Light-dependent signal transduction in the marine diatom *Phaeodactylum tricornutum*

Ananya Agarwal[1,2], Orly Levitan[1,2], Helena Cruz de Carvalho[3,4] & Paul G. Falkowski[1,5]

[1]*Environmental Biophysics and Molecular Ecology Program, Dept. of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901*
[2]*Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ 08901*
[3]*Institut de Biologie de l'ENS (IBENS), Ecole normale supérieure, CNRS, Inserm, Université PSL, 75005 Paris, France*
[4]*Faculté des Sciences et Technologie, Université Paris Est-Créteil (UPEC), 94000 Créteil, France*
[5]*Department of Earth and Planetary Sciences, Rutgers University, Piscataway, NJ 08854*

Corresponding author
Paul G. Falkowski
Email: falko@marine.rutgers.edu
Tel: +1(848)932-6665

Data availability

DESeq's output for all 12,089 genes of the transcriptome is available at the NCBI GEO Accession # GSE133301

Supporting Information:

Additional supporting information may be found in the online version of this article.

Supplemental S1: Figures & Tables

Plasmid Map S1 Plasmid map of the RNAi transformant vector used in this study

Figure S1 Confocal laser microscopy images of low light acclimated Wild Type (WT) *P.tricornutum* and the PTP-33 transformant strain.

Figure S2 Fluorescent microscopic images of transformant line PTP-33 acclimated to low light growth conditions.

Figure S3 Relative expression levels (RT-qPCR) of mRNA Phatr3_J43123 and its cognate Natural Anti-sense Transcript (NAT).

Table S1 Differential expression heat map of Light Harvesting Complex (LHC) genes,

Table S2 Differential expression heat map of Nuclear photosynthetic electron transport components and Photoreceptor genes,

Table S3 Presence of Phatr3_J50052 (LSK) in transcriptome studies of *P.tricornutum* grown under various environmental conditions

Table S4 Primers designed to inserting antisense constructs into pKS plasmid using Gibson Assembly and RT-qPCR

Supplemental S2: Dataset 1

Analyzed data of transcriptome of wild-type *Phaeodactylum tricornutum* cells fully acclimated to low light intensity of 20 μmols photons $m^{-2}s^{-1}$ (LL) as compared to high light intensity of 940 μmols photons $m^{-2}s^{-1}$ (HL).

Supplemental Materials and Methods

Genetic Transformation and Selection of Transformants

Five µg pKS-ShBle-GOI-FA vector was coated onto M17 tungsten particles (1.1 µm) according to the manufacturer's instructions (Bio-Rad). Approximately 5×10^7 WT *P. tricornutum* cells were plated on 1% agar plates (50% F/2) and incubated for one day before the transformation. The cells were bombarded with the DNA-coated M17 particles at 1,550 psi (52) using a PDS-1000/ He Particle Delivery System (Bio-Rad, CA). The plates were incubated at 100 µmol photons $m^{-2}s^{-1}$ constant illumination at 18 °C for 48 hours to recover. Cells were then re-plated onto selective 1% agar plates (50% F/2) with 100 µg/mL Zeocin. Plates were incubated at 40 µmol photons $m^{-2}s^{-1}$ for three-to-four weeks to enable the transformed clones to grow. Single independent transformation events of each plasmid into WT yielded 40-80 colonies each. To screen for putative knockdown strains, each culture was propagated in liquid F/2 supplemented with Zeocin and then split into two cultures – one grown under constant high light (HL) conditions of ~800-950 µmol photons $m^{-2}s^{-1}$ and one under constant low light (LL) conditions of 15-25 µmol photons $m^{-2}s^{-1}$. Of the 638 transformants isolated, the five strains that exhibited the most abnormal light acclimating phenotypes to either/both HL and LL were chosen for additional studies.

Total cellular RNA Extraction, Sequencing and Analysis

Samples for RNA-Seq were harvested and extracted from triplicate sets of cultures acclimated to 20 and 940 µmol photons $m^{-2}s^{-1}$ using TRIzol-Chloroform protocol (53) followed by removal of DNA contamination using TURBO DNA-free kit (AM1907; ThermoFisher Scientific, MA), and cleaning with an RNEasy MinElute Kit (74204; Qiagen, Germany). Integrity of RNA was verified by Polyacrylamide Gel Electrophoresis (54) with reagents treated with DEPC (D5758; Sigma-Aldrich, MO: according to manufacturer's instructions) to eliminate RNAse activity.

TruSeq RNA Library Prep Kit v2 (Illumina, CA) was used to prepare mRNA libraries for each of the six samples according to the manufacturer's instructions. The 250 bp single-indexed libraries were multiplexed and sequenced on an Illumina MiSeq platform. The raw reads were trimmed for adaptor and low-quality sequences and then aligned to *P. tricornutum* version 3.0 which is the reannotation of 12,089 filtered gene models. After aligning the raw data to *P. tricornutum*'s version 3.0 set 12,089 filtered gene models (protists.ensembl.org) files were filtered to retrieve uniquely aligned reads with no more than three mismatches. Gene counts (unique aligned reads per gene) were used for differential expression (DE) analysis carried out using the DESeq R/Bioconductor package, which infers DE based on the negative binomial distribution. For this analysis, we used a cutoff of 5% to control for false detection rate (false positives) and considered only genes that had a $\log_2$ -fold change $\geq \pm 2$, and a false detection rate < 0.05 to be DE. DESeq's output for all 12,089 genes was submitted to the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus under accession no. GSE133301.

Quantification of target gene mRNA copies using Quantitative Real-time PCR (RT-qPCR)

Samples for RT-qPCR were pelleted by centrifuging 6×10^7 cells for 5 min at 6,500 × g at 4 °C. The samples were frozen in liquid N_2 and stored at −80 °C. Total RNA was extracted using TRIzol™ Reagent (ThermoFisher Scientific; MA), followed by cleaning with an RNEasy MinElute Kit (Qiagen, Germany). DNA contamination was removed using Ambion Turbo DNase (Life Technologies, CA). Samples were then run on an RNAse-free polyacrylamide Gel to confirm RNA integrity. Total RNA quantification and quality assessment were made spectrophotometrically on a DS-11 FX+ Series Spectrophotometer/ Fluorometer (DS-11 FX, DeNovix Inc.; DE). Double stranded cDNA was generated using random primers with a High-Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific; MA) and directly used as the template for qPCR. Primers for target genes (Table S4) were designed with Primer Express™ Software v3.0.1 (ThermoFisher Scientific; MA). The PCR reaction was performed using the Applied Biosystems Power SYBR® Green Master Mix (Life Technologies, CA) on a QuantStudio3 (Applied Biosystems, CA). A serial dilution of five orders of magnitude of WT genomic DNA was used to plot a standard curve for copy number calculation with each primer pair. All standard curves had an $R^2 > 0.94$.

Strand-specific cDNA synthesis was carried out as above but with target specific primers (Table S4) also designed with Primer Express. This qPCR reaction was carried out using primer pairs from above with the Applied Biosystems PowerUP™ SYBR® Green Master Mix (Thermo Fisher Scientific; MA) for its increased sensitivity. At least three technical, as well as biological replicates were performed for each observation, and statistical significance was defined as $p<0.05$.

Analytical Methods

Cell densities were determined using a Beckman Multisizer™ 3 Coulter Counter® (Beckman Coulter Life Sciences, IN) as well as a Guava® easyCyte 12HT Sampling Flow Cytometer (EMD Millipore Sigma, MA). Relative chlorophyll fluorescence data, obtained from the Guava, were used for high throughput screening. Based on significantly variable LL/HL ratios of Chl a /cell, the five most interesting transformants were analyzed.

Chlorophyll *a* content per cell (Chl *a*) was measured spectrophotometrically on a Cary 60 UV-Vis Spectrophotometer (Agilent Technologies, CA) from cells collected on GF/F filters (Whatman plc., United Kingdom) and extracted in 90% acetone (55) in a FastPrep-24™ using Lysis Matrix C with modifications to manufacturer's instructions (MP Biomedicals; CA). *In vivo* absorption spectra were measured with an SLM-Aminco™ DW-2000 spectrophotometer (Olis; GA) using optically thin cell suspensions. These values were normalized to Chl *a* and used to calculate the optical absorption cross-sections, referred to as a^* (56, 57).

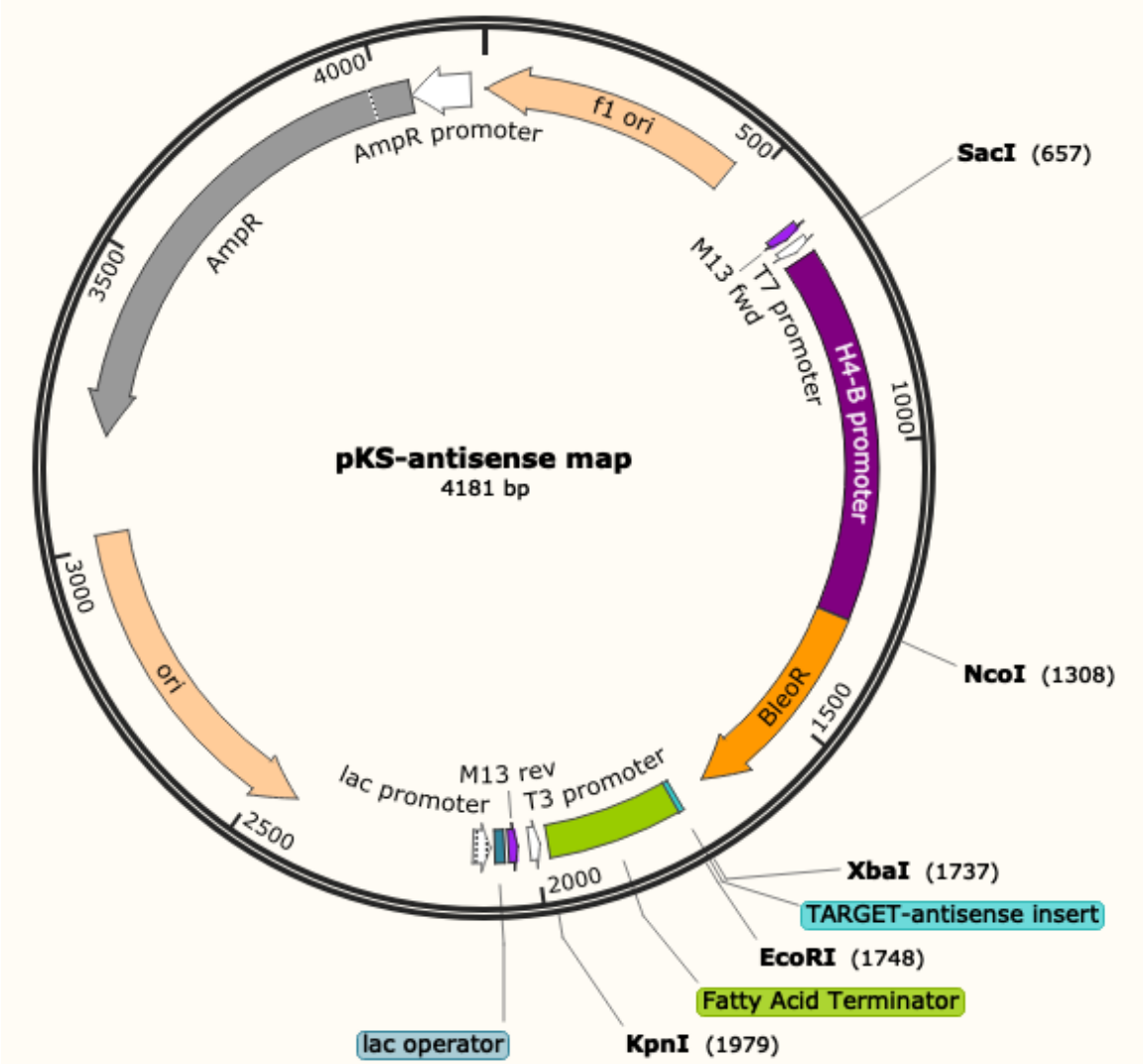
PSII biophysical characteristics were measured on a custom-built fluorescence induction and relaxation instrument (FIRe, Satlantic Inc., Canada; Gorbunov & Falkowski, 2004). The kinetics of the single-turnover saturating flash were analyzed to obtain the maximum quantum efficiency of photochemistry (F_V/F_M) and the functional absorption cross-section of PSII (σ_{PSII}).

Functional metabolic assignments for identified gene transcripts into 16 relevant gene categories was done using DiatomCyc (www.diatomcyc.org), JGI Protist for Phatr2 (genome.jgi.doe.gov/Phatr2), Ensembl for Phatr3 (https://protists.ensembl.org/ Phaeodactylum _tricornutum/Info/Index), and NCBI databases (www.ncbi.nlm.nih.gov).

Information regarding the localization of these target genes was predicted *in silico*, based on gene sequences retrieved from NCBI. HECTAR v1.3 software was used to predict subcellular targeting for heterokont proteins (webtools.sb-roscoff.fr); TMHMM was used to predict transmembrane helices in proteins (www.cbs.dtu.dk/services/TMHMM); SignalP 4.0 Server was used to predict presence and location of signal peptide cleavage sites (www.cbs.dtu.dk/ services/SignalP); and Target P1.1 Server was used to predict the subcellular location of eukaryotic proteins (www.cbs.dtu.dk /services/TargetP).

Identification of putative diatom lncRNAs candidates was based on Coding Potential Calculators, CPC (17) and CPC2 (18) scores and then filtering for transcript length $\geq$ 200 nt and open reading frames (ORFs) < 100 aa (19).

Plasmid Map S1 : RNA antisense vector



CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCA
TTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACT
CCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCA
CCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGG
GAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGG
AAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCG
TAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCA
GGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG
GCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTC
ACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAAT
TGGAGCTCGCATCTCACGCACCAGGCGCTGGAAGGGCAACTTGCGGATGAGAAGGTCCG
TGGACTTCTGGTAACGACGGATCTCACGCAGAGCGACGGTTCCAGGGCGATAACGGTGG
GGCTTCTTGACTCCTCCGGTAGCCGGAGCGGACTTGCGGGCAGCCTTGGTGGCAAGCTG
CTTGCGCGGCGCTTTGCCTCCGGTGGATTTACGGGCGGTTTGCTTGGTTCGGGCCATTTT
GACGGTTTTTTTTTACAAGAGAAGAGTTCTTGAAATTTGTGAGGTTAAAGTGTGTGGCTT
CCGCCGTAGTCAAGGAGCGTGCGGTTGCCGATCGCACCGGTACGTTCTGTAGAAATGAA
CACAGTGTGTTGAATTGAAAGTATGGCGCAGGTATGGTGTGTGATAAGTAGCAGCCGCG
CCGAGACAAACAAACTTTGGTTTCTACGACAATCTCTGTAGACAAGTACTAGAAACCCG

TTTGAACGAGCATAAATCTGCACCGGCAGGCCACCAGACATCGTTTCAACGTAATATTCT
ACGTAACCATTTTATCCCAGGAAACCTACGGCCTGTGAACCATGAAAGGAAGCACTCAC
AATTCGCTCTCGGCAACAACCGACAATAGTCTTACTCACAGTCAATACCGAAAACAAAC
AACAGCCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCGCGCGACGTCGCCG
GAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGAC
TTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGTCCAGGACCAGGT
GGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCG
AGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAG
ATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGT
GCACTTCGTGGCCGAGGAGCAGGACTGACCGACGCCGACCAACACCGCCGGTCCGACGC
GGCCCGACGGGTCCGAGGCCTTCTAGA(TARGET)GAATTCTGAGCTACCTCGACTTTGGC
TGGGACACTTTCAGTGAGGACAAGAAGCTTCAGAAGCGTGCTATCGAACTCAACCAGGG
ACGTGCGGCACAAATGGGCATCCTTGCTCTCATGGTGCACGAACAGTTGGGAGTCTCTA
TCCTTCCTTAAAAATTTAATTTTCATTAGTTGCAGTCACTCCGCTTTGGTTTCACAGTCAG
GAATAACACTAGCTCGTCTTCAGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTG
CGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAA
TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTG
AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG
CTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT
ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA
AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT
GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC
AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC
CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTAT
CCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT
GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTA
AGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG
CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTG
ACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCA
GCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT
TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT
TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC
CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG
TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC
TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT
CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC
TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG
GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTAT
TGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCC
GCGCACATTTCCCCGAAAAGTGCCAC

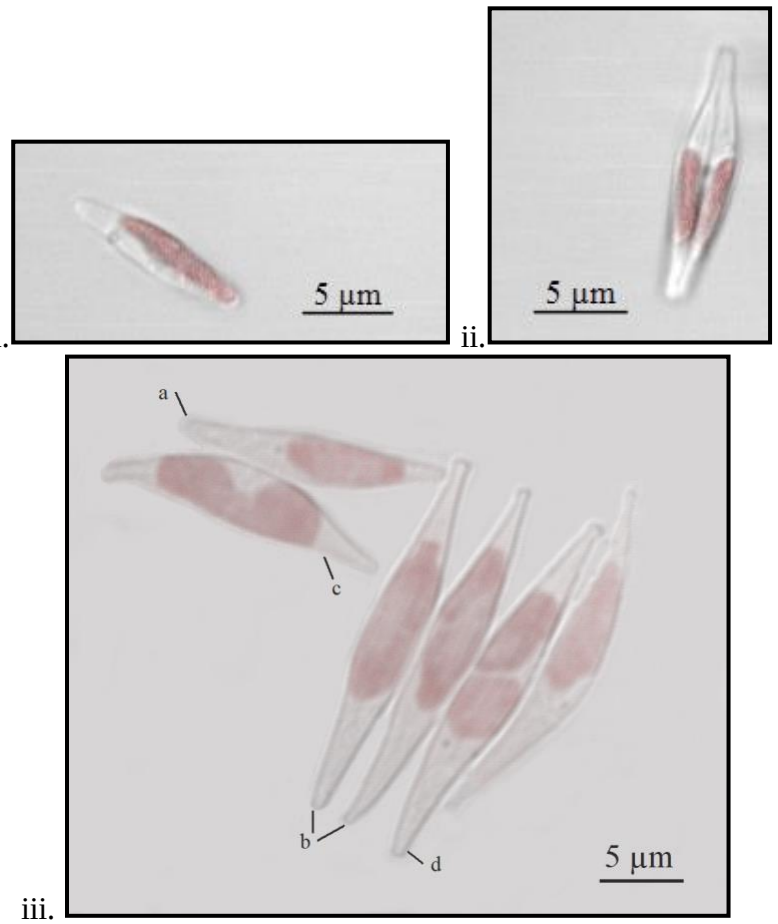


Figure S1:
Confocal laser microscopy images of low light acclimated Wild Type (WT) *P.tricornutum* and the PTP-33 transformant strain. Plastids (pink) showing various stages of growth and division.
(i.) LL (20 µmol photons $m^{-2}s^{-1}$) acclimated WT having a single plastid.
(ii.) WT cell acclimated to LL, undergoing mitosis. The visible septum and the two separated plastids observed are a result of this cell division, one for each daughter cell.
(iii.) PTP-33 transformants acclimated to LL. Normal or WT-like plastid in cell 'a'; abnormally large plastids in cells 'b'; a plastid undergoing fission within the cell 'c' and cell 'd' with two plastid organelles.

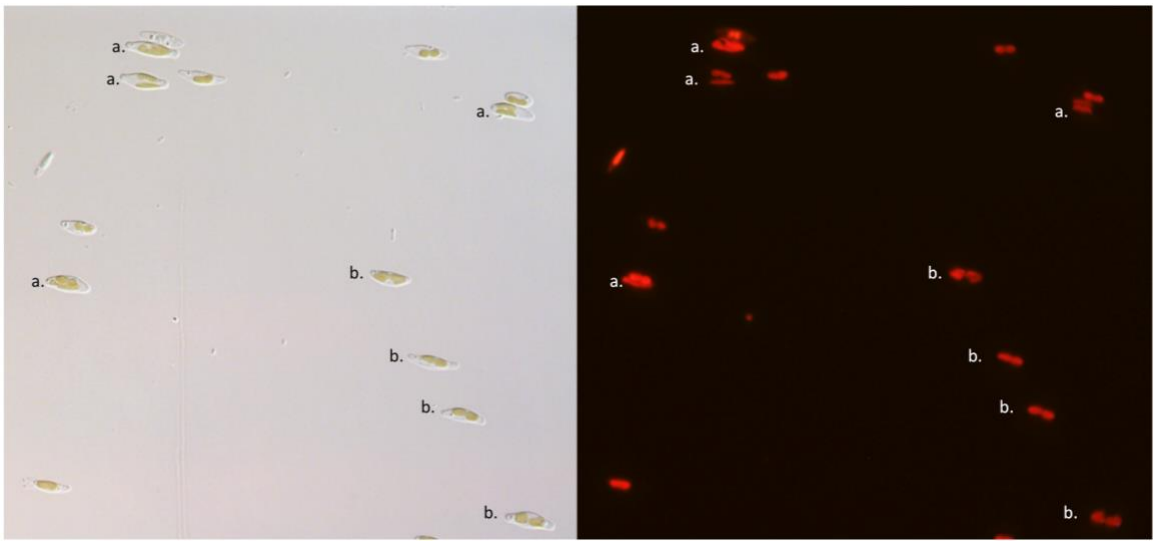


Figure S2 – Fluorescent microscopic images of transformant line PTP-33 acclimated to low light growth conditions

600x magnification microscopic image of late exponential phase PTP-33 transformant acclimated to low light (LL) intensity of 20 μmol photons $m^{-2}s^{-1}$. Brightfield image (left) and fluorescent image capturing chlorophyll autofluorescence (right) clearly show cells that appear to have two plastids per cell. The plastids of cells in the process of mitosis cleave along the same axis as the cell, i.e., a meridional plane (a.). Plastid division independent of mitosis, however, divide perpendicular, along an equatorial plane (b.).

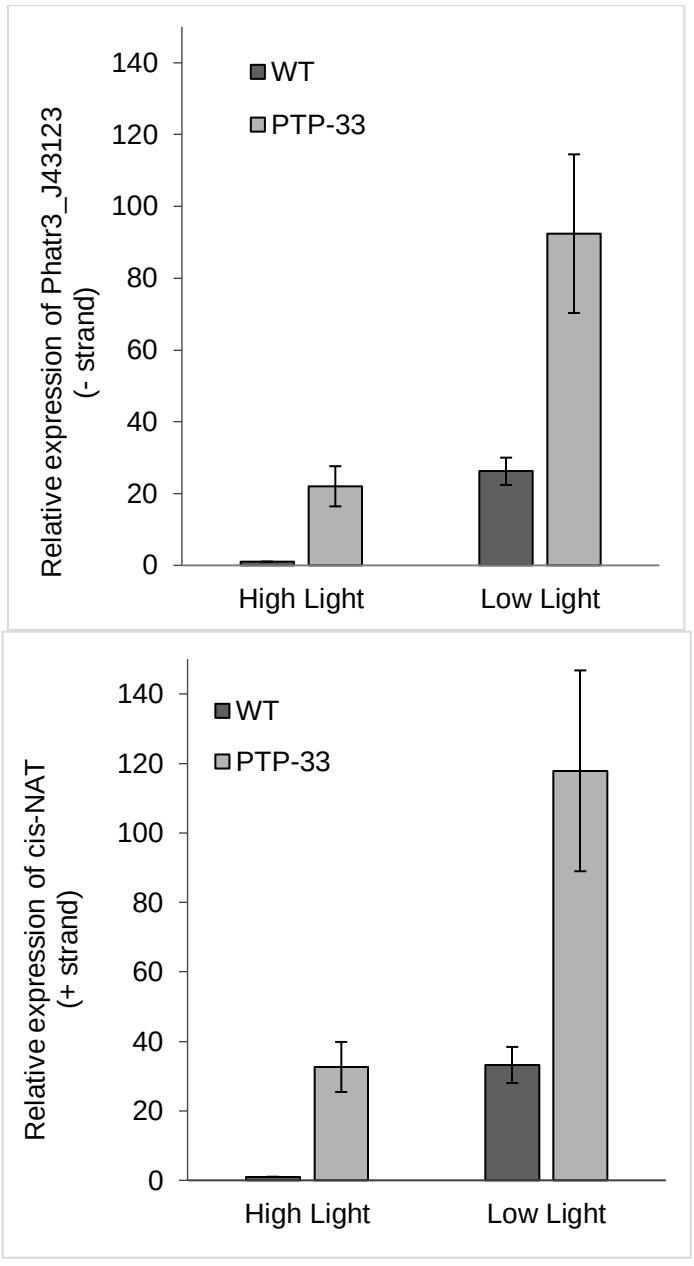


Figure S3:

Relative expression levels (RT-qPCR) of mRNA Phatr3_J43123 and its cognate Natural Anti-sense Transcript (*cis*-NAT) from the opposite strand, normalized to wild-type levels under high light (n = 3; Error bars represent ± SD). For cis-NAT p(WT/PTP-33 in HL) = 0.035; p(WT/PTP-33 in LL) = 0.029; p(LL/HL of WT) = 0.019. For Phatr3_J43123 p(WT/PTP-33 in HL) = 0.037; p(WT/PTP-33 in LL) = 0.007; p(LL/HL of WT) = 0.036.

Phatr2	*Phatr3*	*LL/HL log2 fold change*	*LL/HL fold change*	*p-adj (FDR)*	*Gene name / description*
22680	22680	-	-	-	LHCF13 - fucoxanthin chl a/c protein
27278	27278	5	**32.1**	**4.35E-121**	LHCX1 - fucoxanthin chl a/c protein
48882	48882	4	**21.1**	**6.10E-169**	LHCF15 - fucoxanthin chl a/c protein
30648	30648	4	**19.7**	**5.99E-68**	LHCF5 - fucoxanthin chl a/c protein
22395	22395	2	**3.2**	**5.12E-12**	LHCF8 - fucoxanthin chl a/c protein
50725	9799	2	**2.9**	**4.20E-12**	LHCR3 - fucoxanthin chl a/c protein
47485	13877	1	**2.7**	**8.54E-12**	LHC13877 - fucox - chl a/c protein
48798	EG00416	1	**2.7**	**6.15E-12**	LHCdeviant - fucox - chl a/c protein
54027	54027	1	**2.6**	**3.40E-12**	LHCR12 - fucoxanthin chl a/c protein
25893	25893	1	**2.5**	**1.16E-13**	LHCF14 - fucoxanthin chl a/c protein
44601	11006	1	**2.4**	**1.90E-11**	LHCR1 - fucoxanthin chl a/c protein
17531	17531	1	**2.4**	**4.94E-08**	LHC17531 - fucox - chl a/c protein
25172	25172	1	**2.3**	**n/a**	LHCF2 - fucoxanthin chl a/c protein
38720	38720	1	**2.2**	**1.58E-06**	LHCX4 - fucoxanthin chl a/c protein
22006	22006	1	**2.1**	**2.19E-12**	LHCF10 - fucoxanthin chl a/c protein
34536	34536	1	**2.0**	**3.60E-03**	LHCF16 - fucoxanthin chl a/c protein
30643	29266	1	**2.0**	**5.33E-07**	LHCF7 - fucoxanthin chl a/c protein
23257	23257	1	**2.0**	**4.12E-05**	LHCR11 - fucoxanthin chl a/c protein
30031	EG00427	1	**2.0**	**1.02E-05**	LHCF9 - fucoxanthin chl a/c protein
29266	29266	1	**2.0**	**3.53E-07**	LHCF6 - fucoxanthin chl a/c protein
17766	17766	1	1.9	n/a	LHCR4 - fucoxanthin chl a/c protein
18049	18049	1	1.9	2.18E-05	LHCF1 - fucoxanthin chl a/c protein
47813	14386	1	1.9	3.60E-04	LHCR14 - fucoxanthin chl a/c protein
43522	18180	1	1.8	4.51E-04	LHCR7 - fucoxanthin chl a/c protein
22956	22956	1	1.8	1.52E-05	LHCR2 - fucoxanthin chl a/c protein
6062	6062	1	1.8	2.97E-06	LHC6062 – fucox - chl a/c protein,
43860	10243	1	1.5	1.84E-02	LHCR9 - fucoxanthin chl a/c protein
51230	51230	1	1.5	6.95E-03	LHCF11 - fucoxanthin chl a/c protein
16302	16302	1	1.5	7.96E-02	LHCF12 - fucoxanthin chl a/c protein
14986	14986	0	1.3	3.02E-01	LHCR5 - fucoxanthin chl a/c protein
25168	25168	0	1.1	5.60E-01	LHCF4 - fucoxanthin chl a/c protein
44733	44733	0	-0.9	5.50E-01	LHCX3 - fucoxanthin chl a/c protein
15820	EG02552	0	-1.0	9.83E-01	LHC15820 - fucoxanthin chl protein
50705	50705	0	-1.0	9.04E-01	LHCF3 - fucoxanthin chl a/c protein
56319	14242	0	-1.2	4.43E-01	LHCR6 - fucoxanthin chl a/c protein
24119	24119	0	-1.4	1.79E-01	LHC24119 – fucox - chl a/c protein
56310	EG02221	-1	-1.6	8.26E-03	LHCF17 - fucoxanthin chl a/c protein
-	EG02404	-1	**-2.2**	**5.63E-05**	LHCX2 - fucoxanthin chl a/c protein
42519	42519	-1	**-2.3**	**1.24E-06**	LHC related
32294	32294	-2	**-3.7**	**1.56E-13**	LHCR8 - fucoxanthin chl a/c protein
50086	16481	-2	**-4.7**	**2.96E-29**	LHCR10 - fucoxanthin chl a/c protein

Table S1:

Differential expression heat map of Light Harvesting Complex (LHC) genes observed in *P.tricornutum* grown in low light [LL - 20 µmol photons $m^{-2}s^{-1}$] as compared to high light [HL - 940 µmol photons $m^{-2}s^{-1}$]. (n=3; p-adj (FDR) < 0.05)

Phatr2	*Phatr3*	*LL/HL log2 fold change*	*LL/HL fold change*	*Gene name / description*
Nuclear Photosynthetic Electron Transport Components				
26293	26293	1.35	2.6	PSBU PSII OEC extrinsic subunit
13895	13895	1.35	2.5	HCF136 PSII assembly & repair
9078	9078	1.20	2.3	PSB27 PSII assembly & repair
44056	44056	1.11	2.2	PETJ Cytochrome c6, cytochrome c553
54499	54499	1.03	2.0	PSBQOEE3 Oxygen-evolving enhancer protein 3
20331	20331	1.02	2.0	PSBO Oxygen-evolving enhancer protein 1
Photoreceptors				
8113	8113	2.49	5.6	AUREO1a /Pt aureochrome1
15468	15468	2.11	4.3	AUREO2/ Pt aureochrome 4
54342	54342	1.73	3.3	CryP Cryptochrome -plant like
27429	27429	1.72	3.3	CPF1 Cryptochrome photolyase family 1
54330	54330	0.24	1.2	DPH Diatom Phytochrome1
55037	55037	0.30	1.2	SKP3 Sensor kinase protein 3
34592	34592	-0.93	-1.9	CPF2 Cry-dash cryptochrome/photolyase family

Table S2:

Differential expression heat map of Nuclear photosynthetic electron transport components and Photoreceptor genes, as observed in *P.tricornutum* grown in low light [LL - 20 µmol photons $m^{-2}s^{-1}$] as compared to high light [HL - 940 µmol photons $m^{-2}s^{-1}$]. (n=3)

Conditions	***Phatr3_J50052 differential expression (log2 fold)***	***Citation***
+/- Silica	absent	Sapriel et al. 2009
+/- Cadmium	absent	Brembu et al. 2011
+/- Nitrate	absent	Remmers et al. 2018
WT -Nitrate/+Nitrate	+0.33	Levitan et al. 2015
- Nitrate (after 4 hrs) - Nitrate (after 8 hrs) - Nitrate (after 20 hrs)	+0.47 +0.24 +0.03	Matthijs et al. 2016
- Phosphate (after 36 hrs)	-0.95	Matthijs et al. 2017
Cell cycle (79%G1 20%S) vs. (100%G1)	-1.08	Kim et al. 2017
10 umol Blue Light (1 hr exposure)	-0.83	König et al. 2017
Redox sensitive peptide fragments	absent	Rossenwasser et al. 2014
Iron limitation	absent	Smith et al. 2016

Table S3 – Presence of Phatr3_J50052 (LSK) in transcriptome studies of *P. tricornutum* grown under various environmental conditions

276

	Primer Name	Nucleotide Sequence
1.	Phatr3_EG01529 Gibson - Forward	5’ - CGG GTC CGA GGC CTT CTA GAT AGA TAG TTC CAC AGA ACT ATT C – 3’
2.	Phatr3_EG01529 Gibson - Reverse	5’ - GTC GAG GTA GCT CAG AAT TCA GCG AAC TAA AAC CAT GAC – 3’
3.	Phatr3_J47715 Gibson - Forward	5’ - CGG GTC CGA GGC CTT CTA GAC TCT CTT GAC AGC CTT CAC – 3’
4.	Phatr3_J47715 Gibson - Reverse	5’ - GTC GAG GTA GCT CAG AAT TCA TGA CCA AAA CGA ACA GAA C – 3’
5.	Phatr3_J43123 Gibson - Forward	5’ - CGG GTC CGA GGC CTT CTA GAC ACA GGA CTC TTA CGC TC – 3’
6.	Phatr3_J43123 Gibson - Reverse	5’ - GTC GAG GTA GCT CAG AAT TCA TGC AGC TTT ACC TTG TG – 3’
7.	Phatr3_EG00707 Gibson - Forward	5’ - CGG GTC CGA GGC CTT CTA GAT TAC TCT GAC TTG GTT GGT TC – 3’
8.	Phatr3_EG00707 Gibson - Reverse	5’ - GTC GAG GTA GCT CAG AAT TCA TGT CGA CCC GCT CTT TC – 3’
9.	Phatr3_J37655 Gibson - Forward	5’ - CGG GTC CGA GGC CTT CTA GAA TCC GTA CTT TCC AAC CC – 3’
10.	Phatr3_J37655 Gibson - Reverse	5’ - GTC GAG GTA GCT CAG AAT TCA CGA GAC CCG TCA ACA AC – 3’
11.	Phatr3_J50052 Gibson - Forward	5’ - CGG GTC CGA GGC CTT CTA GAC TCT TGT CCA GTC GTT CC – 3’
12.	Phatr3_J50052 Gibson - Reverse	5’ - GTC GAG GTA GCT CAG AAT TCG TAC AAG GCC ACA CTG AAA TC – 3’
13.	qPCR Phatr3_J43123 - Forward (Used for ssqPCR as well)	5’ - GCC CGC TAT GCC AAT GC – 3’
14.	qPCR Phatr3_J43123 - Reverse	5’ - CGC TTG ATT GGC GGA AAA T – 3’
15.	qPCR Phatr3_J50052_Exon1 - Forward	5’ - TCC ACC ACC GCC CAT TC– 3’
16.	qPCR Phatr3_J50052_Exon1 - Reverse	5’ - TGT AAG ATC CGG TCG ATT CCA– 3’
17.	qPCR Phatr3_J50052_Exon2 - Forward	5’ - CGA CTG GAC AAG AGC GAA AAG – 3’
18.	qPCR Phatr3_J50052_Exon2 - Reverse	5’ - AAT ACG TGT TCA TGG CAA TGC T– 3’
19.	qPCR Phatr3_J10847_RPS - Forward	5’ – CGA AGT CAA CCA GGA AAC CAA – 3’
20.	qPCR Phatr3_J10847_RPS - Reverse (used for ssqPCR as well)	5’ - GTG CAA GAG ACC GGA CAT ACC – 3’

277

278 **Table S4** – Primer sequences related to inserting antisense constructs into pKS plasmid using
279 Gibson Assembly and RT-qPCR