NTNU
Norwegian University of
Science and Technology

# Molecular and physiological responses to shifts in light quality and genomic editing using CRISPR Cas9 system targeting highly regulated genes in light response of diatom phaeodactylum tricornutum

Dechasa Shitaye Gebre

MSc in Biology
Submission date: June 2016
Supervisor: Per Winge, IBI
Co-supervisor: Atle M. Bones, IBI
Marianne Nymark, IBI

Norwegian University of Science and Technology
Department of Biology

**Acknowledge**

The work presented in this thesis was conducted in Norwegian University of Science and Technology (NTNU) at the department of biology.

I would first like to thanks my supervisor Atle Magnar Bones, Per Winge and Marianne Nymark. Atle for trusting and allowing me to work in his research group, cell and molecular biology group. This opened a door for me to have practical experience on molecular level. Per for advising, guiding and encouraging me throughout this thesis. I also want to thanks Marianne Nymark and Torfinn Sparstad for helping me in lab. I feel really lucky to have knowledge and practical experience from both of them. Torfinn for showing me how to harvest sample and isolated RNA. Marianne for helping me throughout this thesis and sharing her amazing knowledge and experience about light response of *Phaeodactylum tricornutum*. I do like to say thanks to Amit Kumar for sharing his experience with me. Amit always takes paper to show me how to conduct. It was easy to understand what Amit explain and show me on paper before I conduct the experiment. This makes me to work independently. Finally I do like to say thanks my sponsor Lanekassen for enabling me to go to school (NTNU) and Tore for helping me helping me in a lab.

Trondheim, 28 of June, 2016

Dechasa Shitaye

**Abstract**

Diatoms are unicellular brown eukaryotic algae belongings to the subphylum of Bacillariophyta under the division of heterokonta (Stramenopiles) and are thought to have arisen 280 million years ago. They are photosynthetic organism having a diplontic life cycle that can adjust to a various light intensity that enables them to adapt to a different environment which enables them to achieve optimum growth. The wide distribution of diatoms in marine ecosystems indicates that they are able to adapt to a broad range of habitats with changing light and nutrient availability. Diatoms uses different mechanism to adjust to different light condition, a central component here are the light harvesting complexes (LHCs) which can be fine-tuned and obtain maximum efficiency as well as adequate protection against photodamage under varying environmental condition. In this study the model diatom, *Phaeodactylum tricornutum* was used to study transcriptional responses to high intensity red and blue light. *P.tricornutum* has at least 42 light harvesting proteins, 41 of them can be divided in four main groups: LHCFs, LHCRs, LHCYs and LHCXs. The expression of a selected set of these genes together with other genes encoding other components of the LHCs were studied under high red and blue light. The Photosynthetic efficiency under normal (white), red and blue light was measured using Aquapen-AP 100. qRT-PCR data showed that genes which is involved in photoprotection and repair were repressed during 48h red light (photodamage) but recovered in subsequent blue light exposure (photoprotection). High intensity blue light initially reduced the photosynthetic efficiency in during the first 3h, but partially recovery after 48h. Subsequent high red light exposure resulted in reduced photosynthetic efficiency indicating that blue light had no longer lasting effects on photoprotection. Based on these results two LHC-genes (LHCF15 and LHCR8) were selected for genomic editing using CRISPR Cas9 technology.

**Abbreviations**

| | |
|---|---|
| BL | Blue light |
| bp | base pair |
| Cas 9 | CRISPR associated protein 9 |
| cDNA | Complementary DNA |
| CRISPR | Clustered regularly interspaced short palindromic repeats. |
| crRNA | CRISPR RNA |
| CWL | Continuous white light |
| dsDNA | Double stranded DNA |
| DSB | Double stranded break |
| DT | Dark treatment. |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxy ribonucleotide triphosphate. |
| gRNA | Guide RNA |
| gDNA | Genomic DNA |
| InDel | Insertion and deletion |
| ML | Milliliter |
| NHEJ | Non-homologous end joining |
| ng | Nanogram |
| PAM | Protospacer Adjacent Motif |
| PCR | Polymerase chain reaction |
| qRT-PCR | quantitative real time PCR |
| RL | Red light |
| RNase | Ribonuclease |
| rRNA | Ribosomal RNA |
| RNA | Ribonucleic acid |
| RIN | RNA integrity number |
| RT | Reverse transcript |
| TALEN | Transcriptional activator like effector nucleases |
| tracrRNA | Trans-activating crRNA. |
| Wt | Wild type |
| ZFN | Zinc finger nuclease |
| $^0$C | Celsius |

| | |
|---|---|
| µL | Microliter |
| µg | Microgram |

# CONTENTS

# 1. INTRODUCTION

## 1.1 **Diatom**

The unicellular brown algae Bacillariophyta (diatoms) is one of the most diverse groups of photosynthetic eukaryotic organism and have a world-wide distribution. Diatoms are phylogenetically placed a subphylum under the division of heterokonta (Stramenopiles) and are believed to arisen approximately 280 million years ago. This subphylum includes some diatom having centric structure and others having a pennate form(Loke et al. 2004, Mann 2010, Seckbach and Kociolek 2011). More than 250 living genera of diatoms are present in today's ocean's and fresh water habitat and with about 100,000 species they are responsible for approximately 25% of world primary production(Scala et al. 2002). They have an elaborate cell wall made from silicon dioxide ($SiO_2$). The cell wall of diatom are composed of two halves of different size(Smol and Stoermer 2010). It gives the diatoms different forms such as polygonal, needle-like and sigmoid shapes. The exoskeleton of diatom, known as the frustule, consists of a mixture of amorphous silica and organic matter. The unique patterning of the frustule can be used to differentiate the different species using light microscopy. However, the differences in valve size and shape observed during the diatom life cycle may complicate the taxonomic process. The difference in valve's shape, at various stages of cell division (Figure 1) is a complicating factor that has resulted in incorrect classification of some diatom as new species (Vrieling et al. 1999). Even though adaptation of diatom to environment is not well known, they can adjust themselves to different light intensity to achieve optimum growth rates. This enable them to colonize a wide range of ecological niches and they are found in benthic, open water and some of them are also terrestrial (Scala et al. 2002). Diatoms are important for biogeochemical cycling of carbon and minerals like silica and they can be also be used as environmental indicators (Scala et al. 2002, Loke et al. 2004).

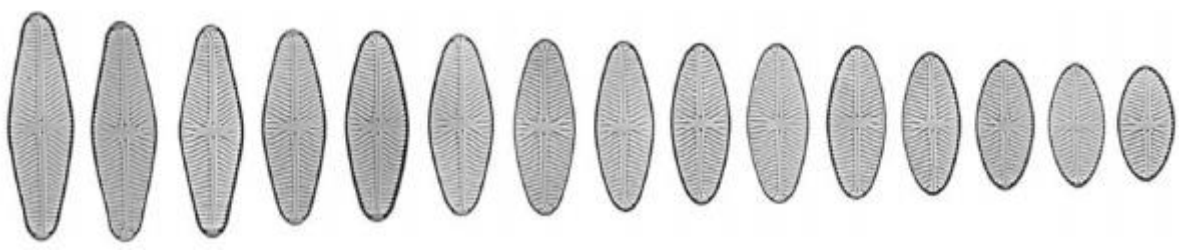


**Figure 1.** Difference in size and shape caused by cell division of diatom *Navicula reinhardtii* (Loke et al. 2004)

## 1.2 Evolution of diatom

As molecular sequence data indicated, the diatoms are believed to be arisen around 280 million years ago. The Stramenopile group, which the diatoms are a subphylum, evolved even earlier as the result of one or more endosymbiotic events (Zurzolo and Bowler 2001). Although the diatom has a complicated evolutionary history, they can be easily identified among eukaryotic group of algae due to the presence of silicified cell wall (frustules) that contain overlying thecae(Armbrust 2009, Seckbach and Kociolek 2011). Several studies have been conducted to throw light upon the origin of the plastids in diatom and the important role they play in the biosphere. The plastid (chloroplast) is an organelle in which photons from sun light are harvested and used for photosynthesis. Several proteins that are used for photosynthesis and rRNA and tRNA are transcribed from circular plastids genome. During evolution the majority of genes from the initial endosymbiont have been transferred to the nucleus. Single celled protists that could not prepare their own food by photosynthesis became photosynthetic organisms by engulfing cyanobacteria(Yoon et al. 2002). After this event, a secondary heterotrophic host cell did engulf the primary endosymbiont. Euglenoid or chlorarachniophytes algae evolved when a green algae got engulfed by an ancient heterotroph. Figure 2 shows the outcome-when a red algae got engulfed by heterotroph and give rise to the cryptophytes, haptophytes, heterokonts or dinoflagellate algae(Seckbach and Kociolek 2011).

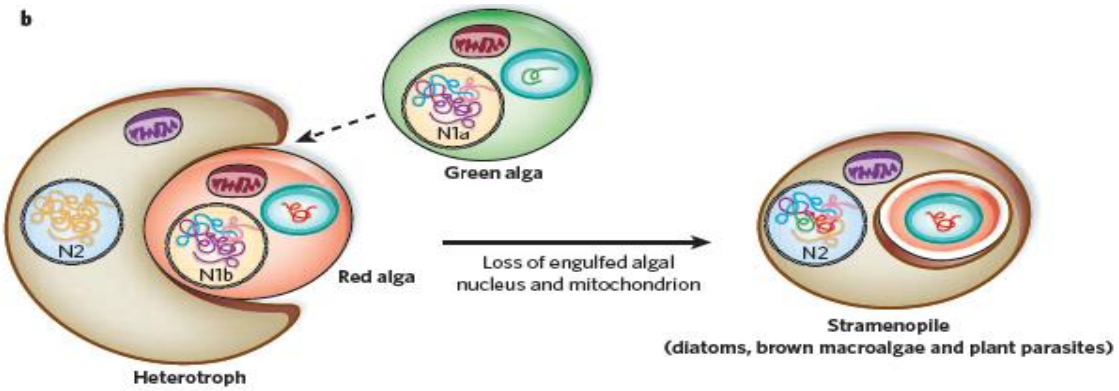



Figure 2**.** Stramenopiles including diatoms are thought to have arisen through an ancient endosymbiotic event, where a heterotroph engulfed a red algae and possibly a green algae (Armbrust et al. 2004, Armbrust 2009)

## 1.3 Diatoms and cell cycle/cell division

Having diplontic life cycle makes diatom different from other algae(Mann 1993). Even though diatoms frustule becomes smaller and smaller in size as the diatom goes through cell division, the regaining of original size can occur through sexual reproduction, through auxospore formation. As the cell size decreases to a threshold limit (30-40% of maximum diameter size), the diatom will go through gametogenesis to produce haploid gametes. Fertilization of both gametes produces a diploid auxospore that is greater in size than both parents. The larger sizes of the offspring have a selective advantage and spread through the population. The produced offspring will go through asexual reproduction until they the size threshold limit which make them inducible by environmental factors or pheromones and thereby start a new round of gametogenesis. Different environmental factors like light, nutrient, salinity and temperature can initiate the sexual phase by shifting from asexual to sexual reproduction when the condition are suitable (Armbrust 1999, Falciatore and Bowler 2002a).

During asexual reproduction, the diploid cell is divided into two (figure 3) and the daughter cells have one valves, the epitheca, that comes from parent cell. Then the daughter cell produces new hypotheca. Before the cell undergoes cell division, the epitheca is pushed from hypotheca by elongation of the cell then the nucleus divides by mitosis. Following nuclear division, the protoplast divides by invagination. This is followed by production of a new hypotheca by the daughter cell.

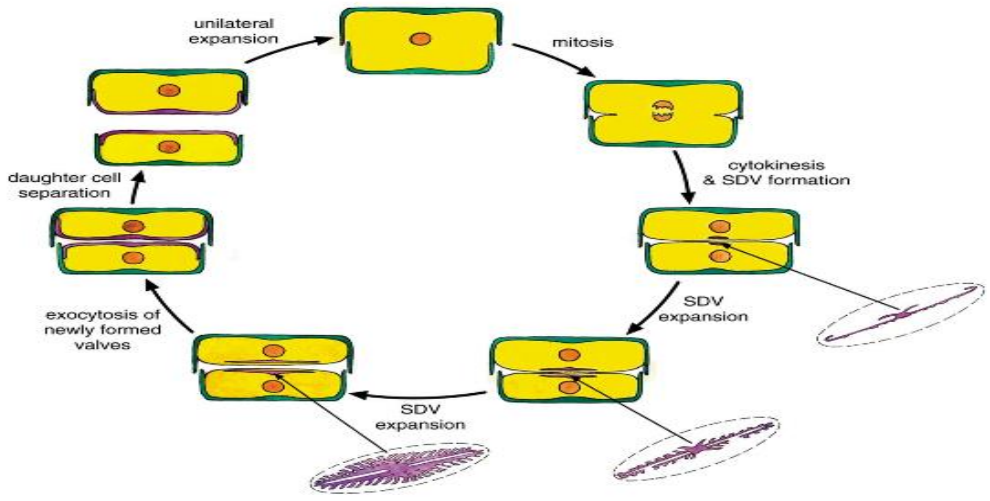



**Figure 3** Mitotic cell division and production of hypovalves in the pennate diatom. SDV is silica deposition vesicle(Falciatore and Bowler 2002b).

## 1.4 Phaeodactylum tricornutum.

Even though *Phaeodactylum tricornutum* is generally accepted to exist is three forms oval, fusiform and triradiate, Douglas P.Wilson reports in 1946, *P.tricornutum* are able to produce four morphotypes such as fusiform, triradiate, oval and cruciform and as Bohlin reports in 1897 new description part of the paper, *P.tricornutum* has two forms ovate and fusiform (He et al. 2014). Bohlin explained a new genus *Phaeodactylum* that contains single species of *P.tricornutum*. It is unicellular having a brown chromatophore located at the central part and there is three form of the cell, ovate, fusiform and triradiate. The ovate is 8µm long and 3 µm wide that contain one silica valve per cell. The fusiform form morphotype is 25-35µm long having two more or less blunt and somewhat bent arms. As long as the *P.tricornutum* having golden brown chromatophores, store oil and leucosin and silica wall are produced by the cell at some stage of development, the cell is certainly placed under Chrysophyta and the taxonomic classification of *P.tricornutum* is as below (Lewin 1958). Because *P.tricornutum* has a small genome size, a short generation of time and there exists methods for genetic transformation, it has become a model organism for diatoms(Scala et al. 2002).

**Taxonomic position.**

| | |
|---|---|
| **Phylum** | Bacillariophyta |
| ***Class*** | Bacillariophyceae |
| ***Order*** | Bacillariales *(*Hendey, *1937)* |
| ***Suborder*** | Phaeodactylineae (nov.subord**).** |
| ***Family*** | Phaeodactylaceae |
| ***Genus*** | Phaeodactylum *(*Bohlin, 1897*)* |
| ***Species*** | *P.tricornutum.* The only know species(Lewin 1958*).* |

### 1.5 Light and photosynthesis in ocean.

The wide distribution of diatoms in marine ecosystems show that they are able to adapt to a broad range of habitats where light conditions and nutrient availability show large variations. The decrease in irradiance with the depth causes a difference in light intensity in the water column (Owens 1986, Zurzolo and Bowler 2001). Sunlight can reach a depth of 1000 meters if the water is clear, but below 1000 meter it is dark, making difficult for phototrophic organism to survive below 100-200 meters as shown in appendix 1a. The wavelength of yellow, blue and green light can go deeper in the ocean while red light can only penetrates 5-10 meters (Gakuran 2008). Although there is different light intensity in the marine environment, phytoplanktonic organisms have different strategies to adjust to the changing conditions. For example cyanobacteria has a method of overcoming variability of light intensity by using complementary chromatic adaptation that adjusts the composition of photosynthetic harvesting apparatus. The second approach that phytoplankton use to respond to light intensity differences is movement toward the light using phototaxis but planktonic diatoms haves no flagella and can only perform passive movement like sinking or be transported through water turbulence. These various mechanism enable them to maximize photosynthesis(Falciatore and Bowler 2002b).

In diatom, the plastid is the organelle that contains photosynthetic apparatus. In the plastids, chlorophyll a/c and carotenoid pigments are bound to light harvesting antenna complexes proteins, chlorophyll a/c binding proteins, also known as fucoxanthin chlorophyll proteins (FCPs). These proteins are used to transport light energy to chlorophyll a in the reaction centres of the photosystems(Falciatore and Bowler 2002b). In photosystem I and II, the brown carotenoid pigments are used for both transmission of energy as well as protecting them against photodamage.

### 1.5.1 Light harvesting complex

Phylogenetic analysis of light harvesting complex superfamily of *P.tricornutum* was done and described using cladogram (appendix 1b). Among 42 light harvesting complex proteins, 41 of them were annotated and placed into four main groups: LHCFs, LHCRs, LHCXs and LHCYs (Nymark et al. 2013).

The pigment and pigment-binding proteins in light harvesting complexes are involved in capturing of photons and transferring them to the reaction centre. These complexes surround the reaction centre and are embedded in the thylakoid membrane together with other protein complexes involved in electron transfer and splitting of water. PSII, cytochrome $b_6f$ ($Cytb_6f$) complex and PSI are the three main complexes involved in photosynthetic electron transport. The light harvesting complex (LHC) capture and transfer the photon to the reaction centre which triggers charge separation and subsequence electron transfer reaction. This creates a strong oxidant in PSII that can split a water molecule to generate two protons ($H^+$), and two electrons, and produce an oxygen molecule. The electrons are transported through the electron transport chain from PSII to PSI through the Cytb6f complex. The strong reductant created by charge separation in PSI is able to reduce ferredoxin which in turn is used by ferrodoxin NADPH reductase to reduce $NADP^+$ **t**o NADPH(Rochaix 2011).

The LHC proteins in diatoms have pigment binding site that can bind chlorophyll a, c1, c2 and various carotenoids. So far it is unknown whether the different classes of LHC proteins bind different combinations of fucoxanthin-chlorophyll molecules. The pigment-protein composition of LHC of green algae and higher plan is different from that of diatoms, mainly due presence of fucoxanthin. There are three forms of protein fucoxanthin (fuco) complexes, each with a little bit difference in light harvesting spectrum, $fuco_{blue}$, $fuco_{green}$ and $fuco_{red}$. $Fuco_{green}$ and $fuco_{red}$ are capable of light harvesting but $fuco_{blue}$ is involved in structural stability and photo protection. Fuco in diatoms is involved in the harvesting of light within the blue-green region (450_550nm), while chlorophyll absorbs the blue light with 400-500nm and red in 600-700nm(Valle et al. 2014).

### 1.5.2 Photo induced damage and repair of PSII

Photosynthesis is the process in which organisms use photons of light to drive endergonic reaction that can produce chemical energy in the form of ATP and NADPH which can be used in the calvin cycle for production of various sugars. Maximum photosynthetic efficiency is the maximum efficiency of the organism to convert absorbed photons to chemical energy(Bolton and HALL 1991). The maximum photosynthetic efficiency of algae exposed to the high intensity of light will normally decrease because this causes the absorption exceeds the ability of the cells to use the photon for photosynthesises that result in acceleration of production of ROS(Nishiyama et al. 2001). ROS can cause damage to proteins involved in photosynthesis targeting particular PSII by affecting protein D1 that binds to D2 protein in the reaction centre of PSII. In normal light the D1 protein damage and degradation is low, but it increases in parallel with higher light intensity (Barber and Andersson 1992, Nymark et al. 2013). Photodamage is caused by singlet oxygen, superoxide, hydrogen peroxide ($H_2O_2$) and hydroxyl radicals, the oxygen evolving complex of PSII singlet oxygen is produced by the reaction of excited chlorophyll a ($P680^*$) with oxygen, when the capacity to transport electron is exceeded. The effect of photodamage can be reduced by partly disassembling the damaged PSII complexes, resynthesizing of new D1 protein and incorporation of a new protein subunits into a reassembled PSII complex. This process is called PSII repair cycle. The cycle contain some steps to completely repair damaged protein: '' i) monomerization of PSII dimer and its localization from grana to stroma, ii) partial disassembly, iii) FtsH protease breaks D1, iv) synthesis of precursor D1 protein, V) D1 matured through C-terminal cleavage of protein , Vi) reassembly of PSII reaction center proteins and Activation of PSII complex'' (Rochaix 2011, Takahashi and Badger 2011).

Photoinhibition is a process where photoinactivation is greater than photoprotection. Photo protection occurs when there is a high turnover of many PSII subunits coupled to high transcription of genes coupled to photorepair (Domingues et al. 2012). Damage of D1 protein and high level of diatoxanthin (diato) results in lowering the maximum quantum yield ($\Phi_{PSII\_max}$) of PSII during exposure of *P. tricornutum* to highlight from low light. Genes (PtLHCR8, PtLHCR10 and PtLHCX3) that involves in photoprotection are induced to encode proteins that for photoprotection. After photodamage, the cell has repair mechanism in which some genes (Psb27, Psb29 and HCF136) involves in assembling and repair of damaged protein. In a previous study, it has been reported that the maximum quantum yield ($\Phi_{PSII\_max}$) became

lower when *P.tricornutum* was exposed to high red light (730µmol $m^{-2}s^{-1}$) for 96h, but showed 100% recovery (when compared with white light) during exposure to low red light (80µmol m2s1) for 72h (Valle et al. 2014). Does the same occur when *P.tricornutum* is exposed to high blue light for 48h or is this an effect only induced by high red light exposure? This is one of the objectives of this study.

### 1.6 Genome editing using CRISPR Cas9 technology

Homologous recombination (HR) is the process in which DNA double stranded breaks (DSBs) is repaired depending on the template. It correctly repair DSB using homologous sequence. Homologous recombination to create knockout has been used in yeast and mice. This is achieved by inserting blastocysts cells with embryonic stem (ES) cells that is modified. This method is time consuming to create a single gene knockout mice. (Li and Heyer 2008, Wang et al. 2013).

Bacteria and archaea have developed a systems that protects them from invading bacteriophages. This defensive strategy against bacteriophages involves among others abortive infection (altruistic cell death) and prevention of phage adsorption. Other systems are targeting directly the nucleic acid (DNA or RNA) that invade the bacteria. Recently 'clustered regularly interspaced short palindromic repeats (CRISPR)-type systems was discovered that works as an adaptive immune system in bacteria and Archea species to protect them against phages and plasmids. These clustered repeats were discovered for the first time in 1987, but the functions of these repeats and associated genes were elusive for many years (Jansen et al. 2002, Horvath and Barrangou 2010). Since then RNA guided Cas9 proteins derived from CRISPR-loci have been developed as a genetic tools which can create specific double stranded target breaks. The ability to target and edit specific genes in the genome has been a major challenge and was used for a long time restricted to a few species like yeast and mice. Initially homologous recombination was used to target specific genes in the genome and produce gene knockouts.
Since then new method have been developed then relies on producing double stranded breaks (DSBs) in DNA, four type of proteins that can bind to DNA are used; mega nuclease, zinc finger nuclease(ZFNs), transcription activator-like effector nucleases(TALENs) and recently

CRISPR/Cas9 nucleases (Hsu et al. 2014, addgene 2016, Nymark et al. 2016). Mega nucleases, ZFNs and TALENs were designed to knockout genes through homologous recombination and have been successfully used in various applications but they have their own limitations. For example meganucleases bind to relative large recognition sites, up to 40 bp, but have still problems with off-target effects (they cut at multiple sites in the genome). ZFNs can be difficult to design, they are relatively expensive and must to be tested in vitro before use. TALENs also require a laborious design process, but are easier to work with ZFNs (Cong et al. 2013, Hruscha et al. 2013). CRISPR/Cas9 nucleases were recently designed for genome editing. The CRISPR was originally discovered as an adaptive immune system in bacteria and archaea that depend on trans-activating crRNA (tracrRNA) and pre-CRISPR RNA (crRNA) to remember and target invading phages (Mali et al. 2013). The CRISPR/Cas9 system are composed of two components, the CRISPR/associated endonuclease (Cas9) and guide RNA (gRNA)(Perez-Pinera et al. 2013).

The tracrRNA is used as a scaffold to bind crRNA to Cas9, where the crRNA is used to indicate the target site which will be cut upstream of the Protospacer Adjacent Motif (PAM). Cas9 binding is directed by PAM sequence having short (3-5) nucleotide sequence that gives signal for Cas9 to bind to target site (Figure 4) (Addgene no date-a).

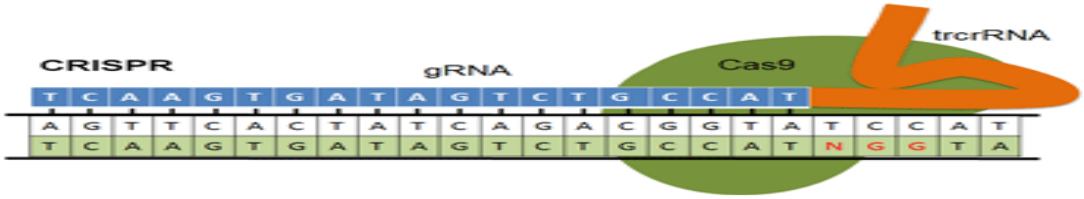



Figure 4. Riboprotein form of gRNA and Cas9 complex bind to the target site of dsDNA to create DSB(addgene no date-b)

The guide RNA (gRNA) is a fusion of tracrRNA and crRNA, where the 20 bp crRNA (spacer), which define the target site, is located at the 5' end of gRNA (shown in figure 5). To change the target site the only modification needed is to change the 5' end (the crRNA part) of the gRNA.

### 1.6.1 Producing a Knock-out using CRISPR/Ca9

By co-expressing a gRNA targeting a specific gene and a nuclear targeted Cas9 protein this system can produce DSBs that are likely to produce a gene knockout. The target site must fulfil two requirements: the sequence should not be present anywhere else in the genome and the

target site must be located upstream of PAM site. After expression, gRNA interact with Cas9 by using its scaffold domain and the positively charged groove of Cas9 to form a riboprotein complex. Electron micrograph and X-crystallography indicates that Cas9 up on binding to gRNA and target dsDNA, undergoes a conformational change that results in change from non-active DNA binding to active DNA binding form and that the spacer (located at the 5' end of the gRNA) interact with the complementary sequence of target DNA_(Doudna and Charpentier 2014). The complex of gRNA and Cas9 can bind to any sequence (as long as there is a suitable PAM-site available), but Cas9 can only cut the target site if there is sufficient match between the target DNA and spacer. After the complex of Cas9-gRNA binds to the assumed target DNA then 3' ends of gRNA start to anneal to the target DNA. The match between spacer of gRNA and target DNA cause the gRNA to anneal from 3'-5' direction. The HNH and RuvC endonuclease domains of Cas9 binds to the complementary and non-complementary dsDNA for cleavage and the nuclease lobe also contain a carboxyl-terminal domain that is used to interact with PAM (Nishimasu et al. 2014). After sufficient homology exist between gRNA spacer and target DNA, cleavage is performed by the Cas9 nuclease. However, mismatches at the 3' or 5' of the spacer can eliminate target site cleavage due to the 'Zipper-like' annealing mechanism initiated by the Cas9 protein. The two nuclease domains undergoes a conformational after binding to the target site that results in DSBs of DNA. The DSBs can be repaired by either Non-homologous or homologous end joining (addgene No date ).

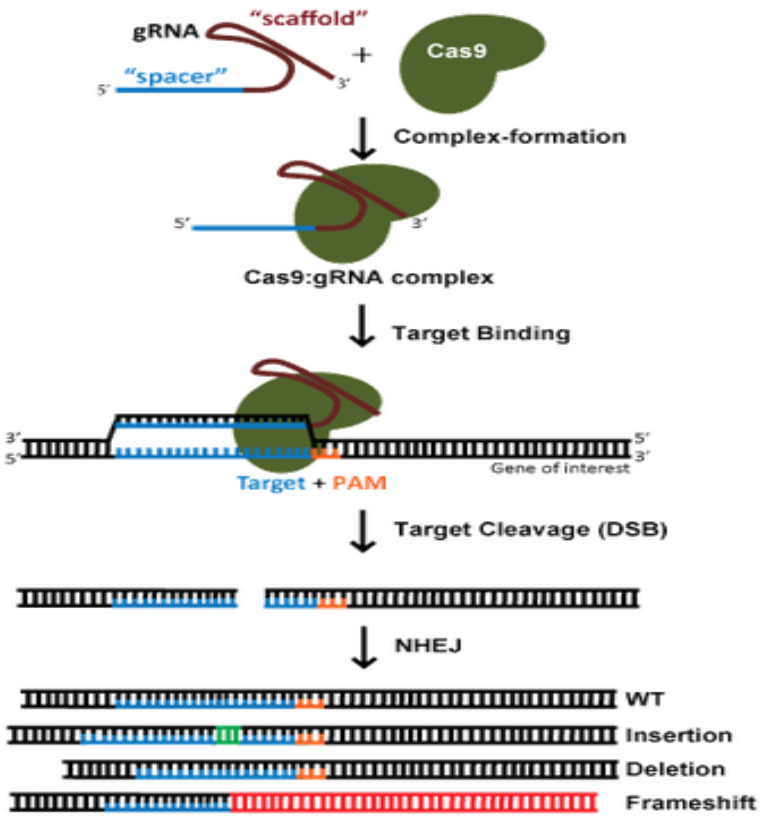



Figure 5. gRNA and Cas9 combined to form riboprotein complex that bind to the target dsDNA to create DSB(Addgene no date-a).

### 1.6.2 Non-homologous end joining of DSB.

Organism has repair mechanism for double strand break of DNA and can be either by non-homologous or homologous end joining. Non-homologous end joining (NHEJ) can cause deletion or insertion (InDels) of the nucleotide at the site of DSB. As shown in figure 6, this results in insertion, deletion or frameshift mutation(Gorbunova and Levy 1997, Addgene no date-a).

## 1.7 Objectives:

1) To investigate whether the photosynthetic efficiency of high red light intensity treated cells will able to recover in high intensity blue light.
2) To investigate whether 48h of high-intensity blue light treatment will protect the cells from photodamage after transfer to high intensity red light.
3) To produce *P.tricornutum* gene knockouts of genes which were highly regulated during light exposure using CRISPR Cas9 gene editing technology.

# 2. METHOD AND MATERIAL

## 2.1. Cultivation and experimental set up

The culture of *P.tricornutum* was incubated in 275 ml sterile tissue flask under 24 h continuous white light fluorescence of 85 $\mu molm^{-2}s^{-1}$ at 20°C in the growth room. New medium (F/2 medium) have been given to them at least once per week. During exponential phase, the algae were taken from growth room for experimental setup. Under this condition, the photosynthetic efficiency (the maximum quatum yield of charge separation PSII, ΦPSII max) of algae was 0.66. As shown in figure 7, three replicas of the flask (275ml sterile tissue flask), were exposed to dark treatment for 48 h at $15^0$C. The cultures were then exposed to the high intensity of red light (RL) 930 $\mu m^{-2}s^{-1}$ for 48 h. After this treatment, they were exposed to 200 $\mu molm^{-2}s^{-1}$ blue light (BL) for 24 h. Photosynthetic efficiency and number of cells were measured during this experiment at continuous white light (CWL), 48 hour dark treatment (48DT), red light (RL) for 0.5 h, 1 h, 2 h, 3 h, 6 h, 24 h, 30 h, 48 h and blue light (BL) for 0.5 h, 1 h,2 h, 3 h,6 h,24 h and 48 h. Light (red and blue) which the culture were exposed to, was measured approximately at the centre of culture. Since *P. tricornutum* live suspended in the seawater, the culture was placed on a shaker while exposed to light (RL 930 $\mu molm^{-2}s^{-1}$ and BL200 $\mu molm^{-2}s^{-1}$) and also during the 48 h dark treatment.

### 2.1.1. Axenization of cells

*P.tricornutum* was provided by Marianne Nymark (Ph.D.). The purpose of axenization of *P. tricornutum* is to get a viable culture that is uncontaminated with other species such as yeast and bacteria. (Guillard 2005).

1. The 0.5 ml of cells from culture was transferred to falcon tube after suspending.
2. 10 ml of f/2 medium was added to the falcon tube. To initiate bacterial growth, a mixture of f/2 (0.5ml) and peptone (1ml, g/l) was added to the tube.
3. The cultures were incubated horizontally for two week at room temperature after wrapped with aluminium foil. *P.tricornutum* culture shows cloudy or opaque when contaminated.

### 2.1.2. Harvesting cell

65 ml of culture was filtered using 0.65 µm polycarbonate filter paper. The algae were transferred from the filter paper to a 1.5 ml Eppendorf tube in a washing step. Cells were re suspended in 1 ml of f/2 medium and vortexed. The re suspended sample was centrifuged at 13000 rpm for 1 minute at 15°C and supernatant was removed and cell pellet was flash frozen in liquid nitrogen and stored in -80°C with 5mm cooled stainless steel beads (QIAGEN) until total RNA was extracted.

### 2.1.3 Photosynthetic efficiency

AquaPen is an instrument that is used to measure photosynthetic efficiency of algal and cyanobacterial suspensions. It contains blue and red LED emitters that works based on excitation lights. Blue light excitation is 455 nm which is used to measure alga's chlorophyll fluorescence in culture. Red excitation light is 620 nm that is used to measure cyanobacteria (AquaPen-AP 100 manual). 2 ml *P.tricornutum* culture from each sample exposed to DT, RL and BL was used for measurement of photosynthetic efficiency. Figure 7 shows the time interval which the samples were harvested for photosynthetic measurement. The PSII efficiency (maximum quatum yield; $F_V/F_M$) was measured at the end of a 3 min dark acclimation period using an AquaPen C (Photosystem instruments).

### 2.1.4. Isolation of total RNA

The quality and quantity of RNA is important and should be analysed before downstream application such as qPCR. Because of RNA is sensitive to degradation, a strict RNA isolation regime was followed to ensure that the quality and integrity of RNA material were good during handling. Isolation of total RNA was done using by spectrum™ plant Total RNA Kit, sigma (Fleige and Pfaffl 2006).

**Procedure**

After the sample of total RNA from -80°C freezer was mechanically disrupted for 2 minutes at 25 HZ using tissue lyser, 500 µl of the lysis/2-ME mixture was added to the sample and disrupted for the second time for 2 minutes and then incubated at 56°C for 4 minutes. The sample was centrifuged at the maximum speed of 13000 rpm for 3 minutes and then the supernatant was pipetted onto a filtration column (blue retainer ring) seated in a 2 ml collection

tube and centrifuged at a maximum speed of 13000 rpm for 1minute. 700 µl of the solution was transferred to a red binding column seated in 2 ml collection tube and centrifuged at a maximum speed of 13000 rpm for 1 minute to bind RNA before 750 µl of the binding solution was added to clarified lysate and mixed thoroughly by pipetting. After 300 µl of washing solution I was pipetted to binding column and centrifuged at maximum speed 13000 rpm for 1 minute, flow through liquid was decanted.

For on column DNase digestion, 10 µl of DNases I was mixed with 70 µl of DNase digestion buffer and 80 µl of the solution was added to the centre of the binding column then incubated for 15 minutes at room temperature. 500 µl of washing solution I was pipetted to the binding column of incubated solution at room temperature then centrifuged at a maximum speed of 13000 rpm and the flow through liquid was discarded. For the second column wash, 500 µl of washing solution II was added two times to the column and centrifuged at high maximum speed (13000 rpm) for 1 minute and 30 seconds then the flow through liquid was discarded. And finally to dry column, the column was centrifuged at maximum speed 13000 rpm for 1 minute and transferred to new collection tube 50 µl of elution buffer was added to the centre of transferred column and subjected to centrifugation for 1 minute at high-speed 13000 rpm to elute. Finally, the concentration of total RNA was measured using Nanodrop 1000 spectrophotometer and the sample was kept in -80$^0$C.

### 2.1.5. Nanodrop 1000 spectrophotometer.

The concentration and purity of total RNA was measured using Nanodrop ND-1000 spectrophotometer. 1ul of sample was loaded on instrument for measurement.

### 2.1.6. Assessment of RNA integrity (RIN).

RIN is the algorithm to shows the integrity values to RNA. This is evaluated using 28S to 18S rRNA ratio. Even though RNA is thermodynamically stable, RNase in the sample can digest RNA molecules that create shorter fragment and this can affect the downstream application of gene expression data. Agilent 2100 bioanalyzer is used to assess the degree of degradation. In microfabricated chips, the small amount of RNA sample is separated based on their weight and detected by laser-induced fluorescence detection. Its result is visualized as electropherogram that shows the fluorescence is related to the amount of RNA in the sample. RNA integrity was done using ‘Agilent RNA 6000 Nano Kit Quick Start Guide’. RNA having RIN value from 6.4-

9 is good quality for diatom. The quality of isolated RNA in this experiment ranges from 6.4-9. Table in appendix 2a shows RIN value of isolated RNA.

### 2.1.7. Complementary DNA (cDNA) synthesis

cDNA is double stranded DNA which is synthesised from single stranded RNA as a template by reverse transcriptase enzyme. RNA template must be copied to cDNA for amplification by RT-PCR. During reverse transcriptase, mRNA template must hybridize to the primer. The primer may be specific, random or oligo (dT).

**Protocol**

Messenger RNA was copied to cDNA using the QuantiTect Reverse Transcription kit (Qiagen) following the recommended protocol. Template RNA was thawed on ice but others reagents like 'gDNA Wipeout Buffer, Quantiscrip reverse Transcriptase, and Quantiscript RT Buffer, Rt Primer Mix and RNase-free water' were thawed at room temperature (15-25°C). The reaction of gDNA elimination was done on ice according to the following table 1 and incubated at $45^0$C for 2 minutes then transferred back to the ice.

Table 1. Genomic DNA elimination reaction

| Component | Volume/reaction |
|---|---|
| Template RNA | Variable (calculated) |
| gDNA Wipeout Buffer, 7x | 3µl |
| RNase free water | Variable (calculated) |
| Total volume | 21µl |

14 µl of gDNA elimination reaction was transferred to 96 well plate. Reverse transcription master mix was also prepared based on table 2 and 6 µl of it was added to the sample to make the total volume 20 µl. The samples were incubated at $45^0$C for 15 minutes and the reaction was inactivated by incubating the samples for 3 minutes at $95^0$C. Samples containing no reverse transcriptase (-RT) were also prepared using the same step following the same procedure to use as genomic DNA control. All cDNA sample were diluted 1:10 before being used as a templated for qRT-PCR analysis.

Table 2. Reverse transcription reaction.

| Component | Volume per reaction |
|---|---|
| Quantiscript Reverse Transcriptase | 1.1µl |
| RT primer Mix | 1.1µl |
| Quantiscript RT Buffer, 5x | 4.4µl |
| Total volume | 6.6µl |

### 2.1.8. Quantitative real-time PCR.

Template DNA or cDNA, can be amplified using gene specific oligonucleotides as primers by polymerase chain reaction (PCR) using a thermostable DNA polymerase. The increase in PCR product (amount of DNA) can be displayed at the end of each cycle of real-time quantitative PCR. Template DNA or cDNA is amplified exponentially. The number of DNA molecule are doubled after each cycle and the amount is quantified by using fluorescence dye that produce the fluorescent signal which is related to the amount of product. Amplification plot is produced by using fluorescent and cycle number of the reaction that shows the amount of the product during PCR run (Logan et al. 2009). Relative quantification of PCR was done by comparing the expression of treated genes with reference genes (control) that was not affected by the treatment (PtDLST, PtGK and PtRsps).

## Protocol

Table 3. Master Mix for qRT-PCR.

| Reagent | Volume per reaction |
|---|---|
| Water | 3 |
| PCR primer, 10 concentration | 0.4 µM |
| Light cycler 480 probe Master, 2x concentration | 10 |
| Total volume | 15 |

Light Cycler 480 SYBR Green I Master kit.

After master mix was prepared using Light Cycler 480 SYBR green I Master kit (Roche) based on table 3, 15µl of it was added to the wells of the PCR plate depending on the number of sample. 5ul of 1:10 diluted cDNA was added to the master mix in the well to make the total volume of the solution 20µl. For the no template controls (NTCs) nuclease free water was added instead of cDNA. Then the plate was spun down at 1500 rpm for 2 minutes to run qRT-PCR. Steps in qRT-PCR and PCR primers sequence are in appendix 2b and 2i respectively. During

the PCR, Ct value of the samples under analysis is the number of the cycle needed to cross threshold. PCR primer bind to the template (single stranded DNA) to form small region of double stranded where SYBR green I bind to the region to give fluorescence. During elongation the more several SYBR green I dye intercalate and this result in increase of fluorescence. More amount of SBR green I intercalate to the end of elongation. PCR products are imported and measured using computer program called LinRegPCR and qBase. Fluorescence data was imported to calculate the efficiency. The cycle values from the qRT-PCR analysis and PCR efficiency values for each primer pair calculated by LinRegPCR (Ramakers, 2003), were imported in to the qRT-PCR analysis software qBase (Biogazelle). The qBase program was used for calculations of expression ratios for the genes of interest and for statistical analysis. Steps in PCR program is described in appendix.

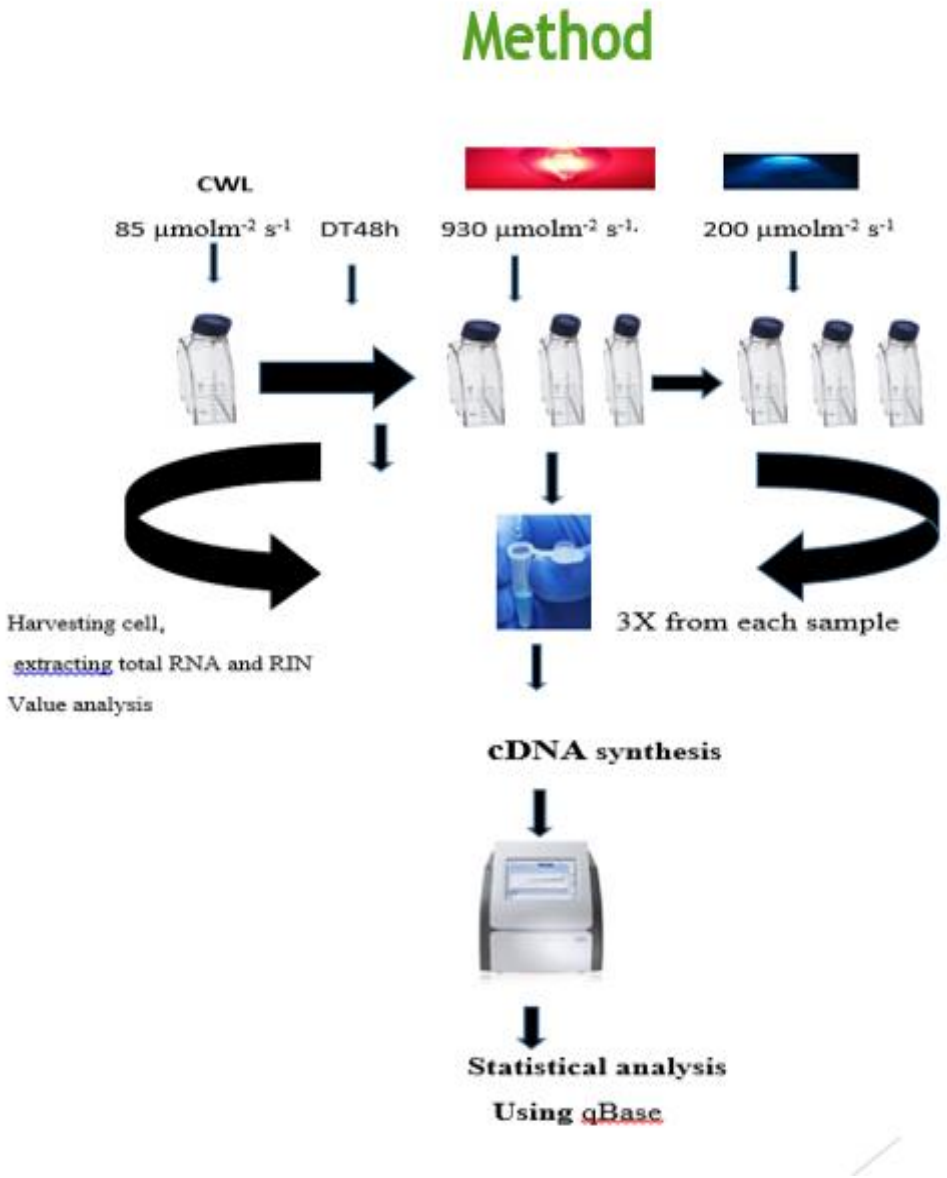



**Figure 6.** Summary of light experimental set up. Sample from continuous white light (CWL) was exposed to high red light (RL48h) and subsequent blue light (BL24h) after 48h dark treatment (DT48) and vice versa (sample from CWL was exposed to BL48h and then RL24h after 48h dark treatment). Samples for total RNA isolation was taken during red light exposure (0.5h, 24h and 48h) and subsequent blue light (0.5h and 24h). In the second experiment the order of the red and blue light exposure was switched, and samples was taken from blue light (0.5h, 24h and 48h) and subsequent red light (0.5h and 24h). After RNA concentration and RIN value was measured, cDNA was synthesized using total RNA as a template. The cDNA was used as qRT-PCR data was analysed using the qRT-PCR software qBase (Biogazelle).

## 2.2. Constructing plasmid for performing targeted gene editing using CRISPR/Cas 9 system.

The two components necessary for the CRISPR/Cas9 system to work, the Cas9 nuclease and the gRNA, can be expressed from the pKS diaCas9-sgRNA vector (Nymark et. 2006). This plasmid contain a LHCF2 promoter that can be recognised by RNA polymerase II of *P.tricornutum* to bind and initiate transcription of diaCas9 gene. The diaCas9 mRNA is translated on cytosolic ribosomes then the Cas9 enzyme translocated to the nucleus where it can cause double strand break of DNA when combined with gRNA. The plasmid also has a gene that confers resistance for ampicillin enabling selection of transformed bacteria. sgRNA contain U6 promoter that is recognised by RNA polymerases III to make RNA from sgRNA by transcription. This RNA combined with Cas9 protein to create double stranded break at specific target site. The plasmid contains two BsaI restriction sites that enable insertion of a short adaptor (target) at the 5' end of the gRNA. Insertion of an adaptor complementary to the gene that it is desired to edit, enables editing of the gene of interest. The adapter is made by annealing complementary oligos and the annealed fragment has overhang at both side. pKS diaCas9-sgRNA plasmid was treated by an restriction enzyme called BsaI-HF which cuts at target site, making it possible to insert the annealed DNA fragment. A new plasmid containing a target specific gRNA was constructed by ligated the annealed adapter in to the restriction enzyme treated vector. Construction of this plasmid was multistep process that involves cutting, annealing and DNA insert ligation reaction. The following is a protocol that show each step of the process.

### 2.2.1. Cutting reaction

2 µg of plasmid DNA was treated with 1 µL of BsaI-HF enzyme (NEB), 5 µL of 10X CutSmart buffer (NEB) and nuclease free water and in total volume of 50 µL. 50µL of mixture was incubated overnight at 37°C for enzyme to work. The next day, the enzyme was inactivated by incubating the mixture for 20 minute at 65°C on heating block to prepare for further reaction (DNA insert ligation reaction).

### 2.2.2. Annealing reaction

To anneal the forward and reverse oligo (1 µg of each), (the concentration and volumes of the forward and reverse oligo used in the reaction are in table 4), they were mixed in a tube with 5 µL of 10XT4 ligase buffer (NEB) with nuclease free water to make a total volume of 50 µL after annealed adapter was diluted 1:10. Then the mixture was mixed and incubated for 10 minutes at 85°C in a heating block followed by slow cooling to room temperature for approximately 1hour

Table 4. Sequence and concentration of Oligo DNA.

| Oligo DNA | concentration | Sequence (5'-3') | Volume |
|---|---|---|---|
| LHCF15_PAM1f | 516.73ng/µl | TCGAGGATTCGCCCCCGCGAAGCC | 1.94µl |
| LHCF15_PAM1r | 486.84ng/µl | AAACGGCTTCGCGGGGGCGAATCC | 2.05µl |
| LHCR8_PAM2f | 942.05ng/µl | TCGAGACTTGCTCCCGACTTTGTA | 1.1µl |
| LHCR8_PAM2r | 979.58ng/µl | AAACTACAAAGTCGGGAGCAAGTC | 1µl |
| LHCR8_PAM1F | 966.03ng/µl | TCGATAGCCTCATGTATATGCGCG | 1µl |
| LHCR8_PAM1R | 1261.58ng/µl | AAACCGCGCATATACATGAGGCTA | 0.7µl |
| Plasmid | 198.6ng/µl | Insert: Vector ratio 20:1 | |

### 2.2.3. Purification of restriction enzyme treated pKS diaCas9-sgRNA vector.

Purification was done as described in Wizard SV Gel and PCR Clean-Up System. 50 µl of membrane binding solution was added to 50 µl of cutting reaction that had been incubated overnight, and the solution was mixed by pipetting up and down. The mixture was transferred to a SV Minicolumn and incubated at room temperature for 1minute, followed by a 1 minute centrifugation at 14,000 rpm. The flow through was discarded. Centrifugation at 16,000Xg for 1 minute was done after SV Minicolumn was washed by using 700 µl of membrane wash solution. The flow through was discarded and Minicolumn was reinserted to repeat washing and centrifugation with 500 µl membrane washing solution. After discarding the flow through the Minicolumn was centrifuged with the lid open to let the ethanol evaporate. Then finally the Minicolumn was safely placed in a clean 1.5 ml Eppendorf tube. After 30 µl nuclease free was added to the Minicolumn and kept at room temperature for 1 minute, the minicolumn was centrifuged for 1minute to elute the purifified and cut plasmid. Then the concentration of purified plasmid was measured by Nanodrop and stored in -20$^0$C.

### 2.2.4. DNA insert Ligation reaction

Insertion to vector ration (20:1) was treated with the 1 µL of enzyme called T4 DNA ligase in the presence of T4 10X DNA ligase buffer (Fermants) (2 µL) and 13.3 µL of nuclease free water to make the total volume of the mixture 20.0 µL. In this mixture the insert was ligated in to the vector by the T4 DNA ligase to create a new plasmid as shown in appendix (pKS diaCas9_sgRNA_LHCF15 plasmid). E.coli was transformed with ligation mix to enable amplification of the newly constructed plasmid.

### 2.2.5. Transforming *E.coli* by heat-shock

Competent DH5α-*E.coli* cells was taken from the -80$^0$C freezer and thawed on ice. 3 µl of the ligation reaction was added to 100 µl of competent cells mixed by flicking the tube and incubated for 30 minutes on ice. The cells were heat shocked by incubating in a water bath for 45 seconds at 42$^{\circ}$C, and then immediately placed on ice for 2 minutes. 1 ml of pre-warmed LB medium was added to the mixture before incubation for 1h at 37$^{\circ}$C with shaking at 220 rpm. Then 100 µl of the transformation mixture was spread on LB agar plates with 100 µg/ml ampicillin and then incubated overnight at 37$^{\circ}$C.

### 2.2.6. Colony Analysis by PCR and Overnight culture

Colony PCR analysis was used to figure out if the bacterial colonies on the agar plates contained a plasmid where the correct adapter had been inserted. PCR tubes containing 5 µl of nuclease free water and falcon round-bottom tubes having 3 ml of LB medium with 100 µl/ml ampicillin, were prepared. A few colonies where picked from each agar plate using sterilization pipette tips. The pipette tip with the bacterial colony was fist dipped and shaken in the PCR tubes containing nuclease free water, before being transferred to the falcon tubes (overnight culture). For being able to find out if the bacterial colony contained a plasmid with the desired adapter, a PCR mixture were prepared that included a forward primer specific for the adapter and a reverse primer specific for the vector. For PCR amplification a master mix was prepared by mixing 2.5 µl of 10 X optimized DYNAZyme buffer (thermoScientific), 2 µl of dNTP mixture, 1 µl of 10 µM forward (adapter specific) and reverse primer KST3, 0.125 µl of DYNAzyme II DNA polymerase (ThermoScientific) and 13.375 µl nuclease free water per reaction. Then 20 µl of master mix was added to the PCR tubes containing 5 µl of template (5 µl of nuclease free water containing the bacterial colony) for each sample to make a final volume 25 µl. The colony

PCR program was called DYNAZ. This program is used for amplification. It has some cycle through which the target is amplified. The step is described in appendix 2 e. The temperature cycle causes lysate the cells at 94 °C for 10 minutes and then melt double stranded DNA. Forward and reverse primer is annealed to melted DNA which is followed by binding of the polymerase. Amplified amplicon by polymerase under goes different steps repeating the previous step. Falcon tubes containing 3 ml LB medium (100 µg/ml ampicillin) and bacterial confirmed by PCR and gel electrophoresis (see section 2.2.8) to contain plasmids with the correct adapter were incubated overnight at 37 °C with shaking at 220rpm.

### 2.2.7. Gel electrophoresis.

Gel electrophoresis is used to separate DNA molecules of different size by applying electric field to an agarose. DNA molecules are negatively charged, and start to migrate to the positive pole through the pores in the gel by the influence of voltage. Separation of DNA fragments is based on their length. The DNA molecule that has smaller molecular size moves faster than the larger through the gel. The concentration of agarose and form of DNA molecule affect the movement of DNA through the gel (Johnson and Grossman 1977, Bjornsti and Megonigal 1999).

Procedure

1. 1g of agarose was mixed with 100ml of 1 x TAE buffer in a flask.
2. After the mixture was heated using microwave at 900W for a while, then all the agarose powder was dissolved.
3. The mixture was then kept at room temperature for a while before 4 µl of Gel Red™ Nucleic Acid Gel Stain was added while shaking slowly.
4. The mixture was poured in to a gel tray that containing a comb and then kept at room temperature to let it solidify.
5. After the gel became solidified, the comb was removed, and it was placed in to gel tank almost filled with 1 x TAE buffer. Pre-prepared sample (PCR product and loading buffer) was loaded to the wells in the gel. GeneRuler 1kb plus (ThermoScientifc) was loaded in the first well for size determination of the PCR product (DNA fragments)
6. After the gel had been run for 45 minutes 90 V, the gel result was analysed on Bio-Rad Gel Doc 2000

### 2.2.8. Plasmid DNA isolation.

Bacterial cultures were harvested by centrifugation at 8,000 rpm for 2 minutes and plasmid was extracted using QIAprep Spin Miniprep Kit (Qiagen) that uses silica membrane technology. The method is on the bases of alkaline lysis of *E.coli* cells in the presence of salt, and plasmid DNA binds to the silica membrane in the column. 250 µl buffer P1 (RNase added) was added to the pelleted cells to re suspend by vortexing. 250 µl buffer P2 was added to the re-suspended solution and mixed by inverting the tube 4-6 times so that the solution became viscous and slightly clear. 350 µl buffer N3 was added to the solution and immediately mixed by inverting the tube 4-6 time so that the solution became cloudy and centrifuged at 13,000 for 10minutes. By decanting, the supernatant was transferred to QIAprep spin column and centrifuged for 1 minute. Centrifugation was done for 1 minute after QIAprep spin column was washed by addition of 750 µl buffer PE. The flow through was discarded and centrifugation was done for 1 minute more. QIAprep column was transferred to a new 1.5 ml Eppendorf tube, and finally the 50 µl nuclease free water was added. After 1 minute incubation in room temperature, the plasmid was eluted by centrifugation for 1 minute. Then the concentration of plasmid DNA was measured using Nanodrop 1000.

### 2.2.9.1. Sequencing reaction.

To confirm that the pKS diaCas9 sgRNA plasmids contained the correct adapter, sequencing was performed. The sequencing reaction was done by mixing 1 µl of Big dye 3.1 (ThermoScientific), 3 µl of sequencing buffer, 1 µl of reverse primer KST (3.2 µM) and 14 µl nuclease free water to prepare master mix. Then 19 µl of master mix was added to 1 µl of plasmid and the final volume became 20 µl. After 20 µl of each sample run with PCR program called RBD31 program. The sequencing reactions were sent to the DNA sequencing Core Facility in Tromsø, for Sanger sequencing. RBD31 program is used to have sequencing reaction for sequencing the PCR product. It contain several step which through which the PCR product is ready for sequencing. Each step contain specific temperature that result in activation of enzyme, melting of double stranded DNA, annealing and elongation as described in appendix 2 f.

## 2.3. Preparation of F/2 agar medium plates for growing of diatom cells.

10 g of agar was mixed with 500 ml of sterile filtered sea water and 490 ml of distilled water in a flask containing a magnet and then autoclaved. After autoclaving was done, the solution

was cooled down to < 55 °C, and F/2 agar plates were prepared by adding 1 ml $NaNO_3$ (75g/L stock solution) , 1 ml $NaH_2PO_4$ x $H_2O$ ( 5 g/L stock solution), 1 ml of trace metals and 0.5 ml of vitamins according to Guillard (1975). 1 ml of 100mg/ml of zeocin (InvivoGen) was added to the F/2 agar solution for preparation of plates that would be used for selection of transformed diatoms. Finally, the F/2 agar solution was poured in to petri dishes and then kept at room temperature until it solidified. After the plates had solidified, they were stored in cold room at 4°C.

### 2.3. 1. Biolistic transformation of *P. tricornutum.*

Transformation of the diatoms cells can be achieved using biolistic bombardment. The biolistic PD-1000/He particle delivery system (Bio-Rad, Hercules, CA, USA) uses high pressure helium gas released by a rupture disk to accelerate DNA coated micro carriers toward cells (Figure 7). DNA coated microcarriers (prepared for transforming the cell with DNA of interest) are loaded on to macrocarriers, but macrocarriers are stopped after a short distance by the stopping screen, whereas microcarriers will continue towards the target cells with high speed penetrating the cell wells of the target organism. The helium pressure, amount of vacuum and distance can influence the speed of microcarriers. Microcarriers that is accelerated to transform the call by penetrating the cell wall can also cause damage to the cell. The DNA is separated from microcarriers and transferred to the interior of the cell and integrated to the nuclear genome(Taylor and Fauquet 2002).

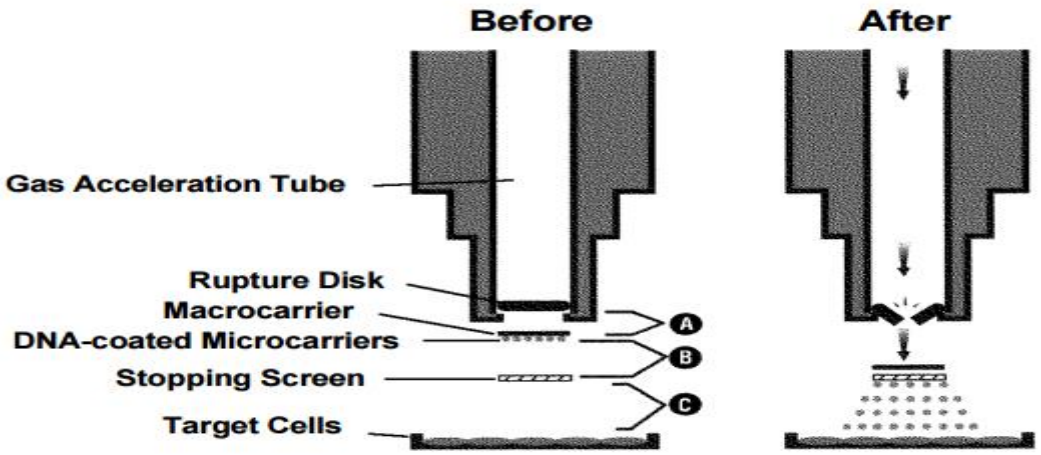



**Figure 7.** Biolistic bombardment. After pressure helium is built, the rupture disk release the gas that result in accelerating microcarriers to ward target cells and causes transformation (Kirupamurthy 2014).

### 2.3.2. Preparing diatom for transformation.

After the concentration of the cell was counted by using Bürker counting chamber, the cells were harvested from 175 ml of *P.tricornutum* culture by centrifugation for 10 minutes at 4500 g. The supernatant was discarded and the pellet was re suspended by adding 4 ml of F/2 medium to achieve a concentration of '$1x10^8$ cells/ml. 500 µl of cells was plated on F/2 agar plates (containing 50% sea water) that did not have antibiotics, and the plate was allowed to dry for a while before being transferred to growth room. The cells were kept for 24 h in the growth room before transformation by biolistic bombardment.

### 2.3.3. Coating microcarriers with DNA.

Tube containing M17 tungsten particles (BioRad) was continuously vortexed for 5 minutes, before adding plasmid DNA, CaCl2, and spermidine (Table 5). The mixture was continuously vortexing for 3 more minutes. After the mixture was kept at room temperature for 1 minute, microcarriers were pelleted by spinning for 2 seconds and the liquid was discarded. 140 µl of 70% ethanol was added to the mixture and removed. Then this step was repeated using 140 µl ethanol. Finally 50 µl of 100% ethanol was added and the pellet was re suspended and vortexed for 3 seconds at low speed. DNA-coated tungsten particles were loaded to macrocarrier and kept at room temperature to dry. All components of the PDS-1000/He instrument, macrocarriers and stopping screens were cleaned with 70% ethanol before performing the biolistic bombardment. After turning on the PDS-1000/He instrument, helium pressure was adjusted to 200 psi over the rupture disk pressure. For biolistic transformation of *P.tricornutum*

1550 psi rupture discs were used. After the target shelf had been placed at appropriate distance (6 cm), the plate with *P.tricornutum* cells were placed on it. The source of vacuum was turned on until it reached a vacuum of 26.5 mm Hg and the sample was bombarded with the microcarriers coated with plasmid DNA. Bombarded cells were taken to growth room for a 24 h to grow. *P.tricornutum* cells were co-transfected with pAF6 plasmid to enable selection of transformants on plates with zeocin. PAF6 plasmid contain gene that makes *P.tricornutum* cells resistant to zeocin.

Table 5. Concentration of reagent and plasmid used to coat tungsten (microcarriers)

| Plasmid types | 2.5M CaCl2 | pAF6 plasmid | 0.1M Spermidine |
|---|---|---|---|
| pKS diaCas9_sgRNA LHCF15 PAM1. 2.5 µg | 50 µl | 2.2 µl | 20 µl |
| pKS diaCas9_sgRNA LHCR8 PAM1. 2.5 µg | 50 µl | 2.2 µl | 20 µl |
| pKS diaCas9_sgRNA LHCR8 PAM2. 2.5 µg | 50 µl | 2.2 µl | 20 µl |

### 2.3.4. Re plating *P.tricornutum* on zeocin medium.

F/2 agar plates with zeocin were placed in the sterile bench for 45 minutes before transformed algae was transferred to it. 1 ml of F/2 was added to the plate having transformed plate and the cells were detached. Then the cells were collected to one side by tilting the plate and then the cell suspension was transferred to two zeocin plates. Additional 0.5 ml of F/2 medium was added to the plate with bombarded cells and then transferred to a third zeocin plate. All the plate was kept in sterile bench until they were dry, they were then sealed and taken to growth room.

## 2.4. Harvesting the cell (Picking colony)

Colonies of transformed *P.tricornutum* cells were visible within 2 weeks after bombardment but picked after three weeks and transferred to liquid media in a 48 well plate.

1. 50ml f/2 medium was mixed with 25 µl 100mg/ml zeocin
2. Each well was filled with 0.5ml of f/2 medium with zeocin.

3. Colony of transformed *P.tricornutum* was picked then transferred to the well containing liquid medium (f/2 medium + Zeocin)
4. The plate was sealed to prevent evaporation and then placed in growth room for almost a week.

### 2.4.1. Genomic DNA extraction

100 µl of transformed *P.tricornutum* was transferred to Eppendorf tube after scraping the bottom of well plate to detach the cells. The supernatant was discarded after spinning for 1 minute and 20 µl of lysation buffer was added to break the membrane. It contain either salt or detergent to break open the membrane. Salts are used normalise osmolarity and acidity the released material. Detergents are used to identify membrane from membrane proteins. After vortexing for 30 seconds, the sample was kept on ice for 15minute and then heated for 10 minute at 85 °C. Then kept at room temperature and diluted 1:5 with nuclease free water. Finally centrifuged for a while to pellet the cell debris and stored at -20 °C

### 2.4.2. Amplification of target region (PCR reaction).

To amplify genomic DNA of *P. tricornutum* around the target site (the area where the mutation was supposed to be created), two special set of primer pairs were designed and ordered for each of the two genes of interest. The first set of primer (LHCF15_P1F, LHCF15_P1R, LHCR8_P2F, and LHCR8_P2R) targets about 700 bp of either the LHCF15 or LHCR8 for amplification. This PCR product was diluted and used as template for High Resolution Melting (HRM) analysis. The second set of primers (LHCF15HRM_P1F, LHCF15HRM_P1R, LHCR8HRM_P2F and LHCRR8HRM_P2R) amplify a smaller PCR product (100bp) with in the first amplicon which was closer to the region where the mutation is expected to be created (appendix 2 g)

**Protocol**

Master Mix was prepared as follows.

After the first set of primer was diluted to 10 µM, 1 µl from both primers (forward and reverse), 2.5 µl of 10X optimized DYNAzyme buffers (ThermoScientific), 2 µl of dNTPs, 0.2 µl of DYNAzyme II DNA polymerase (ThermoScientific) and 15.8 µl of nuclease free water was mixed per reaction to prepare master mix. 22.5 µl of master mix was added to 2.5 µl of template

(lysated *P.tricornutum* cells) in PCR strips. The mixture was amplified by PCR program called DYNAZ (described in section 2.2.6.)

### 2.4.3. High Resolution Melting (HRM) Analysis

HRM is a technique for analysing genetic variation in DNA sequence. It can detect genetic variation without sequencing. The region of interest is amplified using standard PCR techniques and specialized dye that is highly fluorescent when it bind to double stranded DNA and less fluorescent when not binding. Then the temperature is increased to denature the PCR product, this reduce the fluorescent that leads to production of melting profile. To detect the mutation in target gene of *P. tricornutum* (LHCF15 and LHCR8 gene) by HRM, primers designed that flank the target site (where mutation is supposed to happen). The primers are designed for amplification of a 100 bp PCR product. The amplicon is then analysed using HRM software.

#### 2.4.3.1. Preparing sample for HRM

1.4 µl of the template (the 700 bp LHCF15 or LHCR8 PCR product) was diluted in 198.6 µl nuclease free water and using serial dilution, this mixture was diluted to 1:10, 1:100, 1:10000, and 1:100000 and 1:1000000. 15 µl of master mix in table 6 was added to the well of the plate and 5 µl of 1:1000000 dilution of the template (wild type) was also added to the well to make the final volume 20 µl. Then the plate was spun down and amplification was done using the Light Cycler 96 instrument (Roche).

Table 6. Master Mix for HRM.

| Component | Volume | 24 (Number of samples), reaction + 10% |
|---|---|---|
| H2O | 0.6 µl | 15.84 µl |
| HRM master mix | 10 µl | 264 µl |
| MgCl 25 mM | 2.4 µl | 63.36 µl |
| Primer | 2 µl | 52.8 µl |
| Total volume | 15 µl | 396 µl |

The working solution was prepared by mixing 4 µl from both primers (forward and reverse) with concentration of 100 µM, was mixed with 192 µl of nuclease free water to make 200 µl of mixture and master mix was also prepared based on the table 6. Then the well of the plate was filled with 15 µl of master mix and then 5 µl of the template from dilution of 1:100000 was

added to the first three rows of the plate. Well number 1-6 on the fourth row was added with 5 µl of the template from the wild type as a control and 5 µl nuclease free water was added in well number 7-12. Finally, the plate was sealed, spun down for 2 minutes at 1500 g, and then amplification and following HRMA analyses were done using the LightCycler 96 instrument and software (Roche).

### 2.4.4. Purification from Gel.

PCR products used for template in the HRM reaction, which had positive HRM result (suspected to contain mutations) was prepared for gel electrophoresis. Using Wizard SV Gel and PCR Clean-Up system the gel band which shows 700 bp was cut out from the gel and placed in 1.5 ml Eppendorf tube. Equal amount of membrane binding solution was added to the tube and then the gel slice was dissolved at $65^0$C. Purification of cDNA was done as described in section 2.2.3

### 2.4.5. TOPO cloning reaction.

**TOPO** cloning is technique in which DNA fragment is cloned in to the vectors. This reaction does not need DNA ligase. Taq polymerase adds single adenosine (A) to 3' end of amplicon. TOPO cloning reaction was done using TOPO® TA Cloning® Kit. After the reagents in table 7 were mixed gently and incubated at room temperature for 5 minute to prepare TOPO cloning reaction, the reaction was transferred to ice. 2 µl of TOPO cloning was used for transformation of bacterial cells. The transformation, preparation of overnight cultures, and isolation of TOPO vectors with LHCF15 or LHCR8 inserts were performed as described in section 2.2.5, 2.2.6 and 2.2.8. Sequencing reaction of the insert in the TOPO vector were performed as described in section 2.2.9.1 but reverse primer (M13) is used to for sequencing and preparation of master mix is described in appendix 2 c. The sequence of M13 primer is in appendix 2 d.

Table 7. TOPO cloning reaction. Amount of reagent used for TOPO cloning reaction.

| Reagent | Volume |
|---|---|
| Fresh PCR product | 0.5 µl |
| Salt solution | 1 µl |
| water | 3.5 µl |
| TOPO vector | 1 µl |
| Final volume | 6 µl |

# 3. RESULT

## 3.1 Light response

All *P.tricornutum* cell cultures were maintained in continuous white light (CWL 85 µmol $m^{-2}s^{-1}$) for several weeks, and dark treated for 48h before the start of the light quality experiment.

### 3.1.1. Photosynthetic efficiency of *P.tricornutum* that was exposed to high red light and subsequently to blue light.

Dark treated *P.tricornutum* cells were exposed to high intensity red light for 48 h and subsequently to high intensity blue light for 24 h. The sample was taken at each at each time interval indicated in figure 9 and measured using Aquapen-AP 100.

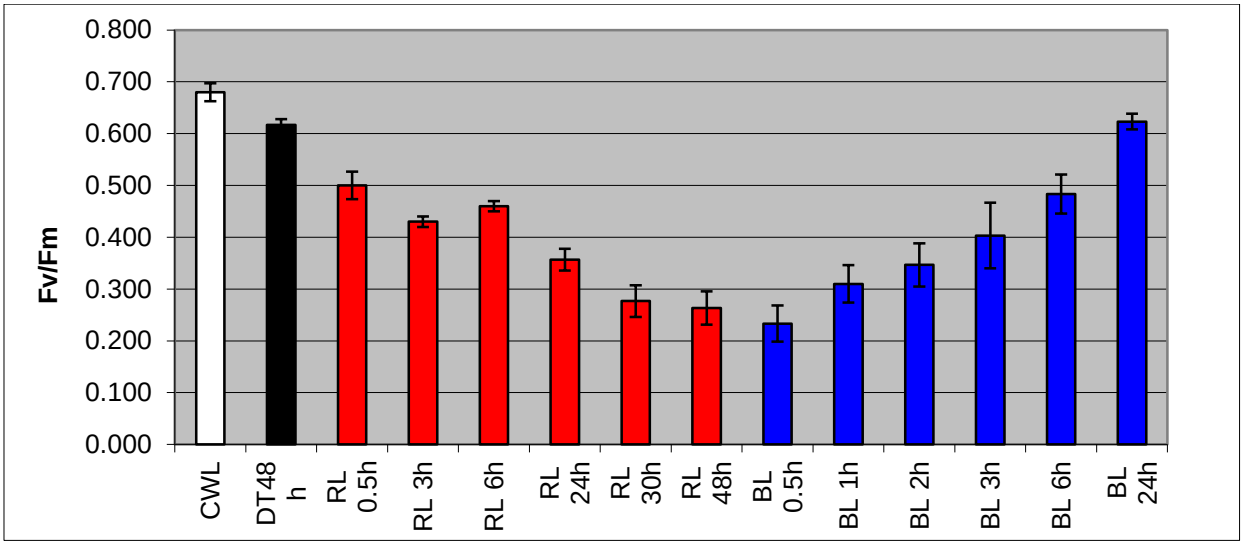



**Figure 8.** Cell **c**ulture from continuous white light (CWL) was exposed to high red light and subsequently to blue light after 48h dark treatment. Photosynthetic efficiency was measured by taking sample from CWL, DT48 h, red light (0.5 h, 3 h, 6 h, 24 h, 30 h, 48 h) and blue light (0.5 h, 1 h, 2 h, 3 h, 6 h and 24 h).

In Figure 9, photosynthetic efficiency of white light was 0.68 that is used as control. Photosynthetic efficiency in 48 h dark treated alga (DT48 h) is almost comparable to those exposed to white light (CWL). During the sample was exposed to high red light after 48 h dark, the Photosynthetic efficiency was decreased from 0.62 (DT48 h)-0.25 (RL48 h) but 6 h exposure showed a little bit instability. As the period of exposure prolonged, the photosynthetic efficiency decreased for 48 h. After 48 h high red light exposure, the cell was exposed to high

blue light for 24 h. During the initial 0.5 h blue light exposure, photosynthetic efficiency was decreased relatively but as the exposure period prolonged for 24 h, p. efficiency shows 100% recovered.

### 3.1.2 Quantitative real time-PCR

*Phaeodactylum tricornutum* from CWL was exposed to highlight (red and then blue) after 48 h dark treatment. Differential gene expression as a response to light intensity can be measured by performing qRT-PCR analysis of material harvested during the light exposure experiment.

The result from qRT-PCR shows differential expression of the genes during exposure to high light (high red and blue). The data produced from qRT-PCR was used for statistical analysis and the result of this analysis is presented in Figure 10. The bar chart visualizes the differential expression of the genes in response to high light (red and blue).

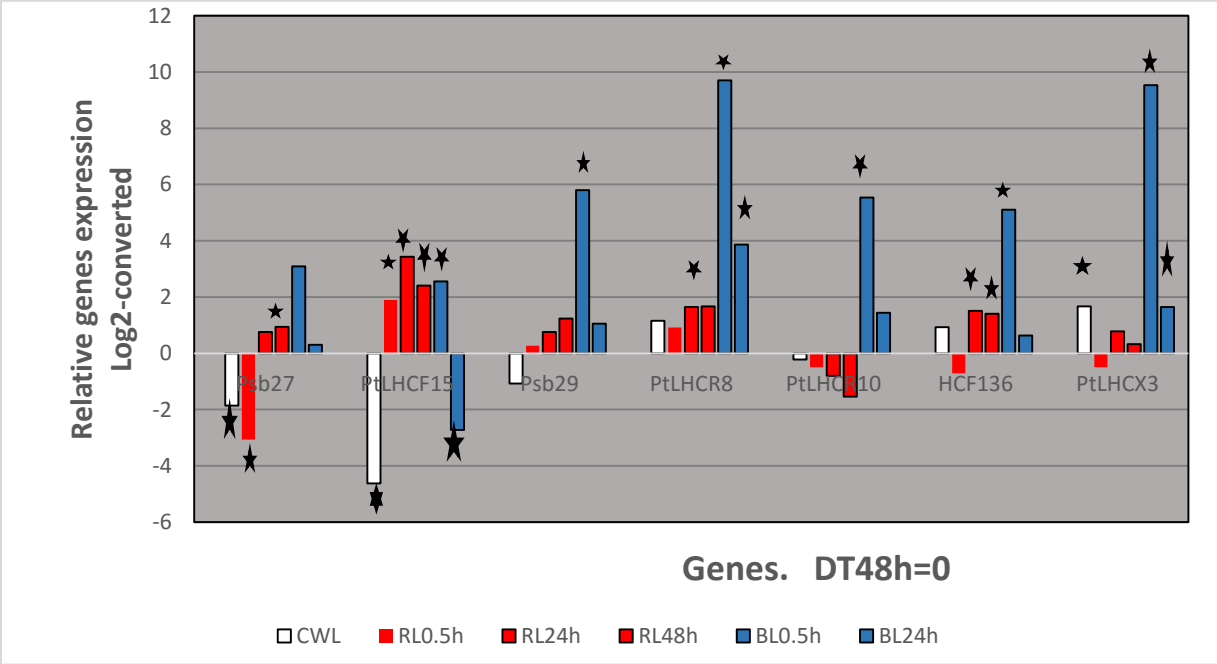



**Figure 9.** Relative gene expression studies in time series of light exposure, using qRT-PCR. The expression of each gene was compared to 48h dark treatment (DT48 h). The sample from continuous white light (CWL) was exposed to red (0.5 h, 24 h and 48 h) and then blue (0.5 h and 24 h) light after 48h dark treatment. The expression of each gene was converted to the log2 value for visualization on graph bar. Differential expression of genes was normalized by using the three reference genes PtDLST, PtGK, and PtRsps. The colour (red and blue) indicates type of light that cells was exposed. Astra indicates a gene that is significantly regulated.

### 3.1.3. Photosystem II (PSII) repair genes.

PSII is repaired by degradation and production of new D1 protein. The Psb27, Psb29 and HFC136 are all encoding proteins assumed to be involved in PSII repair. They are induced in BL after 48h RL exposure that caused photodamage. PSII repair genes are BL dependent and they have similar protein gene expression profile. The proteins encoded from these genes involved in repair of damaged protein caused by RL.

**Psb27 gene.**

Psb27 gene was significantly downregulated in CWL and RL0.5h but as exposure period was prolonged for RL24 and RL48, the expression was slightly upregulated when compared with control (DT48h). Blue light was turned on after RL48h, which causes the expression significantly upregulated in the initial exposure time (0.5h) but as the exposure period prolonged, the gene was slightly expressed when compared with control.

**Psb29 gene**

Psb29 gene was upregulated in all exposure except CWL when compared to DT48h. The gene was downregulated in continuous white light but slightly upregulated when exposed to red light for 0.5h and increased in height as the period of exposure was prolonged to 24h and 48h but none of them are significant. After 48h red light exposure, the gene was significantly upregulated increasing by 4.7 fold in blue light for 0.5h (BL0.5h) but decreased by 5.5 fold as exposure period prolonged to 24h blue light.

**HCF136 gene**

HCF136 gene was upregulated in all exposure except RL0.5h when compared to DT48h and shows significant in RL24h, RL48h, and BL0.5h. In CWL, the gene was slightly upregulated but downregulated during exposure to red light for 0.5h when compared to DT48h. As the exposure period prolonged to 24h and 48h red light, the gene was upregulated significantly and keeps the same height in both periods of exposure. After RL48h, the sample was exposed to blue light for 0.5h and the gene shows upregulated significantly increasing by 3.6 fold but decreased by 8 fold as the exposure period prolonged for 24h.

### 3.1.4. LHC genes.

Light harvesting chlorophyll a/b-binding proteins are encoded by LHC super-gene family. LHC PLHCR8, PtLHCR10, and PtLHCX3 are LHC gene that encode proteins that assumed to be involved in photoprotection. They also have similar gene expression profile. Their expression was down regulated in red light and up regulated in blue light.

**PLHCR8 gene**

PtLHCR8 gene was upregulated significantly in all exposure except CWL and RL0.5h when compared to DT48h. During exposure to CWL, the gene was slightly upregulated and didn't show any difference when exposed to red light for 0.5h but as the period of exposure prolonged to 24h and 48h, the gene shows upregulated significantly and didn't show the difference in both exposure period (RL24 and RL48h). After RL48h exposure, the gene was upregulated significantly increasing by 5.8 fold in blue light for 0.5h (BL0, 5h) but relatively decreased by 2.5 fold as period prolonged to 24h.

**PtLHCR10 gene**

PtLHCR10 gene was down-regulated in CWL and red light but upregulated in blue light when compared to DT48h and none of them are significant except BL0.5h. In CWL, this gene was slightly downregulated and increased a little bit when exposed to red light (0.5, 24 and 48h) but none of them are significant. After RL48h, the gene was exposed to blue light for 0.5h and shows upregulated significantly but as the exposure period increased to 24h, expression of a gene was decreased by 3.8 fold.

**PtLHCX3 gene**

PtLHCX3 gene show upregulated in all exposure except for RL0.5h that is downregulated when compared to DT48h. In CWL, the gene shows upregulated significantly but when exposed to 0.5h red light, it shows slightly downregulated when compared to DT48h. As the exposure period prolonged to RL24h, it shows slightly upregulated but decrease by 2.4 fold when period reached 48h (DT48h). After RL48h exposure, the sample was exposed to blue light for 0.5h and gene was upregulated significantly increasing its height by 30 fold but decreased by 6 fold as exposure period reached to 24h (BL24h).

**PtLHCF15 gene**

PtLHCF15 gene was significantly regulated in all exposure when compared to DT48h. The exposure to CWL caused the gene to significantly downregulated but upregulated when exposed to red light for 0.5h (RL0.5h) and increased by 1.8 fold for 24h exposure but as a period of exposure prolonged the gene shows slightly decreased by 1.4 fold. After RL48h exposure, the gene was upregulated keeping the height almost the same as RL48h in 0.5h blue light but as the period of exposure reached 24h blue light the gene show significantly downregulated

## 3.2. Photosynthetic efficiency of *P.tricornutum* that was exposed to blue light and subsequently to red light.

*P.tricornutum* from continuous white light (CWL) was exposed to high blue light (200 µmolm$^{-2}$s$^{-1}$) for 48 h and subsequently exposed to red light (920 µmolm$^{-2}$s$^{-1}$) for 24h after 48 h dark treatment. During this experiment total, RNA was extracted at the time interval indicated in figure 12 and photosynthetic efficiency was also measured at the time interval indicated in figure 11.

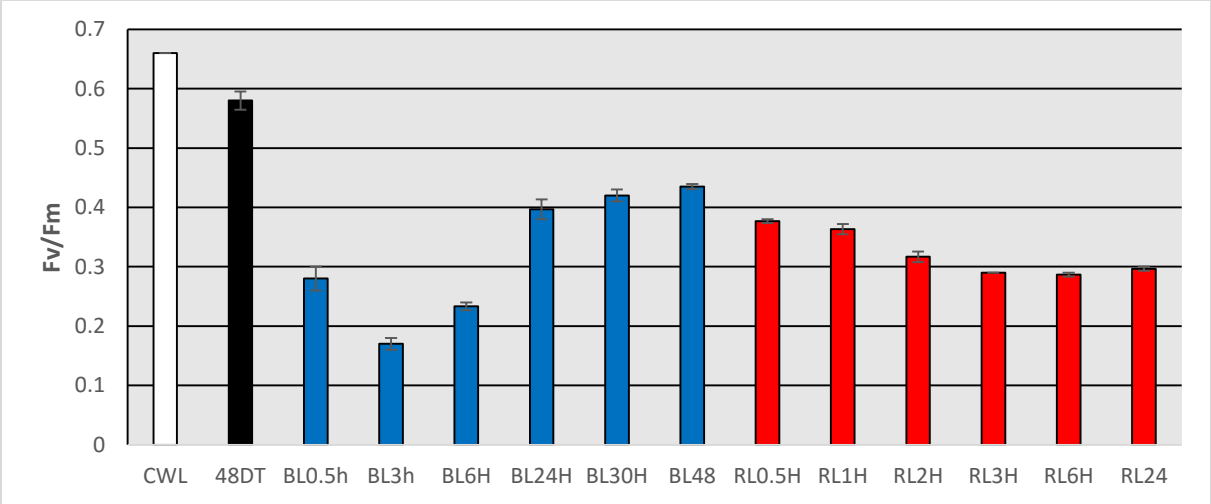



**Figure 10.** Photosynthetic efficiency results from AquePen photometer. Culture from continuous white light was exposed to high blue and then red light intensity respectively. 2ml sample was harvested at time interval indicated in the figure and are treated with dark for 3 minutes then the photosynthetic efficiency was measured by AquePen machine.

Photosynthetic efficiency from continuous white light (CWL) is used as a reference to compare with treatment group (high blue and red light). Photosynthetic efficiency in 48h dark treatment is almost in the same range with continuous white light. After 48h dark treatment (48DT) the

cell was exposed to high blue light and the photosynthetic efficiency was decreased from 0.5h-3h exposure period but as the exposure period increased to 48h, the efficiency also increased. After BL48h exposure, the cell was exposed to high red light that caused the efficiency to decrease as the exposure period increased to 24h (RL24h).

### 3.2.1 qRT-PCR result.

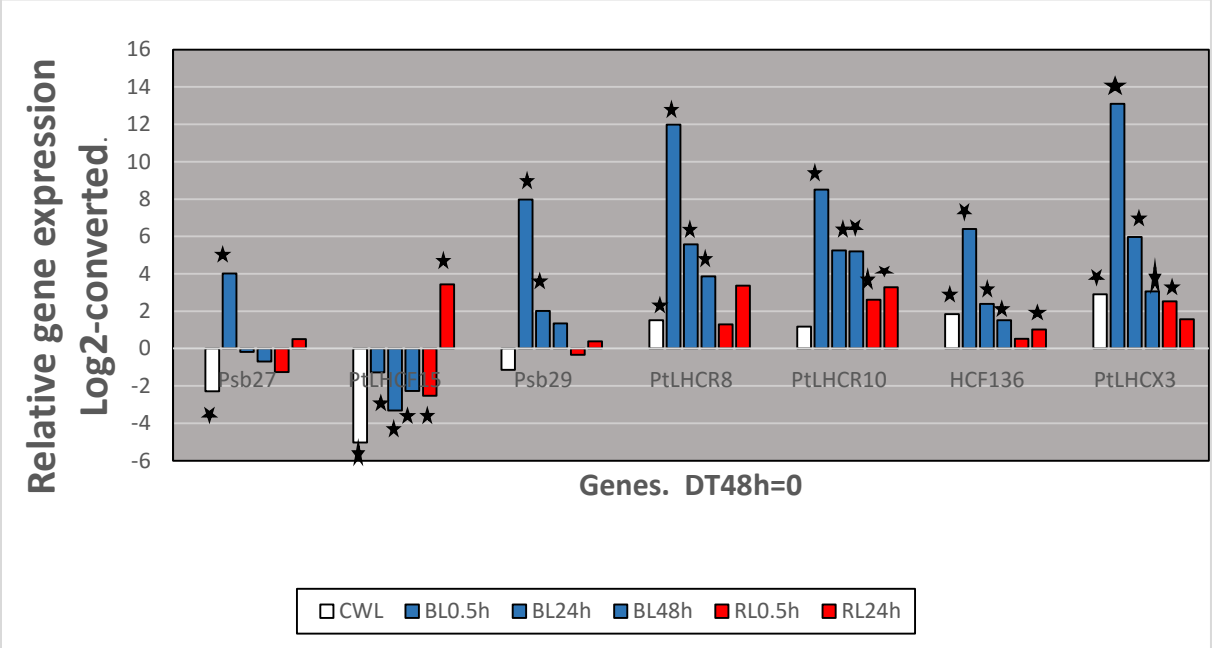



Figure 11. *P.tricornutum* from continuous white light was exposed to high blue light (BL48) and subsequently exposed to high red light intensity (RL24) after 48h dark treatment. The expression of each gene was converted to the log2 value for visualization on graph bar. The sample for RNA isolation was taken from dark 48 (DT48), blue (0.5h, 24h and 48h), red (0.5h and 24h) and continuous white light (CWL). Differential expression of genes was normalized by three gene PtDLST, PtGK, and PtRsps. The colour (blue and red) indicates type of light that cells was exposed. Astra indicates a gene that is significantly regulated

### 3.2.2 PSII repair

**Psb27 gene.**

In figure 12, Psb27 gene was regulated but not significantly in all exposure when compared with 48DT. Even though this gene was downregulated in continuous white light, as they exposed to 0.5h blue light, the gene was upregulated but the expression was a little bit downregulated as the exposure period prolonged to 24h and 48h blue light. As the cell was

exposed to the red light of 0.5h, the gene expression was downregulated but a little bit upregulated in 24h exposure.

**Psb27 gene**

The expression of this gene shows significantly upregulated only in 0.5h and 24h of blue light exposure. In the continuous white light, the gene expression was decreased when compared to 48h dark treatment but significantly upregulated during exposure of blue light 0.5h and 24h. Slightly decreased in 48h blue light exposure relatively. When the cell exposed to red 0.5h from 48h blue light, the gene expression was slightly decreased but increased during 24h red light when compared with 48h dark treatment and none of them are significant.

**HCF136 gene**

This gene expression was upregulated in all treatment but only significant in blue light exposure. The expression of this gene was slightly upregulated in continuous white light and significantly upregulated as the gene exposed to 0.5h blue light but as the exposure period prolonged the expression became decreased in 24h and 48h blue light. After 48h red light, the gene was exposed to red light that caused the expression of the gene decreased almost closer to 48h dark treatment but as the period of exposure increased the expression is slightly increased

### 3.2.3 Protein encoding genes that involved in photoprotection

**PtLHCr8 gene**

The expression of this gene was significantly upregulated in all treatment except in continuous white light and 0.5h red light. Slightly upregulated expression of this gene in continuous white light became significantly upregulated as the cell exposed to 0.5h blue light but start to decrease as the exposure period prolonged to 24h and 48h. After 48h blue light exposure, the cell was exposed to 0.5h red light which caused to decrease the expression of this gene relatively but as the exposure period prolonged the expression of the gene was significantly upregulated.

**PtLHCR10 gene**

The expression of this gene was significantly upregulated in all treatment except in continuous white light and 0.5h red light. Slightly upregulated gene expression in continuous white light was significantly upregulated as the cell exposed to blue light 0.5h but the expression decreased as the period prolonged to 24h and kept the same in the 48h blue light. The exposure to red light

for 0.5h caused the expression of the gene to decrease when compared with 48h blue light but significantly increased relatively.

**PtLHCX3 gene**

The expression of this gene was significantly upregulated in all treatment except in 24h red light exposure. Upregulated gene expression in continuous white light was increased during the exposure to the blue light for 0.5h but start to decrease as the exposure period prolonged to 24h and 48h, and kept decreasing even during the exposure to red light 0.5h and 24h.

**LHCF15 gene.**

This gene was significantly regulated in all exposure. In all exposure except 24h red light, the gene was significantly downregulated when compared with 48h dark treatment. As the exposure period prolonged in blue light, the gene expression is various at different time interval. In the continuous white light, the gene was downregulated but as the cell exposed to blue light for 0.5 the gene expression starts to increase relatively. In prolonged exposure period of blue light 24h, the expression of the start to decrease but slightly decrease in 48h exposure. In red light of 0.5h the expression of the is downregulated as the in 48h blue light but highly upregulated as the exposure period reaches the 24h red light.

## 3.3 Gene editing using CRISPR Cas9 technology.

Pks diaCas9_sgRNA is a plasmid which has sgRNA, U6 promoter, diaCas9, and ampicillin resistance, the origin of replication and insertion site as shown in appendix (pKS diaCas9_sgRNA plasmid). It is designed for genome editing by CRISPR Cas9 system. Taking this as an advantage the forward and reverse oligonucleotide DNA (Table 4) that has 24 base sequence which is complementary with the gene of interest was annealed and then ligated to the insertion site of this plasmid, targeting lhcf15 and the lhcr8 gene that was highly regulated in light exposure. During the light treatment of *Phaeodactylum tricornutum*, lhcf15 gene shows significantly regulated in all treatment and lhcr8 was also shows significantly upregulated during blue light exposure. E coli was transformed with constructed plasmid by heat shock for amplification and then spread on a plate having ampicillin. After constructed plasmid isolated, *Phaeodactylum tricornutum* was transformed with isolated constructed plasmid by the biolistic method to create double strand break (DSB) on DNA at a specific site of genes (lhcf15 and lhcr8 gene) by CRISPR Cas9 technology. Double stranded DNA break can be repaired either

by homologous end joining (HEJ) or non-homologous end joining (NHEJ) repair system. NHEJ can cause deletion or insertion (InDels) of the nucleotide at the site of DSB(Jiang et al. 2014).

The *E. coli* colony was picked randomly from an overnight plate having ampicillin and the insert fragment of plasmid was amplified using a primer (forward and reverse, see appendix 2 d). Gel electrophoresis detected the amplicon that contains oligonucleotide DNA and this was confirmed by sequencing.

### 3.3.1 Mutation detection in genomic DNA using high resolution melting (HRM).

Within two weeks after transformation, the colony of transformed *P.tricornutum* was visible but genomic DNA was extracted in the third week. Target site of genomic DNA of transformed *P. tricornutum* (the lhcf15 and lhcr8 gene) was amplified around 700bp by colony PCR using LHCF15 and LHCR8 primer (forward and reverse, appendix 2j). The PCR product was amplified by light cycler 96 using F15HRMP1 and R8HRMP2 (forward and reverse ) primer that targets 100bp closer to mutation site after normalization of primer melting temperature ($55^0$C and $60^0$C respectively). HRM was discriminate among various mutant gene and wild type by scanning 100bp.

### 3.3.2 Mutation in LHCF15 gene.

100bp amplicon from clone LHCF15_PAM1 was scanned by HRM for the presence of genetic variation at the target site. The total number of sample that was analysed for this clone is 24. As figure 13 indicates, 20 among 24 sample have a melting curve that is the same as wild type but 4 of them shows the difference. An amplicon of the different melting curve has a different sequence of the base that indicates genetic variation when compared with wild type. After HRM analysis, 700bp PCR product from four sample that shows positive HRM result were confirmed by sequencing.

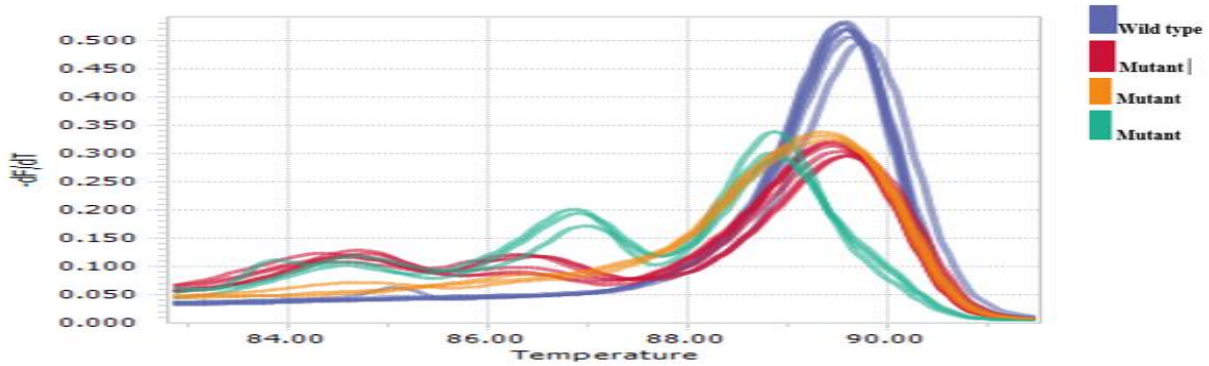



Figure 12. The different melting curve of lhcf15_pam1. Lhcf15_pam1 gene was isolated from genomic DNA of *Phaeodactylum tricornutum* that was transformed by constructed plasmid having sgRNA targeting lhcf15_pam1 gene to cause DSB using CRISPR Cas9 system. The different colour represents the different sample having the different melting curve. Y and X axis shows fluorescence and temperature respectively. Wild-type curve is used as a control and primer temperature is $55^0$C.

**Mutation in LHCR8 gene.**

Amplicon having 100bp from Lhcr8_pam1 and Lhcr8_pam2 construct was scanned by HRM for the presence of the genetic variation. 24 sample from two construct (clones) that target two different site of the same gene(lhcr8 gene) was screened for mutant using HRM but three of them were detected as mutant having different DNA melting curve when compared with wild type. 12 sample from lhcr8_pam 1 construct was homozygous with wild-type having the similar melting curve. Figure 14 shows, 12 sample from lhcr8_pam2 was analysed and different melting curve exists between the wild type and various mutation. This mutation was confirmed

by sequencing (figure 15).melting curve exists between the wild type and various mutation. This mutation was confirmed by sequencing (figure 15).

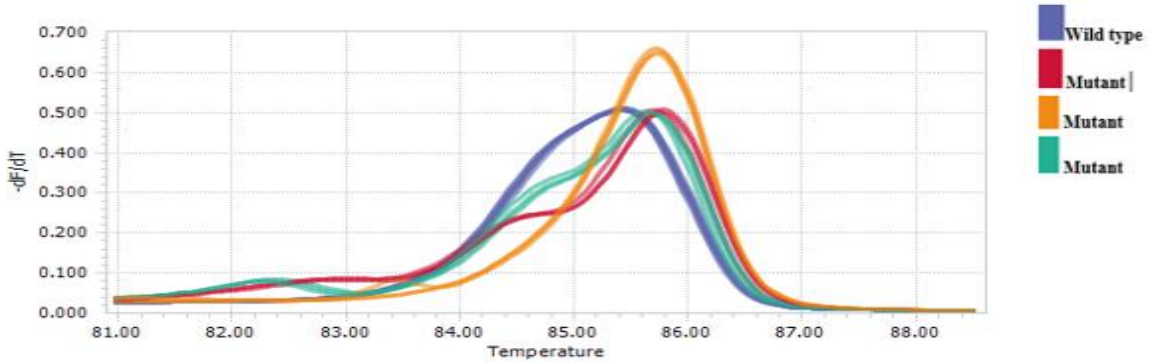



**Figure 13.** HRM curve that shows positive result of lhcr8_pam2 gene. Lhcr8_pam2 gene was isolated from genomic DNA of *Phaeodactylum tricornutum* that was transformed by constructed plasmid having sgRNA targeting lhcr8_pam2 gene to cause DSB using CRISPR Cas9 system. The Y and X axis indicates fluorescence and temperature respectively. Colour indicates type of sample having different melting curve. Wild type is used as control and primer temperature is $60^0$C.

### 3.3.3 Positive result of HRM was confirmed by sequencing.

All sample that shows positive HRM result was taken from both sample and amplified by colony PCR using primer (forward and reverse) which targets 700bp around mutation site. PCR product was used for gel electrophoresis. Gel band was excised, purified, Topo cloned and *E.coli* was transformed by heat shook for amplification. After topocloning plasmid was isolated, sequencing reaction was prepared for sequencing and figure 15, is the result of sequencing that confirmed the positive result of HRM. Figure 15 indicates, DSB of *P.tricornutum* was repaired using NHEJ that created deletion, insertion and substitution of nucleotide at DSB site. It created deletion of nucleotide ranges from 1-23 base at different site in 15 sample, insertion of 5 nucleotide (TCCAA) in one sample and substitution of A by G in 2 sample. In lhcr8 NHEJ created deletion of 17 base in 3 sample, substitution of two base T and G by C and A in the same gene of 4 sample and single base substitution T by C in 8 sample.

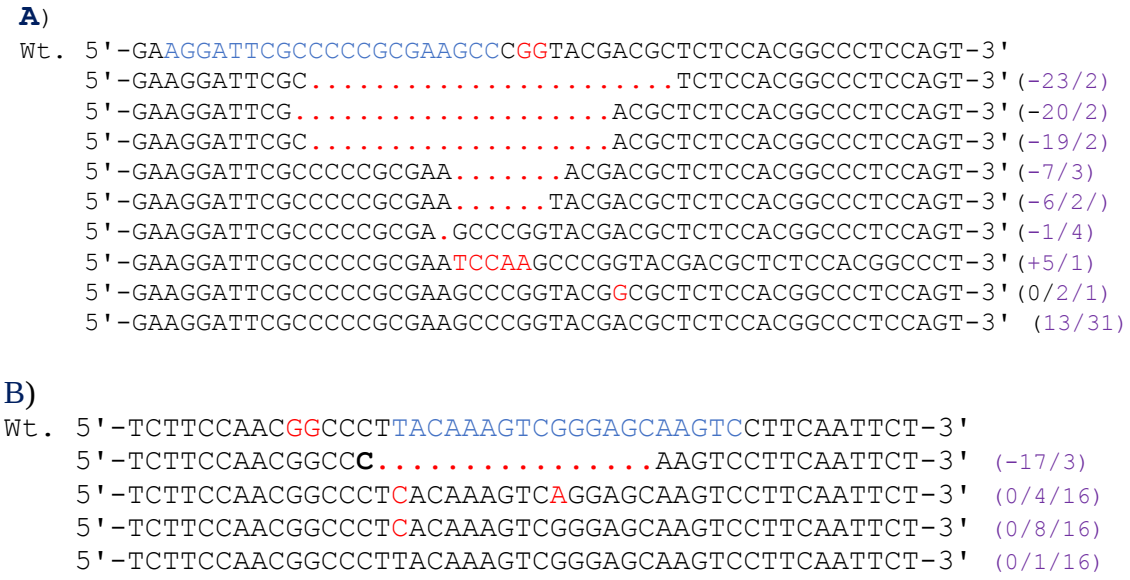



**Figure 14.** LHCF15 and LHCR8 gene were edited using CRISPR Cas 9 by creating DSB that is repair using NHEJ**.** 700bp from both gene was amplified and topo cloned. Then sequencing reaction was prepared and sent for sequencing to confirm positive result that was detected by HRMA. A) LHCF15 gene at DSB was repaired by NHEJ that caused insertion of 5 nucleotide in 1 sample, deletion of several nucleotide ranges from 1-23 nucleotide and substitution of single nucleotide A by G in one sample. B) LHCR8 gene at DSB was repaired by NHEJ that created deletion of 17 nucleotide in 3 sample and substitution of two nucleotide T and G by C and A respectively in 4 sample, substitution of single nucleotide T by C in 8 sample. Mines sign (-) in front of number indicates deletion of nucleotide and plus sign (+) indicates addition of nucleotide, zero (0) indicates substitution. Number indicates number of nucleotide deleted or added/number sample/total sample. Wt indicates wild type.

# 4. Discussion

The result from this study suggest that, the photosynthetic efficiency was recovered in blue light after 48 h high red light treatment but treatment of the cell with 48 h blue light did not protect the cell from photodamage after transferral to high red light (RL) from blue light (BL). In addition the CRISPR Cas9 technology was used to edit two genes that were found to be strongly induced by either the red light treatment (LHCF15) or the blue light treatment (LHCR8).

## 4.1 Photosynthetic efficiency of cells in 48 h high red light exposure.

In figure 9, the photosynthetic efficiency in all experiment drops initially when transferred from darkness/low light to high light that caused photo damaged to PSII as a result of ROS production. In high intensity white light (Nymark et al. 2009), the photosynthetic efficiency starts to increase again when they acclimate to high light. This does not happened in red light. The cells are not able to acclimate to high red light intensity. But when light is switched from red to blue, the photosynthetic efficiency increased. This result is in parallel with other results that showed the recovery of maximum quantum yield (photosynthetic efficiency) in low red light (80 µmol $m^{-2}s^{-1}$), dark treatment and low white light (Domingues *et al*. 2012, Valle *et al*. 2014). Different high light intensity (1250 µmol $m^{-2}s^{-1}$, 930 µmol $m^{-2}s^{-1}$, 500 µmol $m^{-2}s^{-1}$) and type (Red, WL) caused the photosynthetic efficiency of *P.tricornutum* to decrease. This indicates photosynthetic efficiency are able to recover in various light intensity such as red light (red 80 µmol $m^{-2}s^{-1}$) and blue light (200 µmol $m^{-2}s^{-1}$) including dark treatment after the cells were transferred from high light. Based on qRT-PCR result, the reason might be that the expression of several genes encoding proteins important for PSII repair and photoprotection are dependent on blue light.

The excess light energy that was absorbed during *P.tricornutum* exposure to high red light could not be used efficiently for photosynthesis, this results in the production of reactive oxygen species (ROS). Production of ROS in high red light exposure is thought to create photo-oxidation (Ohnishi et al. 2005, Krieger-Liszkay et al. 2008). The result from Northern blot and immunoblotting data analysis indicates that ROS produced in highlight exposure is able to inhibit the mRNA that is transcribed from PsbA gene, and polysomes bound PsbA mRNA that is used for localization are susceptible to ROS. Therefore, Suppression of translational elongation of mRNA by reactive oxygen species generated in high red light exposure which leads to inhibit the repair of PSII during high light exposure. Result from other species

(synechocystis sp) shows singlet oxygen inhibits translation elongation of D1 protein that result in inactivation of PSII (Barber and Andersson 1992, Nishiyama et al. 2004, Nishiyama et al. 2006) . This and lack of photoprotection might be connected with a drop of photosynthetic efficiency during 48h red light exposure of *P.tricornutum*. This is supported by qRT-PCR data that shows genes that involve in photoprotection (LHCR8, LHCR10 and PtLHCX3) and repair (HCF136, Psb27 and Psb29) were moderately affected by high red light during 48 h high red light exposure. The result of qRT-PCR is in parallel with Valle et. al. 2014 and Nymark et al 2009.

### 4.2 Recovery of photosynthetic efficiency after transferred to blue light from 48 h red light.

Repairing photodamage to PSII require production of new proteins (primary D1 protein) and partial degradation of D1 protein. this is called PSII repair cycle (Tyystjärvi 2008, Takahashi and Badger 2011). As the blue light turned on after 48 h red light, the efficiency gradually recovers and finally within 24 h exposure, it was almost the same level with the control which shows that the cells are able to acclimate and protect themselves against photodamage when treated with high intensity blue light (figure 9). This result is confirmed by qRT-PCR data which shows induction of genes that involves in photoprotection (LHCR8, LHCR10 and PtLHCX3) and repair (HCF136, Psb27 and Psb29) in blue light exposure. Result reported by Valle et al 2014 shows that the cells were able to recover in red light when the light intensity was lowed. Now in this result indicates that they are also able to recover in high intensity blue light, meaning that the algae are capable of acclimating and coping with light of this intensity, but the mechanism for photoprotection and acclimation are blue light depend. Induction of genes that encodes proteins involved in photoprotection and repair in blue light after 48 h red light might be connected with the recovery of photosynthetic efficiency. Transcriptional analysis of *P.tricornutum* indicates that the indication of several genes encoding proteins involved in photoprotection and PSII repair are blue light-dependent. These newly produced proteins might be involved in replacing damaged protein and PSII repair which leads to recovery of photosynthetic efficiency in blue light. The LHCR proteins have been suggested to be able to bind increased amounts of photo protective pigments (Lepetit et al. 2010). Therefore, the ability to recover after the transferal to blue light after 48 h red light treatment might be explained by the induction of genes encoding proteins important for PSII repair and photoprotection.

### 4.3 Photosynthetic efficiency of cells that was exposed to high red light (RL) after 48h blue light (BL) treatment.

Photosynthetic efficiency was going down until BL3h but partly recovers as the BL exposure reached 48 h which is in parallel with result reported by Nymark *et al* 2009 that shows drop of photosynthetic efficiency during the initial and intermediate phase of exposure after the cells were transferred from low light (LL) to high light (HL).This indicates that photodamage occur initially (0–BL3h) and photo protection/adaptation to blue light exposure are partly restored after 3 h blue light. Figure 12 shows that qRT-PCR data support partial recovery of photosynthetic efficiency from BL3h-BL48h by indicating, induction of genes (BL0.5h and BL24h) that involved in repairing and protecting the cell from photodamage and this qRT-PCR data (HCF136, Psb27, LHCF15,LHCR8 and LHCX3) are in parallel with the result reported by Nymark *et al* 2009. The protein produced by induction of these genes might be involved replacing the damaged protein in 0-3 h blue light exposure that caused photosynthetic efficiency to drop. This indicates the existence of photodamage during the cells were exposed to blue light for the first 3 h from 48 h dark treatment.

Photosynthetic efficiency was again reduced after exposed to RL from 48 h BL exposure. This result is confirmed by qRT-PCR data showing that, genes expression reduced in RL after BL48h. Reduction of gene expression in RL might be associated with a drop in photosynthetic efficiency. Therefore, in this case (this data) partial recovery of photoprotection (photosynthetic efficiency) in late phase of BL (200 µmol $m^{-2}s^{-1}$) exposure might not protect *P.tricorntum* from photodamage in RL (930 µmol $m^{-2}s^{-1}$) exposure but protein expression of these genes might be required to come up with better conclusion. The amount of mRNA transcribed in this exposure do not tell the amount of proteins translation from this exposure. So protein expression of these genes in blue light from 48 h dark treatment needs further work.

### 4.4 Genes editing using CRISPR Cas9.

From my knowledge, there is no LHCF15 and LHCR8 knockout reported so far in *P.tricornutum,* So that I am going to discuss my result based on the data from HRMA and sequencing. After double strand break (DSB) is created by CRISPR Cas9, the cell has NHEJ repair mechanism which is active in cell cycle to repair at Cas9 target break site. It causes insertion, deletion of base and frameshift mutation(Ran et al. 2013, Hsu et al. 2014) . In LHCR8 gene, only one allele was edited because there was a mismatch between gRNA and other allele's target site. This indicates that to knockout this gene, it must be targeted with other gRNAs. Positive result of HRMA (LHCF15 and LHCR8) are supported by sequencing. After transformation each colony came from a single transformed single cells. Some of the colony contain different genetic variation but some other contain similar mutation.

A similar mutation in different colony created by Cas9 causing DSB at the same site in the genome of the different colony that can be repaired by either error-prone non-homologous end joining (NHEJ) or homologous end joining (HEJ) by using the first mutant as a template to repair the other (second) DSB. Sequencing data shows different mutant are present in a single colony and some other colonies contain a few mutant. This implies that expression of transformed plasmid and subsequent DSB followed by repair has happened after some cell division event that results in different mutant in a single colony but a few mutant in a single colony is the result of transformed plasmid was expressed early to cause DSB that is followed by repair system.

## 5. Conclusion

The photosynthetic efficiency of *P.tricornutum* varies depending on the intensity, type, length of exposure and sequence of light they exposed in time series experiment. High red light fails to induce the expression of a subset of LHC's genes that are believed to be connected to photoprotection. This might be one of the reasons for the photodamage (measured as lower photosynthetic efficiency) in red light. Subsequent blue light exposure induced the expression of the LHC's genes, and a simultaneous recovery of photosynthetic efficiency. Exposure to blue light from dark, induced genes that involves in repair and protection which result in short time photodamage and long period photo protection. Subsequent exposure to red light repress the genes that were induced in blue light that result in photodamage (low photosynthetic efficiency) in red light.. These data implies that negatively affected the photosynthetic efficiency of *P.tricornutum* is recovered in blue light, and lhcf15 and lchr8 genes are editable by CRISPR Cas9 to enable further functional analysis. Lhcf15 and lhcr8 gene are edited by CRISPR Cas9 to enable further functional analysis. The genes (lhcf15 and lhcr8 genes) that involves in photosynthesis are successful edited by CRISPR Cas9.

## 6. Reference

addgene. 2016. CRISPR 101: Introduction to addgenes desktop resourse.

Addgene. no date-a. CRISPR/Cas9 Guide.

addgene. no date-b. Genome Engineering Guide.

addgene. No date CRISPR/Cas9 Guide.

Armbrust, E. V. 1999. Identification of a New Gene Family Expressed during the Onset of Sexual Reproduction in the Centric DiatomThalassiosira weissflogii. Appl Environ Microbiol **65**:3121-3128.

Armbrust, E. V. 2009. The life of diatoms in the world's oceans. Nature **459**:185-192.

Armbrust, E. V., J. A. Berges, C. Bowler, B. R. Green, D. Martinez, N. H. Putnam, S. Zhou, A. E. Allen, K. E. Apt, and M. Bechner. 2004. The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. Science **306**:79-86.

Barber, J. and B. Andersson. 1992. Too much of a good thing: light can be bad for photosynthesis. Trends Biochem Sci **17**:61-66.

Bjornsti, M.-A. and M. D. Megonigal. 1999. Resolution of DNA molecules by one-dimensional agarose-gel electrophoresis. DNA Topoisomerase Protocols: Volume I: DNA Topology and Enzymes:9-17.

Bolton, J. R. and D. O. HALL. 1991. The maximum efficiency of photosynthesis*. Photochemistry and Photobiology **53**:545-548.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, and L. A. Marraffini. 2013. Multiplex genome engineering using CRISPR/Cas systems. Science **339**:819-823.

Domingues, N., A. R. Matos, J. M. da Silva, and P. Cartaxana. 2012. Response of the diatom Phaeodactylum tricornutum to photooxidative stress resulting from high light exposure. PLoS One **7**:e38162.

Doudna, J. A. and E. Charpentier. 2014. The new frontier of genome engineering with CRISPR-Cas9. Science **346**:1258096.

Falciatore, A. and C. Bowler. 2002a. Revealing the molecular secrets of marine diatoms. Annu Rev Plant Biol **53**:109-130.

Falciatore, A. and C. Bowler. 2002b. Revealing the molecular secrets of marine diatoms. Annu Rev Plant Biol **53**:109-130.

Fleige, S. and M. W. Pfaffl. 2006. RNA integrity and the effect on the real-time qRT-PCR performance. Molecular aspects of medicine **27**:126-139.

Gakuran, M. 2008. Colour in the deep sea.

Gorbunova, V. and A. A. Levy. 1997. Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Res **25**:4650-4657.

Guillard, R. R. 2005. Purification methods for microalgae. Algal culturing techniques. Elsevier, New York:117-132.

He, L., X. Han, and Z. Yu. 2014. A rare Phaeodactylum tricornutum cruciform morphotype: Culture conditions, transformation and unique fatty acid characteristics. PLoS One **9**:e93922.

Horvath, P. and R. Barrangou. 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science **327**:167-170.

Hruscha, A., P. Krawitz, A. Rechenberg, V. Heinrich, J. Hecht, C. Haass, and B. Schmid. 2013. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development **140**:4982-4987.

Hsu, P. D., E. S. Lander, and F. Zhang. 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell **157**:1262-1278.

Jansen, R., J. Embden, W. Gaastra, and L. Schouls. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Molecular microbiology **43**:1565-1575.

Jiang, W., B. Yang, and D. P. Weeks. 2014. Efficient CRISPR/Cas9-mediated gene editing in Arabidopsis thaliana and inheritance of modified genes in the T2 and T3 generations. PLoS One **9**:e99225.

Johnson, P. H. and L. I. Grossman. 1977. Electrophoresis of DNA in agarose gels. Optimizing separations of conformational isomers of double-and single-stranded DNAs. Biochemistry **16**:4217-4225.

Kirupamurthy, D. 2014. Studies of light responses and the development of a transformation system for the benthic diatom Seminavis robusta.

Lewin, J. C. 1958. The taxonomic position of Phaeodactylum tricornutum. Microbiology **18**:427-432.

Li, X. and W.-D. Heyer. 2008. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res **18**:99-113.

Logan, J. M., K. J. Edwards, and N. A. Saunders. 2009. Real-time PCR: current technology and applications. Horizon Scientific Press.

Loke, R. E., J. H. du Buf, M. Bayer, and D. G. Mann. 2004. Diatom classification in ecological applications. Pattern recognition **37**:1283-1285.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church. 2013. RNA-guided human genome engineering via Cas9. Science **339**:823-826.

Mann, D. 1993. Patterns of sexual reproduction in diatoms. Pages 11-20 *in* Twelfth International Diatom Symposium. Springer.

Mann, D. G. 2010. Bacillariophytina.

Nishimasu, H., F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata, N. Dohmae, R. Ishitani, F. Zhang, and O. Nureki. 2014. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell **156**:935-949.

Nishiyama, Y., S. I. Allakhverdiev, and N. Murata. 2006. A new paradigm for the action of reactive oxygen species in the photoinhibition of photosystem II. Biochimica et Biophysica Acta (BBA)-Bioenergetics **1757**:742-749.

Nishiyama, Y., S. I. Allakhverdiev, H. Yamamoto, H. Hayashi, and N. Murata. 2004. Singlet oxygen inhibits the repair of photosystem II by suppressing the translation elongation of the D1 protein in Synechocystis sp. PCC 6803. Biochemistry **43**:11321-11330.

Nishiyama, Y., H. Yamamoto, S. I. Allakhverdiev, M. Inaba, A. Yokota, and N. Murata. 2001. Oxidative stress inhibits the repair of photodamage to the photosynthetic machinery. EMBO J **20**:5587-5594.

Nymark, M., A. K. Sharma, T. Sparstad, A. M. Bones, and P. Winge. 2016. A CRISPR/Cas9 system adapted for gene editing in marine algae. Scientific reports **6**.

Nymark, M., K. C. Valle, K. Hancke, P. Winge, K. Andresen, G. Johnsen, A. M. Bones, and T. Brembu. 2013. Molecular and photosynthetic responses to prolonged darkness and subsequent acclimation to re-illumination in the diatom Phaeodactylum tricornutum. PLoS One **8**:e58722.

Owens, T. G. 1986. Light-harvesting function in the diatom Phaeodactylum tricornutum II. Distribution of excitation energy between the photosystems. Plant Physiol **80**:739-746.

Perez-Pinera, P., D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, and K. W. Leong. 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nature methods **10**:973-976.

Ran, F. A., P. D. Hsu, C.-Y. Lin, J. S. Gootenberg, S. Konermann, A. E. Trevino, D. A. Scott, A. Inoue, S. Matoba, and Y. Zhang. 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell **154**:1380-1389.

Rochaix, J.-D. 2011. Assembly of the photosynthetic apparatus. Plant Physiol **155**:1493-1500.

Scala, S., N. Carels, A. Falciatore, M. L. Chiusano, and C. Bowler. 2002. Genome properties of the diatom Phaeodactylum tricornutum. Plant Physiol **129**:993-1002.

Seckbach, J. and P. Kociolek. 2011. The diatom world. Springer Science & Business Media.

Smol, J. P. and E. F. Stoermer. 2010. The diatoms: applications for the environmental and earth sciences. Cambridge University Press.

Takahashi, S. and M. R. Badger. 2011. Photoprotection in plants: a new light on photosystem II damage. Trends Plant Sci **16**:53-60.

Taylor, N. J. and C. M. Fauquet. 2002. Microparticle bombardment as a tool in plant science and agricultural biotechnology. DNA and cell biology **21**:963-977.

Tyystjärvi, E. 2008. Photoinhibition of photosystem II and photodamage of the oxygen evolving manganese cluster. Coordination Chemistry Reviews **252**:361-376.

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

Vrieling, E. G., W. Gieskes, and T. P. Beelen. 1999. Silicon deposition in diatoms: control by the pH inside the silicon deposition vesicle. Journal of phycology **35**:548-559.

Wang, H., H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang, and R. Jaenisch. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell **153**:910-918.

Yoon, H. S., J. D. Hackett, G. Pinto, and D. Bhattacharya. 2002. The single, ancient origin of chromist plastids. Journal of phycology **38**:40-40.

Zurzolo, C. and C. Bowler. 2001. Exploring Bioinorganic Pattern Formation in Diatoms. A Story of Polarized Trafficking. Plant Physiol **127**:1339-1345.

# 7. Appendix.

**Appendix 1 a.** Cladogram

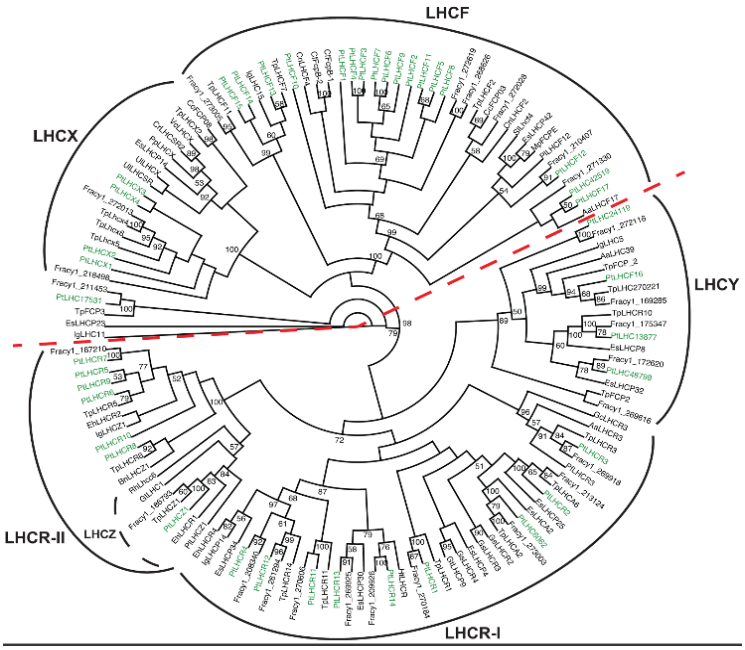



**Figure**. 'Cladogram of the antenna proteins of *P.tricornutum'*. Antenna proteins of *P.tricornutum* that involves in light harvesting complex (LHC) superfamily (Nymark et al. 2013)

**Appendix 1 b**. Spectrum of light in ocean

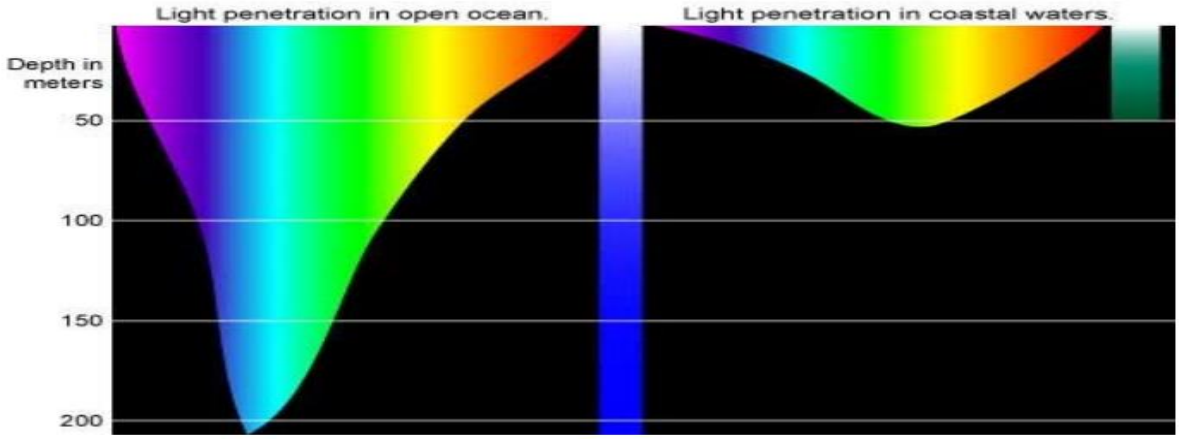



***Figure.*** Penetration of light in coastal water. Different light type has different penetration level in water. Red light has low penetration level. https://www.montereybayaquarium.org/-/m/pdf/education/curriculum/light_in_the_deep_sea.pdf.

**Appendix 2 a**. RIN value of isolated RNA

| CWLI | CWL II | CWL III | BL0.5h I | BL0.5 II | BL0.5h III | Bl24h I | BL24h II | BL24h III | RL48h I | RL48h II | RL48h III |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 8.5 | 9 | 8.5 | 6.5 | 6.9 | 6.9 | 7.9 | 7.6 | 6.4 | 7.4 | 7.5 | 7.5 |

**Appendix 2 b**. PCR program

| Steps | Temperature | Duration |
|---|---|---|
| 1 | 94 | 5 minutes |
| 2 | 94 | 30 seconds |
| 3 | 55 | 30 seconds |
| 4 | 72 | 2 minutes |
| 5 | Repeat 2-4 | 30 times |
| 6 | 4 | |

**Appendix 2 c.** Master Mix for sequencing reaction.

| Reagent | Volume per reaction |
|---|---|
| Big dye | 1 µl |
| Sequence buffer | 3 µl |
| Primer3.2 pmol(M13 rev) | 1 µl |
| Nuclease free water | Variable |
| Template | Variable |
| Total volume | 20 µl |

**Appendix 2d.** Sequence of primer for sequencing reaction.

| Primer | Sequence (5'-3')2 |
|---|---|
| M13 Reverse | CAGGAAACAGCTATGAC |
| KST3 3.2 | ACTAAAGGGAACAAAAGCTG |

**Appendix 2 e.** DYNAZ program

| Temperature | Time |
|---|---|
| 1. 94 0C | **10 minutes** |
| 2. 940C | **30 seconds** |
| 3. 550C | 30 seconds |
| 4. 72 0C | 1:30 minutes |
| 5. Go to step 2 and repeat 29 times | |
| 6. 72 0C | 10 minutes |
| 7. 4 0C | It holds at 4 0C |

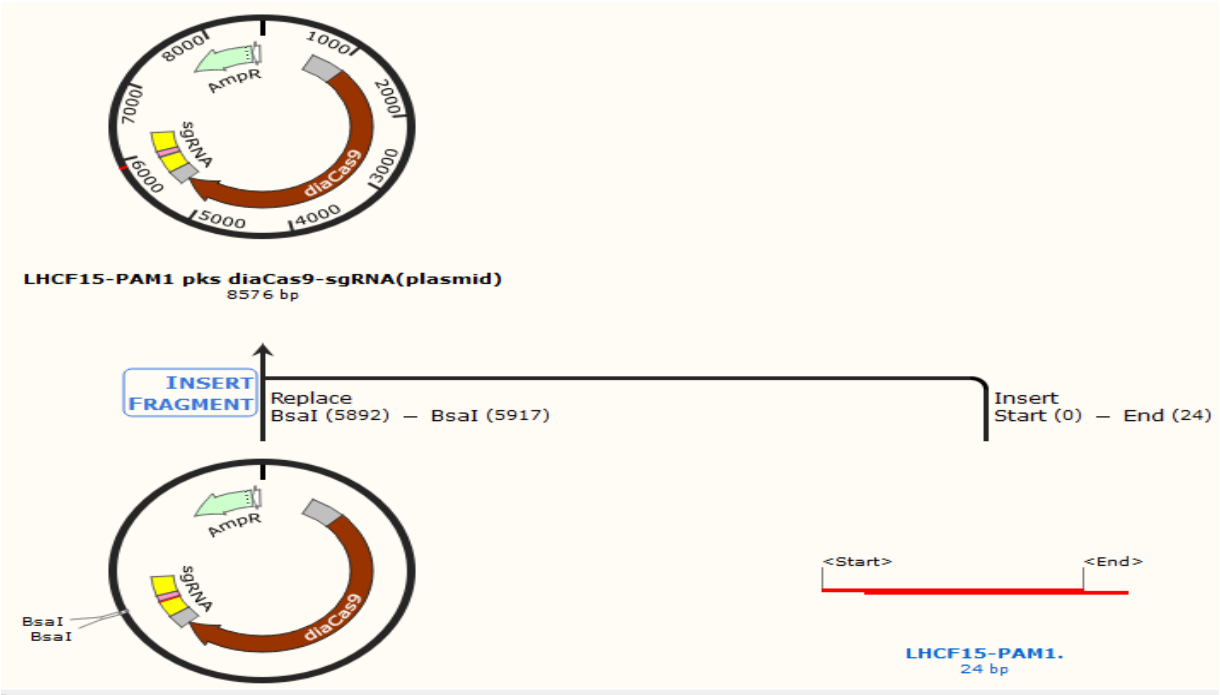
**Appendix 2 f.** RBD31.

| Temperature | Times |
|---|---|
| 1. 96 $^0$C | 5 minutes |
| 2. 96 $^0$C | 10 seconds |
| 3.50 $^0$C | 5 seconds |
| 4. 60 $^0$C | 4 minutes |
| 5. Repeat step 2-4, 25 times. | |

**Appendix 2 g.** HRM primer to amplify PCR product (The second set of primer that targets 100 bp away from mutation was supposed to happen)

| Genes | Target site | Primers | Sequence (5'-3') |
|---|---|---|---|
| LHCF15 | LHCF15_PAM1 | Forward | CAGACAGTCATGAAGTTTGC |
| | | Reverse | GTAGCTCGACGAGGACTGGA |
| LHCR8 | LHCR8_PAM1 | Forward | CTGGAGACGTCGGCTTTGAC |
| | | Reverse | CATGGCCAAACGGCAATGCT |
| | LHCR8_PAM2 | Forward | TGCAGCCTATTGCCGATGC |
| | | Reverse | TCCCTGCAAGCGTGACAAC |

**Appendix 2 h**. pKS diaCas9_sgRNA (LHCF15) plasmid




**Figure,** Annealed forward and reverse primer was inserted to the vector and ligated by T4DNA ligase. The ligation of annealed oligo fragment in a vector produce pKs diaCas9_sgRNA (LHCF15) plasmid. That is ready for amplification using *E.coli*. For annealed LHCR8 fragment also inserted and ligated to vector in the same way to produce plasmid.

**Appendix 2 i.** PCR primer

| Genes | Primer | Sequence 5'-3' |
|---|---|---|
| Psb29 | Forward | GGATGAAGTCTACCCGTTATG |
| | Reverse | CATGGTCTCCATCATGTCAAGC |
| Psb27 | Forward | GCTGGCAAACCCTCTTTCAGTA |
| | Reverse | AGTGTCCAGCGAGTACGTTGAT |
| LHCX3 | Forward | TTCTACAGACCAAGGAACTCCA |
| | **Reverse** | TGTGCAACGTTACGGGAAACT |
| HFC 136 | Forward | ACTGGTGGATACGGCATTACTG |
| | Reverse | CGTCAAAGCTCACGTACATCGT |
| LHCF15 | Forward | AAAGATATTCCCAACGGCATCG |
| | Reverse | GGTTATCCGAATTGAGGTAGTC |
| LHCR10 | Forward | GGAAGATTCCCGTGGAATACTG |
| | **Reverse** | CCGACGTTTCTGAATGTTAGC |
| | | |
| | | |
| | | |
| | | |

**Appendix 2 j.** Primer (The first set of primer) that targets 700 bp away from DSB

| Gene | Primers | Sequence 5'-3' |
|---|---|---|
| LHCF15 | Forward | ACGAGAAAATGCATGTAAGCTG |
| | Reverse | GGAACAATGTACCCAATGACAG |
| LHCR8 | Forward | AACGAATACGTAGACCCTTAGA |
| | Reverse | CAGTTCTTTGGTCTGCATAGCG |