev Evol)* **322**: 65–72
- Djouani-Tahri EB, Christie JM, Sanchez-Ferandin S, Sanchez F, Bouget FY, Corellou F** (2011) A eukaryotic LOV-histidine kinase with circadian clock function in the picoalga *Ostreococcus*. *Plant Journal* **65**: 578–588
- Dorrell R, Lui F, Bowler C** (2022) Reconstructing Dynamic Evolutionary Events in Diatom Nuclear and Organelle Genomes. *Molecular Life of Diatoms*, Springer. pp 147–177
- Dorrell RG, Gile G, McCallum G, Méheust R, Baptiste EP, Klinger CM, Brillet-Guéguen L, Freeman KD, Richter DJ, Bowler C** (2017) Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. *eLife* **6**: 1–45
- Dortch Q, Clayton JR, Thoresen SS, Ahmed SI** (1984) Species differences in accumulation of nitrogen pools in phytoplankton. *Marine Biology* **81**: 237–250
- Dröge-Laser W, Snoek BL, Snel B, Weiste C** (2018) The Arabidopsis bZIP transcription factor family — an update. *Current Opinion in Plant Biology* **45**: 36–49
- Duanmu D, Bachy C, Sudek S, Wong C-H, Jimenez V, Rockwell NC, Martin SS, Ngan CY, Reistetter EN, van Baren MJ, et al** (2014) Marine algae and land plants share conserved phytochrome signaling systems. *Proceedings of the National Academy of Sciences* **111**: 15827–15832
- Ehn JK, Mundy CJ** (2013) Assessment of light absorption within highly scattering bottom sea ice from under-ice light measurements: Implications for Arctic ice algae primary production. *Limnol Oceanogr* **58**: 893–902

- Eilertsen H, Sandberg S, Tøllefsen H** (1995) Photoperiodic control of diatom spore growth; a theory to explain the onset of phytoplankton blooms. *Mar Ecol Prog Ser* **116**: 303–307
- Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H** (2014) Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem Rev* **114**: 126–163
- Essen L-O, Franz S, Banerjee A** (2017) Structural and evolutionary aspects of algal blue light receptors of the cryptochrome and aureochrome type. *Journal of Plant Physiology* **217**: 27–37
- Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C** (1999) Transformation of Nonselectable Reporter Genes in Marine Diatoms. *Marine Biotechnology* **1**: 239–251
- Falciatore A, Jaubert M, Bouly J-P, Bailleul B, Mock T** (2020) Diatom Molecular Research Comes of Age: Model Species for Studying Phytoplankton Biology and Diversity. *Plant Cell* **32**: 547–572
- Fernández MB, Tossi V, Lamattina L, Cassia R** (2016) A Comprehensive Phylogeny Reveals Functional Conservation of the UV-B Photoreceptor UVR8 from Green Algae to Higher Plants. *Frontiers in Plant Science* **7**: 1–6
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P** (1998) Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* **281**: 237–240
- Finkel OM, Bèjà O, Belkin S** (2013) Global abundance of microbial rhodopsins. *ISME J* **7**: 448–451
- Fortunato AE, Annunziata R, Jaubert M, Bouly J-P, Falciatore A** (2015) Dealing with light: the widespread and multitasking cryptochrome/photolyase family in photosynthetic organisms. *J Plant Physiol* **172**: 42–54
- Fortunato AE, Jaubert M, Enomoto G, Bouly J-P, Raniello R, Thaler M, Malviya S, Bernardes JS, Rappaport F, Gentili B, et al** (2016) Diatom Phytochromes Reveal the Existence of Far-Red-Light-Based Sensing in the Ocean. *Plant Cell* **28**: 616–628
- Foster RA, Kuypers MMM, Vagner T, Paerl RW, Musat N, Zehr JP** (2011) Nitrogen fixation and transfer in open ocean diatom – cyanobacterial symbioses. *The ISME Journal* **5**: 1484–1493
- Frankenberg N, Mukougawa K, Kohchi T, Lagarias JC** (2001) Functional Genomic Analysis of the HY2 Family of Ferredoxin-Dependent Bilin Reductases from Oxygenic Photosynthetic Organisms. *Plant Cell* **13**: 965–978
- Franklin KA, Quail PH** (2010) Phytochrome functions in Arabidopsis development. *Journal of Experimental Botany* **61**: 11–24
- Froehlich AC, Noh B, Vierstra RD, Loros J, Dunlap JC** (2005) Genetic and Molecular Analysis of Phytochromes from the Filamentous Fungus *Neurospora crassa*. *Eukaryot Cell* **4**: 2140–2152
- Fujita Y, Ohki K** (2004) On the 710 nm Fluorescence Emitted by the Diatom *Phaeodactylum tricornutum* at Room Temperature. *Plant and Cell Physiology* **45**: 392–397
- Gabed N, Verret F, Peticca A, Kryvoruchko I, Gastineau R, Bosson O, Séveno J, Davidovich O, Davidovich N, Witkowski A, et al** (2022) What Was Old Is New Again: The Pennate Diatom *Haslea ostrearia* (Gaillon) Simonsen in the Multi-Omic Age. *Marine Drugs* **20**: 234
- Galachyants YP, Zakharova YuR, Petrova DP, Morozov AA, Sidorov IA, Marchenkov AM, Logacheva MD, Markelov ML, Khabudaev KV, Likhoshway YeV, et al** (2015) Sequencing of the complete genome of an araphid pennate diatom *Synedra acus* subsp. *radians* from Lake Baikal. *Doklady Biochemistry and Biophysics* **461**: 84–88

- Gan F, Bryant DA** (2015) Adaptive and acclimative responses of cyanobacteria to far-red light: Responses of cyanobacteria to far-red light. *Environ Microbiol* **17**: 3450–3465
- Gan F, Shen G, Bryant D** (2014) Occurrence of Far-Red Light Photoacclimation (FaRLiP) in Diverse Cyanobacteria. *Life* **5**: 4–24
- Garcia-Pichel F, Bebout B** (1996) Penetration of ultraviolet radiation into shallow water sediments: high exposure for photosynthetic communities. *Mar Ecol Prog Ser* **131**: 257–262
- George J, Kahlke T, Abbriano RM, Kuzhiumparambil U, Ralph PJ, Fabris M** (2020) Metabolic Engineering Strategies in Diatoms Reveal Unique Phenotypes and Genetic Configurations With Implications for Algal Genetics and Synthetic Biology. *Front Bioeng Biotechnol* **8**: 513
- Gietz RD, Schiestl RH** (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**: 31–34
- Giovagnetti V, Ruban AV** (2021) The mechanism of regulation of photosystem I cross-section in the pennate diatom *Phaeodactylum tricornutum*. *Journal of Experimental Botany* **72**: 561–575
- Giraud E, Fardoux J, Fourrier N, Hannibal L, Genty B, Bouyer P, Dreyfus B, Verméglio A** (2002) Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria. *Nature* **417**: 202–205
- Giraud E, Verméglio A** (2008) Bacteriophytochromes in anoxygenic photosynthetic bacteria. *Photosynthesis Research* **97**: 141–153
- Giraud E, Zappa S, Vuillet L, Adriano JM, Hannibal L, Fardoux J, Berthomieu C, Bouyer P, Pignol D, Verméglio A** (2005) A new type of bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in *Rhodospseudomonas palustris*. *Journal of Biological Chemistry* **280**: 32389–32397
- Glantz ST, Carpenter EJ, Melkonian M, Gardner KH, Boyden ES, Wong GKS, Chow BY** (2016) Functional and topological diversity of LOV domain photoreceptors. *Proceedings of the National Academy of Sciences of the United States of America* **113**: E1442–E1451
- Godhe A, Kremp A, Montresor M** (2014) Genetic and microscopic evidence for sexual reproduction in the centric diatom *Skeletonema marinoi*. *Protist* **165**: 401–416
- Goett-Zink L, Klocke JL, Bögeholz LAK, Kottke T** (2020) In-cell infrared difference spectroscopy of LOV photoreceptors reveals structural responses to light altered in living cells. *J Biol Chem* **295**: 11729–11741
- Goldman JC, McGillicuddy DJrJ** (2003) Effect of large marine diatoms growing at low light on episodic new production. *Limnol Oceanogr* **48**: 1176–1182
- Grossman AR** (2003) A molecular understanding of complementary chromatic adaptation. *Photosynthesis Research* **76**: 207–215
- Grote M, Engelhard M, Hegemann P** (2014) Of ion pumps, sensors and channels - perspectives on microbial rhodopsins between science and history. *Biochim Biophys Acta* **1837**: 533–545
- Gruber A, Rocap G, Kroth PG, Armbrust EV, Mock T** (2015) Plastid proteome prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant J* **81**: 519–528
- Gschloessl B, Guermeur Y, Cock JM** (2008) HECTAR: A method to predict subcellular targeting in heterokonts. *BMC Bioinformatics* **9**: 393

- Guérin N, Ciccarella M, Flamant E, Frémont P, Mangenot S, Istace B, Noel B, Romac S, Bachy C, Gachenot M, et al** (2021) Genomic adaptation of the picoeukaryote *Pelagomonas calceolata* to iron-poor oceans revealed by a chromosome-scale genome sequence. doi: [10.1101/2021.10.25.465678](https://doi.org/10.1101/2021.10.25.465678)
- Guita Niang, Hoebeke M, Meng A, Liu X, Scheremetjew M, Finn R, Pelletier E, Corre E** (2020) METdb: A GENOMIC REFERENCE DATABASE FOR MARINE SPECIES. doi: [10.7490/F1000RESEARCH.1118000.1](https://doi.org/10.7490/F1000RESEARCH.1118000.1)
- Haardt H, Nielsen Gae** (1980) Attenuation measurements of monochromatic light in marine-sediments. *oceanologica acta* **3**: 333–338
- Haimovich-Dayan M, Garfinkel N, Ewe D, Marcus Y, Gruber A, Wagner H, Kroth PG, Kaplan A** (2013) The role of C<sub>4</sub> metabolism in the marine diatom *P. haeodactylum tricorutum*. *New Phytol* **197**: 177–185
- Han X, Chang X, Zhang Z, Chen H, He H, Zhong B, Deng XW** (2019) Origin and Evolution of Core Components Responsible for Monitoring Light Environment Changes during Plant Terrestrialization. *Molecular Plant* **12**: 847–862
- Harnstrom K, Ellegaard M, Andersen TJ, Godhe A** (2011) Hundred years of genetic structure in a sediment revived diatom population. *Proceedings of the National Academy of Sciences* **108**: 4252–4257
- Hedtke M, Rauscher S, Röhrig J, Rodríguez-Romero J, Yu Z, Fischer R** (2015) Light-dependent gene activation in *Aspergillus nidulans* is strictly dependent on phytochrome and involves the interplay of phytochrome and white collar-regulated histone H3 acetylation. *Molecular Microbiology* **97**: 733–745
- Heintz U, Schlichting I** (2016) Blue light-induced LOV domain dimerization enhances the affinity of Aureochrome 1a for its target DNA sequence. *eLife*. doi: [10.7554/eLife.11860](https://doi.org/10.7554/eLife.11860)
- Hepp S, Trauth J, Hasenjäger S, Bezold F, Essen L-O, Taxis C** (2020) An Optogenetic Tool for Induced Protein Stabilization Based on the *Phaeodactylum tricorutum* Aureochrome 1a Light-Oxygen-Voltage Domain. *J Mol Biol* **432**: 1880–1900
- Herbstová M, Bína D, Kaňa R, Vácha F, Litvín R** (2017) Red-light phenotype in a marine diatom involves a specialized oligomeric red-shifted antenna and altered cell morphology. *Sci Rep* **7**: 11976
- Herbstová M, Bína D, Koník P, Gardian Z, Vácha F, Litvín R** (2015) Molecular basis of chromatic adaptation in pennate diatom *Phaeodactylum tricorutum*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**: 534–543
- Herman E, Kottke T** (2015) Allosterically regulated unfolding of the A'α helix exposes the dimerization site of the blue-light-sensing aureochrome-LOV domain. *Biochemistry* **54**: 1484–1492
- Herman E, Sachse M, Kroth PG, Kottke T** (2013) Blue-light-induced unfolding of the Jα helix allows for the dimerization of aureochrome-LOV from the diatom *Phaeodactylum tricorutum*. *Biochemistry* **52**: 3094–3101
- Hisatomi O, Furuya K** (2015) A light-regulated bZIP module, photozipper, induces the binding of fused proteins to the target DNA sequence in a blue light-dependent manner. *Photochem Photobiol Sci* **14**: 1998–2006
- Hongo Y, Kimura K, Takaki Y, Yoshida Y, Baba S, Kobayashi G, Nagasaki K, Hano T, Tomaru Y** (2021) The genome of the diatom *Chaetoceros tenuissimus* carries an ancient integrated fragment of an extant virus. *Sci Rep* **11**: 22877
- Hopes A, Nekrasov V, Kamoun S, Mock T** (2016) Editing of the urease gene by CRISPR-Cas in the diatom *Thalassiosira pseudonana*. *Plant Methods* **12**: 1–12



- Huang K, Beck CF** (2003) Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci USA* **100**: 6269–6274
- Hughes J, Lamparter T** (1999) Prokaryotes and Phytochrome. The Connection to Chromophores and Signaling. *Plant Physiology* **121**: 1059–1068
- Hughes J, Lamparter T, Mittmann F, Hartmann E, Gärtner W, Wilde A, Börner T** (1997) A prokaryotic phytochrome. *Nature* **386**: 663–663
- Huysman MJJ, Fortunato AE, Matthijs M, Costa BS, Vanderhaeghen R, Van den Daele H, Sachse M, Inzé D, Bowler C, Kroth PG, et al** (2013a) AUREOCHROME1a-Mediated Induction of the Diatom-Specific Cyclin *dsCYC2* Controls the Onset of Cell Division in Diatoms (*Phaeodactylum tricornutum*). *The Plant Cell* **25**: 215–228
- Huysman MJJ, Fortunato AE, Matthijs M, Costa BS, Vanderhaeghen R, Van den Daele H, Sachse M, Inzé D, Bowler C, Kroth PG, et al** (2013b) AUREOCHROME1a-Mediated Induction of the Diatom-Specific Cyclin *dsCYC2* Controls the Onset of Cell Division in Diatoms (*Phaeodactylum tricornutum*). *The Plant Cell* **25**: 215–228
- Ishikawa M, Takahashi F, Nozaki H, Nagasato C, Motomura T, Kataoka H** (2009) Distribution and phylogeny of the blue light receptors aureochromes in eukaryotes. *Planta* **230**: 543–552
- Jackson C, Knoll AH, Chan CX, Verbruggen H** (2018) Plastid phylogenomics with broad taxon sampling further elucidates the distinct evolutionary origins and timing of secondary green plastids. *Sci Rep* **8**: 1523
- Jaubert M, Bouly JP, Ribera d’Alcalà M, Falciatore A** (2017) Light sensing and responses in marine microalgae. *Current Opinion in Plant Biology* **37**: 70–77
- Jaubert M, Lavergne J, Fardoux J, Hannibal L, Vuillet L, Adriano J-M, Bouyer P, Pignol D, Giraud E, Verméglio A** (2007) A Singular Bacteriophytochrome Acquired by Lateral Gene Transfer. *Journal of Biological Chemistry* **282**: 7320–7328
- Jenkins GI** (2014) The UV-B Photoreceptor UVR8: From Structure to Physiology. *Plant Cell* **26**: 21–37
- Jin X, Gruber N, Dunne JP, Sarmiento JL, Armstrong RA** (2006) Diagnosing the contribution of phytoplankton functional groups to the production and export of particulate organic carbon, CaCO<sub>3</sub>, and opal from global nutrient and alkalinity distributions: DIAGNOSING PHYTOPLANKTON FUNCTIONAL GROUPS. *Global Biogeochem Cycles* **20**: n/a-n/a
- Johansson ON, Töpel M, Pinder MIM, Kourtchenko O, Blomberg A, Godhe A, Clarke AK** (2019) *Skeletonema marinoi* as a new genetic model for marine chain-forming diatoms. *Scientific Reports* **9**: 5391
- Johnson LK, Alexander H, Brown CT** (2019) Re-assembly, quality evaluation, and annotation of 678 microbial eukaryotic reference transcriptomes. *GigaScience*. doi: [10.1093/gigascience/giy158](https://doi.org/10.1093/gigascience/giy158)
- Juhas M, Von Zadow A, Spexard M, Schmidt M, Kottke T, Büchel C** (2014a) A novel cryptochrome in the diatom *Phaeodactylum tricornutum* influences the regulation of light-harvesting protein levels. *FEBS Journal* **281**: 2299–2311
- Juhas M, von Zadow A, Spexard M, Schmidt M, Kottke T, Büchel C** (2014b) A novel cryptochrome in the diatom *Phaeodactylum tricornutum* influences the regulation of light-harvesting protein levels. *FEBS J* **281**: 2299–2311
- Jung J-H, Domijan M, Klose C, Biswas S, Ezer D, Gao M, Khattak AK, Box MS, Charoensawan V, Cortijo S, et al** (2016) Phytochromes function as thermosensors in *Arabidopsis*. *Science* **354**: 886–889

- Jungandreas A, Schellenberger Costa B, Jakob T, von Bergen M, Baumann S, 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
- Jupin H, Giraud G** (1971) MODIFICATION DU SPECTRE D'ABSORPTION DANS LE ROUGE LOINTAIN D'UNE DIATOMÉE CULTIVÉE EN LUMIÈRE ROUGE. *Biochimica et Biophysica Acta* **226**: 98–102
- Kamp A, de Beer D, Nitsch JL, Lavik G, Stief P** (2011) Diatoms respire nitrate to survive dark and anoxic conditions. *Proceedings of the National Academy of Sciences* **108**: 5649–5654
- 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**: 1–10
- Karniol B, Wagner JR, Walker JM, Vierstra RD** (2005) Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors. *Biochemical Journal* **392**: 103–116
- Keeling PJ** (2013) The Number, Speed, and Impact of Plastid Endosymbioses in Eukaryotic Evolution. *Annu Rev Plant Biol* **64**: 583–607
- Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, Armbrust EV, Archibald JM, Bharti AK, Bell CJ, 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 Biology*. doi: [10.1371/journal.pbio.1001889](https://doi.org/10.1371/journal.pbio.1001889)
- Kehoe DM, Grossman AR** (1996) Similarity of a Chromatic Adaptation Sensor to Phytochrome and Ethylene Receptors. *Science* **273**: 1409–1412
- Kemp AES, Pike J, Pearce RB, Lange CB** (2000) The 'Fall dump' \* a new perspective on the role of a 'shade' oraa in the annual cycle of diatom production and export #ux. 26
- Kemp AES, Villareal TA** (2018) The case of the diatoms and the muddled mandalas: Time to recognize diatom adaptations to stratified waters. *Progress in Oceanography* **167**: 138–149
- Kerruth S, Ataka K, Frey D, Schlichting I, Heberle J** (2014) Aureochrome 1 illuminated: structural changes of a transcription factor probed by molecular spectroscopy. *PLoS One* **9**: e103307
- Kirk JTO** (2011) *Light and photosynthesis in aquatic ecosystems*, 3rd Editio. Cambridge University Press. doi: [10.1017/S0025315400044180](https://doi.org/10.1017/S0025315400044180)
- Kleine T, Lockhart P, Batschauer A** (2003) An Arabidopsis protein closely related to *Synechocystis* cryptochrome is targeted to organelles: Organelle-targeted cryptochrome. *The Plant Journal* **35**: 93–103
- Klose C, Venezia F, Hussong A, Kircher S, Schäfer E, Fleck C** (2015) Systematic analysis of how phytochrome B dimerization determines its specificity. *Nature Plants* **1**: 1–8
- Kobayashi I, Nakajima H, Hisatomi O** (2020) Molecular Mechanism of Light-Induced Conformational Switching of the LOV Domain in Aureochrome-1. *Biochemistry* **59**: 2592–2601
- Koester JA, Brawley SH, Karp-Boss L, Mann DG** (2007) Sexual reproduction in the marine centric diatom *Ditylum brightwellii* (Bacillariophyta). *European Journal of Phycology* **42**: 351–366

- König S, Eisenhut M, Bräutigam A, Kurz S, Weber APM, Büchel C** (2017) The influence of a cryptochrome on the gene expression profile in the diatom *Phaeodactylum tricornutum* under blue light and in darkness. *Plant Cell Physiol* **58**: 1914–1923
- Kooistra WHCF, Gersonde R, Medlin LK, Mann D** (2007) The origin and evolution of the diatoms: their adaptation to a planktonic existence. In: Falkowski PG, Knoll AH, eds. *Evolution of planktonic photoautotrophs*. Burlington, MA: Academic Press. doi: [10.1016/B978-0-12-370518-1.50012-6](https://doi.org/10.1016/B978-0-12-370518-1.50012-6)
- Kooß S, Lamparter T** (2017) Cyanobacterial origin of plant phytochromes. *Protoplasma* **254**: 603–607
- Kroth PG, Matsuda Y** (2022) *Carbohydrate Metabolism. Molecular Life of Diatoms*, Springer. pp 465–492
- Kroth PG, Wilhelm C, Kottke T** (2017) An update on aureochromes: Phylogeny - mechanism - function. *J Plant Physiol* **217**: 20–26
- Kühl M, Jørgensen BB** (1992) Spectral light measurements in microbenthic phototrophic communities with a fiber-optic microprobe coupled to a sensitive diode array detector. *Limnology and Oceanography* **37**: 1813–1823
- Kühl M, Jørgensen BB** (1994) The light field of microbenthic communities: Radiance distribution and microscale optics of sandy coastal sediments. *Limnol Oceanogr* **39**: 1368–1398
- Kühl M, Lassen C, Jørgensen B** (1994) Light penetration and light intensity in sandy marine sediments measured with irradiance and scalar irradiance fiber-optic microprobes. *Mar Ecol Prog Ser* **105**: 139–148
- Kumazawa M, Nishide H, Nagao R, Inoue-Kashino N, Shen J, Nakano T, Uchiyama I, Kashino Y, Ifuku K** (2022) Molecular phylogeny of fucoxanthin-chlorophyll *a / c* proteins from *Chaetoceros gracilis* and Lhcq/Lhcf diversity. *Physiologia Plantarum*. doi: [10.1111/ppl.13598](https://doi.org/10.1111/ppl.13598)
- Kume K, Zylka MJ, Sriram S, Shearman LP, Weaver DR, Jin X, Maywood ES, Hastings MH, Reppert SM** (1999) mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* **98**: 193–205
- Lagarias JC, Rapoport H** (1980) Chromopeptides from phytochrome. The structure and linkage of the PR form of the phytochrome chromophore. *J Am Chem Soc* **102**: 4821–4828
- Lamparter T, Michael N, Mittmann F, Esteban B** (2002) Phytochrome from *Agrobacterium tumefaciens* has unusual spectral properties and reveals an N-terminal chromophore attachment site. *Proc Natl Acad Sci USA* **99**: 11628–11633
- Lassen C, Ploug H, Jørgensen BB** (1992) A fibre-optic scalar irradiance microsensor: application for spectral light measurements in sediments. *FEMS Microbiology Ecology* **9**: 247–254
- Lavaud J, Lepetit B** (2013) An explanation for the inter-species variability of the photoprotective non-photochemical chlorophyll fluorescence quenching in diatoms. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1827**: 294–302
- Lavaud J, Strzepek RF, Kroth PG** (2007) Photoprotection capacity differs among diatoms: Possible consequences on the spatial distribution of diatoms related to fluctuations in the underwater light climate. *Limnol Oceanogr* **52**: 1188–1194
- Leblanc K, Quéguiner B, Diaz F, Cornet V, Michel-Rodriguez M, Durrieu De Madron X, Bowler C, Malviya S, Thyssen M, Grégori G, et al** (2018) Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. *Nature Communications* **9**: 1–12

- Legris M, Ince YÇ, Fankhauser C** (2019) Molecular mechanisms underlying phytochrome-controlled morphogenesis in plants. *Nat Commun* **10**: 5219
- Legris M, Klose C, Burgie ES, Rojas CCR, Neme M, Hiltbrunner A, Wigge PA, Schäfer E, Vierstra RD, Casal JJ** (2016) Phytochrome B integrates light and temperature signals in *Arabidopsis*. *Science* **354**: 897–900
- Lepetit B, Campbell DA, Lavaud J, Büchel C, Goss R, Bailleul B** (2022) Photosynthetic Light Reactions in Diatoms II. The Dynamic Regulation of the Various Light Reactions. *Molecular Life of Diatoms*, Springer. pp 423–464
- Leu E, Mundy CJ, Assmy P, Campbell K, Gabrielsen TM, Gosselin M, Juul-Pedersen T, Gradinger R** (2015) Arctic spring awakening – Steering principles behind the phenology of vernal ice algal blooms. *Progress in Oceanography* **139**: 151–170
- Lewin JC, Lewin RA** (1960) AUXOTROPHY AND HETEROTROPHY IN MARINE LITTORAL DIATOMS. *Can J Microbiol* **6**: 127–134
- Li F, Melkonian M, Rothfels CJ, Villarreal JC, Stevenson DW, Graham SW, Wong GK, Pryer KM, Mathews S** (2015a) Phytochrome diversity in green plants and the origin of canonical plant phytochromes. *Nature Communications* **6**: 1–12
- Li F-W, Rothfels CJ, Melkonian M, Villarreal JC, Stevenson DW, Graham SW, Wong GK-S, Mathews S, Pryer KM** (2015b) The origin and evolution of phototropins. *Front Plant Sci*. doi: [10.3389/fpls.2015.00637](https://doi.org/10.3389/fpls.2015.00637)
- Li H, Burgie ES, Gannam ZTK, Li H, Vierstra RD** (2022) Plant phytochrome B is an asymmetric dimer with unique signalling potential. *Nature* **604**: 127–133
- Liu G, Shan Y, Zheng R, Liu R, Sun C** (2021) Growth promotion of a deep-sea bacterium by sensing infrared light through a bacteriophytochrome photoreceptor. *Environmental Microbiology* **23**: 4466–4477
- Lommer M, Specht M, Roy AS, Kraemer L, Andreson R, Gutowska MA, Wolf J, Bergner SV, Schilhabel MB, Klostermeier UC, et al** (2012) Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biology*. doi: [10.1186/gb-2012-13-7-r66](https://doi.org/10.1186/gb-2012-13-7-r66)
- Longobardi L, Dubroca L, Margiotta F, Sarno D, Zingone A** (2022) Photoperiod-driven rhythms reveal multi-decadal stability of phytoplankton communities in a highly fluctuating coastal environment. *Sci Rep* **12**: 3908
- Lund-Hansen L, Hawes I, Hancke K, Salmansen N, Nielsen J, Balslev L, Sorrell B** (2020) Effects of increased irradiance on biomass, photobiology, nutritional quality, and pigment composition of Arctic sea ice algae. *Mar Ecol Prog Ser* **648**: 95–110
- Madhuri S, Río Bártulos C, Serif M, Lepetit B, Kroth PG** (2019) A strategy to complement PtAUREO1a in TALEN knockout strains of *Phaeodactylum tricornutum*. *Algal Research* **39**: 101469
- Malviya S, Scalco E, Audic S, Vincent F, Veluchamy A, Poulain J, Wincker P, Iudicone D, de Vargas C, Bittner L, et al** (2016) Insights into global diatom distribution and diversity in the world's ocean. *Proceedings of the National Academy of Sciences* **113**: E1516–E1525
- Man D, Wang W, Sabeji G, Aravind L, Post AF, Massana R, Spudich EN, Spudich JL, Bèjà O** (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**: 1725–1731
- Mancinelli AL** (1994) The physiology of phytochrome action. In RE KENDRICK, G.H.M.KRONENBERG, eds, *Photomorphogenesis in Plants*. Springer Netherlands, Dordrecht, pp 211–269

- Mann DG, Vanormelingen P** (2013) An inordinate fondness? the number, distributions, and origins of diatom species. *Journal of Eukaryotic Microbiology* **60**: 414–420
- Mann KH, Lazier JRN** (2006) *Dynamics of marine ecosystems: biological-physical interactions in the oceans*, 3. ed., [Nachdr.]. Blackwell Publ, Malden, Mass.
- Mann M, Serif M, Jakob T, Kroth PG, Wilhelm C** (2017) PtAUREO1a and PtAUREO1b knockout mutants of the diatom *Phaeodactylum tricornutum* are blocked in photoacclimation to blue light. *J Plant Physiol* **217**: 44–48
- Marchetti A, Catlett D, Hopkinson BM, Ellis K, Cassar N** (2015) Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *ISME J* **9**: 2745–2748
- Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, Morales R, Allen AE, Armbrust EV** (2012) Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci U S A* **109**: E317–325
- Margalef R** (1978) *Life Forms of Phytoplankton as Survival Alternatives in an Unstable Environment*. *Oceanology Acta* **1**:
- Marques da Silva J, Cruz S, Cartaxana P** (2017) Inorganic carbon availability in benthic diatom communities: photosynthesis and migration. *Phil Trans R Soc B* **372**: 20160398
- Martin K, Schmidt K, Toseland A, Boulton CA, Barry K, Beszteri B, Brussaard CPD, Clum A, Daum CG, Eloe-Fadrosh E, et al** (2021) The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole. *Nat Commun* **12**: 5483
- Materna AC, Sturm S, Kroth PG, Lavaud J** (2009) FIRST INDUCED PLASTID GENOME MUTATIONS IN AN ALGA WITH SECONDARY PLASTIDS: psbA MUTATIONS IN THE DIATOM PHAEODACTYLUM TRICORNUTUM (BACILLARIOPHYCEAE) REVEAL CONSEQUENCES ON THE REGULATION OF PHOTOSYNTHESIS1: PHOTOSYNTHESIS IN psbA (D1) DIATOM MUTANTS. *Journal of Phycology* **45**: 838–846
- Mathews S, McBreen K** (2008) Phylogenetic relationships of B-related phytochromes in the Brassicaceae: Redundancy and the persistence of phytochrome D. *Molecular Phylogenetics and Evolution* **49**: 411–423
- McLachlan DH, Brownlee C, Taylor AR, Geider RJ, Underwood GJC** (2009) Light-induced motile responses of the estuarine benthic diatoms *Navicula perminuta* and *Cylindrotheca closterium* (bacillariophyceae). *Journal of Phycology* **45**: 592–599
- McLachlan DH, Underwood GJC, Taylor AR, Brownlee C** (2012) CALCIUM RELEASE FROM INTRACELLULAR STORES IS NECESSARY FOR THE PHOTOPHOBIC RESPONSE IN THE BENTHIC DIATOM NAVICULA PERMINUTA ( BACILLARIOPHYCEAE ) 1. **681**: 675–681
- McQuoid MR, Hobson LA** (1996) Diatom Resting Stages. *Journal of Phycology* **32**: 889–902
- Medlin LK** (2016) Evolution of the diatoms: Major steps in their evolution and a review of the supporting molecular and morphological evidence. *Phycologia* **55**: 79–103
- Merz E, Dick GJ, Beer D, Grim S, Hübener T, Littmann S, Olsen K, Stuart D, Lavik G, Marchant HK, et al** (2021) Nitrate respiration and diel migration patterns of diatoms are linked in sediments underneath a microbial mat. *Environ Microbiol* **23**: 1422–1435
- Middelburg JJ, Barranguet C, Boschker HTS, Herman PMJ, Moens T, Heip CHR** (2000) The fate of intertidal microphytobenthos carbon: An in situ <sup>13</sup>C-labeling study. *Limnol Oceanogr* **45**: 1224–1234

- Mock T, Hodgkinson K, Wu PG, Moulton V, Duncan A, van Oosterhout C, Pichler M** (2022) Structure and Evolution of Diatom Nuclear Genes and Genomes. *Molecular Life of Diatoms*, Springer. pp 111–145
- Mock T, Otilar RP, Strauss J, McMullan M, Paajanen P, Schmutz J, Salamov A, Sanges R, Toseland A, Ward BJ, et al** (2017) Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**: 536–540
- Moeys S, Frenkel J, Lembke C, Gillard JTF, Devos V, Van Den Berge K, Bouillon B, Huysman MJJ, De Decker S, Scharf J, et al** (2016) A sex-inducing pheromone triggers cell cycle arrest and mate attraction in the diatom *Seminavis robusta*. *Scientific Reports* **6**: 1–13
- Möglich A, Yang X, Ayers RA, Moffat K** (2010) Structure and Function of Plant Photoreceptors. *Annual Review of Plant Biology* **61**: 21–47
- Moore ER, Bullington BS, Weisberg AJ, Jiang Y, Chang J, Halsey KH** (2017) Morphological and transcriptomic evidence for ammonium induction of sexual reproduction in *Thalassiosira pseudonana* and other centric diatoms. *PLoS ONE* **12**: 1–18
- Moore J, Villareal T** (1996) Buoyancy and growth characteristics of three positively buoyant marine diatoms. *Mar Ecol Prog Ser* **132**: 203–213
- Moosburner M, Allen AE, Daboussi F** (2022) Genetic Engineering in Marine Diatoms: Current Practices and Emerging Technologies. *Molecular Life of Diatoms*, Springer. pp 743–773
- Morel A, Gentili B, Chami M, Ras J** (2006) Bio-optical properties of high chlorophyll Case 1 waters and of yellow-substance-dominated Case 2 waters. *Deep Sea Research Part I: Oceanographic Research Papers* **53**: 1439–1459
- Mouget JL, Gastineau R, Davidovich O, Gaudin P, Davidovich NA** (2009) Light is a key factor in triggering sexual reproduction in the pennate diatom *Haslea ostrearia*. *FEMS Microbiology Ecology* **69**: 194–201
- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D** (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324**: 1724–1726
- Multamäki E, Nanekar R, Morozov D, Lievonen T, Golonka D, Wahlgren WY, Stucki-Buchli B, Rossi J, Hytönen VP, Westenhoff S, et al** (2021) Comparative analysis of two paradigm bacteriophytochromes reveals opposite functionalities in two-component signaling. *Nat Commun* **12**: 4394
- Muramoto T, Kohchi T, Yokota A, Hwang I, Goodman HM** (1999) The *Arabidopsis* Photomorphogenic Mutant *hy1* Is Deficient in Phytochrome Chromophore Biosynthesis as a Result of a Mutation in a Plastid Heme Oxygenase. *Plant Cell* **11**: 335–347
- Nakajima H, Kobayashi I, Adachi Y, Hisatomi O** (2021) Transmission of light signals from the light-oxygen-voltage core via the hydrophobic region of the  $\beta$ -sheet surface in aureochrome-1. *Sci Rep* **11**: 11995
- Nakatani Y, Hisatomi O** (2018) Quantitative analyses of the equilibria among DNA complexes of a blue-light-regulated bZIP module, Photozipper. *Biophys Physicobiol* **15**: 8–17
- Nakov T, Beaulieu JM, Alverson AJ** (2018) Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (Diatoms, Bacillariophyta). *New Phytologist* **219**: 462–473
- Nelson DM, Tréguer P, Brzezinski MA, Leynaert A, Quéguiner B** (1995) Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles* **9**: 359–372

- Nelson DR, Hazzouri KM, Lauersen KJ, Jaiswal A, Chaiboonchoe A, Mystikou A, Fu W, Daakour S, Dohai B, Alzahmi A, et al** (2021) Large-scale genome sequencing reveals the driving forces of viruses in microalgal evolution. *Cell Host & Microbe* **29**: 250-266.e8
- Ni W, Xu S-L, González-Grandío E, Chalkley RJ, Huhmer AFR, Burlingame AL, Wang Z-Y, Quail PH** (2017) PPKs mediate direct signal transfer from phytochrome photoreceptors to transcription factor PIF3. *Nat Commun* **8**: 15236
- Nultsch W** (1971) Phototactic and Photokinetic Action Spectra of the Diatom *Nitzschia Communis*. *Photochemistry and Photobiology* **14**: 705–712
- 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**: 6–11
- Ogura A, Akizuki Y, Imoda H, Mineta K, Gojobori T, Nagai S** (2018) Comparative genome and transcriptome analysis of diatom, *Skeletonema costatum*, reveals evolution of genes for harmful algal bloom. *BMC genomics* **19**: 1–12
- Oka K, Ueno Y, Yokono M, Shen JR, Nagao R, Akimoto S** (2020) Adaptation of light-harvesting and energy-transfer processes of a diatom *Phaeodactylum tricornutum* to different light qualities. *Photosynthesis Research*. doi: [10.1007/s11120-020-00714-1](https://doi.org/10.1007/s11120-020-00714-1)
- Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J** (2018) Genome size of chrysophytes varies with cell size and nutritional mode. *Org Divers Evol* **18**: 163–173
- Oliver A, Podell S, Pinowska A, Traller JC, Smith SR, McClure R, Beliaev A, Bohutskyi P, Hill EA, Rabines A, et al** (2021) Diploid genomic architecture of *Nitzschia inconspicua*, an elite biomass production diatom. *Sci Rep* **11**: 15592
- Onodera A, Kong S-G, Doi M, Shimazaki K-I, Christie J, Mochizuki N, Nagatani A** (2005) Phototropin from *Chlamydomonas reinhardtii* is functional in *Arabidopsis thaliana*. *Plant and Cell Physiology* **46**: 367–374
- Osuna-Cruz CM, Bilcke G, Vancaester E, De Decker S, Bones AM, Winge P, Poulsen N, Bulankova P, Verhelst B, Audoor S, et al** (2020) The *Seminavis robusta* genome provides insights into the evolutionary adaptations of benthic diatoms. *Nat Commun* **11**: 3320
- Ozturk N** (2017) Phylogenetic and Functional Classification of the Photolyase/Cryptochrome Family. *Photochem Photobiol* **93**: 104–111
- Palmer JD, Round FE** (1967) Persistent, Vertical-Migration Rhythms in Benthic Microflora. VI. The Tidal and Diurnal Nature of the Rhythm in the Diatom *Hantzschia virgata*. *Biological Bulletin* **132**: 44–55
- Pelusi A, Margiotta F, Passarelli A, Ferrante MI, Ribera d'Alcalà M, Montresor M** (2020) Density-dependent mechanisms regulate spore formation in the diatom *CHAETOCEROS SOCIALIS*. *Limnol Oceanogr Letters* **5**: 371–378
- Petroutsos D, Tokutsu R, Maruyama S, Flori S, Greiner A, Magneschi L, Cusant L, Kottke T, Mittag M, Hegemann P, et al** (2016) A blue-light photoreceptor mediates the feedback regulation of photosynthesis. *Nature* **537**: 563–566
- Pierella Karlusich JJ, Bowler C, Biswas H** (2021) Carbon Dioxide Concentration Mechanisms in Natural Populations of Marine Diatoms: Insights From Tara Oceans. *Front Plant Sci* **12**: 657821

- Pierella Karlusich JJ, Ibarbalz FM, Bowler C** (2020) Phytoplankton in the *Tara* Ocean. *Annu Rev Mar Sci* **12**: 233–265
- Pollak B, Matute T, Nuñez I, Cerda A, Lopez C, Vargas V, Kan A, Bielinski V, von Dassow P, Dupont CL, et al** (2020) Universal loop assembly: open, efficient and cross-kingdom DNA fabrication. *Synthetic Biology* **5**: ysaa001
- Port JA, Parker MS, Kodner RB, Wallace JC, Armbrust EV, Faustman EM** (2013) Identification of G protein-coupled receptor signaling pathway proteins in marine diatoms using comparative genomics. *BMC Genomics* **14**: 503
- Poulsen N, Chesley PM, Kröger N** (2006) MOLECULAR GENETIC MANIPULATION OF THE DIATOM THALASSIOSIRA PSEUDONANA (BACILLARIOPHYCEAE). *J Phycol* **42**: 1059–1065
- Poulsen N, Davutoglu MG, Suchanova JZ** (2022) Diatom Adhesion and Motility. *Molecular Life of Diatoms*, Springer. pp 367–393
- Purschwitz J, Müller S, Fischer R** (2009) Mapping the interaction sites of *Aspergillus nidulans* phytochrome FphA with the global regulator VeA and the White Collar protein LreB. *Molecular Genetics and Genomics* **281**: 35–42
- Purschwitz J, Müller S, Kastner C, Schöser M, Haas H, Espeso EA, Atoui A, Calvo AM, Fischer R** (2008) Functional and Physical Interaction of Blue- and Red-Light Sensors in *Aspergillus nidulans*. *Current Biology* **18**: 255–259
- Pushkarev A, Inoue K, Larom S, Flores-Uribe J, Singh M, Konno M, Tomida S, Ito S, Nakamura R, Tsunoda SP, et al** (2018) A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* **558**: 595–599
- Rastogi A, Vieira FRJ, Deton-Cabanillas AF, Veluchamy A, Cantrel C, Wang G, Vanormelingen P, Bowler C, Piganeau G, Hu H, et al** (2019) A genomics approach reveals the global genetic polymorphism, structure, and functional diversity of ten accessions of the marine model diatom *Phaeodactylum tricornutum*. *ISME Journal* **14**: 347–363
- Rayko E, Maumus F, Maheswari U, Jabbari K, Bowler C** (2010) Transcription factor families inferred from genome sequences of photosynthetic stramenopiles. *New Phytol* **188**: 52–66
- Ren H, Junmin L, Qiuqi L, Boping H** (2008) Study of light signal receptor of *Stephanopyxis palmeriana* during sexual reproduction. *Chinese Journal of Oceanology and Limnology* **23**: 330–334
- Roberts WR, Downey KM, Ruck EC, Traller JC, Alverson AJ** (2020) Improved Reference Genome for *Cyclotella cryptica* CCMP332, a Model for Cell Wall Morphogenesis, Salinity Adaptation, and Lipid Production in Diatoms (Bacillariophyta). *G3 Genes|Genomes|Genetics* **10**: 2965–2974
- Rockwell NC, Duanmu D, Martin SS, Bachy C, Price DC, Bhattacharya D, Worden AZ, Lagarias JC** (2014) Eukaryotic algal phytochromes span the visible spectrum. *Proceedings of the National Academy of Sciences* **111**: 3871–3876
- Rockwell NC, Lagarias JC** (2017) Ferredoxin-dependent bilin reductases in eukaryotic algae: Ubiquity and diversity. *Journal of Plant Physiology* **217**: 57–67
- Rockwell NC, Lagarias JC** (2020) Phytochrome evolution in 3D: deletion, duplication, and diversification. *New Phytol* **225**: 2283–2300



- Rozanska M, Gosselin M, Poulin M, Wiktor J, Michel C** (2009) Influence of environmental factors on the development of bottom ice protist communities during the winter–spring transition. *Mar Ecol Prog Ser* **386**: 43–59
- Róžańska M, Poulin M, Gosselin M** (2008) Protist entrapment in newly formed sea ice in the Coastal Arctic Ocean. *Journal of Marine Systems* **74**: 887–901
- Rüdiger W, Thümmeler F** (1994) The phytochrome chromophore. *In* RE Kendrick, GHM Kronenberg, eds, *Photomorphogenesis in Plants*. Springer Netherlands, Dordrecht, pp 51–69
- Sabir JSM, Theriot EC, Manning SR, Al-Malki AL, Khiyami MA, Al-Ghamdi AK, Sabir MJ, Romanovicz DK, Hajrah NH, El Omri A, et al** (2018) Phylogenetic analysis and a review of the history of the accidental phytoplankter, *Phaeodactylum tricornutum* Bohlin (Bacillariophyta). *PLoS ONE* **13**: e0196744
- Sanyal A, Larsson J, Wirdum F, Andrén T, Moros M, Lönn M, Andrén E** (2021) Not dead yet: Diatom resting spores can survive in nature for several millennia. *Am J Bot* **ajb2.1780**
- 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
- Sauzède R, Claustre H, Jamet C, Uitz J, Ras J, Mignot A, D’Ortenzio F** (2015) Retrieving the vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: A method based on a neural network with potential for global-scale applications. *J Geophys Res Oceans* **120**: 451–470
- Schaller GE, Shiu SH, Armitage JP** (2011) Two-component systems and their co-option for eukaryotic signal transduction. *Current Biology* **21**: R320–R330
- Schellenberger Costa B, Jungandreas A, Jakob T, Weisheit W, Mittag M, Wilhelm C** (2013a) Blue light is essential for high light acclimation and photoprotection in the diatom *Phaeodactylum tricornutum*. *J Exp Bot* **64**: 483–493
- Schellenberger Costa B, Sachse M, Jungandreas A, Bartulos CR, Gruber A, Jakob T, Kroth PG, Wilhelm C** (2013b) Aureochrome 1a is involved in the photoacclimation of the diatom *Phaeodactylum tricornutum*. *PloS One* **8**: e74451
- Schumacher J** (2017) How light affects the life of *Botrytis*. *Fungal Genetics and Biology* **106**: 26–41
- Seeleuthner Y, Mondy S, Lombard V, Carradec Q, Pelletier E, Wessner M, Leconte J, Mangot J-F, Poulain J, Labadie K, et al** (2018) Single-cell genomics of multiple uncultured stramenopiles reveals underestimated functional diversity across oceans. *Nature Communications* **9**: 310
- Serif M** (2017) Characterization of Aureochromes in the diatom *Phaeodactylum tricornutum*. PhD thesis. University of Konstanz
- Serif M, Dubois G, Finoux A-L, Teste M-A, Jallet D, Daboussi F** (2018) One-step generation of multiple gene knock-outs in the diatom *Phaeodactylum tricornutum* by DNA-free genome editing. *Nature Communications* **9**: 3924
- Serôdio J, Catarino F** (2000) Modelling the primary productivity of intertidal microphytobenthos: time scales of variability and effects of migratory rhythms. *Mar Ecol Prog Ser* **192**: 13–30

- Ševčíková T, Horák A, Klimeš V, Zbránková V, Demir-Hilton E, Sudek S, Jenkins J, Schmutz J, Příbyl P, Fousek J, et al** (2015) Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? *Sci Rep* **5**: 10134
- 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. *Sci Rep* **8**: 14401
- Shih PM, Matzke NJ** (2013) Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. *Proc Natl Acad Sci USA* **110**: 12355–12360
- Shikata T, Iseki M, Matsunaga S, Higashi S, Kamei Y, Watanabe M** (2011) Blue and Red Light-Induced Germination of Resting Spores in the Red-Tide Diatom *Leptocylindrus danicus*†. *Photochemistry and Photobiology* **87**: 590–597
- Shikata T, Nukata A, Yoshikawa S, Matsubara T, Yamasaki Y, Shimasaki Y, Oshima Y, Honjo T** (2009) Effects of light quality on initiation and development of meroplanktonic diatom blooms in a eutrophic shallow sea. *Mar Biol* **156**: 875–889
- Shin A-Y, Han Y-J, Baek A, Ahn T, Kim SY, Nguyen TS, Son M, Lee KW, Shen Y, Song P-S, et al** (2016) Evidence that phytochrome functions as a protein kinase in plant light signalling. *Nat Commun* **7**: 11545
- Shinomura T, Nagatani a, Hanzawa H, Kubota M, Watanabe M, Furuya M** (1996) Action spectra for phytochrome A- and B-specific photoinduction of seed germination in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 8129–8133
- Shrestha RP, Hildebrand M** (2015) Evidence for a Regulatory Role of Diatom Silicon Transporters in Cellular Silicon Responses. *Eukaryotic Cell* **14**: 29–40
- Sibbald SJ, Archibald JM** (2020) Genomic Insights into Plastid Evolution. *Genome Biol Evol* **13**
- Sims P a., Mann DG, Medlin LK** (2006) Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia* **45**: 361–402
- Singler HR, Villareal TA** (2005) Nitrogen inputs into the euphotic zone by vertically migrating *Rhizosolenia* mats. *Journal of Plankton Research* **27**: 545–556
- Sorhannus U** (2007) A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. *Marine Micropaleontology* **65**: 1–12
- Sorokina M, Barth E, Zulfiqar M, Kwantes M, Pohnert G, Steinbeck C** (2022) Draft genome assembly and sequencing dataset of the marine diatom *Skeletonema cf. costatum* RCC75. *Data in Brief* **41**: 107931
- Stal LJ, Bolhuis H, Cretoiu MS** (2019) Phototrophic marine benthic microbiomes: the ecophysiology of these biological entities. *Environ Microbiol* **21**: 1529–1551
- Stal LJ, van Gernerden H, Krumbein WE** (1985) Structure and development of a benthic marine microbial mat. *FEMS Microbiology Ecology* **1**: 111–125
- Stiller JW, Schreiber J, Yue J, Guo H, Ding Q, Huang J** (2014) The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nat Commun* **5**: 5764
- Stomp M, Huisman J, Stal LJ, Matthijs HCP** (2007) Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. *ISME J* **1**: 271–282

- Strassert JFH, Irisarri I, Williams TA, Burki F** (2021) A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat Commun* **12**: 1879
- Streng C, Hartmann J, Leister K, Krauß N, Lamparter T, Frankenberg-Dinkel N, Weth F, Bastmeyer M, Yu Z, Fischer R** (2021) Fungal phytochrome chromophore biosynthesis at mitochondria. *EMBO J*. doi: [10.15252/emboj.2021108083](https://doi.org/10.15252/emboj.2021108083)
- Struyf E, Smis A, Van Damme S, Meire P, Conley DJ** (2009) The Global Biogeochemical Silicon Cycle. *Silicon* **1**: 207–213
- Studier FW** (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41**: 207–234
- Sullivan S, Petersen J, Blackwood L, Papanatsiou M, Christie JM** (2016) Functional characterization of *Ostreococcus tauri* phototropin. *New Phytol* **209**: 612–623
- Sun Y-F, Xu J-G, Tang K, Miao D, Gärtner W, Scheer H, Zhao K-H, Zhou M** (2014) Orange fluorescent proteins constructed from cyanobacteriochromes chromophorylated with phycoerythrobilin. *Photochem Photobiol Sci* **13**: 757
- Takahashi F, Yamagata D, Ishikawa M, Fukamatsu Y, Ogura Y, Kasahara M, Kiyosue T, Kikuyama M, Wada M, Kataoka H** (2007) AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *Proc Natl Acad Sci U S A* **104**: 19625–19630
- Tanaka T, Maeda Y, Veluchamy A, Tanaka M, Abida H, Maréchal E, Bowler C, Muto M, Sunaga Y, Tanaka M, et al** (2015) Oil Accumulation by the Oleaginous Diatom *Fistulifera solaris* as Revealed by the Genome and Transcriptome. *The Plant Cell Online* **27**: 162–176
- Tateyama S, Kobayashi I, Hisatomi O** (2018) Target Sequence Recognition by a Light-Activatable Basic Leucine Zipper Factor, Photozipper. *Biochemistry* **57**: 6615–6623
- Tian H, Trozzi F, Zoltowski BD, Tao P** (2020) Deciphering the Allosteric Process of the *Phaeodactylum tricornutum* Aureochrome 1a LOV Domain. *J Phys Chem B* **124**: 8960–8972
- Tilbrook K, Dubois M, Crocco CD, Yin R, Chappuis R, Alloreant G, Schmid-Siegert E, Goldschmidt-Clermont M, Ulm R** (2016) UV-B Perception and Acclimation in *Chlamydomonas reinhardtii*. *Plant Cell* **28**: 966–983
- Tirichine L, Rastogi A, Bowler C** (2017) Recent progress in diatom genomics and epigenomics. *Current Opinion in Plant Biology* **36**: 46–55
- Toyooka T, Hisatomi O, Takahashi F, Kataoka H, Terazima M** (2011) Photoreactions of aureochrome-1. *Biophys J* **100**: 2801–2809
- Traller JC, Cokus SJ, Lopez DA, Gaidarenko O, Smith SR, McCrow JP, Gallaher SD, Podell S, Thompson M, Cook O, et al** (2016) Genome and methylome of the oleaginous diatom *Cyclotella cryptica* reveal genetic flexibility toward a high lipid phenotype. *Biotechnology for Biofuels* **9**: 1–20
- Trebilco R, Baum JK, Salomon AK, Dulvy NK** (2013) Ecosystem ecology: size-based constraints on the pyramids of life. *Trends in Ecology & Evolution* **28**: 423–431
- Tréguer P, Bowler C, Moriceau B, Dutkiewicz S, Gehlen M, Aumont O, Bittner L, Dugdale R, Finkel Z, Iudicone D, et al** (2017) Influence of diatom diversity on the ocean biological carbon pump. *Nature Geoscience* **11**: 27–37

- Tréguer PJ, De La Rocha CL** (2013) The World Ocean Silica Cycle. *Annu Rev Mar Sci* **5**: 477–501
- Trippens J, Greiner A, Schellwat J, Neukam M, Rottmann T, Lu Y, Kateriya S, Hegemann P, Kreimer G** (2012) Phototropin Influence on Eyespot Development and Regulation of Phototactic Behavior in *Chlamydomonas reinhardtii*. *Plant Cell* **24**: 4687–4702
- Underwood GJC, Kromkamp J** (1999) Primary Production by Phytoplankton and Microphytobenthos in Estuaries. *Advances in Ecological Research*. Elsevier, pp 93–153
- Underwood GJC, Paterson DM** (2003) The importance of extracellular carbohydrate production by marine epipellic diatoms. *Advances in Botanical Research*. Elsevier, pp 183–240
- Valle KC, Nymark M, Aamot I, Hancke K, Winge P, Andresen K, Johnsen G, Brembu T, Bones AM** (2014) System Responses to Equal Doses of Photosynthetically Usable Radiation of Blue, Green, and Red Light in the Marine Diatom *Phaeodactylum tricornutum*. *PLOS ONE* **37**
- Van Vlierberghe M, Di Franco A, Philippe H, Baurain D** (2021) Decontamination, pooling and dereplication of the 678 samples of the Marine Microbial Eukaryote Transcriptome Sequencing Project. *BMC Res Notes* **14**: 306
- Vieira S, Cartaxana P, Máguas C, Marques da Silva J** (2016) Photosynthesis in estuarine intertidal microphytobenthos is limited by inorganic carbon availability. *Photosynth Res* **128**: 85–92
- Vuillet L, Kojadinovic M, Zappa S, Jaubert M, Adriano JM, Fardoux J, Hannibal L, Pignol D, Verméglio A, Giraud E** (2007) Evolution of a bacteriophytochrome from light to redox sensor. *EMBO Journal* **26**: 3322–3331
- Wagner JR, Brunzelle JS, Forest KT, Vierstra RD** (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**: 325–331
- Wiltbank LB, Kehoe DM** (2019) Diverse light responses of cyanobacteria mediated by phytochrome superfamily photoreceptors. *Nat Rev Microbiol* **17**: 37–50
- Xie W-H, Zhu C-C, Zhang N-S, Li D-W, Yang W-D, Liu J-S, Sathishkumar R, Li H-Y** (2014) Construction of Novel Chloroplast Expression Vector and Development of an Efficient Transformation System for the Diatom *Phaeodactylum tricornutum*. *Mar Biotechnol* **16**: 538–546
- Xue P, Bai Y, Rottwinkel G, Averbukh E, Ma Y, Roeder T, Scheerer P, Krauß N, Lamparter T** (2021) Phytochrome Mediated Responses in *Agrobacterium fabrum*: Growth, Motility and Plant Infection. *Curr Microbiol*. doi: [10.1007/s00284-021-02526-5](https://doi.org/10.1007/s00284-021-02526-5)
- Yeh K-C, Wu S-H, Murphy JT, Lagarias JC** (1997) A Cyanobacterial Phytochrome Two-Component Light Sensory System. *Science* **277**: 1505–1508
- Yu Z, Ali A, Igbalajobi OA, Streng C, Leister K, Krauß N, Lamparter T, Fischer R** (2019) Two hybrid histidine kinases, TcsB and the phytochrome FphA, are involved in temperature sensing in *Aspergillus nidulans*. *Molecular Microbiology* **0**: 1–17
- Yu Z, Armant O, Fischer R** (2016) Fungi use the SakA (HogA) pathway for phytochrome-dependent light signalling. *Nature Microbiology* **1**: 1–7
- Yu Z, Fischer R** (2018) Light sensing and responses in fungi. *Nature Reviews Microbiology*. doi: [10.1038/s41579-018-0109-x](https://doi.org/10.1038/s41579-018-0109-x)
- Zhao C, Gan F, Shen G, Bryant DA** (2015) RfpA, RfpB, and RfpC are the Master Control Elements of Far-Red Light Photoacclimation (FaRLiP). *Front Microbiol* **6**: 1303

# ANNEXE



# Sensing and Signalling in Diatom Responses to Abiotic Cues

Marianne Jaubert, Carole Duchêne, Peter G. Kroth, Alessandra Rogato, Jean-Pierre Bouly, and Angela Falciatore

## Abstract

Diatoms are prominent microalgae that proliferate in a wide range of aquatic environments. Still, fundamental questions regarding their biology, such as how diatoms sense and respond to environmental variations, remain largely unanswered. In recent years, advances in the molecular and cell biology of diatoms and the increasing availability of genomic data have made it possible to explore sensing and signalling pathways in these algae. Pivotal studies of photosensory perception have highlighted the great capacity of diatoms to accurately detect environmental variations by sensing differential light signals and adjust their physiology accordingly. The characterization of photoreceptors and light-dependent processes described in this review, such as plastid signalling and diel regulation, is unveiling sensing systems which are unique to these algae, reflecting their complex evolutionary history and adaptation to aquatic life. Here, we also describe putative sensing components involved in the responses to nutrient, osmotic changes, and fluid motions. Continued elucidation of the molecular systems processing endogenous and environmental cues and their interactions with other biotic and abiotic stress signalling pathways is expected to greatly increase our understanding of the mechanisms controlling the

---

M. Jaubert (✉) · C. Duchêne · J.-P. Bouly · A. Falciatore (✉)  
Laboratoire de Biologie du chloroplaste et perception de la lumière chez les microalgues,  
UMR7141, Centre national de la recherche scientifique, Sorbonne Université, Institut de Biologie  
Physico-Chimique, Paris, France  
e-mail: [marianne.jaubert@ibpc.fr](mailto:marianne.jaubert@ibpc.fr); [angela.falciatore@ibpc.fr](mailto:angela.falciatore@ibpc.fr)

P. G. Kroth  
Fachbereich Biologie, Universität Konstanz, Konstanz, Germany

A. Rogato  
Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Naples, Italy  
Institute of Biosciences and BioResources, IBBR/CNR, Naples, Italy

abundance and distribution of the highly diverse diatom communities in marine ecosystems.

---

## Abbreviations

AUREO	aureochrome
bHLH	basic helix–loop–helix domain
BV	biliverdin
CA	carbonic anhydrase
CCREs	CO <sub>2</sub> /cAMP-responsive elements
CPD	cyclobutane pyrimidine dimer
CPF	cryptochrome/photolyase family
CRY	cryptochrome
DCMU	3-(3', 4'-dichlorophenyl)-1, 1-dimethylurea
DPH	diatom phytochrome
FAD	flavin adenine dinucleotide
FMN	flavin-mononucleotide
FR	far-red light
MTHF	5,10-methenyltetrahydrofolic acid
N	nitrogen
NAGK	N-acetyl-L-glutamate kinase
NAT	natural antisense transcript
ncRNA	non coding RNA
NPF	nitrate transporter 1/peptide transporter family
P	phosphorus
PAS	Per-Arnt-Sim domain (from “period”, “aryl hydrocarbon receptor nuclear transporter” and “single-minded” proteins)
PCM	photosensory core module
PHR	photolyase-related domain
PHY	phytochrome
PL	photolyase
R	red light
ROS	reactive oxygen species
Ser/Thr	serine/threonine
snRNA	small nuclear RNA
sRNA	small RNA
TE	transposable element
TF	transcription factor
TOR	target of rapamycin
UV	ultra-violet

## 1 Introduction

The life history of diatoms in aquatic environments is strongly influenced by a multitude of abiotic factors, including light, nutrients, salinity, and temperature (Mann and Lazier 2006). Physical factors such as circulation and turbulent mixing of water, sea level changes, and proximity to coastal regions can cause drastic changes in the characteristics of abiotic resources which, if limiting or in excess, can quickly become a source of stress and adversely affect diatom growth and productivity (Seckbach and Gordon 2019).

Before the advent of the genomic era, several pioneering studies offered the first indications that diatoms are able to sense their environment and respond to a variety of signals. Observations of migrating diatom mats between deep nutrient pools and the ocean surface (Villareal et al. 1999), or of photoperiodic control of diatom spore growth (Eilertsen et al. 1995), implied that active responses are controlling their physiology and life strategies. This observation also challenged historical paradigms that considered diatom growth and distribution only influenced by the availability of the resources (Sverdup et al. 1942).

When diatoms became accessible to genetic transformation, the analysis of calcium changes in transgenic *Phaeodactylum tricorutum* cells expressing an aequorin reporter revealed the existence of calcium-dependent signal transduction mechanisms involved in responses to fluid motion, osmotic stress, and iron limitation (Falciatore et al. 2000). It was also shown that calcium mediates desensitization to prolonged exposure to stimuli such as turbulence or infochemicals (Falciatore et al. 2000; Vardi et al. 2006), suggesting that these physiological responses are controlled by “sense-process-respond” chains with specific receptors and feedback mechanisms, similar to those known in terrestrial plants or animals. However, compared to the latter systems, our understanding of diatom signal perception mechanisms is still very fragmentary, and mostly limited to light sensing. Molecular aspects of environmental perception remain largely uncharacterized in diatoms, although the increasing availability of data from cells exposed to different cues clearly indicates that cellular responses involve significant changes in gene expression, likely driven by the processing of external and internal signals.

The aim of this chapter is to provide an overview of abiotic signal perception by diatoms, with a particular focus on the regulators involved in the responses to light changes, including light-absorbing photoreceptors, circadian rhythm regulation, and plastid-derived signals. We also provide an initial appraisal of nutrient sensing systems and of possible components of the diatom response systems to osmotic changes and fluid motion. Finally, the ecological implication of these processes and future challenges are also discussed.


































## 2 Light Sensing and Responses in Diatoms

Being photosynthetic organisms, diatoms require light as an essential source for growth. Light provides also a significant amount of information about the surrounding environment. From sunrise to sunset, the intensity, orientation, and spectrum of sunlight vary because the sun's rays hit the atmosphere at different angles, and they are selectively attenuated before reaching the earth's surface. In addition to such daily light/dark changes, there is periodic variation in day length with the changing seasons across latitudes each year. The intensity and relative ratio of different wavebands can also be affected when radiation is filtered by a layer of light-absorbing elements. For example, photosynthetic organisms absorb blue and red light, so the light transmitted or reflected by them is enriched in green and far-red bands. In plants, perception of such low red/far-red ratios allows them to detect surrounding phototrophs competing for light and trigger an appropriate shade avoidance response (Kendrick 1994). In aquatic environments, variations in intensity and spectral composition occur because of the strong and selective absorption of light (Kirk 2011). Light is absorbed by water itself, mostly in the red part of the spectrum, so the blue and green bands penetrate the water column more deeply. Suspended particles and dissolved chromophoric organic matter in water scatter light and/or differentially absorb wavelengths. Therefore, the light field in aquatic habitats is structured in depth and it can vary significantly with the diverse concentrations of constituents occurring in open ocean or coastal regions, estuaries, lakes, or intertidal zones (Kirk 2011).

Sensing of light relies on pigment-bound proteins known as photoreceptors that absorb light of specific wavelengths and activate specific signalling cascades (Duanmu et al. 2017). However, in the cells of phototrophs, light is also absorbed by various photosynthetic pigments collectively absorbing a larger fraction of the incident light than the photoreceptors. As will be described in this chapter, photosynthetic activities in plastids also regulate the responses to light via retrograde pathways from plastids to the nucleus. Therefore, it is often difficult to disentangle information on the relative influence of individual sensors and signalling pathways involved in the responses to light. Experimental evidence is needed to characterize the dependency of these responses on light quality and quantity. When a response is only induced by a particular waveband, we can deduce that a photoreceptor that specifically perceives, say, blue, green, or red light must be involved. Alternatively, when a response is triggered by both blue and red light, it can be supposed that it is regulated by photosynthesis or/and the action of different photoreceptors. Photoreceptor-mediated responses can be activated by very low light fluences and are independent of photosynthesis, and plastid signalling is abolished in conditions that perturb plastid function (Kendrick 1994). Photoreceptor activity can also be identified in acute light response experiments using pulses of light after a prolonged dark treatment that abolishes both photosynthesis and circadian clock regulation (Kendrick 1994).

Many facets of diatom life are known to be tightly linked to light conditions as summarized in Table 1 and described in more detail in many chapters of this book.

**Table 1** Effective wavebands triggering physiological responses in diatoms

	Responses, diatom species	Photosensing system	References
	Cell cycle progression, <i>P. tricornutum</i>	AUREO1a	Huysman et al. 2013
	Cell cycle progression, <i>Navicula</i> sp.		Cao et al. 2013
	Growth rate, <i>D. brightwellii</i>	DPH ?	Lipps 1973
	Auxospore formation, <i>H. ostreraria</i>	DPH ?	Mouget et al. 2009
	Sexual reproduction, <i>S. palmeriana</i>	DPH ?	Ren et al. 2005
	Sexual reproduction, <i>C. didymus</i>	CPF ?, Rhodopsin ?	Baatz 1941
	Resting spore germination and growth, <i>Chaetoceros</i> sp, <i>S. costatum</i> , <i>T. minima</i>		Shikata et al. 2009
	Resting spore germination, <i>L. danicus</i>	Plastid ?	Shikata et al. 2011
	Chloroplast assembly, <i>P. laevis</i>		Furukawa et al. 1998
	Chloroplast dispersion, <i>P. laevis</i>		Furukawa et al. 1998
	Upward migration in sediment, environmental samples, <i>G. fasciola</i> , <i>N. spartinetensis</i>	CPF ? AUREO ?	Barnett et al. 2020
	Upward migration in sediment, Wadden sea populations		Wenderoth and Rhiel 2004
	Sinking rate, <i>T. weissflogii</i>		Fisher et al. 1996
	Photophobotaxis, <i>N. communis</i>		Nultsch 1971
	Photophobotaxis, <i>N. ostrearia</i> , <i>N. gregaria</i> , <i>N. communis</i>	CPF ? AUREO ?	Wenderoth 1979
	Photophobotaxis, <i>N. peregrina</i> , <i>P. viridis</i> , <i>N. sigmaidea</i>	Plastid ?	Wenderoth 1979
	Photophobotaxis, <i>C. cuspidata</i> , <i>Nitzschia cf linearis</i> , <i>P. viridis</i>		Cohn and Weitzell 1996
	Photophobotaxis, <i>S. phoenicenteron</i>		Cohn and Weitzell 1996
	Photophobotaxis and phototopotaxis, <i>N. communis</i>		Nultsch 1971
	Phototopotaxis, <i>N. perminuta</i>		McLachlan et al. 2009
	Possible repelling, <i>N. linearis</i> , <i>C. cuspidate</i> , <i>P. viridis</i>		Cohn and Weitzell 1996
	Possible repelling, <i>N. communis</i>		Nultsch 1971
	Possible repelling, <i>S. phoenicenteron</i>		Cohn and Weitzell 1996
	Photokinesis, <i>C. closterium</i>		McLachlan et al. 2009
	Photokinesis, <i>N. perminuta</i>		McLachlan et al. 2009
	Photokinesis, <i>N. communis</i>		Nultsch 1971
	Chromatic acclimation, <i>P. tricornutum</i> , <i>N. closterium</i>		Fujita and Ohki 2004; Herbstová et al. 2015
	High light acclimation, <i>P. tricornutum</i>	AUREO1a, AUREO1b	Schellenberger Costa et al. 2013a; Mann et al. 2017
	Photoprotection, <i>P. tricornutum</i>	AUREO1a, CPF	Coesel et al. 2009; Schellenberger Costa et al. 2013b; Juhas et al. 2014
	Pigment production, <i>H. ostreraria</i>		Mouget et al. 2005

Information on the effective wavebands has been extracted from the specific conditions used in the reported experiments. Note that the illumination conditions light sources, irradiance are not comparable between each study. Phototopotaxis, movement towards or away from the light source; photophobotaxis, reversing the direction of movement caused by a sudden decrease or increase in light intensity; photokinesis, acceleration of movement by light; repulsion, migration away from light

For instance, cell cycle progression (Chapter “Cellular Hallmarks and Regulation of the Diatom Cell Cycle”), germination of resting spores and sexual reproduction (Chapter “Life-Cycle Regulation”) have been observed to occur at certain light intensities and photoperiods. Movement in several diatom species is also affected by the light intensity applied (Chapter “Adhesion and Motility”). The strong impact of light on diatom gene expression has been also evidenced by independent genome-wide expression analyses in different diatom species exposed to different light regimes (e.g. Nymark et al. 2009; Chauton et al. 2013; Nymark et al. 2013; Ashworth et al. 2013; Valle et al. 2014; Smith et al. 2016; Mann et al. 2020; Bilcke et al. 2021). A wide palette of light colours from violet to far-red has been observed to trigger different responses, with blue and red light being the most effective bands. Light properties triggering each of these responses vary among diatom species, which could reflect the species-specific optimization of spectral light quality for growth and the involvement of multiple signalling pathways regulating acclimation to particular environments.

In the next session, we review the different light sensors controlling responses to these light changes in diatoms discovered through genomic, genetic, and functional characterization in model species.

## 2.1 Diatom Photoreceptors for Sensing the Light Environment

A variety of photoreceptors have been found by exploring the protein sequences encoded in diatom genomes (Jaubert et al. 2017) (Table 2). Some of these proteins are similar to photoreceptors found in bacteria, plants, or animals, while others have distinct combinations of protein domains (Fig. 1). This evidence supports the hypothesis that diatoms evolved their light-sensing abilities from different ancestors via complex routes to fully exploit light cues in their environment. The spectral and functional properties are known for several members of the blue light cryptochrome/photolyase and aureochrome photoreceptor families, and for the red/far-red light phytochromes (Fig. 1b). It should be noted that UV-light-sensing photoreceptors have not been described in diatoms. While some gene products have been found that show similarities to plant UVR8 (Duanmu et al. 2017), they lack the specific tryptophan residues needed to act as a chromophore, so it is unlikely that they may function as photoreceptors.

**The Cryptochrome/Photolyase Family.** The cryptochrome/photolyase family proteins (CPFs) are widespread blue light-absorbing flavoproteins with similar primary sequences, but very diverse functions. Most CPFs non-covalently bind a flavin adenine dinucleotide (FAD) cofactor as a specific prosthetic chromophore. Other chromophores such as 5,10-methenyltetrahydrofolic acid (MTHF), 8-hydroxy-5-deazaflavin, or flavin-mononucleotide (FMN) may also be associated with some CPFs as light antennae (Essen et al. 2017). CPF members have a characteristic conserved photolyase-related (PHR) domain, but amino and carboxy terminal extensions are highly variable in both length and primary sequence. Photolyases (PL) are blue light-activated enzymes repairing UV-induced DNA

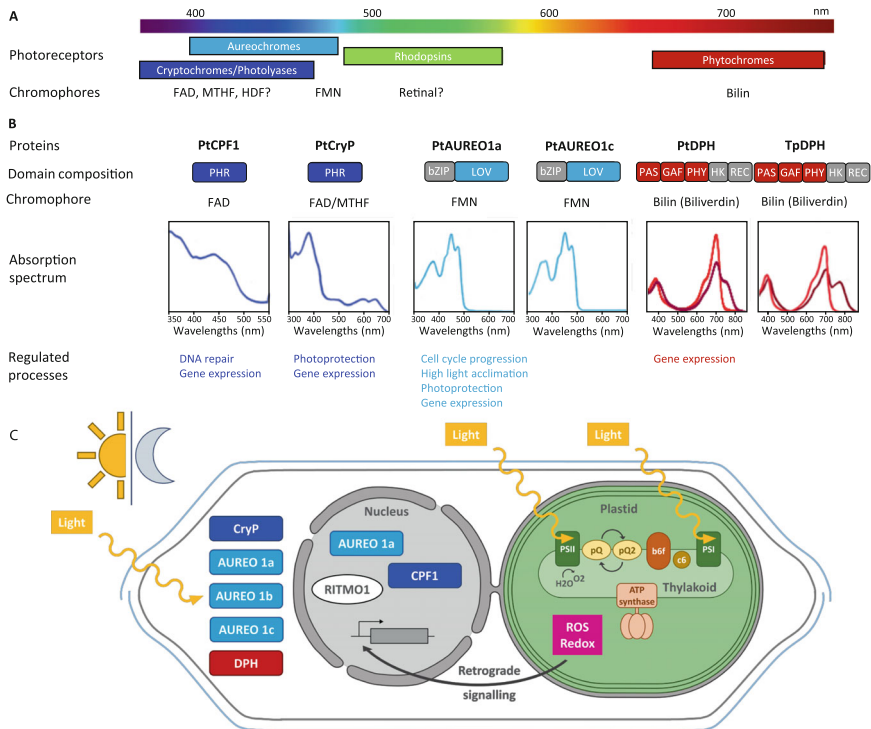
**Table 2** Photoreceptors in diatom genome sequence resources

Species	ID prefix	6-4 Photolysase/ PCPFI	PlantCRY-like	CryDASH	AUREO1a	AUREO1b	AUREO1c	Protoe- rhodopsin	Heliorhodopsin	DPH
<i>Tp</i> centric	Thaps3l	262946	7368	23500, 268988, 35005	33340	20065	37198			22848
<i>To</i> centric	EJK	58421.1	58052.1 62994.1	51584.1, 57652.1, 70255.1 <sup>a</sup> , 57807.1	65981.1	47540.1	75561.1			
<i>Cc</i> centric		g1243.tl		g11904.tl, g3306.tl	g18270.tl	g20747.tl	g23579.tl			g17859.tl
<i>Pt</i> pennate	Phatt3_	J27429-p1	J54342.p1	J55091.p1, J34592.p1	J8113.p1	J15977.p1	J51933.p1		J43903.p1	J54330.p1
<i>Fc</i> pennate	OEU	16032.1	23746.1	18916.1, 15165.1, 21140.1	18451.1	21654.1	06878.1, 17811.1 <sup>a</sup>	10445.1, 23745.1 <sup>a</sup>		
<i>Sr</i> pennate	Sro	82_g043740.1	580_g170080.1	147_g068020.1, 2726_g335660.1 <sup>a</sup>	5_g004540.1, 624_g177380.1	1194_g251270.1	971_g226430.1		149_g068360.1, 149_g068370.1	219_g090570.1, 537_g162340.1, 537_g162360.1, 537_g162370.1
<i>Pme</i> pennate	Psemu11	133684	107549, 107549	28554.1, 190760	172818	294561	202247			
<i>Pmu</i> pennate	VEU	36634.1	42420.1, 44083.1	40361.1, 41071.1	?	34876.1, 34877.1 <sup>a</sup>	?			
<i>Fs</i> pennate	GAX	14982.1, 23745.1 <sup>a</sup>		14744.1, 19840.1 <sup>a</sup> , 23334.1, 28242.1 <sup>a</sup>	11767.1, 20834.1 <sup>a</sup>	09413.1, 20262.1 <sup>a</sup>	16809.1, 13918.1 <sup>a</sup>			
<i>Sa</i> pennate	Sael	19225	1435, 23010	6021	24026, 25403 <sup>a</sup>	24510	?		16855, 16858 <sup>a</sup>	18038, 1907, 1973, 22009, 24689

Identification numbers are given for photoreceptor sequences found in the indicated source

Species abbreviations and resources sources: *Thalassiosira pseudonana*, *Tp*, JGI; *Thalassiosira oceanica*, *To*, GenBank; *Cyclotella cryptica*, *Cc*, <http://genomes.mdb.ucla.edu/Cyclotella/download.html>; *Phaeodactylum triconutum*, *Pt*, Ensembl Genomes; *Fragilariopsis cylindrus*, *Fc*, GenBank; *Seminavis robusta*, *Sr*, <https://bioinformatics.psb.ugent.be/orcae/overview/Semro>; *Pseudo-nitzschia multiseriata*, *Pmu*, GenBank; *Fistulifera solaris*, *Fs*, GenBank; *Synedra acus*, *Sa*, <http://www.lin.irk.ru/sacus/index.php?i=site/page&view=downloads>. (Armbrust et al. 2004; Bowler et al. 2008; Lommer et al. 2012; Galachyants et al. 2015; Tanaka et al. 2015; Traller et al. 2016; Mock et al. 2017; Basu et al. 2017; Osuna-Cruz et al. 2020)

<sup>a</sup>Indicate uncertainty whether the sequence is another gene or an allelic variant of the same gene



**Fig. 1** Sensing of light cues in diatoms. (a) Photoreceptor classes identified in diatom genomes. Cryptochrome-Photolyase, Aureochrome, Rhodopsin, and Phytochrome families are positioned schematically according to the range of wavelengths they usually absorb. The nature of the chromophores is indicated below the respective photoreceptor boxes, with question marks indicating putative chromophore. (b) Summary of characterized photoreceptors from *P. tricornutum* (Pt) and *T. pseudonana* (Tp). Domain architecture, chromophore, absorption spectrum, and regulated processes are indicated for: cryptochrome photolyase 1 (PtCPF1) (Coesel et al. 2009); plant-like Cry (PtCryP) (Juhász et al. 2014); aureochromes PtAUREO1a (Heintz and Schlichting 2016) and PtAUREO1c (Bannister et al. 2019) and diatom phytochromes (PtDPH and TpDPH) (Fortunato et al. 2016). Domains involved in the light sensing are indicated in blue or red. (c) Schematic representation of the characterized *P. tricornutum* regulators processing external and internal cues and controlling nuclear gene expression in response to light changes. They include photoreceptors, the regulator of cellular rhythmicity RITMO1, and plastid to nucleus retrograde signaling pathways related to the photosynthetic processes (ROS and redox signals). In the nucleus are shown the factors whose cellular localization or DNA binding has been experimentally validated

lesions, such as cyclobutane pyrimidine dimer (CPD PL) or (6-4) pyrimidine-pyrimidone photoproducts. The cryptochromes (CRY) of plants and some insects do not have any DNA repair activity, but are either blue light photoreceptors or light-independent components of the central circadian oscillator (Chaves et al. 2011). The activation of light-dependent CRY has been well studied. The photoperception process starts with a very fast photoinduced reduction of FAD. This redox- and

light-dependent change at the core of the protein leads to conformational changes, allowing specific interactions with other protein partners. Globally, both light-dependent and -independent CRY, are involved in transcriptional regulation, respectively, by inhibiting transcriptional repressors or by inactivating transcription.

Phylogenetic relationships in the CPF family helped to identify five major super-classes (sc) which do not necessarily converge functionally. For instance, sc1 includes (6-4) PLs but also light-dependent animal CRYs and light-independent CRYs involved in the transcription/translation feedback loop of the circadian clock. Class I and III CPD photolyases group also together phylogenetically with proteins with different functions such as light dependent plant photoreceptor CRY and plant-like photoreceptor CRY in sc3, but are separate from class II CPD photolyases which are all found in sc4. The last two super-classes (sc2 and sc5) include all CRY-DASH (named after *Drosophila*, *Arabidopsis*, *Synechocystis*, Human) and the proteobacterial PL/CRYs (Fortunato et al. 2015; Ozturk 2017). A surprising result from genome sequence analyses was that diatoms do not possess canonical plant CRY photoreceptors, even though blue light is preponderant in the ocean and CRY regulate so many physiological processes in plants. However, further genomic and functional investigations in diatoms revealed novel CPF variants, including the animal-like CPF1 and plant-like CRYs (CryP) (see below) and several DASH CRYs (Table 2). The biological function of Cry DASH is not yet clearly defined in diatoms or indeed in other organisms, but several studies suggest that they might have a single-stranded DNA CPD PL activity, a signalling role, or be involved in the regulation of metabolism, consistent with their organellar localization (Kleine et al. 2003; Froehlich et al. 2010).

Only two diatom CPFs have been characterized in detail to date: CPF1 and CryP of *P. tricornutum* (Fig. 1b). Initially discovered in diatoms, these proteins are extensively represented in metatranscriptomic data derived from marine environments (Coesel et al. 2021). CPF1 belongs to sc1, is localized in the nucleus, and has (6-4) PL DNA repair activity both *in vitro* (Coesel et al. 2009) and *in vivo* (De Riso et al. 2009). However, like plant CRY, CPF1 regulates the transcription of several genes acting in pathways modulating photoprotection, cell division, nutrient assimilation, etc., under blue light (Coesel et al. 2009), so is likely to be also a photoreceptor. CPF1 is also able to bind mammalian CLOCK protein in heterologous cells where it partially represses the CLOCK/BMAL1 heterodimer (Coesel et al. 2009) in the positive loop of the circadian clock (Kume et al. 1999).

CryP belongs to the group of Plant-like photoreceptors CRY (sc3). Recombinant *P. tricornutum* CryP, produced in *E. coli*, binds FAD and MTHF chromophores (Juhas et al. 2014). Moreover, FAD photoreduction, a common mechanism in CPFs that bind FAD, is also active *in vitro*, suggesting that activation of CryP is light-dependent. Originally, CryP was identified as a blue light regulator of light harvesting proteins directly involved in photoprotection (LHCX) (Juhas et al. 2014). However, subsequent analyses of transcriptional responses to illumination by blue light after prolonged dark, revealed that gene expression was already deregulated in the dark in *CryP* knockdown lines compared to wild type (König et al. 2017). Therefore, despite the presence of chromophores, CryP may not be a

major blue light photoreceptor but rather a component involved in the global modulation of transcription, requiring other blue light photoreceptors to signal the light responses.

**The Aureochromes.** The aureochromes (AUREOs) are unique blue light photoreceptors that possess both an FMN-binding light-oxygen-and-voltage (LOV) domain (Crosson et al. 2003), and a bZIP domain typical for bZIP transcription factors (TFs) (Dröge-Laser et al. 2018) (Fig. 1). The blue light phototropin photoreceptors of green algae and plants also possess two LOV domains, but utilize a serine/threonine (Ser/Thr) kinase domain for signal transduction instead (Christie 2007). AUREOs were originally discovered in the xanthophyte alga *Vaucheria frigida* in 2007 (Takahashi et al. 2007). Their name refers to “aurum” (Latin for gold), because of the golden-brownish colour of most stramenopiles. Takahashi et al. (2007) identified two orthologs in *V. frigida*, VfAUREO1, and VfAUREO2. Using an RNA interference approach to silence these two genes individually, they demonstrated that both AUREOs are involved in the regulation of photomorphogenic responses. Meanwhile, AUREOs have been identified in other stramenopiles (Ishikawa et al. 2009; Jungandreas et al. 2014) and in a raphidophyte (Ji et al. 2017), but not in non-photosynthetic oomycetes (Kroth et al. 2017). AUREOs are not present in red algae, which are considered to represent the endosymbiotic ancestors of stramenopile plastids (Archibald 2015). This indicates that the ancestral AUREO gene with its unique combination of LOV and bZIP domains may have been provided either by the putative host cell of the secondary endosymbiosis event, or that it evolved very early within the stramenopiles, possibly via domain shuffling (Di Roberto and Peisajovich 2014).

Stramenopile AUREOs differ in their structures. While AUREO1 proteins possess the typical LOV domain, AUREO2 proteins have a mutation within the LOV domain, which prevents non-covalent binding of flavin needed for light absorbance in the blue range. The reason is a steric hindrance from a methionine residue within the binding cavity (Banerjee et al. 2016a). AUREO2, therefore, is not a real photoreceptor, but could still be involved in light regulation, e.g. by forming a dimer with a light-sensing AUREO1 protein. In reciprocal experiments, genetic modification of AUREO2 from *P. tricornutum*-restored flavin binding (Serif 2017), while introducing a point mutation at the same site in PtAUREO1a, led to loss of flavin binding (Banerjee et al. 2016a). Based on this distinction and on phylogenetic analyses, aureochromes in diatoms and other organisms have been classified as either AUREO1-type (a/b/c etc.) or as AUREO2 (Schellenberger Costa et al. 2013b). The algae studied so far all possess one AUREO2 protein and one or more AUREO1 isoforms (Table 2).

Blue light absorption, both in aureochromes and phototropins, causes the formation of an adduct between the flavin and a nearby cysteine within a few microseconds, starting the signalling cascade (Toyooka et al. 2011; Kerruth et al. 2014). The domain topology of AUREOs is inverted as compared to most other characterized LOV proteins because the sensory domain of AUREOs is at the C-terminus of the receptor. The  $\alpha$  helix of the AUREO-LOV domain, allosterically regulates the fold of the N-terminally flanking A' $\alpha$  helix (Herman and Kottke 2015).

Subsequent unfolding of the A $\alpha$  helix exposes a high affinity dimerization site and enables the formation of the light state dimer of LOV (Herman et al. 2013; Herman and Kottke 2015; Heintz and Schlichting 2016). Indeed, in full-length AUREO, the J $\alpha$  helix plays a crucial role in the formation of the light state of the receptor (Banerjee et al. 2016b). The activation of the LOV domain results in a loss of the helical secondary structure of the bZIP domain (Banerjee et al. 2016b), indicating that there is a direct communication between the two domains. If a DNA binding site is available, the helical fold of the bZIP domain is increased by light (Banerjee et al. 2016b), resulting in rigidification of the domain (Heintz and Schlichting 2016). In a complementary approach, Tian and colleagues showed the importance of a previously overlooked C $\alpha$  helix in promoting the conformational protein changes (Tian et al. 2020). bZIP domains have a general tendency to dimerize, and are only capable of binding DNA as dimers (Tateyama et al. 2018). For some AUREOs, dimerization and DNA binding have been shown to be induced by blue light (Hisatomi and Furuya 2015; Banerjee et al. 2016b; Nakatani and Hisatomi 2018; Nakajima et al. 2021). PtAUREO1a occurs in the dark as a dimer/higher oligomer (Banerjee et al. 2016b) or as a monomer in equilibrium with a dimer (Heintz and Schlichting 2016). Light induces the dimerization of LOV domains and the association of the monomers (Kobayashi et al. 2020; Goett-Zink et al. 2020), which is the rate-limiting step in the process of DNA binding (Akiyama et al. 2016). Based on these properties, AUREOs could potentially be used as an optogenetic tool, for instance, to increase protein stability under blue light conditions (Hepp et al. 2020) or directly as a light-driven gene switch.

AUREO function has been extensively characterized in *P. tricornutum*, especially for PtAUREO1a that acts as a key regulator of the diatom cell cycle (Table 1; Chapter "Cellular Hallmarks and Regulation of the Diatom Cell Cycle"). Evidence from silenced and knockout lines of *P. tricornutum* further indicates that PtAUREO1a and PtAUREO1b might be involved in regulation of photoacclimation (Table 1; Schellenberger Costa et al. 2013b; Serif et al. 2017; Mann et al. 2017; Madhuri et al. 2019). PtAUREO1c might be a high light sensor in vivo because it recovers faster and is much less sensitive to light than PtAUREO1a (Bannister et al. 2019). Recent studies indicate that AUREOs may have a large impact on the cells. Changes in transcriptomes in response to a shift from red to blue light were analysed in wild-type *P. tricornutum* cells, in *PtAUREO1a* knockout and in *PtAUREO1a* complemented lines (Mann et al. 2020). Wild-type cells react within minutes by up- or down-regulating 75% of the genes, while this massive change in gene expression is mostly inhibited in *PtAUREO1a* knockout strains (Mann et al. 2020). PtAUREO1a, therefore, must have a specific function in cellular regulation that cannot be complemented by other AUREOs. This raises the question of how a single photoreceptor can affect such a large number of genes. Possibly, a cascade of TF transduces the initial response of PtAUREO1a to blue light. This is supported by findings that the transcript abundance of a large number of diatom TFs and photoreceptors (Rayko et al. 2010) is strongly and rapidly affected by blue light in wild-type cells, but not in *PtAUREO1a* knockout mutants (Mann et al. 2020). When common gene expression patterns are analysed, the aureochromes are placed in



different clusters (Ait-Mohamed et al. 2020), supporting the idea that AUREO1 isoforms may have different roles. There also is some evidence that AUREOs might be involved in regulation of the diel cycle and the expression of some AUREOs follows a different diurnal pattern (Banerjee et al. 2016b). The diurnal rhythmic expression of *PtAUREO1a* and *1c* can still be detected when the cells are kept in the dark, while *PtAUREO1b* expression appears to be light activated, and *PtAUREO2* oscillates only weakly throughout the day. The recently discovered diatom clock component RITMO1/PtbHLH1a (Annunziata et al. 2019) is strongly induced by blue light in wild-type cells, but not in the *PtAUREO1a* mutant, indicating that *PtAUREO1a* might be involved in triggering the diatom clock in response to blue light.

Beside the aureochromes, *P. tricornutum* was previously reported to possess putative LOV-Histidine kinase photoreceptors (Bowler et al. 2008). However, closer examination of the corresponding genes showed that although they encode a recognizable PAS-Histidine kinase structure, the PAS signal sensor domain lacks the critical residues required for LOV photosensing, putting into question its function as a photoreceptor.

**The Diatom Phytochromes (DPHs).** Originally discovered in terrestrial plants as red/far-red light (R/FR) sensors, phytochromes (PHY) form a family of modular proteins that have a conserved N-terminal photosensory core module (PCM) for light perception and a variable C-terminal output module involved in signal transduction (Rockwell and Lagarias 2020). The PCM encompasses a conserved architecture of PAS, GAF, and PHY domains. The photosensitivity of phytochromes is conferred by a linear tetrapyrrole, covalently bound to a cysteine residue in the PAS or GAF domains, and derived from the cleavage of heme to biliverdin (BV), which can be further reduced by specific bilin reductases (Rockwell and Lagarias 2020). PHYs have been found in cyanobacteria, photosynthetic and non-photosynthetic bacteria, fungi, and diverse algae (Rockwell and Lagarias 2020). Canonical PHYs exhibit the unique capacity to switch between R and FR-absorbing forms (Pr and Pfr, respectively) in a photoreversible way, with bacterial and fungal BV-binding PHY absorbing at longer R and FR wavelengths compared to the more reduced chromophore-binding PHYs from cyanobacteria, plant, or algae.

Among marine algae, PHYs have been found in Glaucophyta, Prasinophyta, Streptophyta, Cryptophyta, and Ochrophyta, but not in Haptophyta, Rhodophyta, and Chlorophyta (Rockwell and Lagarias 2020). Recombinant PHYs from some Glaucophyta and Prasinophyta microalgae and the multicellular *Ectocarpus siliculosus* exhibit a large spectral diversity with some of them being able to sense orange, green, or blue light (Rockwell et al. 2014). Possibly molecular evolution of PHY absorption properties has been driven by the particular spectral composition in aquatic environments (Rockwell et al. 2014).

DPHs predicted from genomic sequences (Table 2) all exhibit the typical output domain architecture of bacterial PHYs (Fig. 1b). Probably complex processes led to DPH evolution. Based on the phylogeny of the PCM, diatom and other heterokont phytochromes are a sister clade of fungal PHYs, more similar to bacterial PHYs, and separate from other algal PHYs, suggesting a different origin (Duanmu et al. 2014;

Li et al. 2015; Fortunato et al. 2016). Pennate DPHs seem to be closer to those from Phaeophyta than to those from centric diatoms. Moreover, predicted PHYs found in some brown algal viruses are also phylogenetically close to DPHs, suggesting that horizontal gene transfer events may have contributed to the evolution and diversification of this photoreceptor class in marine diatoms.

Recombinant DPHs from the centric diatom *T. pseudonana* and from the pennate *P. tricornutum* can both bind BV as chromophore and exhibit a genuine R/FR bacteriophytochrome absorption spectrum, shifted towards longer wavebands compared to the PHYs found in plants or the other microalgae mentioned (Fortunato et al. 2016). This feature is indeed surprising given the rapid absorption of the R and FR wavelengths in the water column. Another particularity is the poor sequence similarity of the N-terminal part of DPHs of pennate diatoms with known PAS domains. An additional cysteine in the GAF domain, at a position conserved with archaeplastidal PHYs, has been found in some DPHs such as those from the pennate *Amphora coffeaformis* (Fortunato et al. 2016). Some PHYs from the Phaeophyceae *E. siliculosus* also have these two conserved cysteines. One of these PHYs has been spectrally characterized and has a green/FR photocycle in the presence of phytochromobilin (Rockwell et al. 2014). Therefore, we cannot exclude the possibility that some DPHs might also bind phytochromobilin or other BV-derived chromophores and display different light-sensing abilities.

In recent years, functional studies in *P. tricornutum* have provided initial information on DPH activities. Comparing gene expression in wild-type and *dph* knockout mutants in dark adapted cells exposed to either low R or FR treatments revealed a specific deregulation of FR-induced genes in knockout mutants, while responses to R were unaltered. This study demonstrated the activity of DPH as a photoreceptor triggering FR light signal propagation in diatoms (Fortunato et al. 2016). The discovery that FR sensing occurs in the ocean is significant, although the function of this photoreceptor in diatom life is still unexplained. In *P. tricornutum*, DPH-regulated genes include some involved in transcription and signalling, but the majority are of unknown function. No physiological defects have been identified in the *P. tricornutum* knockout lines so far. In other diatoms, R- and FR-induced phenomena have been described; for example, the induction of sexual reproduction in *Stephanopyxis palmeriana* (Table 1). In the absence of genomic data in most of these species, it is still not possible to confidently propose that DPHs are involved in the regulation of these processes.

PHY signalling mechanisms usually involve two-component systems in bacteria (Auldridge and Forest 2011), and PHY interacting factors (PIF) in land plants (Chen and Chory 2011). In the model fungi *Aspergillus nidulans*, two different pathways for PHY signalling have been investigated. One signalling cascade involves the SakA pathway, which starts with a two-component like phosphorelay, followed by a mitogen-activated protein kinase cascade and activation of the TF (Yu et al. 2016). In another pathway, fungal PHYs interacts with chromatin regulation complexes, affecting gene expression via histone modification (Hedtke et al. 2015). In *P. tricornutum*, DPH has been shown to auto-phosphorylate in response to FR, suggesting a phosphorelay mechanism, but the rest of the signalling cascade

activating the observed gene expression changes is still unknown (Fortunato et al. 2016). No homologs of PIFs have been identified in diatoms.

Exploring the diversity of diatoms has yielded some interesting observations on DPHs. First, not all diatoms have a DPH. Several species such as *Thalassiosira oceanica*, *Fragilariopsis cylindrus*, *Pseudo-nitzschia multiestrata* and *multiseries* do not possess the *DPH* gene, indicating that this photoreceptor is not essential for diatom biology. Moreover, differences have been also found between pennate and centric species. In addition to forming phylogenetically distinct clades, *DPH* genes in centric diatoms are always found as a single gene copy, while gene multiplication has occurred in several pennate species which may possess up to five *PHY* genes (Table 2). Data are yet too scarce to determine when multiplication occurred. However, the expansion of the DPH family observed in some biofilm-forming motile diatoms (Osuna-Cruz et al. 2020) suggests that *DPH* duplication might be the result of adaptation to specific environments and life strategies. Extensive mining of genomic and metagenomic data, coupled with careful functional analyses of *DPHs* selected from various diatom species, will hopefully provide enough evidence to resolve how *PHYs* arose in heterokont microalgae and explain the cryptic functions in species that possess DPHs.

**The Rhodopsins.** These light-sensing integral membrane proteins found in Archaea, bacteria, and eukaryotes share a topology of seven transmembrane alpha helices within which a retinal chromophore is covalently bound (Ernst et al. 2014). Rhodopsins exhibit a wide variety of spectral tuning in the blue-green part of the spectrum depending on the nature of a few influential amino acids interacting with the retinal (Man 2003; Ernst et al. 2014). Rhodopsins function as light-driven ion pumps, ion channels, or light sensors (Grote et al. 2014). The discovery that H<sup>+</sup>-pump rhodopsin converts light into ATP has challenged the assumed monopoly of photosynthesis as a phototrophy-enabling mechanism (Béjà et al. 2001; Finkel et al. 2013). Recently, a distinct group of microbial rhodopsins, the heliorhodopsins, has been identified after analysing environmental genomic samples (Pushkarev et al. 2018). Heliorhodopsins do not have the capacity for light-triggered ion transport but they do have a long photocycle, suggesting that they could act as signalling photoreceptors (Pushkarev et al. 2018). H<sup>+</sup>-pump rhodopsins and heliorhodopsins are present in bacteria, Archaea, and algae and are highly represented in environmental genomic data (Pushkarev et al. 2018; Coesel et al. 2021).

Diatom rhodopsin-like sequences falling into the H<sup>+</sup>-pump group, based on conservation of key amino acids and phylogeny, were first identified in the transcriptome of *Pseudo-nitzschia granii* (Marchetti et al. 2012), then in those of other species (Marchetti et al. 2015), and in the genome of *F. cylindrus* (Mock et al. 2017) (Table 2). Because *P. granii* rhodopsin-like transcripts are highly abundant in low iron conditions, it has been hypothesized that this proteorhodopsin-like protein could be involved in energy production under conditions of iron deficiency that affect photosynthesis (Marchetti et al. 2012). It is noteworthy that a gene homologous to heliorhodopsin has been identified in the genomes of *P. tricorutum* and other pennate diatoms, but not in centric ones (Pushkarev et al. 2018) (Table 2). G-protein-coupled receptor rhodopsin-like genes, homologs of receptors that transduce

a wide range of stimuli including light, hormones, volatile molecules, glycoproteins, nucleotides, and chemokines in eukaryotes (Costanzi et al. 2009), have also been identified in various diatom genomes (Port et al. 2013), with evidence of expansion of this gene family in some species (Osuna-Cruz et al. 2020). However, information about the spectral and functional properties of these and other diatom rhodopsins is still lacking, so their function as light sensors remains to be established.

## 2.2 Metabolic Signalling

As for all photosynthetic eukaryotes, diatom genes encoding plastid and mitochondrial proteins are located in the nuclear and in the organellar genomes. Accordingly, organelle biogenesis and function require strong coordination of gene expression in different compartments. In plants and green algae, it is well established that nuclear-encoded proteins participate in the regulation of organellar gene expression and functions (anterograde signalling) and that mitochondrial and plastid factors control the expression of nuclear encoded genes (retrograde signalling) (Hernández-Verdeja and Strand 2018). The coordination of these different signalling pathways is essential to adjust cellular metabolism to environmental changes. The plastid itself is considered as an internal sensor of environmental signals because most of the metabolic reactions occurring in this organelle are extremely sensitive to external changes. Different intermediates of plastid processes, such as tetrapyrroles or carotenoids, and products of the photosynthetic process can act as indicators of the plastid metabolism and nutrient and redox states, and modify the expression of nuclear genes involved in stress responses (Hernández-Verdeja and Strand 2018).

In organisms with a secondary plastid such as diatoms, the information about retrograde signalling is still very limited. However, studies of the photosynthetic processes (Chapters "Photosynthetic Light Reactions in Diatoms. I. The Lipids and Light-harvesting Complexes of the Thylakoid Membrane" and "Photosynthetic Light Reactions in Diatoms. II. The Dynamic Regulation of the Various Light Reactions") are providing initial evidence that different metabolites are involved in the relay of information from plastid to nucleus. In particular, it has been shown that light-induced expression of several nuclear genes involved in photoprotection, such as *LHCX1* and 2, can be altered by artificially modulating the redox state of plastoquinone with drugs such as DCMU (3-(3', 4'-dichlorophenyl)-1, 1-dimethylurea) and DBMIB (2, 5-dibromo-3-methyl-6-isopropyl-p-benzoquinone) that mimic low light and high light phenotypes, respectively, as originally shown in green algae (Escoubas et al. 1995). The effect of these inhibitors has been observed for only some of the *LHCX* genes and only under specific light conditions, supporting the idea that multiple signalling pathways control their expression (Lepetit et al. 2013; Taddei et al. 2016). Plastid-to-nucleus signals have also been identified by genome-wide expression analyses performed in *P. tricornutum* exposed to equal doses of photosynthetic usable radiation of blue, green, and red light (Valle et al. 2014). This study identified light-quality independent responses for many photosynthetic-associated nuclear genes, which are altered by inhibiting

photosynthetic electron transport with DCMU. This study also showed blue light-dependent control of some genes implicated to be involved in photoprotection and PSII repair. Functional studies of CPF1 and AUREO photoreceptors, support the involvement of blue light receptors in the regulation of photoacclimation processes. On the contrary, FR-induced nuclear genes, regulated by DPH, are not affected by the inhibition of photosynthetic electron transport (Fortunato et al. 2016), as expected for wavebands at the extreme end of the visible light spectrum that have a minimal input to photosynthesis.

Nutrient stress can alter plastid metabolism and in turn nuclear gene expression (Chapters “Comparative and Functional Genomics of Macronutrient Utilization in Marine Diatoms” and “Molecular Mechanisms Underlying Micronutrient Utilization in Marine Diatoms”). Genes implicated in carbon fixation are tightly controlled by both light and nutrient signalling. It has been proposed that the CO<sub>2</sub> concentration itself elicits the physiological modifications observed upon changes of the external CO<sub>2</sub> levels, by activating a signalling cascade involving the TF PtbZIP11 and cAMP (cyclic adenosine monophosphate) as a second messenger (Harada et al. 2006; Ohno et al. 2012). TF PtbZIP11 binds to CO<sub>2</sub>/cAMP-responsive elements (CCREs) in CO<sub>2</sub>-regulated genes such as the *P. tricornutum* carbonic anhydrase genes (*PtCAs*) (Harada et al. 2005; Ohno et al. 2012). These genes are also regulated by light activated pathways that crosstalk with CO<sub>2</sub>/cAMP signals by acting on the same CCREs in promoters (Kikutani et al. 2012; Tanaka et al. 2016). It has been shown that different photosynthetically active monochromatic lights confer similar transcriptional activation of *PtCAs* genes when CO<sub>2</sub> is limiting, while an artificial electron acceptor from the reduction side of PSI efficiently inhibits their expression. These data support the idea that there is redox control within CO<sub>2</sub>/light crosstalk (Tanaka et al. 2016). At the same time, the direct involvement of photoreceptors in *PtCAs* activation has not been ruled out.

Recent evidence also indicates that reactive oxygen species (ROS), which accumulate under exposure to excess light, can also play a signalling role in diatoms, acting as secondary messengers, as is known to occur in other organisms (Mittler et al. 2011). As described in detail in Chapter “An Ocean of Signals: Intracellular and Extracellular Signalling”, a compartmentalized redox-sensitive signalling network has been characterized in *P. tricornutum* by analysing the redox proteome as well as redox-sensitive biosensors (e.g. reduction-oxidation sensitive green fluorescent protein) localized in different cellular compartments (Rosenwasser et al. 2014; Graff van Creveld et al. 2015). Single-cell analysis of responses to environmental stresses that generate ROS revealed that changes in the chloroplast redox state play a critical role in controlling cell fate decisions toward acclimation or cell death (Mizrachi et al. 2019). It is still unclear which molecules act as plastid redox signal(s) or how signals are transmitted to the nucleus in diatoms, plants, or green algae.

### 2.3 Regulation of Diatom Biological Rhythms: Integrating External and Internal Signals

Like most other organisms, diatoms experience periodic variation in light/dark cycles because of the Earth's rotation. Numerous laboratory- and field-based investigations have clearly indicated synchrony of algal life processes with these cycles. As in many microalgae, cell cycle progression is strongly linked to light and dark alternations (Vaulot et al. 1986), and progression through various cell cycle phases is tightly controlled by both light and nutrient availability (Chapter "Cellular Hallmarks and Regulation of the Diatom Cell Cycle"). Temporal variations in key metabolites have also been described. Synthesis of photosynthetic pigments (Ragni and d'Alcalà 2007), fixed carbons ( $\beta$ -glucans), or lipids (Chauton et al. 2013; Ashworth et al. 2013) normally increases during the light to sustain other energy-demanding cellular activities at night. Diatoms also vary in their sensitivity to oxidative stress over a diurnal cycle, due to the light-dependent production of antioxidants (e.g. glutathione GSH cellular pool), which decrease during the dark period (Volpert et al. 2018). Rhythmic movements of benthic diatoms dependent on diel cycles and on tides have been well documented over the years (Round and Happpy 1965). It has been shown more recently that this migration strategy relates to the decoupling of cell division, which occurs in darkness deep in the sediment, and photosynthesis, which occurs at the surface of the sediment during light periods (Saburova and Polikarpov 2003; Barnett et al. 2020).

At a molecular level, diverse diurnal transcript profiling experiments performed in the diatom *T. pseudonana* (Ashworth et al. 2013), *P. tricornutum* (Chauton et al. 2013; Smith et al. 2016; Annunziata et al. 2019), and *Seminavis robusta* (Bilcke et al. 2021) revealed a strict temporal separation of diatom transcriptional gene networks with day/night cycles. A 24-h periodicity in transcriptional responses have also been reported in diatoms in situ from the metatranscriptome analysis of environmental populations over multiple diurnal cycles (Kolody et al. 2019; Coesel et al. 2021). The data show robust diel transcriptional patterns and the temporal partitioning of gene expression implicated in key biological processes such as photosynthesis, cell cycle, and light perception. Moreover, semidiurnal (12h) gene expression periodicity has been described in *S. robusta* (Bilcke et al. 2021), which may reflect adaptations to the occurrence of tides in coastal benthic ecosystems or might be linked to the twice daily vertical migration of diatoms into the sediment as previously described.

It is still largely unknown how diatoms orchestrate their life according to light/dark cycles. Processes occurring during the light phase could be directly activated by photoreceptor-mediated signalling pathways. Other light-driven processes such as photosynthesis may also contribute to temporal segregation of cellular processes and metabolism, as in other phototrophs (Haydon et al. 2013; Dodd et al. 2015). Several studies indicate that diatom cellular rhythmicity is also regulated by an endogenous circadian clock. This has been demonstrated by monitoring rhythms of cellular fluorescence (Annunziata et al. 2019) or photosynthesis and migration behaviours (Round and Happpy 1965; Harding et al. 1981) during light/dark cycles compared to

constant light. The persistence of these rhythms for several days in the absence of periodic cues clearly support the presence of an autonomous timekeeper regulating circadian rhythms in diatoms, as in most organisms (Dunlap 1999). Nevertheless, the diatom circadian clock system remains largely uncharacterized because no orthologs of bacterial, animal, or plant circadian clock components have yet been found in the diatom genomes, except for CRYs and casein kinases. The first regulator of diatom biological rhythms named RITMO1 has been identified recently in *P. tricornutum* by searching within all diatom TFs for those showing robust rhythmic expression profiles in light/dark cycles and in the dark (Annunziata et al. 2019). RITMO1 is a nuclear localized protein containing bHLH and PAS domains, and alteration of its expression levels and timing by ectopic overexpression resulted in lines with deregulated diurnal gene expression profiles compared to wild-type cells (Annunziata et al. 2019). Gene expression fluctuations were also reduced in these lines in continuous darkness, showing that the regulation of rhythmicity by RITMO1 is independent of light input. Also, cell fluorescence rhythms were perturbed by deregulation of RITMO1 expression in light/dark cycles and, even more so in constant light, strongly supporting its role in circadian regulation.

Several lines of evidence indicate a partial conservation of the regulatory program generating rhythmicity in animals and diatoms. RITMO1 contains bHLH and PAS protein domains that are also present in the CLOCK and BMAL proteins, key components of the mammalian central oscillator (Dunlap 1999). This raises novel questions on the evolution of circadian clock in diatoms and in many other phytoplankters, in which RITMO-like proteins have been found (Annunziata et al. 2019). bHLH-PAS RITMO-like proteins might have independently acquired a function in rhythm regulation by convergent evolution. However, it has also been hypothesized that this function might be present in the ancient heterotrophic host that subsequently acquired plastids via endosymbiosis. Remarkably, an expansion of bHLH-PAS TFs has been found in the genome of the benthic diatom *S. robusta*, with seven out of ten genes showing rhythmic expression either at dawn or broadly during the day (Bilcke et al. 2021). It has been proposed that the expansion and expression divergence of these putative circadian genes represents an adaptation of diatoms to the highly rhythmic benthic environment. Thus, these factors currently represent ideal entry points to further characterize the mechanisms generating cellular rhythmicity in diatoms and address their relevance for life in the marine ecosystem that is governed by a multitude of rhythms of different periods, including daily, tidal, lunar, or seasonal cycles.

---

### 3 Nutrient Sensing Pathways

The availability of macronutrients such as nitrogen, silicon, phosphorus, and carbon, or micronutrients such as iron are essential for diatom growth. Whether freely living in the water column or attached to substrates, diatoms experience frequent variations in nutrient concentrations due to the movement of currents, water mixing, or tides. It is well known that these microalgae have a strong capacity to adjust their physiology

and metabolism to the nutrients available, due to genomic and metabolic features that have been described in detail in Chapters “Comparative and Functional Genomics of Macronutrient Utilization in Marine Diatoms” and “Molecular Mechanisms Underlying Micronutrient Utilization in Marine Diatoms”. However, compared to other established model systems (Chantranupong et al. 2015), we still have very limited information on how diatoms sense and respond to nutrient changes in their environment (see also “Chapter Comparative and Functional Genomics of Macronutrient Utilization in Marine Diatoms”).

Diatoms have evolved an array of nutrient transporters, including diverse families of channels, carriers, and pump proteins, that are modulated in accordance with nutrient demand and availability. A comprehensive description of these transporters is beyond the scope of this review. However, it is noteworthy that in plants it has been demonstrated that some of these proteins, known as tranceptors, play a signalling role in addition to their transport function (Gojon et al. 2011). In particular, the plant low affinity nitrate transporters (NPF) can bind and transport different substrates, such as phytohormones controlling organogenesis according to nitrate availability (Fan et al. 2017). NPF genes have been found in the genomes of several diatom species (Rogato et al. 2015). Considering the low concentration of nitrate in the oceans, it has been hypothesized that the diatom NPFs, predicted to be located in the plasma membrane, could bind alternative substrates such as oligopeptides or hormones, and could be involved in the control of different processes such as interactions between diatoms and bacteria (Santin et al. 2021). A similar function in nutrient sensing has been proposed for the silicic acid transporter (SIT) of *T. pseudonana* (Shrestha and Hildebrand 2015). By generating and analysing knock-down mutants of *SIT1* and *SIT2* genes, it was demonstrated that in growth conditions with sufficient silicic acid, the transport role of the SITs is relatively minor. It has been proposed that under these conditions, their primary role is to sense silicic acid levels to evaluate whether the cell can proceed with its cell wall formation and division processes.

A recent study integrating transcriptomic and phosphoproteomic analyses in *P. tricornutum* highlighted that protein phosphorylation is an important part of regulation in systemic nutrient sensing (Tan et al. 2020). Proteins that are conditionally phosphorylated as a function of diel cycling and iron and nitrogen (N) availability have been identified, as well as a number of stress related kinases (e.g. calcium-dependent protein kinase, calcium calmodulin-dependent protein kinase, inositol hexakisphosphate/diphosphoinositol-pentakisphosphate kinase, and Ser/Thr kinase) that might be responsible for phosphorylating them. TOR (target of rapamycin) protein kinases have also been found in diatoms (Prioretti et al. 2017). As the name indicates, TOR Ser/Thr kinases are inactivated by selective ATP-competitive inhibitors such as rapamycin. These widespread proteins play a central regulatory role in cell homeostasis because they connect information about the quantity and quality of different environmental cues (including nutrients) to metabolic processes (Dobrenel et al. 2016). A current working model for TOR signalling proposes that these kinases relay a permissive signal to downstream targets only in the presence of sufficient nutrients to fuel protein synthesis. Compared to mammalian TOR, the *TOR*



gene described in *P. tricornutum* (Prioretti et al. 2017) possesses highly conserved motifs and amino acids of the ATP-binding pocket in the TOR kinase domain. In *P. tricornutum*, the inhibition of TOR with inhibitors developed for mammalian TOR, resulted in a nutrient starvation response (Prioretti et al. 2017) as in other organisms (Ingargiola et al. 2020). It has been shown that the inhibition of the TOR activity reduced but did not halt cell proliferation unlike growth under N starvation, and also improved triacylglycerol (TAG) productivity, strongly suggesting that this pathway is somehow involved in nutrient sensing and metabolic remodelling.

The PII proteins are signalling components widely distributed in all domains of life that help to coordinate carbon and N assimilation, but they have not been found in diatoms (Rogato et al. 2015). However, in the genome of *T. pseudonana* and *P. tricornutum*, the gene *ARGB* encodes the enzyme N-acetyl-L-glutamate kinase (NAGK) (Chellamuthu et al. 2013), which in plants and cyanobacteria has been demonstrated to form a complex with PII in response to N starvation. NAGK is a member of the amino acid kinase family that control a step-in arginine synthesis. Arginine is needed to make proteins and, together with its precursor ornithine, to make polyamines and urea. Therefore, it is reasonable to hypothesize that diatom NAGKs are involved in the response to urea, which is indeed part of N metabolism, and that a diatom specific-mechanism of N signalling might have evolved in these microalgae (Chellamuthu et al. 2013; Rogato et al. 2015). Phylogenetically, diatom NAGK cluster with alphaproteobacterial NAGKs (Chellamuthu et al. 2013). It is thus possible that during the evolution of Stramenopiles, the *argB* gene was acquired by horizontal gene transfer, while the original cyanobacterial/plant NAGK and PII genes were lost.

The recent discovery of a novel  $\text{Ca}^{2+}$  signalling pathway involved in environmental phosphorus (P) sensing and acclimation has provided new insights on the peculiar nutrient sensory system of diatoms (Helliwell et al. 2021a). By using the fluorescent  $\text{Ca}^{2+}$  biosensor R-GECO1, it has been shown that centric and pennate diatoms grown under P-limited conditions exhibit rapid and transient rise of cytosolic  $\text{Ca}^{2+}$  within seconds following resupply of environmentally relevant concentrations of a range of P forms. These responses are not detected in P-replete cells. Remarkably, this study also highlights the existence of rapid cross-talk between P and N metabolism based on P- $\text{Ca}^{2+}$  signalling. In particular, primary diatom responses to P limitations involves substantial increase in the assimilation of N, a major constituent of proteins, nucleic acids, and chlorophyll, thus exploiting another vital nutrient for cellular recovery and restoration of growth. This coordinated regulation of P and N metabolism could enhance diatom ability to compete for resources in dynamic nutrient environments.

Although the molecular machinery underpinning nutrient sensing and signalling remains largely uncharacterized, it is well established that diatom responses to nutrient changes require massive changes in gene expression that ultimately controls nutrient transport, mobilization, or storage. How the environment impacts genome expression is still largely unknown, but it is likely that this requires the integrated action of signalling components and genetic and epigenetic factors that we are just starting to discover in these microalgae (Chapter “Epigenetic Control of Diatom

Genomes: An Overview from *in silico* Characterisation to Functional Studies"). TFs that are differentially expressed under different nutrient conditions have been identified in many transcriptomic studies. However, only in few cases, the involvement of TFs in the regulation of nutrient metabolism has been demonstrated. Examples are the novel N starvation-inducible RING-domain protein called RING-GAF-Gln-containing protein (RGQ1), which regulates the response of *P. tricornutum* to N starvation (Matthijs et al. 2016), and the leucine zipper bZIP14 which has been proposed to regulate remodelling of the TCA cycle and N metabolism (Matthijs et al. 2017). Massive changes in the expression of different TF gene families have also been described in response to phosphate fluctuations (Cruz de Carvalho et al. 2016). A regulator of P limitation in the Myb TF family called PtPSR has recently been characterized in *P. tricornutum* (Kumar Sharma et al. 2020). *PtPSR* knockout mutants show reduced cell growth under P limitation and reduced induction of the P-stress response genes containing PtPSR recognition motifs. PtPSR, therefore, has a pivotal role in *P. tricornutum* adaptation to P limitation (Kumar Sharma et al. 2020).

Genes involved in nutrient stress responses are likely to be regulated also by mechanisms involving non coding RNAs (ncRNAs). High-throughput sequencing of small RNAs (sRNAs), extracted from *P. tricornutum* cells grown under different conditions of light and iron starvation, showed the unexpected complexity of sRNA populations in these algae (Rogato et al. 2014). Some sRNAs were identified and experimentally validated as being specifically expressed under iron limitation, including U2 snRNA-derived sRNAs and tRNA-derived sRNAs. A similar analysis in *T. pseudonana* and *F. cylindrus* (Lopez-Gomollon et al. 2014) revealed that a significant number of tRNA-derived sRNAs were up-regulated under specific conditions of Si and Fe limitation. Moreover, both studies reported a huge subset of sRNAs associated to repetitive transposable elements (TEs) marked by DNA methylation. Interestingly, expression of some *P. tricornutum*-specific Ty1/copia-like retrotransposons has been shown to be induced under N deficiency (Maumus et al. 2009), and cytosine methylation is thought to control the mobility of these TEs, as they are hypomethylated when expressed (Veluchamy et al. 2015). Altogether, these data suggest that complex small RNA-dependent DNA methylation processes may contribute to controlling the expression of TEs in diatoms and regulating gene expression and alterations in genome structure when certain nutrients are limiting. In addition to sRNAs, long intergenic non-coding RNA (lincRNA) and natural anti-sense transcript (NAT) classes have been described in *P. tricornutum* (Cruz de Carvalho et al. 2016; Cruz de Carvalho and Bowler 2020). Remarkably, the P-regulated noncoding transcriptome of NATs and lincRNAs is almost equivalent in number to the coding transcriptome (messenger RNA), providing compelling evidence of the relevance of ncRNAs in the regulation of environmental stress responses in diatoms (Cruz de Carvalho and Bowler 2020).

## 4 Diatom Sensing of Other Abiotic Stimuli

Besides the light and nutrient-sensing pathways previously described, some information is also emerging on the mechanisms controlling diatom responses to other abiotic cues. It is well established that rapid shifts in osmolarity in coastal and estuarine regions, as well as in sea-ice environments, can trigger specific metabolic and morphological responses in diatoms (Kirst 1990). Several studies in *P. tricornutum* revealed a key role of intracellular  $\text{Ca}^{2+}$  signalling in diatom osmoregulation (see also Chapter “An Ocean of Signals: Intracellular and Extracellular Signalling”). Elevation of cytosolic  $\text{Ca}^{2+}$  detected with the bioluminescence aequorin reporter were described in response to hypo-, but not hyper-osmotic shock (Falciatore et al. 2000). More recent studies employing novel single-cell imaging techniques revealed intracellular spatiotemporal patterns of osmotic-induced  $\text{Ca}^{2+}$  elevations, regulating  $\text{K}^+$  efflux. These processes enable the fine-tuned cell volume regulation across the whole cell, preventing cell bursting (Helliwell et al. 2021b). Additionally,  $\text{Ca}^{2+}$  independent pathways controlling organic osmolytes efflux also contribute to the diatom osmotic stress tolerance (Kirst 1990; Helliwell et al. 2021b).

There are also evidences in diatoms of sophisticated sensing systems for detecting and responding to fluid motion (shear stress), which also involve transient changes in cytosolic  $\text{Ca}^{2+}$  (Falciatore et al. 2000). In *P. tricornutum*, the response to fluid motion is rapid (seconds) and highly controlled: the cells lose their ability to respond with their original sensitivity, when a second stimulus is applied shortly after the first. With this short-term adaptation or desensitization process, diatoms can reversibly adjust their sensitivity to the level of the stimulus (Falciatore et al. 2000). Differential gene expression analyses and microscopical investigations also revealed morphological and transcriptional changes of the chain-forming diatom *Chaetoceros decipiens* in response to measured levels of turbulence in nutrient replete conditions (Amato et al. 2017). These active responses support the hypothesis that turbulence not only favours nutrient uptake, but it can also act as a signal, that diatoms exploit to reorganize their physiology in particular environmental conditions.

---

## 5 Future Directions

Experimental and *in silico* studies on diatom signalling pathways described in this review are revealing intriguing facts about the environmental sensing abilities of these microalgae. Biophysical and functional characterization of AUREOs, CPFs, and DPHs have provided initial information on the role of these photoreceptors in the control of key biological processes in diatoms. There is still insufficient information to demonstrate the importance of the CRY-DASHs, which are highly represented in the genome of all organisms, including diatoms, and possibly have conserved functions. Similarly, several hypotheses exist on how rhodopsins could be involved in the regulation of diatom metabolism in response to nutritional stress, but

more evidence is needed to understand the pertinence of known rhodopsin variants. Many novel protein domains potentially involved in light sensing have been predicted from diatom genomes, but they are linked to other protein domains in combinations which have not been described before (Bowler et al. 2008). It is important to mention that possible novel light sensors have also been recently identified in planktonic organisms after extensive exploration of metagenomic data (Coesel et al. 2021), which could now be characterized in diatom model species (Falcatore et al. 2020).

Significant work remains to characterize the signalling pathways downstream of the receptors. Intriguingly, equivalents of the components of light signalling pathways of plants (Chen and Chory 2011) have not been found to be encoded in diatom genomes, suggesting differences in the regulatory cascades between these organisms. It is interesting to note that diatom genomes do have several genes predicted to encode bacterial-like proteins with features of two-component phospho-relay systems. Therefore, it is likely that previously unreported combinations of bacterial and eukaryotic signalling systems will be identified by studying receptor-interacting factors and downstream targets with biochemical and genetic approaches.

Much remains to be done to identify sensing and signalling components involved in the responses to other abiotic cues of relevance for diatoms. Diatoms are particularly abundant at high latitudes and show specific temperature acclimation and adaptation responses (Mock et al. 2017; Liang et al. 2019). It is still unclear if these responses are regulated by a true temperature sensor or by the modification of the physical properties of enzymes, proteins, or membranes in the cell caused by temperature changes. There is also the possibility that the signalling pathways transducing temperature information in marine environments overlap with those of other cues (e.g. light), as shown in plants (Lamers et al. 2020). Therefore, the exploration of how diatoms sense and respond to temperature changes will be particularly informative, also to assess the impact of global warming on their growth and distribution. Moreover, though circadian and seasonal rhythms have been described in diatoms, the genetic basis regulating physiological adaptation to varying latitudinal and seasonal photoperiods remain completely unknown. In terrestrial organisms, the endogenous circadian clock has a key role in the regulation of photoperiodism (Millar 2016), coordinating physiological responses not only to photoperiodic cues through specific photoreceptors but also to temperature and other abiotic stresses (Greenham et al. 2017). Uncoupling temperature and photoperiod signals can have a dramatic consequence on the reproductive success of species adapted to specific habitat by altering physiology or predator–prey mismatch. Therefore, as the most advanced molecular phytoplankton model species, diatoms represent the ideal system for the identification and characterization of novel signalling components processing external and internal signals to anticipate and respond to environmental variations, and also to address the relevance of these pathways in the adaptation of phytoplankton to specific latitudinal and temporal niches.

Another challenging question for the future will be to address the ecological relevance of the sensing systems in the real environment, where diatoms are exposed

to a multitude of different signals and stresses. A strong crosstalk between different pathways, like the light-activated processes described here, would be expected to optimize diatom physiology and metabolism in changing environments. We also expect there to be significant interactions between regulatory pathways implicated in the perception of abiotic and biotic cues and ensuing responses. Master regulators of multiple signalling pathway are likely to be discovered by comparing responses to individual and combined signals and stresses. Omics techniques applied to selected mutants could help in identifying common players in environmental perception and responses. An important contribution to these studies could derive from targeted exploration of the vast amount of environmental data that are now available for marine plankton. Analysing the expression of a specific receptor in a particular environment or the identification of its genetic functional variant correlating with specific abiotic or biotic parameters could help to pinpoint novel regulators of diatom life in complex ecosystems. The remarkable diversity of diatoms could be exploited as an additional tool to gain insight into the function of environmental sensors. As noted in this chapter, DPHs and rhodopsins have only been found in some diatom species. By further investigating the diversity and distribution of diatom species possessing a particular receptor or modified receptor-like properties, it will be possible to identify novel regulators of diatom phenotypic plasticity and to assess how they equip these algae to colonize highly diverse marine environments.

**Acknowledgments** We thank Alessandro Manzotti for support for graphic design. Research was supported by funding from the Fondation Bettencourt-Schueller (Coups d'élan pour la recherche française-2018 to A.F.), the grant "DYNAMO" (ANR-11-LABX-0011-01- to A.F.), and the ANR grant CLIMA-CLOCK (ANR-20-CE20-0024 TO A.F.), the Betty Moore Foundation Marine Microbial Initiative (grant GBMF4981.01 to M.J.), a DFG grant (KR1661/19-1) to P.G.K. and the international ANR-DFG Grant "DiaRhythm" to A.F and P.G.K. (KR1661/20-1)

---

## References

- Ait-Mohamed O, Novák Vanclová AMG, Joli N, Liang Y, Zhao X, Genovesio A, Tirichine L, Bowler C, Dorrell RG (2020) PhaeoNet: a holistic RNAseq-based portrait of transcriptional coordination in the model Diatom *Phaeodactylum tricornutum*. *Front Plant Sci* 11:590949. <https://doi.org/10.3389/fpls.2020.590949>
- Akiyama Y, Nakasone Y, Nakatani Y, Hisatomi O, Terazima M (2016) Time-resolved detection of light-induced dimerization of monomeric Aureochrome-1 and change in affinity for DNA. *J Phys Chem B* 120:7360–7370. <https://doi.org/10.1021/acs.jpcc.6b05760>
- Amato A, Dell'Aquila G, Musacchia F, Annunziata R, Ugarte A, Maillat N, Carbone A, Ribera d'Alcalà M, Sanges R, Iudicone D, Ferrante MI (2017) Marine diatoms change their gene expression profile when exposed to microscale turbulence under nutrient replete conditions. *Sci Rep* 7:3826. <https://doi.org/10.1038/s41598-017-03741-6>
- Annunziata R, Ritter A, Fortunato AE, Manzotti A, Cheminant-Navarro S, Agier N, Huysman MJJ, Winge P, Bones AM, Bouget F-Y, Cosentino Lagomarsino M, Bouly J-P, Falciatore A (2019) bHLH-PAS protein RITMO1 regulates diel biological rhythms in the marine diatom *Phaeodactylum tricornutum*. *Proc Natl Acad Sci* 116:13137–13142. <https://doi.org/10.1073/pnas.1819660116>

- Archibald JM (2015) Endosymbiosis and eukaryotic cell evolution. *Curr Biol CB* 25:R911–R921. <https://doi.org/10.1016/j.cub.2015.07.055>
- Armbrust EV et al (2004) The genome of the diatom *Thalassiosira pseudonana* ecology, evolution, and metabolism. *Science* 306:79–86. <https://doi.org/10.1126/science.1101156>
- Ashworth J, Coesel S, Lee A, Armbrust EV, Orellana MV, Baliga NS (2013) Genome-wide diel growth state transitions in the diatom *Thalassiosira pseudonana*. *Proc Natl Acad Sci* 110:7518–7523. <https://doi.org/10.1073/pnas.1300962110>
- Auldridge ME, Forest KT (2011) Bacterial phytochromes: more than meets the light. *Crit Rev Biochem Mol Biol* 46:67–88. <https://doi.org/10.3109/10409238.2010.546389>
- Baatz I (1941) Die bedeutung der lichtqualität für wachstum und stoffproduktion planktonischer meeresdiatomeen. *Planta* 31:726–766
- Banerjee A, Herman E, Kottke T, Essen L-O (2016a) Structure of a native-like Aureochrome 1a LOV domain dimer from *Phaeodactylum tricornutum*. *Structure* 24:171–178. <https://doi.org/10.1016/j.str.2015.10.022>
- Banerjee A, Herman E, Serif M, Maestre-Reyna M, Hepp S, Pokorny R, Kroth PG, Essen L-O, Kottke T (2016b) Allosteric communication between DNA-binding and light-responsive domains of diatom class I aureochromes. *Nucleic Acids Res* 44:5957–5970. <https://doi.org/10.1093/nar/gkw420>
- Bannister S, Böhm E, Zinn T, Hellweg T, Kottke T (2019) Arguments for an additional long-lived intermediate in the photocycle of the full-length aureochrome 1c receptor: a time-resolved small-angle X-ray scattering study. *Struct Dyn Melville N* 6:034701. <https://doi.org/10.1063/1.5095063>
- Barnett A, Méléder V, Dupuy C, Lavaud J (2020) The vertical migratory rhythm of intertidal microphytobenthos in sediment depends on the light photoperiod, intensity, and spectrum: evidence for a positive effect of blue wavelengths. *Front Mar Sci* 7:212. <https://doi.org/10.3389/fmars.2020.00212>
- Basu S et al (2017) Finding a partner in the ocean: molecular and evolutionary bases of the response to sexual cues in a planktonic diatom. *New Phytol* 215:140–156. <https://doi.org/10.1111/nph.14557>
- Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF (2001) Proteorhodopsin phototrophy in the ocean. *Nature* 411:786–789. <https://doi.org/10.1038/35081051>
- Bilcke G et al (2021) Mating type specific transcriptomic response to sex inducing pheromone in the pennate diatom *Seminavis robusta*. *ISME J* 15:562–576. <https://doi.org/10.1038/s41396-020-00797-7>
- Bowler C et al (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456:239–244. <https://doi.org/10.1038/nature07410>
- Cao S, Wang J, Chen D (2013) Settlement and cell division of diatom *Navicula* can be influenced by light of various qualities and intensities: settlement and cell division of diatom *Navicula*. *J Basic Microbiol* 53:884–894. <https://doi.org/10.1002/jobm.201200315>
- Chantranupong L, Wolfson RL, Sabatini DM (2015) Nutrient-sensing mechanisms across evolution. *Cell* 161:67–83. <https://doi.org/10.1016/j.cell.2015.02.041>
- Chauton MS, Winge P, Brembu T, Vadstein O, Bones AM (2013) Gene regulation of carbon fixation, storage, and utilization in the diatom *Phaeodactylum tricornutum* acclimated to light/dark cycles. *Plant Physiol* 161:1034–1048. <https://doi.org/10.1104/pp.112.206177>
- Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T, Brettel K, Essen L-O, van der Horst GTJ, Batschauer A, Ahmad M (2011) The cryptochromes: blue light photoreceptors in plants and animals. *Annu Rev Plant Biol* 62:335–364. <https://doi.org/10.1146/annurev-arplant-042110-103759>
- Chellamuthu VR, Alva V, Forchhammer K (2013) From cyanobacteria to plants: conservation of PII homologs during plastid evolution. *Planta* 237:451–462. <https://doi.org/10.1007/s00425-012-1801-0>
- Chen M, Chory J (2011) Phytochrome signaling mechanisms and the control of plant development. *Trends Cell Biol* 21:664–671. <https://doi.org/10.1016/j.tcb.2011.07.002>

- Christie JM (2007) Phototropin blue-light receptors. *Annu Rev Plant Biol* 58:21–45. <https://doi.org/10.1146/annurev.arplant.58.032806.103951>
- Coesel S, Mangogna M, Ishikawa T, Heijde M, Rogato A, Finazzi G, Todo T, Bowler C, Falciatore A (2009) Diatom PtCPF1 is a new cryptochrome/photolyase family member with DNA repair and transcription regulation activity. *EMBO Rep* 10:655–661. <https://doi.org/10.1038/embor.2009.59>
- Coesel SN, Durham BP, Groussman RD, Hu SK, Caron DA, Morales RL, Ribalet F, Armbrust EV (2021) Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. *Proc Natl Acad Sci* 118:e2011038118. <https://doi.org/10.1073/pnas.2011038118>
- Cohn SA, Weitzell RE (1996) Ecological considerations of diatom cell motility. I. Characterization of motility and adhesion in four diatom species. *J Phycol* 32:928–939. <https://doi.org/10.1111/j.0022-3646.1996.00928.x>
- Costanzi S, Siegel J, Tikhonova IG, Jacobson KA (2009) Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr Pharm Des* 15:3994–4002. <https://doi.org/10.2174/138161209789824795>
- Crosson S, Rajagopal S, Moffat K (2003) The LOV domain family: photoresponsive signaling modules coupled to diverse output domains. *Biochemistry* 42:2–10. <https://doi.org/10.1021/bi026978i>
- Cruz de Carvalho MH, Bowler C (2020) Global identification of a marine diatom long noncoding natural antisense transcripts (NATs) and their response to phosphate fluctuations. *Sci Rep* 10:14110. <https://doi.org/10.1038/s41598-020-71002-0>
- Cruz de Carvalho MH, Sun H-X, Bowler C, Chua N-H (2016) Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations. *New Phytol* 210:497–510. <https://doi.org/10.1111/nph.13787>
- De Riso V, Raniello R, Maumus F, Rogato A, Bowler C, Falciatore A (2009) Gene silencing in the marine diatom *Phaeodactylum tricorutum*. *Nucleic Acids Res* 37:e96. <https://doi.org/10.1093/nar/gkp448>
- Di Roberto RB, Peisajovich SG (2014) The role of domain shuffling in the evolution of signaling networks: signaling networks evolution. *J Exp Zool B Mol Dev Evol* 322:65–72. <https://doi.org/10.1002/jez.b.22551>
- Dobrenel T, Caldana C, Hanson J, Robaglia C, Vincenz M, Veit B, Meyer C (2016) TOR signaling and nutrient sensing. *Annu Rev Plant Biol* 67:261–285. <https://doi.org/10.1146/annurev-arplant-043014-114648>
- Dodd AN, Belbin FE, Frank A, Webb AAR (2015) Interactions between circadian clocks and photosynthesis for the temporal and spatial coordination of metabolism. *Front Plant Sci* 6:245. <https://doi.org/10.3389/fpls.2015.00245>
- Dröge-Laser W, Snoek BL, Snel B, Weiste C (2018) The Arabidopsis bZIP transcription factor family—an update. *Curr Opin Plant Biol* 45:36–49. <https://doi.org/10.1016/j.pbi.2018.05.001>
- Duanmu D et al (2014) Marine algae and land plants share conserved phytochrome signaling systems. *Proc Natl Acad Sci* 111:15827–15832. <https://doi.org/10.1073/pnas.1416751111>
- Duanmu D, Rockwell NC, Lagarias JC (2017) Algal light sensing and photoacclimation in aquatic environments: Bilin and light signalling in eukaryotic algae. *Plant Cell Environ* 40:2558–2570. <https://doi.org/10.1111/pce.12943>
- Dunlap JC (1999) Molecular bases for circadian clocks. *Cell* 96:271–290. [https://doi.org/10.1016/S0092-8674\(00\)80566-8](https://doi.org/10.1016/S0092-8674(00)80566-8)
- Eilertsen HC, Sandberg S, Tøllefsen H, Tellefsen H (1995) Photoperiodic control of diatom spore growth: a theory to explain the onset of phytoplankton blooms. *Mar Ecol Prog Ser* 116:303–307
- Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H (2014) Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem Rev* 114:126–163. <https://doi.org/10.1021/cr4003769>

- Escoubas JM, Lomas M, LaRoche J, Falkowski PG (1995) Light intensity regulation of *cab* gene transcription is signaled by the redox state of the plastoquinone pool. *Proc Natl Acad Sci* 92: 10237–10241. <https://doi.org/10.1073/pnas.92.22.10237>
- Essen L-O, Franz S, Banerjee A (2017) Structural and evolutionary aspects of algal blue light receptors of the cryptochrome and aureochrome type. *J Plant Physiol* 217:27–37. <https://doi.org/10.1016/j.jplph.2017.07.005>
- Falciatore A, d'Alcalà MR, Croot P, Bowler C (2000) Perception of environmental signals by a marine diatom. *Science* 288:2363–2366
- Falciatore A, Jaubert M, Bouly J-P, Bailleul B, Mock T (2020) Diatom molecular research comes of age: model species for studying phytoplankton biology and diversity. *Plant Cell* 32:547–572. <https://doi.org/10.1105/tpc.19.00158>
- Fan X, Naz M, Fan X, Xuan W, Miller AJ, Xu G (2017) Plant nitrate transporters: from gene function to application. *J Exp Bot* 68:2463–2475. <https://doi.org/10.1093/jxb/erx011>
- Finkel OM, Béjã O, Belkin S (2013) Global abundance of microbial rhodopsins. *ISME J* 7:448–451. <https://doi.org/10.1038/ismej.2012.112>
- Fisher AE, Berges JA, Harrison PJ (1996) Does light quality affect the sinking rates of marine diatoms? *J Phycol* 32:353–360. <https://doi.org/10.1111/j.0022-3646.1996.00353.x>
- Fortunato AE, Annunziata R, Jaubert M, Bouly J-P, Falciatore A (2015) Dealing with light: the widespread and multitasking cryptochrome/photolyase family in photosynthetic organisms. *J Plant Physiol* 172:42–54. <https://doi.org/10.1016/j.jplph.2014.06.011>
- Fortunato AE et al (2016) Diatom phytochromes reveal the existence of far-red-light-based sensing in the ocean. *Plant Cell* 28:616–628. <https://doi.org/10.1105/tpc.15.00928>
- Froehlich AC, Chen C-H, Belden WJ, Madeti C, Roenneberg T, Meroow M, Loros JJ, Dunlap JC (2010) Genetic and molecular characterization of a cryptochrome from the filamentous fungus *Neurospora crassa*. *Eukaryot Cell* 9:738–750. <https://doi.org/10.1128/EC.00380-09>
- Fujita Y, Ohki K (2004) On the 710 nm fluorescence emitted by the diatom *Phaeodactylum tricornutum* at room temperature. *Plant Cell Physiol* 45:392–397
- Furukawa T, Watanabe M, Shihira-Ishikawa I (1998) Green- and blue-light-mediated chloroplast migration in the centric diatom *Pleurosira laevis*. *Protoplasma* 203:214–220. <https://doi.org/10.1007/BF01279479>
- Galachyants YP et al (2015) Sequencing of the complete genome of an araphid pennate diatom *Synedra acus* subsp. *radians* from Lake Baikal. *Dokl Biochem Biophys* 461:84–88. <https://doi.org/10.1134/S1607672915020064>
- Goett-Zink L, Klocke JL, Bögeholz LAK, Kottke T (2020) In-cell infrared difference spectroscopy of LOV photoreceptors reveals structural responses to light altered in living cells. *J Biol Chem* 295:11729–11741. <https://doi.org/10.1074/jbc.RA120.013091>
- Gojon A, Krouk G, Perrine-Walker F, Laugier E (2011) Nitrate transporter(s) in plants. *J Exp Bot* 62:2299–2308. <https://doi.org/10.1093/jxb/erq419>
- Graff van Creveld S, Rosenwasser S, Schatz D, Koren I, Vardi A (2015) Early perturbation in mitochondria redox homeostasis in response to environmental stress predicts cell fate in diatoms. *ISME J* 9:385–395. <https://doi.org/10.1038/ismej.2014.136>
- Greenham K, Lou P, Puzey JR, Kumar G, Arnevik C, Farid H, Willis JH, McClung CR (2017) Geographic variation of plant circadian clock function in natural and agricultural settings. *J Biol Rhythm* 32:26–34. <https://doi.org/10.1177/0748730416679307>
- Grote M, Engelhard M, Hegemann P (2014) Of ion pumps, sensors and channels—Perspectives on microbial rhodopsins between science and history. *Biochim Biophys Acta BBA—Bioenerg* 1837:533–545. <https://doi.org/10.1016/j.bbabi.2013.08.006>
- Harada H, Nakatsuma D, Ishida M, Matsuda Y (2005) Regulation of the expression of intracellular beta-carbonic anhydrase in response to CO<sub>2</sub> and light in the marine diatom *Phaeodactylum tricornutum*. *Plant Physiol* 139:1041–1050. <https://doi.org/10.1104/pp.105.065185>
- Harada H, Nakajima K, Sakaue K, Matsuda Y (2006) CO<sub>2</sub> sensing at ocean surface mediated by cAMP in a marine diatom. *Plant Physiol* 142:1318–1328. <https://doi.org/10.1104/pp.106.086561>



- Harding LW, Meeson BW, Prézélin BB, Sweeney BM (1981) Diel periodicity of photosynthesis in marine phytoplankton. *Mar Biol* 61:95–105. <https://doi.org/10.1007/BF00386649>
- Haydon MJ, Mielczarek O, Robertson FC, Hubbard KE, Webb AAR (2013) Photosynthetic entrainment of the *Arabidopsis thaliana* circadian clock. *Nature* 502:689–692. <https://doi.org/10.1038/nature12603>
- Hedtke M, Rauscher S, Röhrig J, Rodríguez-Romero J, Yu Z, Fischer R (2015) Light-dependent gene activation in *Aspergillus nidulans* is strictly dependent on phytochrome and involves the interplay of phytochrome and white collar-regulated histone H3 acetylation. *Mol Microbiol* 97:733–745. <https://doi.org/10.1111/mmi.13062>
- Heintz U, Schlichting I (2016) Blue light-induced LOV domain dimerization enhances the affinity of Aureochrome 1a for its target DNA sequence. *elife* 5:e11860. <https://doi.org/10.7554/eLife.11860>
- Helliwell KE et al (2021a) A novel Ca<sup>2+</sup> signaling pathway coordinates environmental phosphorus sensing and nitrogen metabolism in marine diatoms. *Curr Biol* 31:978–989. <https://doi.org/10.1016/j.cub.2020.11.073>
- Helliwell KE, Kleiner FH, Hardstaff H, Chrachri A, Gaikwad T, Salmon D, Smirnoff N, Wheeler GL, Brownlee C (2021b) Spatiotemporal patterns of intracellular Ca<sup>2+</sup> signalling govern hypo-osmotic stress resilience in marine diatoms. *New Phytol* 230:155–170. <https://doi.org/10.1111/nph.17162>
- Hepp S, Trauth J, Hasenjäger S, Bezold F, Essen L-O, Taxis C (2020) An optogenetic tool for induced protein stabilization based on the *Phaeodactylum tricorutum* Aureochrome 1a light-oxygen-voltage domain. *J Mol Biol* 432:1880–1900. <https://doi.org/10.1016/j.jmb.2020.02.019>
- Herbstová M, Bína D, Koník P, Gardian Z, Vácha F, Litvín R (2015) Molecular basis of chromatic adaptation in pennate diatom *Phaeodactylum tricorutum*. *Biochim Biophys Acta Bioenerg* 1847:534–543. <https://doi.org/10.1016/j.bbabi.2015.02.016>
- Herman E, Kottke T (2015) Allosterically regulated unfolding of the A'α helix exposes the dimerization site of the blue-light-sensing aureochrome-LOV domain. *Biochemistry* 54:1484–1492. <https://doi.org/10.1021/bi501509z>
- Herman E, Sachse M, Kroth PG, Kottke T (2013) Blue-light-induced unfolding of the Jα helix allows for the dimerization of aureochrome-LOV from the diatom *Phaeodactylum tricorutum*. *Biochemistry* 52:3094–3101. <https://doi.org/10.1021/bi400197u>
- Hernández-Verdeja T, Strand Å (2018) Retrograde signals navigate the path to chloroplast development. *Plant Physiol* 176:967–976. <https://doi.org/10.1104/pp.17.01299>
- Hisatomi O, Furuya K (2015) A light-regulated bZIP module, photozipper, induces the binding of fused proteins to the target DNA sequence in a blue light-dependent manner. *Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol* 14:1998–2006. <https://doi.org/10.1039/c5pp00178a>
- Huysman MJJ et al (2013) AUREOCHROME1a-mediated induction of the diatom-specific cyclin dsCYC2 controls the onset of cell division in diatoms (*Phaeodactylum tricorutum*). *Plant Cell* 25:215–228. <https://doi.org/10.1105/tpc.112.106377>
- Ingargiola C, Turqueto Duarte G, Robaglia C, Leprince A-S, Meyer C (2020) The plant target of rapamycin: a conduit TOR of nutrition and metabolism in photosynthetic organisms. *Genes* 11:1285. <https://doi.org/10.3390/genes11111285>
- Ishikawa M, Takahashi F, Nozaki H, Nagasato C, Motomura T, Kataoka H (2009) Distribution and phylogeny of the blue light receptors aureochromes in eukaryotes. *Planta* 230:543–552. <https://doi.org/10.1007/s00425-009-0967-6>
- Jaubert M, Bouly J-P, Ribera d'Alcalà M, Falciatore A (2017) Light sensing and responses in marine microalgae. *Curr Opin Plant Biol* 37:70–77. <https://doi.org/10.1016/j.pbi.2017.03.005>
- Ji N, Li L, Lin L, Lin S (2017) Identification and expression analysis of blue light receptor aureochrome in the harmful alga *Heterosigma akashiwo* (raphidophyceae). *Harmful Algae* 61:71–79. <https://doi.org/10.1016/j.hal.2016.11.016>

- Juhas M, von Zadow A, Spexard M, Schmidt M, Kottke T, Büchel C (2014) A novel cryptochrome in the diatom *Phaeodactylum tricoratum* influences the regulation of light-harvesting protein levels. *FEBS J* 281:2299–2311. <https://doi.org/10.1111/febs.12782>
- Jungandreas A, Schellenberger Costa B, Jakob T, von Bergen M, Baumann S, Wilhelm C (2014) The acclimation of *Phaeodactylum tricoratum* to blue and red light does not influence the photosynthetic light reaction but strongly disturbs the carbon allocation pattern. *PLoS One* 9: e99727. <https://doi.org/10.1371/journal.pone.0099727>
- Kendrick RE (ed) (1994) *Photomorphogenesis in plants*, 2nd. edn. Kluwer Academic Publ, Dordrecht
- Kerruth S, Ataka K, Frey D, Schlichting I, Heberle J (2014) Aureochrome 1 illuminated: structural changes of a transcription factor probed by molecular spectroscopy. *PLoS One* 9:e103307. <https://doi.org/10.1371/journal.pone.0103307>
- Kikutani S, Tanaka R, Yamazaki Y, Hara S, Hisabori T, Kroth PG, Matsuda Y (2012) Redox regulation of carbonic anhydrases via thioredoxin in chloroplast of the marine diatom *Phaeodactylum tricoratum*. *J Biol Chem* 287:20689–20700. <https://doi.org/10.1074/jbc.M111.322743>
- Kirk JTO (2011) *Light and photosynthesis in aquatic ecosystems*, 3rd edn. Cambridge University Press
- Kirst GO (1990) Salinity tolerance of eukaryotic marine algae. *Annu Rev Plant Physiol Plant Mol Biol* 41:21–53. <https://doi.org/10.1146/annurev.pp.41.060190.000321>
- Kleine T, Lockhart P, Batschauer A (2003) An Arabidopsis protein closely related to Synechocystis cryptochrome is targeted to organelles: organelle-targeted cryptochrome. *Plant J* 35:93–103. <https://doi.org/10.1046/j.1365-313X.2003.01787.x>
- Kobayashi I, Nakajima H, Hisatomi O (2020) Molecular mechanism of light-induced conformational switching of the LOV domain in aureochrome-1. *Biochemistry* 59:2592–2601. <https://doi.org/10.1021/acs.biochem.0c00271>
- Kolody B, McCrow JP, Allen LZ, Aylward FO, Fontanez KM, Moustafa A, Moniruzzaman M, Chavez FP, Scholin CA, Allen EE, Worden AZ, Delong EF, Allen AE (2019) Diel transcriptional response of a California Current plankton microbiome to light, low iron, and enduring viral infection. *ISME J* 13:2817–2833. <https://doi.org/10.1038/s41396-019-0472-2>
- König S, Eisenhut M, Bräutigam A, Kurz S, Weber APM, Büchel C (2017) The influence of a cryptochrome on the gene expression profile in the diatom *Phaeodactylum tricoratum* under blue light and in darkness. *Plant Cell Physiol* 58:1914–1923. <https://doi.org/10.1093/pcp/pcx127>
- Kroth PG, Wilhelm C, Kottke T (2017) An update on aureochromes: phylogeny – mechanism – function. *J Plant Physiol* 217:20–26. <https://doi.org/10.1016/j.jplph.2017.06.010>
- Kumar Sharma A, Mühlroth A, Jouhet J, Maréchal E, Alipanah L, Kissen R, Brembu T, Bones AM, Winge P (2020) The Myb-like transcription factor phosphorus starvation response (PtPSR) controls conditional P acquisition and remodelling in marine microalgae. *New Phytol* 225: 2380–2395. <https://doi.org/10.1111/nph.16248>
- Kume K, Zylka MJ, Sriram S, Shearman LP, Weaver DR, Jin X, Maywood ES, Hastings MH, Reppert SM (1999) mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* 98:193–205. [https://doi.org/10.1016/S0092-8674\(00\)81014-4](https://doi.org/10.1016/S0092-8674(00)81014-4)
- Lamers J, van der Meer T, Testerink C (2020) How plants sense and respond to stressful environments. *Plant Physiol* 182:1624–1635. <https://doi.org/10.1104/pp.19.01464>
- Lepetit B, Sturm S, Rogato A, Gruber A, Sachse M, Falciatore A, Kroth PG, Lavaud J (2013) High light acclimation in the secondary plastids containing diatom *Phaeodactylum tricoratum* is triggered by the redox state of the plastoquinone pool. *Plant Physiol* 161:853–865. <https://doi.org/10.1104/pp.112.207811>
- Li F-W, Melkonian M, Rothfels CJ, Villarreal JC, Stevenson DW, Graham SW, Wong GK-S, Pryer KM, Mathews S (2015) Phytochrome diversity in green plants and the origin of canonical plant phytochromes. *Nat Commun* 6:7852. <https://doi.org/10.1038/ncomms8852>

- Liang Y, Koester JA, Liefer JD, Irwin AJ, Finkel ZV (2019) Molecular mechanisms of temperature acclimation and adaptation in marine diatoms. *ISME J* 13:2415–2425. <https://doi.org/10.1038/s41396-019-0441-9>
- Lipps MJ (1973) The determination of the far-red effect in marine phytoplankton. *J Phycol* 9:237–242. <https://doi.org/10.1111/j.1529-8817.1973.tb04087.x>
- Lommer M et al (2012) Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol* 13:R66. <https://doi.org/10.1186/gb-2012-13-7-r66>
- Lopez-Gomollon S, Beckers M, Rathjen T, Moxon S, Maumus F, Mohorianu I, Moulton V, Dalmay T, Mock T (2014) Global discovery and characterization of small non-coding RNAs in marine microalgae. *BMC Genomics* 15:697. <https://doi.org/10.1186/1471-2164-15-697>
- Madhuri S, Río Bártulos C, Serif M, Lepetit B, Kroth PG (2019) A strategy to complement PtAUREO1a in TALEN knockout strains of *Phaeodactylum tricorutum*. *Algal Res* 39: 101469. <https://doi.org/10.1016/j.algal.2019.101469>
- Man D (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* 22:1725–1731. <https://doi.org/10.1093/emboj/cdg183>
- Mann KH, Lazier JRN (2006) Dynamics of marine ecosystems: biological-physical interactions in the oceans. Blackwell Pub, Malden, MA
- Mann M, Serif M, Jakob T, Kroth PG, Wilhelm C (2017) PtAUREO1a and PtAUREO1b knockout mutants of the diatom *Phaeodactylum tricorutum* are blocked in photoacclimation to blue light. *J Plant Physiol* 217:44–48. <https://doi.org/10.1016/j.jplph.2017.05.020>
- Mann M, Serif M, Wrobel T, Eisenhut M, Madhuri S, Flachbart S, Weber APM, Lepetit B, Wilhelm C, Kroth PG (2020) The aureochrome photoreceptor PtAUREO1a is a highly effective blue light switch in diatoms. *iScience* 23:101730. <https://doi.org/10.1016/j.isci.2020.101730>
- Marchetti A, Schrueth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, Morales R, Allen AE, Armbrust EV (2012) Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci U S A* 109:E317–E325. <https://doi.org/10.1073/pnas.1118408109>
- Marchetti A, Catlett D, Hopkinson BM, Ellis K, Cassar N (2015) Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *ISME J* 9:2745–2748. <https://doi.org/10.1038/ismej.2015.74>
- Matthijs M, Fabris M, Broos S, Vyverman W, Goossens A (2016) Profiling of the early nitrogen stress response in the diatom *Phaeodactylum tricorutum* reveals a novel family of RING-domain transcription factors. *Plant Physiol* 170:489–498. <https://doi.org/10.1104/pp.15.01300>
- 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 tricorutum*. *EMBO J* 36:1559–1576. <https://doi.org/10.15252/embj.201696392>
- Maumus F, Allen AE, Mhiri C, Hu H, Jabbari K, Vardi A, Grandbastien M-A, Bowler C (2009) Potential impact of stress activated retrotransposons on genome evolution in a marine diatom. *BMC Genomics* 10:624. <https://doi.org/10.1186/1471-2164-10-624>
- McLachlan DH, Brownlee C, Taylor AR, Geider RJ, Underwood GJC (2009) Light-induced motile responses of the estuarine benthic diatoms *Navicula perminuta* and *Cylindrotheca closterium* (bacillariophyceae). *J Phycol* 45:592–599. <https://doi.org/10.1111/j.1529-8817.2009.00681.x>
- Millar AJ (2016) The intracellular dynamics of circadian clocks reach for the light of ecology and evolution. *Annu Rev Plant Biol* 67:595–618. <https://doi.org/10.1146/annurev-arplant-043014-115619>
- Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F (2011) ROS signaling: the new wave? *Trends Plant Sci* 16: 300–309. <https://doi.org/10.1016/j.tplants.2011.03.007>
- Mizrachi A, Graff van Creveld S, Shapiro OH, Rosenwasser S, Vardi A (2019) Light-dependent single-cell heterogeneity in the chloroplast redox state regulates cell fate in a marine diatom. *elife* 8:e47732. <https://doi.org/10.7554/eLife.47732>

- Mock T et al (2017) Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* 541:536–540. <https://doi.org/10.1038/nature20803>
- Mouget J-L, Rosa P, Vachoux C, Tremblin G (2005) Enhancement of marennine production by blue light in the diatom *Haslea ostrearia*. *J Appl Phycol* 17:437–445. <https://doi.org/10.1007/s10811-005-0561-7>
- Mouget J-L, Gastineau R, Davidovich O, Gaudin P, Davidovich NA (2009) Light is a key factor in triggering sexual reproduction in the pennate diatom *Haslea ostrearia*: light induction of sexual reproduction in diatoms. *FEMS Microbiol Ecol* 69:194–201. <https://doi.org/10.1111/j.1574-6941.2009.00700.x>
- Nakajima H, Kobayashi I, Adachi Y, Hisatomi O (2021) Transmission of light signals from the light-oxygen-voltage core via the hydrophobic region of the  $\beta$ -sheet surface in aureochrome-1. *Sci Rep* 11:11995. <https://doi.org/10.1038/s41598-021-91497-5>
- Nakatani Y, Hisatomi O (2018) Quantitative analyses of the equilibria among DNA complexes of a blue-light-regulated bZIP module, Photozipper. *Biophys Physicobiology* 15:8–17. [https://doi.org/10.2142/biophysico.15.0\\_8](https://doi.org/10.2142/biophysico.15.0_8)
- Nultsch W (1971) Phototactic and photokinetic action spectra of the diatom *Nitzschia communis*. *Photochem Photobiol* 14:705–712. <https://doi.org/10.1111/j.1751-1097.1971.tb06209.x>
- 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 tricorutum*. *PLoS One* 4:e7743. <https://doi.org/10.1371/journal.pone.0007743>
- Nymark M, Valle KC, Hancke K, Winge P, Andresen K, Johnsen G, Bones AM, Brembu T (2013) Molecular and photosynthetic responses to prolonged darkness and subsequent acclimation to re-illumination in the diatom *Phaeodactylum tricorutum*. *PLoS One* 8:e58722. <https://doi.org/10.1371/journal.pone.0058722>
- Ohno N, Inoue T, Yamashiki R, Nakajima K, Kitahara Y, Ishibashi M, Matsuda Y (2012) CO<sub>2</sub>-cAMP-responsive cis-elements targeted by a transcription factor with CREB/ATF-like basic zipper domain in the marine diatom *Phaeodactylum tricorutum*. *Plant Physiol* 158:499–513. <https://doi.org/10.1104/pp.111.190249>
- Osuna-Cruz CM et al (2020) The *Seminavis robusta* genome provides insights into the evolutionary adaptations of benthic diatoms. *Nat Commun* 11:3320. <https://doi.org/10.1038/s41467-020-17191-8>
- Ozturk N (2017) Phylogenetic and functional classification of the photolyase/cryptochrome family. *Photochem Photobiol* 93:104–111. <https://doi.org/10.1111/php.12676>
- Port JA, Parker MS, Kodner RB, Wallace JC, Armbrust E, Faustman EM (2013) Identification of G protein-coupled receptor signaling pathway proteins in marine diatoms using comparative genomics. *BMC Genomics* 14:503. <https://doi.org/10.1186/1471-2164-14-503>
- Prioretti L, Avilan L, Carrière F, Montané M-H, Field B, Grégori G, Menand B, Gontero B (2017) The inhibition of TOR in the model diatom *Phaeodactylum tricorutum* promotes a get-fat growth regime. *Algal Res* 26:265–274. <https://doi.org/10.1016/j.algal.2017.08.009>
- Pushkarev A et al (2018) A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* 558:595–599. <https://doi.org/10.1038/s41586-018-0225-9>
- Ragni M, d'Alcalà MR (2007) Circadian variability in the photobiology of *Phaeodactylum tricorutum*: pigment content. *J Plankton Res* 29:141–156. <https://doi.org/10.1093/plankt/fbm002>
- Rayko E, Maumus F, Maheswari U, Jabbari K, Bowler C (2010) Transcription factor families inferred from genome sequences of photosynthetic stramenopiles. *New Phytol* 188:52–66. <https://doi.org/10.1111/j.1469-8137.2010.03371.x>
- Ren H, Junmin L, Qiuqi L, Boping H (2005) Study of light signal receptor of *Stephanopyxis palmeriana* during sexual reproduction. *Chin J Oceanol Limnol* 23:330–334. <https://doi.org/10.1007/BF02847156>
- Rockwell NC, Lagarias JC (2020) Phytochrome evolution in 3D: deletion, duplication, and diversification. *New Phytol* 225:2283–2300. <https://doi.org/10.1111/nph.16240>

- Rockwell NC, Duanmu D, Martin SS, Bachy C, Price DC, Bhattacharya D, Worden AZ, Lagarias JC (2014) Eukaryotic algal phytochromes span the visible spectrum. *Proc Natl Acad Sci* 111: 3871–3876. <https://doi.org/10.1073/pnas.1401871111>
- Rogato A, Richard H, Sarazin A, Voss B, Cheminant Navarro S, Champeimont R, Navarro L, Carbone A, Hess WR, Falcitore A (2014) The diversity of small non-coding RNAs in the diatom *Phaeodactylum tricoratum*. *BMC Genomics* 15:698. <https://doi.org/10.1186/1471-2164-15-698>
- Rogato A, Amato A, Iudicone D, Chiurazzi M, Ferrante MI, d'Alcalà MR (2015) The diatom molecular toolkit to handle nitrogen uptake. *Mar Genomics* 24:95–108. <https://doi.org/10.1016/j.margen.2015.05.018>
- Rosenwasser S, Graff van Creveld S, Schatz D, Malitsky S, Tzfadia O, Aharoni A, Levin Y, Gabashvili A, Feldmesser E, Vardi A (2014) Mapping the diatom redox-sensitive proteome provides insight into response to nitrogen stress in the marine environment. *Proc Natl Acad Sci U S A* 111:2740–2745. <https://doi.org/10.1073/pnas.1319773111>
- Round FE, Hapley CM (1965) Persistent, vertical-migration rhythms in benthic microflora: part IV a diurnal rhythm of the epipelagic diatom association in non-tidal flowing water. *Br Phycol Bull* 2: 463–471. <https://doi.org/10.1080/00071616500650081>
- Saburova M, Polikarpov I (2003) Diatom activity within soft sediments: behavioural and physiological processes. *Mar Ecol Prog Ser* 251:115–126. <https://doi.org/10.3354/meps251115>
- Santin A, Caputi L, Longo A, Chiurazzi M, Ribera d'Alcalà M, Russo MT, Ferrante MI, Rogato A (2021) Integrative omics identification, evolutionary and structural analysis of low affinity nitrate transporters in diatoms, diNPFs. *Open Biol* 11:200395. <https://doi.org/10.1098/rsob.200395>
- Schellenberger Costa B, Jungandreas A, Jakob T, Weisheit W, Mittag M, Wilhelm C (2013a) Blue light is essential for high light acclimation and photoprotection in the diatom *Phaeodactylum tricoratum*. *J Exp Bot* 64:483–493. <https://doi.org/10.1093/jxb/ers340>
- Schellenberger Costa B, Sachse M, Jungandreas A, Bartulos CR, Gruber A, Jakob T, Kroth PG, Wilhelm C (2013b) Aureochrome 1a is involved in the photoacclimation of the diatom *Phaeodactylum tricoratum*. *PLoS One* 8:e74451. <https://doi.org/10.1371/journal.pone.0074451>
- Seckbach J, Gordon R (eds) (2019) *Diatoms: fundamentals and applications*, 1st edn. Wiley
- Serif M (2017) *Characterization of Aureochromes in the diatom Phaeodactylum tricoratum*. Dissertation, University of Konstanz
- Serif M, Lepetit B, Weißert K, Kroth PG, Rio Bartulos C (2017) A fast and reliable strategy to generate TALEN-mediated gene knockouts in the diatom *Phaeodactylum tricoratum*. *Algal Res* 23:186–195. <https://doi.org/10.1016/j.algal.2017.02.005>
- Shikata T, Nukata A, Yoshikawa S, Matsubara T, Yamasaki Y, Shimasaki Y, Oshima Y, Honjo T (2009) Effects of light quality on initiation and development of meroplanktonic diatom blooms in a eutrophic shallow sea. *Mar Biol* 156:875–889. <https://doi.org/10.1007/s00227-009-1131-3>
- Shikata T, Iseki M, Matsunaga S, Higashi S, Kamei Y, Watanabe M (2011) Blue and red light-induced germination of resting spores in the red-tide diatom *Leptocylindrus danicus*. *Photochem Photobiol* 87:590–597. <https://doi.org/10.1111/j.1751-1097.2011.00914.x>
- Shrestha RP, Hildebrand M (2015) Evidence for a regulatory role of diatom silicon transporters in cellular silicon responses. *Eukaryot Cell* 14:29–40. <https://doi.org/10.1128/EC.00209-14>
- Smith SR et al (2016) Transcriptional orchestration of the global cellular response of a model pennate diatom to diel light cycling under iron limitation. *PLoS Genet* 12:e1006490. <https://doi.org/10.1371/journal.pgen.1006490>
- Sverdup HU, Johnson MW, Fleming RH (1942) *The oceans: their physics, chemistry, and general biology*. Prentice-Hall
- Taddei L et al (2016) Multisignal control of expression of the LHCX protein family in the marine diatom *Phaeodactylum tricoratum*. *J Exp Bot* 67:3939–3951. <https://doi.org/10.1093/jxb/erw198>

- Takahashi F, Yamagata D, Ishikawa M, Fukamatsu Y, Ogura Y, Kasahara M, Kiyosue T, Kikuyama M, Wada M, Kataoka H (2007) AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *Proc Natl Acad Sci* 104:19625–19630. <https://doi.org/10.1073/pnas.0707692104>
- Tan MH, Smith SR, Hixson KK, Tan J, McCarthy JK, Kustka AB, Allen AE (2020) The importance of protein phosphorylation for signaling and metabolism in response to diel light cycling and nutrient availability in a marine diatom. *Biology* 9:155. <https://doi.org/10.3390/biology9070155>
- Tanaka T, Maeda Y, Veluchamy A, Tanaka M, Abida H, Maréchal E, Bowler C, Muto M, Sunaga Y, Tanaka M, Yoshino T, Taniguchi T, Fukuda Y, Nemoto M, Matsumoto M, Wong PS, Aburatani S, Fujibuchi W (2015) Oil accumulation by the oleaginous diatom *Fistulifera solaris* as revealed by the genome and transcriptome. *Plant Cell* 27:162–176. <https://doi.org/10.1105/tpc.114.135194>
- Tanaka A, Ohno N, Nakajima K, Matsuda Y (2016) Light and CO<sub>2</sub>/cAMP signal cross talk on the promoter elements of chloroplastic β-carbonic anhydrase genes in the marine diatom *Phaeodactylum tricorutum*. *Plant Physiol* 170:1105–1116. <https://doi.org/10.1104/pp.15.01738>
- Tateyama S, Kobayashi I, Hisatomi O (2018) Target sequence recognition by a light-activatable basic leucine zipper factor, photozipper. *Biochemistry* 57:6615–6623. <https://doi.org/10.1021/acs.biochem.8b00995>
- Tian H, Trozzi F, Zoltowski BD, Tao P (2020) Deciphering the allosteric process of the *Phaeodactylum tricorutum* aureochrome 1a LOV domain. *J Phys Chem B* 124:8960–8972. <https://doi.org/10.1021/acs.jpcc.0c05842>
- Toyooka T, Hisatomi O, Takahashi F, Kataoka H, Terazima M (2011) Photoreactions of aureochrome-1. *Biophys J* 100:2801–2809. <https://doi.org/10.1016/j.bpj.2011.02.043>
- Traller JC et al (2016) Genome and methylome of the oleaginous diatom *Cyclotella cryptica* reveal genetic flexibility toward a high lipid phenotype. *Biotechnol Biofuels* 9:258. <https://doi.org/10.1186/s13068-016-0670-3>
- Valle KC, Nymark M, Aamot I, Hancke K, Winge P, Andresen K, Johnsen G, Brembu T, Bones AM (2014) System responses to equal doses of photosynthetically usable radiation of blue, green, and red light in the marine diatom *Phaeodactylum tricorutum*. *PLoS One* 9:e114211. <https://doi.org/10.1371/journal.pone.0114211>
- Vardi A, Formiggini F, Casotti R, De Martino A, Ribalet F, Miralto A, Bowler C (2006) A stress surveillance system based on calcium and nitric oxide in marine diatoms. *PLoS Biol* 4:e60. <https://doi.org/10.1371/journal.pbio.0040060>
- Vaulot D, Olson RJ, Chisholm SW (1986) Light and dark control of the cell cycle in two marine phytoplankton species. *Exp Cell Res* 167:38–52. [https://doi.org/10.1016/0014-4827\(86\)90202-8](https://doi.org/10.1016/0014-4827(86)90202-8)
- Veluchamy A et al (2015) An integrative analysis of post-translational histone modifications in the marine diatom *Phaeodactylum tricorutum*. *Genome Biol* 16:102. <https://doi.org/10.1186/s13059-015-0671-8>
- Villareal TA, Pilskaln C, Brzezinski M, Lipschultz F, Dennett M, Gardner GB (1999) Upward transport of oceanic nitrate by migrating diatom mats. *Nature* 397:423–425. <https://doi.org/10.1038/17103>
- Volpert A, Graff van Creveld S, Rosenwasser S, Vardi A (2018) Diurnal fluctuations in chloroplast GSH redox state regulate susceptibility to oxidative stress and cell fate in a bloom-forming diatom. *J Phycol* 54:1–13. <https://doi.org/10.1111/jpy.12638>
- Wenderoth K (1979) Photophobische reaktionen von diatomeen im monochromatischen licht. *Berichte Dtsch Bot Ges* 92:313–321
- Wenderoth K, Rhiel E (2004) Influence of light quality and gassing on the vertical migration of diatoms inhabiting the Wadden Sea. *Helgol Mar Res* 58:211–215. <https://doi.org/10.1007/s10152-004-0187-1>
- Yu Z, Armant O, Fischer R (2016) Fungi use the SakA (HogA) pathway for phytochrome-dependent light signalling. *Nat Microbiol* 1:1–7. <https://doi.org/10.1038/nmicrobiol.2016.19>



## **[Perception de la lumière dans l'océan : étude des phytochromes des diatomées]**

Résumé :

La lumière est une source essentielle d'énergie et d'information pour les organismes photosynthétiques. Dans l'environnement marin, les lumières rouge et rouge lointain sont rapidement atténuées dans la colonne d'eau par rapport au bleu et au vert. Les algues marines telles que les diatomées possèdent un large éventail de photorécepteurs de lumière bleue et verte, mais aussi des phytochromes (DPH) qui perçoivent la lumière rouge (R) et rouge lointain (RL), capable de réguler l'expression des gènes en réponse à la lumière RL chez la diatomée modèle *Phaeodactylum tricornerutum* (Pt). Cependant, la fonction biologique de ce photorécepteur est encore inconnue. En mettant en place un système rapporteur permettant de suivre l'activité de PtDPH *in vivo*, j'ai pu caractériser ses propriétés photochimiques, modéliser son activité dans différents champs lumineux marins, et montrer que les DPH peuvent en détecter les variations liées à la profondeur et la concentration de phytoplancton. Via des approches bioinformatiques, j'ai recherché les DPH dans les génomes et transcriptomes disponibles de diverses diatomées et analysé leur distribution dans l'environnement en utilisant les données méta-omiques générées au cours de l'expédition *Tara Oceans*. Cela a révélé que les diatomées planctoniques du groupe des centriques possédant des DPH sont présentes dans les zones polaires et tempérées, tandis que les diatomées pennées, vivant dans les sédiments peuvent présenter une duplication du gène DPH. Nous avons montré que ces gènes dupliqués ont des propriétés spectrales différentes, et que certaines diatomées benthiques montrent une adaptation spécifique à la lumière R qui pourrait être régulée par DPH. Enfin, j'ai commencé à aborder la voie de signalisation de PtDPH en examinant la synthèse de son chromophore, sa localisation et ses partenaires d'interaction. Ce travail apporte de nouvelles connaissances sur les mécanismes de perception de la lumière chez les diatomées, et leurs importances pour coloniser différentes niches environnementales.

Mots clés : [Diatomées, phytochrome, photorécepteur, lumière, Ocean]

## **[Light sensing in the Ocean : studying diatom phytochrome photoreceptors]**

Abstract :

Light is an essential source of energy and information for photosynthetic organisms. In the marine environment, red and far-red lights are quickly attenuated in the water column compared to blue and green light. Accordingly, predominant marine algae such as diatoms possess a wide array of blue and green light photoreceptors, but also red (R)/far-red (FR) light sensing phytochrome photoreceptors (DPH), capable of regulating gene expression in response to FR light in the model diatom *Phaeodactylum tricornerutum* (Pt). However, the biological function of this photoreceptor is still unknown. By setting up a reporter system to monitor PtDPH activity *in vivo*, I was able to characterize its photochemical properties, model its activity in different marine light fields, and show that DPHs can detect variations related to depth and phytoplankton concentration. Using bioinformatics approaches, I looked for DPH in the available genomes and transcriptomes of diverse diatoms and analyzed their distribution in the environment using the meta-omics data from the *Tara Oceans* expedition. This revealed that planktonic diatoms of the centric group possessing DPH are present in temperate and polar regions. In these species DPH may work as sensor of depth and phytoplankton concentration. Pennate diatoms living in sediments can present duplications of the *DPH* gene. We showed that these duplicated genes can have different spectral properties, and that some benthic diatoms show a specific adaptation to R light that could be regulated by DPH. Finally, I started to address the PtDPH signaling pathway, by investigating its chromophore synthesis, subcellular localization and interacting partners. This work brings new insights into DPH-mediated light perception mechanisms in diatoms, and their significance for colonizing various environmental niches.

Keywords : [Diatom, phytochrome, photoperception, marine light, sediment]