RESEARCH PAPER

Multisignal control of expression of the LHCX protein family in the marine diatom *Phaeodactylum tricornutum*

Lucilla Taddei1,*, Giulio Rocco Stella1,2,*, Alessandra Rogato1,3,4,*, Benjamin Bailleul5, Antonio Emidio Fortunato1, Rossella Annunziata1, Remo Sanges4, Michael Thaler1, Bernard Lepetti6, Johann Lavaud7, Marianne Jaubert1, Giovanni Finazzi8, Jean-Pierre Bouly1 and Angela Falciatore1,†

1 Sorbonne Universités, UPMC, Institut de Biologie Paris-Seine, CNRS, Laboratoire de Biologie Computationnelle et Quantitative, 15 rue de l’Ecole de Médecine, 75006 Paris, France
2 Department of Biotechnology, University of Verona, Strada Le Grazie, I-37134 Verona, Italy
3 Institute of Biosciences and BioResources, CNR, Via P. Castellino 111, 80131 Naples, Italy
4 Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy
5 Institut de Biologie Physico-Chimique, UMR 7141 CNRS-UPMC, 13 rue Pierre et Marie Curie, 75005 Paris, France
6 Zukunftskolleg, Department of Plant Ecophysiology, University of Konstanz, D-78457 Konstanz, Germany
7 UMI 3376 TAKUVIK, CNRS/Université Laval, Département de Biologie, Pavillon Alexandre-Vachon, 1045 avenue de la Médecine, Québec (Québec) G1V 0A6, Canada
8 Laboratoire de Physiologie Cellulaire et Végétale, UMR 5168, Centre National de la Recherche Scientifique (CNRS), Institut National Recherche Agronomique (INRA), Université Grenoble Alpes, Commissariat à l’Energie Atomique et aux Énergies Alternatives (CEA), Institut de Biosciences et Biotechnologies de Grenoble, (BIG), CEA Grenoble, F-38054 Grenoble cedex 9, France

* These authors contributed equally to this work.
† Correspondence: angela.falciatore@upmc.fr

Received 26 February 2016; Accepted 26 April 2016

Editor: Markus Teige, University of Vienna

Abstract

Diatoms are phytoplanktonic organisms that grow successfully in the ocean where light conditions are highly variable. Studies of the molecular mechanisms of light acclimation in the marine diatom *Phaeodactylum tricornutum* show that carotenoid de-epoxidation enzymes and LHCX1, a member of the light-harvesting protein family, both contribute to dissipate excess light energy through non-photochemical quenching (NPQ). In this study, we investigate the role of the other members of the LHCX family in diatom stress responses. Our analysis of available genomic data shows that the presence of multiple LHCX genes is a conserved feature of diatom species living in different ecological niches. Moreover, an analysis of the levels of four *P. tricornutum* LHCX transcripts in relation to protein expression and photosynthetic activity indicates that LHCXs are differentially regulated under different light intensities and nutrient starvation, mostly modulating NPQ capacity. We conclude that multiple abiotic stress signals converge to regulate the LHCX content of cells, providing a way to fine-tune light harvesting and photoprotection. Moreover, our data indicate that the expansion of the LHCX gene family reflects functional diversification of its members which could benefit cells responding to highly variable ocean environments.

Key words: Dark, gene expression, iron starvation, LHCX, light, marine diatom, nitrogen starvation, non-photochemical quenching.
Introduction

The perception of environmental signals and the activation of appropriate responses to external stimuli are of major importance in the growth and survival of all organisms. At the cellular level, this requires the presence of complex signal perception and transduction networks, triggering changes in nuclear gene expression (Lee and Yaffe, 2014). External cues such as light, temperature, and nutrient availability strongly affect the physiology and metabolism of photosynthetic organisms, so acclimation mechanisms are needed to cope efficiently with short- and long-term environmental changes to maintain photosynthetic performances (Walters, 2005; Eberhard et al., 2008). In eukaryotic phototrophs, chloroplast biogenesis and activity are integrated in broader regulatory programmes, requiring coordination between the nucleus and chloroplast genomic systems (Rochaix, 2011; Jarvis and Lopez-Juez, 2013). The nucleus responds to stimuli inducing the synthesis of regulatory proteins that modulate chloroplast responses. In turn, molecules originating from the chloroplast activity (e.g. redox state of the photosynthetic electron carriers, reactive oxygen species, plastid gene transcription, tetrapyrroles, and other metabolites) provide a retrograde signal feeding back to the nucleus (Woodson and Chory, 2008).

Marine photosynthesis is dominated by unicellular phytoplanktonic organisms, which are passive drifters in the water column and often experience drastic changes in their surrounding environment (Falkowski et al., 2004; Depauw et al., 2012). Diatoms are among the most abundant and diversified groups of photosynthetic organisms. They are particularly adapted to growing in very dynamic environments such as turbulent coastal waters and upwelling areas, as well as in polar oceans (Margalef, 1978; Field et al., 1998; Kooistra et al., 2007; Arrigo et al., 2012). Several species can survive for long periods at depths where light is limiting for growth, and quickly reactivate their metabolism after returning to the photic zone (Sicko-Goad et al., 1989; Reeves et al., 2011). The adaptive capacity of such algae suggests that they have sophisticated mechanisms to perceive and rapidly respond to environmental variations. Consistent with this notion, genome sequence information of representative diatom model species such as Thalassiosira pseudonana and Phaeodactylum tricornutum (Armbrust et al., 2004; Montsant et al., 2007; Bowler et al., 2008; Rayko et al., 2010), and the availability of transcriptomic and proteomic data in various species exposed to different stimuli and stresses (Nymark et al., 2009; Dyhrman et al., 2012; Thamatrakoln et al., 2012; Ashworth et al., 2013; Nymark et al., 2013; Keeling et al., 2014; Valle et al., 2014; Alipanah et al., 2015; Muhseen et al., 2015) have highlighted the existence of some diatom-specific adaptive strategies, pinpointing molecular regulators of environmental change responses. Several photoreceptors for efficient light colour sensing (Huysman et al., 2013; Schellenberger Costa et al., 2013; Fortunato et al., 2015, 2016) have been identified in diatoms. Peculiar iron acquisition and concentration mechanisms are also known (Allen et al., 2008; Marchetti et al., 2012; Morrissey et al., 2015), which contribute to their survival in iron-limited waters and to their rapid proliferation when iron becomes available (de Baar et al., 2005). Diatoms have peculiar gene sets implicated in nitrogen metabolism, such as a complete urea cycle, that could be used as temporary energy storage or as a sink for photorespiration (Allen et al., 2011). Eventually, diatoms optimize their photosynthesis via extensive energetic exchanges between plastids and mitochondria (Bailleul et al., 2015).

The ecological dominance of diatoms also relies on their capacity to cope with light stresses, thanks to very efficient photoprotective mechanisms. Diatoms possess a high capacity to dissipate excess light energy as heat through high energy quenching (qE) that, together with the photoinhibitory quenching (qI), can be visualized via the non-photochemical quenching (NPQ) of Chl a fluorescence (Lavaud and Goss, 2014; Goss and Lepetit, 2015). The xanthophyll diatoxanthin (Dt) pigment, synthesized from the de-epoxidation of diadinoxanthin (Dd) during illumination (Goss and Jakob, 2010; Lavaud et al., 2012), and the LH CX1 protein, a member of the light-harvesting protein family (Bailleul et al., 2010), have been identified as key components of the qE process in diatoms. P. tricornutum cells with deregulated LH CX1 expression display a significantly reduced NPQ capacity and a decreased fitness, demonstrating a key role for this protein in light acclimation (Bailleul et al., 2010), similarly to the light harvesting complex stress-related (LHCSR) proteins of green algae and mosses (Alboresi et al., 2010; Ballottari et al., 2016).

Multiple nuclear-encoded and plastid-localized LH CX family members have been identified in the genomes of the diatoms P. tricornutum and T. pseudonana. Scattered information derived from independent gene expression analyses indicated that some LH CX isoforms are constitutively expressed while others are expressed in response to stress (Becker and Rhiel, 2006; Allen et al., 2008; Nymark et al., 2009; Zhu and Green, 2010; Bailleul et al., 2010; Beer et al., 2011; Lepetit et al., 2013), similarly to what is observed for the two LHCSR proteins in Physcomitrella patens (Gerotto et al., 2011). In this study, we have extended the characterization of the four P. tricornutum LH CXs, by combining detailed gene expression analysis in cells exposed to different conditions with in vivo analysis of photosynthetic parameters. The result of this analysis revealed a complex regulatory landscape, suggesting that the expansion of the LH CXs reflects a functional diversification of these proteins and may contribute to the regulation of the chloroplast physiology in response to diverse extracellular and intracellular signals.

Materials and methods

Analysis of the LH CXs in the diatom genomes

Phaeodactylum tricornutum and T. pseudonana LH CX gene model identifiers were retrieved from the diatom genomes, respectively, on P. tricornutum Phatr2 and T. pseudonana Thaps3 in the JGI database (http://genome.jgi.doe.gov/). Phaeodactylum tricornutum LH CX proteins were used as query to perform BlastP searches on the Pseudo-nitzschia multiseries (http://genome.jgi.doe.gov/Psemi1/Psemi1.home.html) and Thalassiosira oceanica (http://protists.
Diatom growth conditions

The *P. tricornutum* (Pt1 8.6, CCMP2561) cultures, obtained from the Provasoli-Guillard National Center for Culture of Marine Phytoplankton, were used for the gene expression and photophysiology analyses. Cells were grown in ventilated flasks in f/2 medium (Guillard, 1975) at 18 °C, in a 12 h light/12 h dark photoperiod using white fluorescence neon lamps (Philips TL-D 90), at 30 μmol m⁻² s⁻¹ (low light). High light treatments were performed by irradiating the cells with 500 μmol m⁻² s⁻¹ for 5 h, 2 h after the onset of light, with the same light sources. Dark adaptation treatments were performed for 60 h. Blue light (450 nm, 1 μmol m⁻² s⁻¹) was applied for 10 min, 30 min, and 1 h on dark-adapted cells in the absence and presence of 2 μM DCMU [3-(3,4-dichlorophenyl)-1,1-dimethylurea]. In the iron starvation experiments, *P. tricornutum* cells at an initial concentration of 2 × 10⁶ cells ml⁻¹ were grown in f/2 artificial sea water medium (Allen et al., 2008) modified to contain either 11 μM iron (iron-replete) or 5 nM iron with the addition of 100 μM of the Fe²⁺ chelator FerroZine™ (iron-limited) (Stokey, 1970). Cells were harvested after 3 d to perform the analyses. Nitrogen starvation was achieved by diluting *P. tricornutum* cells to 2 × 10⁵ cells ml⁻¹ in f/2 medium containing 1 mM nitrate (NO₃-replete) or 50 μM nitrate (NO₃-limited). When cells attained a concentration of 1 × 10⁶ cells ml⁻¹, they were re-diluted to 2 × 10⁵ cells ml⁻¹ in their respective media and harvested after 3d, 2 h after the onset of light, and then used for experiments.

**Generation of transgenic lines overexpressing the LHXC proteins**

Vectors for LHXC overexpression were generated by cloning the full-length cDNA sequences of the four LHXC genes in the pKSFcpBpA1-C-3HA vector (Siaut et al., 2007), using the EcoRI and NolI restriction sites. The LHXC cDNAs were amplified by PCR using the primers described in Supplementary Table S1 at JXB online. Each vector was co-transformed with the pFCPF-p-Shble vector for antibiotic selection into *P. tricornutum* Pt4 cells (DQ085804; De Martino et al., 2007) by microparticle bombardment (Falcicatore et al., 1999). Transgenic lines were selected on 100 μg ml⁻¹ phleomycin (Invitrogen) and screened by PCR using primers specific for the four LHXC isoforms (Supplementary Table S1). Transgenic lines overexpressing the LHXC4 isoform in the Pt1 ecotype were also generated, as for the Pt4 ecotype.

**RNA extraction and qRT-PCR analysis**

Total RNA was isolated from 10⁸ cells with TriPure isolation reagent (Roche Applied Science, IN, USA) according to the manufacturer’s instructions. Quantitative real-time PCR (qRT-PCR) was performed on wild-type cells and on the LHXC-overexpressing clones as described in De Riso et al. (2009). The relative quantification of the different LHXC transcripts was obtained using RPS5 (ribosomal protein small subunit 3OS; ID10847) and H4 (histone H4; ID34971) as reference genes, and by averaging of two reference genes using the geometric mean and the fold changes calculated with the 2⁻ΔΔCt Livak method (Livak and Schmittgen, 2001). Primer sequences used in qRT-PCR analysis are reported in Supplementary Table S1.

**Protein extraction and western blot analysis**

Western blot analyses were performed on total cell protein extracts prepared as in Bailleul et al. (2010), and resolved on 14% LDS-PAGE gels. Proteins were detected with different antibodies: anti-LHCSR (gift of G. Peers, University of California, Berkeley, CA, USA) (1:5000); anti-D2 (gift of J-D. Rochaix, University of Geneva, Switzerland) (1:10 000); anti-PsaF (1:1000) and anti-βCF1 (1:10 000) (gift of F.-A. Wollman, Institut de Biologie Physico-Chimique, Paris, France); and anti-HA primary antibody (Roche) (1:2000). Proteins were revealed with Clarity reagents (Bio-Rad) and an Image Quant LAS4000 camera (GE Healthcare, USA).
of all the available diatom species to 17 members have been found in the genome of the polar species, compared with five in \textit{P. multiseries}, four in \textit{P. tricornutum}, and up to 17 members have been found in the genome of the polar species \textit{Fragilariopsis cylindrus} (B. Green and T. Mock, personal communication). Analysis of the intron–exon structure of the \textit{LHCX} genes revealed a variable number of introns (from zero to three) as well as variable intron and exon lengths (Table 1).

### Light versus dark regulation of expression of the LHCX genes

Independent gene expression studies performed in \textit{P. tricornutum} cells suggest that the \textit{LHCX} gene family is regulated by light via multiple regulatory pathways. To explore the mechanisms controlling the light responses of the \textit{LHCX} genes further, we analysed mRNA and protein contents in cells exposed to different light conditions. We first monitored the expression of the \textit{LHCX} genes in cells grown in low light (LL) and then exposed to high light (HL). In line with previous studies (Bailleul et al., 2010; Lepetit et al., 2013), qRT-PCR and western blot analyses (Fig. 1A and B, respectively) showed that \textit{LHCX1} is expressed at very high levels in LL-adapted cells, and that HL treatment slightly increases the \textit{LHCX1} content. Conversely, the isoforms 2 and 3 showed different responses to the LL to HL shift.

The \textit{LHCX2} transcripts, which are significantly less abundant than that of \textit{LHCX1} in LL, rapidly increased following HL stress, reaching levels comparable with those of \textit{LHCX1} after 1 h HL exposure (Fig. 1A). This translated into an increase of the \textit{LHCX2} protein observed by western blot (Fig. 1B). However, the increase in the protein content was lower than that of the transcript, possibly because of a low affinity of the LHCSR antibody (Peers et al., 2009) for the \textit{LHCX2} isoform. The \textit{LHCX3} transcripts that were expressed at very low levels in LL quickly rose upon HL treatment, peaking after 30 min and starting to decrease after 1 h of light stress. Conversely, a different mRNA expression profile was found in the case of \textit{LHCX4}, which, unlike the other isoforms, was barely detectable in both LL and HL conditions (Fig. 1A). The \textit{LHCX3} and \textit{LHCX4} proteins, having very similar molecular weights (22.8 kDa and 22.2 kDa, respectively), cannot be discriminated by western blot analysis. Based on the different transcriptional regulation of \textit{LHCX3} and \textit{LHCX4} by light, it is tempting to propose that the light-induced protein of ~22.8 kDa reflects the accumulation of the \textit{LHCX3} isoform (Fig. 1B). However, in contrast to the transient induction of the \textit{LHCX3} mRNAs, this protein is gradually accumulated during the LL to HL shift and it remains stable over the treatment. This discrepancy between transcript and protein expression profiles could be explained assuming that: (i) some post-transcriptional modifications regulate the accumulation of \textit{LHCX3} in the light; or (ii) the light-induced protein isoform at 22.8 kDa also comprises the \textit{LHCX4} protein, which could be present in HL-exposed cells, along with \textit{LHCX3}.

### Table 1. List of the LHCXs identified in the diatom genomes

<table>
<thead>
<tr>
<th>Species Name</th>
<th>Name</th>
<th>ID</th>
<th>Chromosomal localization</th>
<th>Length (no. of amino acids)</th>
<th>No. of introns</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Thalassiosira pseudonana}</td>
<td>\textit{LHCX1}</td>
<td>264921</td>
<td>chr\textsubscript{23}:365603–366232 (-)</td>
<td>209</td>
<td>0</td>
</tr>
<tr>
<td>\textit{LHCX2}</td>
<td>38879</td>
<td>chr\textsubscript{23}:368273–368902 (+)</td>
<td>209</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX4}</td>
<td>270228</td>
<td>chr\textsubscript{5}:1446306–1447125 (-)</td>
<td>231</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX5}</td>
<td>31128</td>
<td>chr\textsubscript{1}:2849139–2850176 (-)</td>
<td>236</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX6}</td>
<td>12097</td>
<td>chr\textsubscript{23}:366611–367378 (+)</td>
<td>255</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>\textit{Phaeodactylum tricornutum}</td>
<td>\textit{LHCX1}</td>
<td>27278</td>
<td>chr\textsubscript{7}:996379–997300 (+)</td>
<td>206</td>
<td>1</td>
</tr>
<tr>
<td>\textit{LHCX2}</td>
<td>56312</td>
<td>chr\textsubscript{1}:2471232–2472170 (+)</td>
<td>238</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX3}</td>
<td>44733</td>
<td>chr\textsubscript{5}:76676–77606 (+)</td>
<td>206</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX4}</td>
<td>12097</td>
<td>chr\textsubscript{23}:366611–367378 (+)</td>
<td>255</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>\textit{Pseudo-nitzschia multiseries}</td>
<td>–</td>
<td>–</td>
<td>scaffold\textsubscript{189}:181982–182948 (+)</td>
<td>201</td>
<td>1</td>
</tr>
<tr>
<td>\textit{LHCX1}</td>
<td>238335</td>
<td>scaffold\textsubscript{95}:121459–122306 (-)</td>
<td>202</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX2}</td>
<td>257821</td>
<td>scaffold\textsubscript{246}:124909–125745 (+)</td>
<td>197</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX3}</td>
<td>264022</td>
<td>scaffold\textsubscript{1353}:8720–9877 (+)</td>
<td>206</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>\textit{LHCX4}</td>
<td>283956</td>
<td>scaffold\textsubscript{38}:284133–284828 (-)</td>
<td>231</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>\textit{Thalassiosira oceanica}</td>
<td>–</td>
<td>–</td>
<td>SuperContig To\textsubscript{1}g10869: 4.331–5.040 (-)</td>
<td>210</td>
<td>1</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>SuperContig To\textsubscript{1}g15184: 10.800–11.435 (-)</td>
<td>172</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>SuperContig To\textsubscript{1}g45669: 1–1.025 (-)</td>
<td>205</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>SuperContig To\textsubscript{1}g46152: 5.664–6.285 (-)</td>
<td>180</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

For \textit{T. pseudonana} (Thaps3), \textit{P. tricornutum} (Phatr2), and \textit{P. multiseries} (Psemi1), IDs refer to the genome annotation in the JGI database (http://genome.jgi.doe.gov/). For \textit{T. oceanica} (ThaOc_1.0), ID refers to the Ensembl Protist database (http://protists.ensembl.org/Thalassiosira_oceanica/Info/Index?db=core). (+) and (–) indicate the forward and reverse chromosomal or scaffolds, respectively. The protein length and intron numbers are also indicated.
Previous reports (Lepetit et al., 2013; Nymark et al., 2013) indicate that the LHCX4 transcript is induced in dark-adapted cells. Therefore, we extended the analysis of the expression of the four LHCX genes to cells adapted to prolonged darkness (60 h). In these conditions, we observed a significant increase only of the LHCX4 mRNAs (Fig. 1C). For the same reason as described above, we attributed the band of ~22 kDa observed in the dark-adapted cells to the LHCX4 protein, although LHCX3 could also be present. In the dark, cells were also showing a decreased NPQ capacity (Fig. 1E) and a slightly reduced PSII maximal quantum yield and overall photosynthetic electron flow capacity (Table 2). Other studies have revealed that blue light photoreceptors (Coesel et al., 2009; Juhás et al., 2014) and the redox state of the chloroplast (Lepetit et al., 2013) could both contribute to the light regulation of LHCX1, 2, and 3 gene expression. Thus, we tested the possible role of these processes in the inhibition of LHCX4 expression upon light exposure. We irradiated dark-adapted cells with low intensity blue light (1 µmol m⁻² s⁻¹) during 1 h, in the presence or absence of the PSII inhibitor DCMU (Fig. 1F). The analysis revealed that the LHCX4 expression is repressed even at such low light irradiance. Moreover, this repression is lost by poisoning photosynthesis with DCMU, suggesting that this process plays an active role in the light-induced repression of LHCX4.

A recent study in Chlamydomonas reinhardtii showed that the activity of the protein LHCSR3 is regulated by the reversible protonation of three specific amino acidic residues following luminal pH acidification in the light (Ballottari et al., 2016). In order to assess if this mechanism is conserved in diatoms, we analysed the P. tricornutum LHCX protein sequences. We found that LHCX1, 2, and 3 possess two of the three protonatable residues conserved in diatoms (Fig. 1G).
LHCX expression in iron starvation

Besides light, nutrient availability also affects chloroplast activity (Wilhelm et al., 2006; Gross, 2012). In many oceanic regions, iron is a major limiting factor for diatom distribution. A general down-regulation of photosynthesis has been reported in iron starvation in several diatom species (Laroche et al., 1995; Allen et al., 2008; Hohner et al., 2013), with a consequent decrease of the carbon fixation reactions, growth rate, and cell size. Since increased NPQ was previously observed in iron-starved *P. tricornutum* cells (Allen et al., 2008), we compared the expression of the different LHCX isoforms in cells grown under Fe-replete and Fe-limited conditions. We found that while a slight induction of the other LHCX proteins was seen, the LHCX2 transcript was greatly induced in iron-limited cells (Fig. 2A), leading to a significant accumulation of the LHCX2 protein (Fig. 2B). Fe limitation also enhanced NPQ, while slowing down its kinetics (Fig. 2C), possibly because of a slower diadinoxanthin de-epoxidation rate. We also observed a severe impairment of the photosynthetic capacity in iron limitation as indicated by the decrease in *F*0/*F*m (Table 2) and in the PSII maximal electron transport rate (rETR<sub>PSII</sub>) (Fig. 2D; see also Allen et al., 2008). The decreased maximal rETR<sub>PSII</sub> was probably caused by a diminished capacity for carbon fixation. Moreover, in agreement with previous studies (Allen et al., 2008; Thamatrakoln et al., 2013), we observed a decrease in the amount of PSI (PsaF), which is the complex with the highest Fe content. This complex has been already shown to represent the first target of Fe limitation (Moseley et al., 2002). We also found a significant decrease in PSII (D2 protein), which was probably degraded because of sustained photo-inhibition (see also Allen et al., 2008) (Fig. 2B).

### Table 2. Photosynthetic parameters of the *P. tricornutum* wild type and transgenic lines

<table>
<thead>
<tr>
<th>Strain</th>
<th>Conditions</th>
<th><em>F</em>0/<em>F</em>m</th>
<th>rETR&lt;sub&gt;PSII&lt;/sub&gt;</th>
<th>NPQ max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pt1 LL</td>
<td>0.66 ± 0.03</td>
<td>74.4 ± 1.2</td>
<td>2.1 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Pt1 Dark</td>
<td>0.60 ± 0.02</td>
<td>67.2 ± 4.0</td>
<td>1.0 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Pt1 +Fe</td>
<td>0.65 ± 0.003</td>
<td>72.6 ± 4.8</td>
<td>2.2 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>Pt1 –Fe</td>
<td>0.20 ± 0.004</td>
<td>25.7 ± 1.8</td>
<td>4.4 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>Pt1 +N</td>
<td>0.65 ± 0.006</td>
<td>74.4 ± 2.4</td>
<td>2.1 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Pt1 –N</td>
<td>0.40 ± 0.008</td>
<td>27.5 ± 3.2</td>
<td>3.2 ± 0.4</td>
<td></td>
</tr>
<tr>
<td>Pt4 WT</td>
<td>0.68 ± 0.01</td>
<td>79.3 ± 2.6</td>
<td>0.83 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>Pt4 EVL</td>
<td>0.66 ± 0.01</td>
<td>72.5 ± 3.5</td>
<td>0.82 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE1</td>
<td>0.67 ± 0.01</td>
<td>70.7 ± 2.6</td>
<td>1.00 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE2.5</td>
<td>0.67 ± 0.01</td>
<td>70.0 ± 1.6</td>
<td>1.06 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE2.20</td>
<td>0.66 ± 0.01</td>
<td>69.3 ± 1.6</td>
<td>1.02 ± 0.08</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE12</td>
<td>0.66 ± 0.01</td>
<td>75.8 ± 4.9</td>
<td>1.04 ± 0.07</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE33</td>
<td>0.68 ± 0.03</td>
<td>71.0 ± 3.1</td>
<td>1.00 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE4.11</td>
<td>0.59 ± 0.01</td>
<td>68.7 ± 1.9</td>
<td>1.03 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Pt4 OE4.13</td>
<td>0.58 ± 0.02</td>
<td>69.4 ± 1.3</td>
<td>1.07 ± 0.04</td>
<td></td>
</tr>
</tbody>
</table>

PSII efficiency (*F*0/*F*m) and relative electron transport rate (rETR<sub>PSII</sub>) in different growth conditions are reported. rETR<sub>PSII</sub> was measured at 260 µmol photons m<sup>−2</sup> s<sup>−1</sup> light intensity and calculated as: rETR<sub>PSII</sub>=PSII actinic light intensity. Non-photochemical quenching (NPQ) was measured with an actinic light intensity of 950 µmol photons m<sup>−2</sup> s<sup>−1</sup> and calculated as in Maxwell and Johnson (2000). Data are the average of three biological replicates ±SD.

### Fig. 2. Effect of iron starvation on *P. tricornutum* LHCX expression and photophysiology. Experiments were performed on cells grown in iron-replete (11 µM, +Fe) or iron-limited (6 nM iron+100 µM Ferrozine™, –Fe) conditions: (A) qRT-PCR analysis of LHCX transcripts in –Fe, normalized against the +Fe condition and using *RPS* and *H4* as reference genes. (B) Immunoblot analysis of the LHCX2, D2, and PsaF proteins, using βCF1 as loading control. NPQ capacity (C) and relative electron transfer rates (rETR<sub>PSII</sub>) (D) of cells grown in +Fe or –Fe. The horizontal bar in (C) indicates when the actinic light was on (white) or off (black). rETR<sub>PSII</sub> was measured at different light intensities (20, 170, 260, 320, 520, and 950 µmol m<sup>−2</sup> s<sup>−1</sup>). In (A), (C), and (D), error bars represent ±SD of three biological replicates. (This figure is available in colour at JXB online).
Fig. 3. Effect of nitrogen starvation on P. tricornutum LHCX expression and photophysiology. Experiments were performed on cells grown in nitrogen-replete (1 mM, +NO\textsubscript{3}) or nitrogen starvation (50 µM, −NO\textsubscript{3}) conditions: (A) qRT-PCR analysis of LHCX transcripts in −NO\textsubscript{3}, normalized against the values in the +NO\textsubscript{3} condition, and using RPS and H4 as reference genes. (B) Immunoblot analysis of the LHCX, D2, and PsAF proteins, using βCF1 as loading control. NQ capacity (C) and relative electron transfer rates (rETR\textsubscript{PSII}) (D) of cells grown in +NO\textsubscript{3}– and −NO\textsubscript{3}– conditions. The horizontal bar in (C) indicates when the actinic light was on (white) or off (black). rETR\textsubscript{PSII} was measured at different light intensities (20, 170, 260, 320, 520, and 950 µmol m\textsuperscript{-2} s\textsuperscript{-1}). In (A), (C), and (D), error bars represent ±SD of three biological replicates. (This figure is available in colour at JXB online).

Table 3. The identified regulatory motifs and their occurrence in the P. tricornutum LHCX non-coding sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>LHCX1</th>
<th>LHCX2</th>
<th>LHCX3</th>
<th>LHCX4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOTIF 1</td>
<td>TCA[CT][AT]GTCA</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>MOTIF 2</td>
<td>CGAACCTTGG</td>
<td>–</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>MOTIF 3</td>
<td>CCT[GC]TCCGTA</td>
<td>–</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>MOTIF 4</td>
<td>GAGTCATCG</td>
<td>–</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>MOTIF 5</td>
<td>CGATCACGGC</td>
<td>–</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>MOTIF 6</td>
<td>[TA]TGACTG</td>
<td>–</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CCRE-1</td>
<td>TGACGT</td>
<td>1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CCRE-2</td>
<td>ACGTCA</td>
<td>1</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>CCRE-3</td>
<td>TGACCG</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>

Analysis of the LHCX non-coding regions

Due to the observed transcriptional responses of LHCX genes in different light and nutrient conditions, we searched for known and potentially novel regulatory motifs in the 5′-flanking regions and the intronic sequences of the four isoforms (see Table 3; Supplementary Fig. S1). Because of their involvement in integrating light signals with CO\textsubscript{2}/cAMP-induced transcriptional responses, we searched for the three CO\textsubscript{2}/cAMP-responsive cis-regulatory elements (CCREs) identified in Ohno et al. (2012) and further characterized in Tanaka et al. (2016). Interestingly, we found CCRE-1 in the 5′-flanking sequences of LHCX1, and CCRE-2 in 1, 2, and 4, and CCRE-3 in LHCX4. These cis-regulatory elements may participate in the light-mediated regulation of the four LHCX genes.

In contrast, the two P. tricornutum iron-responsive elements identified in Yoshinaga et al. (2014) are not present in the analysed non-coding regions, suggesting that a different transcription factor should be involved in modulating the LHCX2 transcriptional response to iron availability. Similarly, we could not find the two P. tricornutum motifs identified as responsive to short-term nitrogen deprivation (from 4 h to 20 h) in Matthijs et al. (2016), suggesting that distinct regulatory circuits may act in the short- and long-term acclimation to nitrogen deprivation.

To pinpoint possible novel regulatory motifs, we also scanned the non-coding sequences of the four isoforms using the MEME Suite program (Bailey et al., 2009). The analysis revealed six motifs repeated at least twice in each isoform and/or shared by more than one isoform (Table 3; Supplementary Fig. S1). None of the identified motifs corresponds to a known transcription factor-binding site. This suggests that these motifs could represent novel diatom-specific cis-regulatory elements. In order to examine the potential involvement of the identified motifs in the long-term nitrate deprivation transcriptional responses of LHCX genes, we analysed a published microarray data set performed on 48 h and 72 h nitrogen-deprived P. tricornutum cells (Alipanah et al., 2015). We compared the frequency of the six identified motifs in the 5′-flanking sequences of responsive and unresponsive transcripts. Interestingly, motif 6 ([T-A]TGACTG) was significantly enriched \(p=0.035\) in the 5′-flanking sequences of genes up-regulated in response to nitrogen starvation compared with down-regulated genes. The result suggests that motif 6 may be involved in gene transcriptional regulation in cells exposed to prolonged nitrogen starvation.

Modulation of LHCX gene expression in P. tricornutum transgenic lines

A role in the regulation of the NPQ in P. tricornutum has been proven for the LHCX1 protein by characterizing transgenic...
lines with a modulated content of LHCX1 by either gene silencing or gene overexpression (Bailleul et al., 2010). Unfortunately, all the attempts to down-regulate the expression of LHCX2, 3, or 4 have been unsuccessful. Therefore, to explore their function, we opted for the strategy used in Bailleul et al. (2010), and tried to rescue the intrinsically lower NPQ capacity of the Pt4 ecotype. To this end, independent transgenic Pt4 lines were generated, bearing a vector in which the LHCX2, 3, or 4 genes were expressed under the control of the P. tricornutum FCPB (LHCF2) promoter. A HA-tag was fused to the C-terminal end of the LHCX transgenes to allow the specific detection of the transgenic proteins. qRT-PCR and western blot analyses (Fig. A, C, E) on independent transgenic lines confirmed the expression of the transgenic LHCX isoforms. NPQ analyses (Fig. B, D, F) showed that the overexpression of each LHCX isoform generated a modest, but statistically significant, increase in the NPQ capacity compared with the Pt4 wild type as well as compared with a transgenic line transformed only with the antibiotic resistance gene and used as control. Strikingly, we found that all the transgenic lines showed a similar NPQ increase, regardless of which isoform was overexpressed and the different overexpression levels.

We also checked the possible effect of LHCX overexpression on growth and photosynthetic capacity. For the lines overexpressing the LHCX2 and LHCX3 proteins, we did not observe any altered phenotype (Table 2). In contrast, the Pt4 lines overexpressing LHCX4 showed a reduced PSII efficiency (Table 2). By performing a growth curve analysis, we also observed that these overexpressing lines showed a lag phase lasting 2–3 d (Fig. 4G), which was not the case in wild-type cells. A similar effect on growth was also observed in transgenic lines in which the LHCX4 gene was overexpressed in the Pt1 ecotype (Fig. 4G).

Discussion

The presence of multiple LHCX genes in all the diatom genomes analysed to date strongly suggests that the expansion of this gene family is a common feature of these algae and may represent an adaptive trait to cope with highly variable environmental conditions. To investigate this scenario, in
this work we correlated LHCX expression profiles with the photosynthetic and photoprotective performances in variable experimental conditions, including changes in light irradiance and nutrient availability. These analyses revealed that the four *P. tricornutum* LHCX genes respond differently to various environmental cues, as summarized in Fig. 5.

The analyses of the mRNA and protein responses indicate that amounts of the different LHCXs are tightly regulated at the transcriptional, and probably also the post-translational level. As LHCX3 and LHCX4 have a similar size, it was not possible to quantify the amount of these two proteins under the different stresses using one-dimensional electrophoresis. However, considering the transcript and biochemical analyses together (in the case of LHCX2 and LHCX1), it seems that LHCX1 is always expressed at high levels even in non-stress conditions, which is consistent with it having a pivotal role in NPQ regulation and light acclimation as proposed previously (Bailleul et al., 2010).

LHCX2 and 3 are induced following high light stress, where they may contribute to increase the diatom photoprotection capacity. Their induction, as well as the accumulation of LHCX1, may result from the integration of different signals. Two members of the blue light-sensing cryptochrome photolyase family, CPF1 (Coesel et al., 2009) and CRYP (Juhas et al., 2014), modulate the light-dependent expression of LHCX1, LHCX2, and LHCX3. Also, the recently identified Aureochrome 1a blue light photoreceptor, which regulates *P. tricornutum* photoacclimation (Schellenberger Costa et al., 2013), may affect how much of each LHCX there is in a cell. Moreover, chloroplast activity, through the redox state of the plastoquinone pool, may also regulate *LHCX1* and *LHCX2* gene expression in HL (Lepetit et al., 2013).

A different regulation pattern is seen in the case of LHCX4, the only isoform which is induced in the absence of light. The amount of *LHCX4* mRNA rapidly decreases following a dark to light transition, and this repression is lost when photosynthesis is halted with the PSII inhibitor DCMU. This suggests that chloroplast-derived signals could participate in inhibiting gene expression, even at very low light irradiance, by an as yet unknown process. The peculiar trend observed in the LHCX4 light response suggests a possible role for this protein in *P. tricornutum* photoacclimation. The increased LHCX4 transcript and possibly protein content is mirrored by a decrease in NPQ capacity and a slightly reduced $F_{v}/F_{m}$ in the dark-adapted cells, compared with cells grown in the light (Fig. 1E; Table 2). Moreover, reduced PSII efficiency and slightly altered growth were observed in cells overexpressing LHCX4 in the light (Fig. 4G), suggesting that LHCX4 could have a negative impact on chloroplast physiology. Indeed, a comparative analysis of the *P. tricornutum* LHCX protein sequences indicates that LHCX4 lacks key protonatable residues that in *Chlamydomonas* are involved in NPQ onset when the lumen acidifies (Ballottari et al., 2016). These residues are, however, conserved in the LHCX1, 2, and 3 isoforms. According to the model established in green algae for the protein LHCSR3, these residues diminish their electrostatic repulsion upon protonation, allowing a rearrangement of the protein structure and pigment orientation and enhancement of the quenching capacity (Ballottari et al., 2016). The substitution in LHCX4 of the acidic residues (aspartate and

---

**Fig. 5.** Model of the *P. tricornutum* LHCX regulation. Scheme summarizing the multiple external signals and stresses that differentially regulate the expression of the four LHCXs. The LHCX genes are shown in the nucleus and the LHCX proteins in the chloroplast. + and – boxes indicate positive and negative transcriptional regulation, respectively, in response to white light (yellow), blue light (blue, through the cryptochromes, Cry, and aureochromes, Aureo, photoreceptors), darkness (black), chloroplast signals (green), iron starvation (grey), and nitrogen starvation (orange). In the *P. tricornutum* cell: N, nucleus; C, chloroplast; M, mitochondrion. In the chloroplast: PSI and PSII, photosystem I and II, respectively; PSI* and PSII*, excited photosystems; PQ, plastoquinone pool; $b_{6}/f$, cytochrome $b_{6}/f$ complex; $\Delta$P, proton gradient; NADPH, redox potential.
 glutamate) with non-protonatable residues (asparagine and glycine) would prevent such regulation. Instead, LHCX4 could contribute to the observed capacity of *P. tricornutum* to survive long periods in the dark and its repression could be needed for a rapid acclimation following re-illumination (Nymark et al., 2013). Consistent with this, high LHCX gene expression has also been observed in sea-ice algal communities dominated by diatoms that have adapted to the polar night (Pearson et al., 2015).

Besides the light and redox signals discussed above, our study also shows that differences in the availability of iron and nitrogen strongly affect the expression of the different LHCXs. The signalling cascades controlling these responses are still largely unknown, but they probably involve multiple regulatory pathways into the nucleus and chloroplast, considering that these nutrients are essential for diatom photosynthesis and growth (Table 2; Fig. 5). Nitrogen starvation induces a general increase of all the LHCX isoforms, including LHCX4 that is normally repressed in light-grown cells (Fig. 3). We can hypothesize that the general increase of the LHCX content is needed to protect the photosynthetic apparatus, which is strongly affected by nitrogen deprivation, as shown by the drastically reduced $F_{v}/F_{m}$ (Table 2), the lower maximal rETR$_{PSII}$ (Fig. 3D), and the reduction of PSI and PSII protein content. Interestingly, an opposite trend is observed for the main enzymes of the xanthophyll cycle, which are either not induced or are repressed in cells grown in similar nitrogen stress conditions (Supplementary Fig. S3). Thus, in nitrogen starvation, the LHCXs could represent the major contributors to the observed NPQ increase (Fig. 3C).

At variance with nitrogen starvation, iron starvation has a more specific effect on LHCX expression. A strong induction of the LHCX2 mRNA and protein levels (Fig. 2) compared with the other isoforms was observed, pinpointing this isoform as the most likely regulator of the increased NPQ capacity observed in iron stress (Fig. 2C). NPQ in iron-limiting conditions is characterized by a slow induction and a complete relaxation in the dark. These slow induction kinetics might reflect either lower concentrations of the pH-activated de-epoxidase enzyme or its cofactor ascorbate (Grouneva et al., 2006) or slower acidification of the thylakoid lumen due to a reduced photosynthetic activity. Indeed, the photosynthetic capacity is severely impaired when iron is limiting, as demonstrated by the reduction in PSI and PSII subunits (Fig. 2B), but also the lower $F_{v}/F_{m}$ (Table 2) and rETR$_{PSII}$ (Fig. 2D). The decreased electron flow per PSI could also reflect a decrease in the iron-containing cytochrome $b_{5}f$ complex, as previously shown for other iron-limited diatoms (Strzepek and Harrison, 2004; Thamatrakoln et al., 2013).

The observations made in this and in previous studies about the complex LHCX regulation in response to different signals prompted us to explore their possible functions in *P. tricornutum*, by modulating their expression in a natural P14 strain characterized by constitutive lower NPQ levels (Fig. 4). We observed that the increased expression of all the tested isoforms generates a small but still consistent increase in the NPQ levels, suggesting a potential involvement of the diverse proteins in NPQ modulation, as previously shown for LHCX1 (Bailleul et al., 2010). However, we also noticed that different overexpressing lines with different transcript and protein levels showed a similar NPQ increase. It is difficult to interpret these first results, especially in the case of lines over-expressing LHCX4, whose endogenous expression is inhibited by light (Fig. 1F). They probably reflect the complexity of NPQ regulation in diatoms, where the presence of multiple players (e.g. several LHCXs and enzymes of the xanthophyll cycle) possibly tend to reduce the consequences on NPQ of genetic modifications of the qE machinery.

Finally, the exploration of the 5'-flanking regions and intronic sequences of the LHCX genes revealed the presence of known and potentially novel cis-regulatory elements that may contribute to the transcriptional regulation of the different isoforms in stress conditions. We revealed an uneven distribution of the CCREs (Ohno et al., 2012; Tanaka et al., 2016) in the four LHCX genes that may be linked to their different light-mediated transcriptional responses. In addition, we identified a 7 bp motif in the non-coding sequences of LHCX2, 3, and 4. Using genome-wide transcriptomic data, we found this motif specifically enriched in long-term nitrogen starvation-induced genes, suggesting a possible involvement in the regulation of gene expression in response to nitrogen fluctuations. Although additional studies are required to demonstrate the functionality of these motifs, their discovery may represent a starting point for the identification of the LHCX regulators in the diatom acclimation mechanisms to stress.

**Outlook**

Here we discovered that the four *P. tricornutum* LHCXs are regulated in a sophisticated way (Fig. 5). Different and probably interconnected regulatory pathways activated by different signals and stresses tightly control the amount of each LHCX isoform in the cell. By narrowing down the specific growth conditions in which the different LHCXs are required, our results set the basis for future work to define the function of each isoform in the regulation of chloroplast physiology. The generation of new transgenic lines in which the content of each LHCX isoform is specifically modulated will be instrumental in assessing whether they act with the NPQ regulator LHCX1, or play other specific roles. Considering the robustness of LHCX1 expression in all the conditions tested, future studies will probably require the use of new LHCX1 loss-of-function diatom strains. Additional information about the association of LHCXs with photosynthetic complexes and pigments will also be necessary to understand the role played by the expanded LHCX gene family in the efficient acclimation of diatoms to environmental changes.

**Supplementary data**

Supplementary data are available at *JXB* online.

**Figure S1.** Localization of the enriched motifs in non-coding regions of *P. tricornutum* LHCX genes.

**Figure S2.** Alignment of the LHCX proteins and three-dimensional model of LHCX1.
Figure S3. Expression of the *P. tricornutum* xanthophyll cycle genes in nitrogen starvation.

Table S1. List of the oligonucleotides used in this work.

Acknowledgements

This work was supported by Marie-Curie ITNs CALIPSO (ITN 2013 GA 607607) and AcchiPhot (ITN 2012 GA 316427) grants and a Gordon and Betty Moore Foundation grant (GBMF 4966) to AF, the French ANR ‘DiaDomOil’ PROGRAMME BIO-MATIERES & ENERGIES to AF and GF, and the Marie Curie Zukunftskolleg Incoming Fellowship and a Zukunftskolleg Interim Grant to BL.

References


acclimation of iron deficiency responses in *Chlamydomonas reinhardtii* as revealed by proteomics based hierarchical clustering and reverse genetics. Molecular and Cellular Proteomics 12, 2774–2790.


Environmental modulation of diatom LHCX multigene family

β-carbonic anhydrase genes in the marine diatom *Phaeodactylum tricornutum*. Plant Physiology **170**, 1105–1116.


