Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

Andreas Holm Loriente

Investigating heavy metal tolerance in the diatom *Phaeodactylum tricornutum*

Creating mutants of three putative metal transporters using CRISPR/Cas9 technology

Master's thesis in Biotechnology
Supervisor: Atle M. Bones
Co-supervisor: Ralph Kissen & Marianne Nymark
May 2022

NTNU
Kunnskap for en bedre verden

Andreas Holm Loriente

Investigating heavy metal tolerance in the diatom *Phaeodactylum tricornutum*

Creating mutants of three putative metal transporters using CRISPR/Cas9 technology

Master's thesis in Biotechnology
Supervisor: Atle M. Bones
Co-supervisor: Ralph Kissen & Marianne Nymark
May 2022

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

NTNU

Kunnskap for en bedre verden

Acknowledgements

This thesis has been the culmination of a 5-year master's programme in biotechnology at the Norwegian University of Science and Technology (NTNU). The lab work presented in this thesis was carried out as part of the Cell, Molecular Biology and Genomics (CMBG) group at the Department of Biology.

I would like to thank my supervisor professor Atle M. Bones, head of the CMBG group, for giving me the opportunity to work on such an interesting and rewarding project. I wish to express my deepest gratitude to my co-supervisor Ralph Kissen for tutoring me in the laboratory, for answering all my questions, for his patience and for his company. You have been a vital source of guidance and assistance throughout this entire project. My sincerest thanks my co-supervisor Marianne Nymark for her help with designing vectors, for answering any questions I had, and for helping me in the laboratory whenever I needed it. I am very grateful to the rest of the CMBG group and my fellow master students for their insights, for helping me whenever needed and for making me feel included.

These university years would also not have been the same without all the good times spent amongst friends and classmates, both on and off campus, so thank you all. Finally, I wish to thank my family for their continuous support and encouragement throughout all my years of study and my work on this thesis. I would never have gotten to this point without you.

Andreas Holm Loriente

Abbreviations

A1	*ATPase5-1B* mutant line 1
A2	*ATPase5-1B* mutant line 2
ATP	Adenosine triphosphate
bp	Base pair
C1	Conjugation protocol 1
C2	Conjugation protocol 2
Cas	CRISPR associated protein
Chl	Chlorophyll
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
DSB	Double stranded break
EDTA	Ethylenediaminetetraacetic acid
E_k	Minimum saturating irradiance
gRNA	Guide RNA
HCl	Hydrochloric acid
HDR	Homology-directed repair
HRM	High-resolution melting
ICE	Inference of CRISPR edits
ICP-MS	Inductively coupled plasma mass spectrometry
Indel	Insertion or deletion
K'_{max}	Maximum specific growth rate
LB	Luria-Bertani
NEB	New England Biolabs
NHEJ	Nonhomologous end-joining
PAM (fluorometer)	Pulse amplitude modulation
PAM (sequence)	Protospacer adjacent motif
PCR	Polymerase chain reaction

PC	Phytochelatin
PSI/II	Photosystem I/II
qPCR	Quantitative PCR
$rETR_{max}$	Maximum relative electron transport
RLC	Rapid light curve
ROS	Reactive oxygen species
SDS	Sodium dodecyl sulfate
TMD	Transmembrane domain
tracrRNA	Trans activating crRNA
Tris	Tris(hydroxymethyl)aminomethane
U	Unit
V1	*VIT1/VIT2* mutant line 1
V2	*VIT1/VIT2* mutant line 2
VIT	Vacuolar iron transporter
WT	Wild type
Φ_{PSII}	Effective quantum yield of PSII

Abstract

Industrialisation has led to heavy metal pollution from anthropogenic sources becoming a global problem, as these elements often represent a serious risk to human health and to the environment. Despite having received much attention in recent decades, currently utilized technologies for heavy metal waste removal are considered economically prohibitive, inefficient under many industrial conditions and often environmentally harmful themselves. The use of organisms to remove heavy metal pollutants, referred to as bioremediation, has emerged as a potentially cheap, effective and sustainably alternative. Bioremediation using microalgae, also referred to as phycoremediation, is among the most promising candidates for the development of such a next-generation waste management technology and has been the subject of much research in later years. Many microalgal species display high innate tolerance to heavy metals and the ability to effectively accumulate. Advancements in molecular biology and genetic analysis technology has allowed scientists to better understand the molecular basis of these traits, opening opportunities to further improve phycoremediation technologies and the potential of creating hypertolerant and hyperaccumulating mutant strains through genetic engineering.

The goal of this project was to investigate the role of three genes - *VIT1*, *VIT2* and *ATPase5-1B* - encoding predicted metal transporter proteins play in the innate heavy metal tolerance of the marine diatom *Phaeodactylum tricornutum*, and thereby contributing to this body of knowledge. Due to their sequence homology with metal transporters described in other organisms, it has been predicted that the VIT1 and VIT2 transporters help store cytosolic heavy metals in vacuoles while the ATPase5-1B transporter facilitates efflux from the cell. Their functions were investigated in this project by creating knockouts of these genes using the CRISPR/Cas9 gene editing technology and observing if mutant lines displayed any changes in tolerance to zinc and cadmium.

Plasmids containing the CRISPR-Cas9 endonuclease and a gRNA targeting either *ATPase5-1B* or both *VIT1* and *VIT2* simultaneously, were generated and used to transform *P. tricornutum* by bacterial conjugation and biolistic particle bombardment. CRISPR/Cas9-induced mutations were thus obtained, but due to time constraints the screening and purification of mutant lines had to be concluded prematurely. The four most promising mutant lines were selected for further assays: two lines with knockout mutations for *ATPase5-1B*, and two lines with mutations leading to a knockout of *VIT2* and knockdown of *VIT1.* Measurement of the maximum specific growth rates K'_{max} and variable Chl *a* fluorescence indicated that cultivation in 5 mg/L Zn^{2+} or 1.25 mg/L Cd^{2+} did not induce disproportional toxicity in any of the mutant lines. The long-term growth of *ATPase5-1B* knockout when exposed to 1.25 mg/L Cd^{2+} was more affected than that of WT, indicating that this transporter might be involved in adaptive heavy metal tolerance in *P. tricornutum.* Further studies, including changes in gene expression and heavy metal accumulation in such knockout mutants, are however necessary to determine the role that these metal transporters play in the heavy metal tolerance of *P. tricornutum* and their potential relevance for the development of a sustainable bioremediation technology.

Sammendrag

Industrialisering har gjort antropogen tungmetallforurensning til et globalt problem, da utslippet av disse elementene ofte representerer en alvorlig risiko for menneskehelse og miljøet. Til tross for at dette har mottatt mye oppmerksomhet de siste tiårene, er de nåværende teknologiene benyttet for fjerning av tungmetaller ansett som dyre, ineffektive under mange industrielle kondisjoner og ofte selv miljøskadelige. Bruken av organismer for å fjerne tungmetaller, kalt bioremediering, har dukket opp som et potensielt billig, effektivt og bærekraftig alternativ. Bioremediering som benytter seg av microalger, referert til som phycoremediering, er blant de mest lovende kandidatene for utviklingen av slik nestegenerasjons avfallsbehandling og har i de senere år mottatt mye oppmerksomhet. Mange mikroalgearter demonstrerer høye naturlige toleranser mot tungmetaller og evnen til å effektivt akkumulere dem. Fremskritt innenfor molekylær biologi og genetisk analyseteknologi har tillat forskere å bedre forstå det molekylære grunnlaget for disse trekkene, hvilket åpner muligheter for å videre forbedre phycoremedieringsteknologi og potensielt skape hypertolerante og hyperakkumulerende mutanter igjennom genteknologi.

Målet for dette prosjektet var å undersøke rollene tre gener - *VIT1*, *VIT2* og *ATPase5-1B* - som koder for antatte metaltransportør-proteiner i den naturlige tungmetalltoleransen til den marine diatomeen (kiselalgen) *Phaeodactylum tricornutum*, og dermed bidra til dette kunnskapsgrunnlaget. På grunn av deres sekvens-homologi med metatransportører beskrevet i andre organismer har det blitt forutsett at *VIT1* og *VIT2* transporterer tungmetaller fra cytosol inn i vakuoler, mens *ATPase5-1B* pumper tungmetaller ut av cellen. Deres funksjoner ble undersøkt i dette prosjektet ved å lage knockouts av disse genene ved hjelp av CRISPR/Cas9 gen editering, og deretter undersøke om mutantlinjene demonstrerte endringer i sink- og kadmiumtoleranse.

Plasmider som kodet for CRISPR-Cas9 endonukleasen og et gRNA som gjenkjente enten *ATPase5-1B* eller både *VIT1* og *VIT2*, ble skapt og benyttet til å transformere *P. tricornutum* cellelinjer via bakteriell konjugering og biolistisk partikkelbombardering. CRISPR/Cas9-induserte mutasjoner ble oppnådd, men isolasjonen av rene mutantlinjer måtte avbrytes tidlig på grunn av manglende tid. De fire mest lovende mutantlinjene ble valgt for videre undersøkelser: to linjer med knockout-mutasjoner for *ATPase5-1B*, og to linjer med mutasjoner som førte til knockout av *VIT2* og knockdown av *VIT1*. Målinger av maksimum spesifikk vekstrate K'_{max} variable klorofyll *a* fluorescens indikerte at verken kultivering i 5 mg/L Zn^{2+} eller 1.25 mg/L Cd^{2+} induserte uforholdsmessig toksisitet i noen av mutantlinjene. Langtidsveksten til *ATPase5-1B*-knockoutlinjene ved utsettelse for 1.25 mg/L Cd^{2+} var mer påvirket enn den for villtype, hvilket indikerer at denne transportøren kan være involvert i adaptiv tungmetalltoleranse i *P. tricornutum*. Videre undersøkelser, inkludert av endringer i genekspresjon og tungmetallakkumulasjon i slike knockoutmutanter, er nødvendige for å fastslå rollene disse metalltransportørene spiller i tungemetalltoleransen til *P. tricornutum* og hvilken betydning de kan ha for utviklingen av bærekraftig bioremedieringsteknologi.

Table of contents

1 Introduction

1.1 Diatoms

Diatoms are a diverse group of photosynthetic, unicellular eukaryotic algae belonging to the class Bacillariophyceae, considered to be of great ecological significance. They can be present in the environment as single cells or in different forms of colonies depending on species (Marella et al., 2020). Their sizes range from 5µm for smaller-celled species to upwards of 5mm for giant diatoms, and they exhibit a great variety of shapes (Falciatore & Bowler, 2002). While estimations of the number of diatom species vary greatly, more recent estimations range from 12.000-30.000 species (Guiry, 2012; Mann & Vanormelingen, 2013). They are considered robust organisms and can be found in almost all aquatic environments, including freshwater rivers and lakes, marine ecosystems, as well as in moist terrestrial environments like soils, vegetation and lichen. (Bozarth et al., 2009; Pfister et al., 2017). Marine diatoms are particularly common in nutrient-rich polar and coastal regions, while less common in the warmest and most hypersaline environments (Malviya et al., 2016). Smaller diatoms (5–50 µm) are most abundant when light conditions are ideal and nutrient availability is high during spring and autumn but form aggregates that sink out of the photic zone in under light- and nutrient-deprived conditions. Large diatoms (which can reach sizes of 2-5mm) are ubiquitous in all oceans and show less seasonal variability in their abundance (Falciatore & Bowler, 2002). Aquatic diatoms can be either planktonic and free floating in open water, or benthic and attached to waterbeds and other fixed surfaces. Some benthic diatoms are also motile, being adapted to gliding locomotion (Pfister et al., 2017).

Marine diatoms are significant contributors to biogeochemical cycles in several ways. They account for about 40% of annual marine biomass production, and thus play an important role in carbon cycling (Bozarth et al., 2009). Furthermore, they are photoautotrophic and contribute significantly to global carbon dioxide (CO^2)-fixation and oxygen production, representing about 20% of total primary production on earth (Malviya et al., 2016). This means that diatoms account for as much photosynthetic activity as all the earth's rainforests combined (Benoiston et al., 2017). Diatoms have been found to have symbiotic relationships with nitrogen-fixating cyanobacteria in warm oligotrophic oceans, contributing a significant amount of nitrogen to the environment (Falciatore & Bowler, 2002; Fiore et al., 2010). Diatoms also play an important role in global biochemical cycling of silica, due to it being required for the synthesis of their cell walls (Benoiston et al., 2017).

Diatoms are characteristic for their synthesis of ornamental cell wall structures made up of amorphous silica [(SiO2)n(H2O)] known as frustules (Falciatore & Bowler, 2002). The name diatom (from Greek, meaning "cut in half") is a reference to this structure comprised of two halves (referred to as theca), where the smaller half (the hypotheca) fits into the larger half (the epitheca) like a petri dish (Armbrust, 2009; Falciatore & Bowler, 2002). Diatoms are typically classified as either centric or pennate, based on the symmetry of their frustules (De Tommasi et al., 2017). Centric diatoms are characterized by radially symmetrical frustules and are usually planktonic, while pennate diatoms have unilaterally symmetrical frustules and are mainly benthic (De Tommasi et al., 2017).

Diatoms have several other noteworthy traits and advantages over other microalgae that are of scientific, commercial and industrial interest. They have a fast growth rate (reaching up to 2-4 cell divisions/day), large storage vacuoles, high surface-to volume ratios, are highly adaptable to changes in environment and light conditions, are more productive than other eukaryotes of similar size and are often pollution tolerant (Marella et al., 2020). Diatoms also possess urea cycles which, unlike the nitrogen-excreting urea cycle seen in the mammals, serve to repackage, distribute and store carbon and nitrogen, allowing recovery from nutrient withdrawal and more efficient metabolic activity (Allen et al., 2011). Scientific work is underway to explore the potential applications of diatoms for biofuel production, pharmaceuticals, healthy and sustainable foods, nanotechnology and biomolecule production (Bozarth et al., 2009). Furthermore, the potential diatoms present for cost-effective and eco-friendly bioremediation of contaminated waters has received much attention (Bozarth et al., 2009; Marella et al., 2020).

1.1.1 The diatom *Phaeodactylum tricornutum*

Phaeodactylum tricornutum is a pennate diatom and the only current member of the genus *Phaeodactylum* (Martin-Jézéquel & Tesson, 2012). It was first described by Knut Bohlin in 1897, but due to its unusual characteristics it was not formally acknowledged as a diatom until 1958 (Bohlin, 1897; Lewin, 1958). It is a brackish water species with a worldwide distribution, and can be found primarily in unstable coastal environments like estuaries and rock pools (Martin-Jézéquel & Tesson, 2012; Martino et al., 2007). It is known as a robust species that is able to grow at high pH values compared to the majority of other microalgae (upper limit 10.3) and in a wide salinity range (5 - 65 g/kg) (Martin-Jézéquel & Tesson, 2012).

P. tricornutum is considered an atypical diatom due to its pleiomorphism, having at least three distinct morphotypes: fusiform, oval and the less common triradiate form (Martino et al., 2007). The fusiform and triradiate types are planktonic and have higher buoyancy, while the oval form is preferentially benthic (Lewin, 1958; Martin-Jézéquel & Tesson, 2012; Martino et al., 2007). Rarer cruciform and lunate morphologies have sometimes been described (He et al., 2014; Lewin, 1958). The three commonly observed morphotypes are shown in Figure 1.3 (Vardi et al., 2008). Its morphological plasticity is related to the unusual cell wall of *P. tricornutum,* which is poorly silicified compared to other diatoms and thus less rigid (Martino et al., 2007). This lack of cell wall silicification makes *P. tricornutum* notable among diatoms for not having an obligate growth requirement for silica (Martino et al., 2007). Changes in morphology occur in response to environmental changes and are thought to represent different ecophenotypes adapted for growth under specific conditions (Martin-Jézéquel & Tesson, 2012; Martino et al., 2007). The fusiform (25-35µm) is the most common morphology found in nature and dominates in laboratory conditions with adequate light and nutrition (Lewin, 1958; Martino et al., 2007). It lacks silicified cell walls and has not been observed to grow better in media containing silica (Lewin, 1958; Martino et al., 2007) The oval form (8x3µm) represents a resistance form to cellular stress, as it appears under conditions of nutrient depletion, low light exposure, hyposalinity and low temperatures (Galas et al., 2021; Lewin, 1958; Martino et al., 2007). It is the only morphotype that has

been observed to survive for prolonged periods under these conditions (Martin-Jézéquel & Tesson, 2012). It is also the only morphotype that will incorporate silica into a typical frustule structure when grown in the presence of silicic acid (Martino et al., 2007). The absence of silicic acid in media has been shown to reduce the growth rate of the oval form by 50%, as opposed to the fusiform and triradiate form for which no difference in growth rate was observed. (Martino et al., 2007). The oval morphotypes tend to aggregate into clusters of hundreds to thousands of cells and excrete adhesive extracellular polymeric substances, which are thought to be physiological adaptation mechanisms meant to protect cells during nutrient starvation or from other stressors (Martin-Jézéquel & Tesson, 2012; Martino et al., 2007). Observed maximum average growth rates have been lower for the oval and triradiate morphotypes than the fusiform when *P. tricornutum* is grown in artificial sea water (Martino et al., 2007).

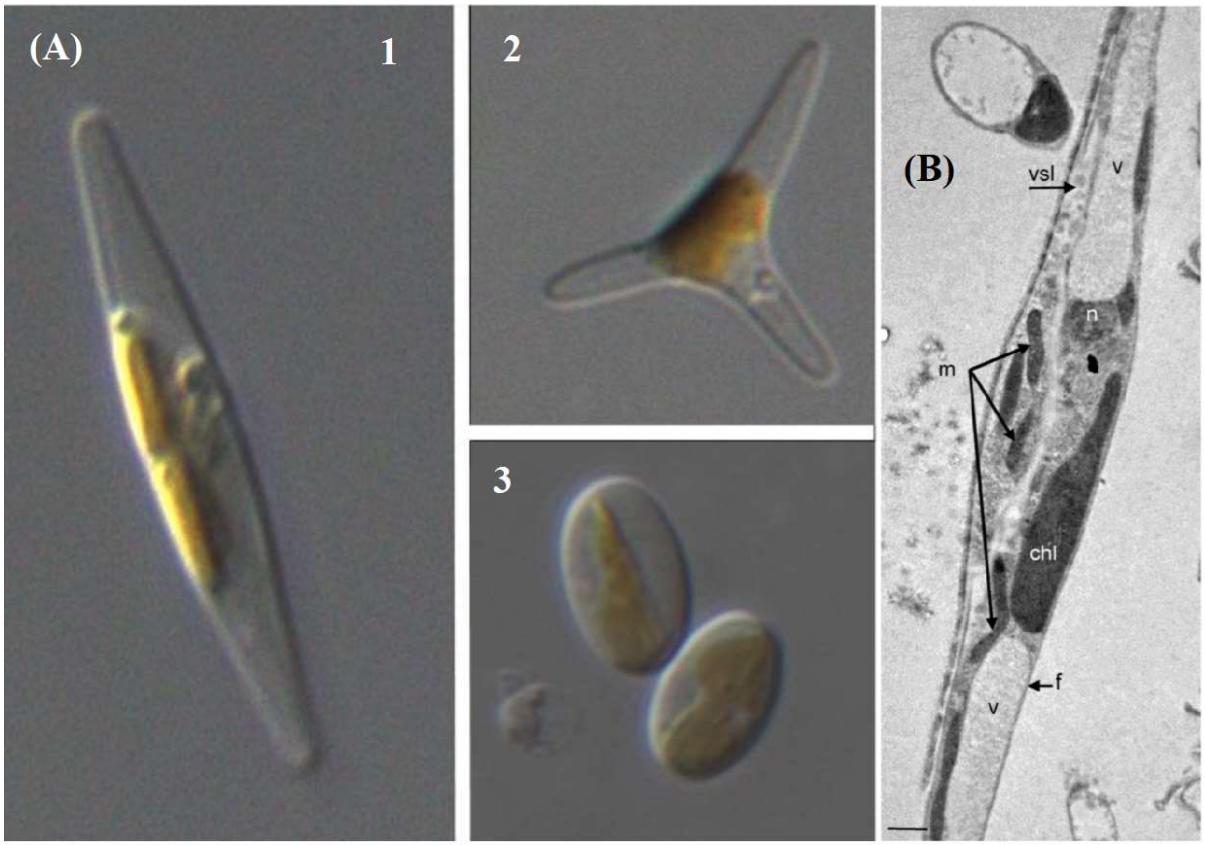


Figure 1.3: The diatom *Phaeodactylum tricornutum*. **(A)**: Light micrographs showing the three common morphotypes of *P. tricornutum.* **1**: Fusiform; **2**: Triradiate; **3**: Oval. Image is reproduced from Vardi et al. (2008). **(B)**: Transmission Electron micrograph of *P. tricornutum* fusiform morphotype. chl: plastid; f: frustule; m: mitochondria; n: nucleus; vsl: vesicle; v: vacuole. Scale bar 1µm. Image is reproduced from Galas et al. (2021).

P. tricornutum has been considered as a potential diatom model organism for some time (Falciatore & Bowler, 2002). Its fast growth rate, robustness and lack of silica requirements make it easy and practical to cultivate under laboratory conditions (Falciatore & Bowler, 2002; Martino et al., 2007). Its genome was fully sequenced in 2008 by Bowler et al. and shown to be approximately 27.4 megabases in size. The genome is predicted to contain some 10,402 genes, for which identification and functional analysis has been facilitated by the identification of more than 130,000 expressed sequence tags (ESTs) generated from cells grown under 16 different conditions (Bowler et al., 2008). This has contributed to a deeper understanding of the molecular biology of *P. tricornutum,* with which molecular tools and transformation methods for the diatom have been developed (Galas et al., 2021; Siaut et al., 2007). Genetic engineering utilising methods such as gene silencing (De Riso et al., 2009), meganucleases and TALENs (Transcription Activator-Like Effector Nucleases) (Daboussi et al., 2014) as well as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 (Nymark et al., 2016; Serif et al., 2018) have been shown to be effective for *P. tricornutum* (Galas et al., 2021). These tools are expected to help further decipher the cellular processes of *P. tricornutum* and facilitate its commercial possibilities in the near future (Galas et al., 2021).

1.2 Heavy metal pollution

The term heavy metal lacks a clear scientific definition. Heavy metals are often characterized as metals having an elemental density greater than five times that of water ($\rho < 5$ g/cm^{-3}), although the term has also been used to describe both metalloids and metals with lower densities in the literature. (Ali & Khan, 2018). Other definitions commonly cite heavy metals as having an atomic weight larger that sodium, or an atomic number larger than 20 (Duffus, 2002). In ecotoxicology the term is used to describe metallic and metalloid pollutants that display toxic effects on humans and the environment (Briffa et al., 2020). There are therefore ongoing discussions as to whether an alternative to the disputed term should be established (Duffus, 2002).

Heavy metals are naturally occurring trace elements that are found throughout the earth's crust, and are released into the environment by natural phenomena like volcanic activity and the weathering of metal-bearing soil and rocks (Tchounwou et al., 2012). However, industrial production and agriculture have made anthropogenic heavy metal pollution a global environmental problem. Major contributors include mining activities, fossil fuel combustion, agricultural pesticide use, nanomaterial production, industrial electroplating and manufacture of metal and textile products (Ali et al., 2019; Wu et al., 2019). Heavy metals enter the environment through leaching from solid waste and products into soils and water bodies and through the deposition of airborne pollutants (Vareda et al., 2019).

1.2.1 Toxicology of heavy metals

Certain heavy metals (e.g. Zn, Ni, Fe, Mn, Cu) are necessary for living organisms in small quantities and are therefore classified as essential heavy metals. These metals nevertheless display toxicity when concentrations exceed levels necessary for proper biological function. Non-essential heavy metals (e.g. Pb, Cd, Hg) are not necessary for life, and are toxic even at low concentrations (Ali et al., 2019). Furthermore, heavy metal toxicity is dependent on its bioavailability. Bioavailability depends on physical factors like pH, temperature, adsorption and sequestration. Furthermore, the various chemical forms heavy metals can take, or molecules they can be part, of changes their bioavailability by influencing solubility, partitions in different media and ability to form complexes (Tchounwou et al., 2012). Heavy metals bound to organic groups forming lipophilic complexes can be particularly toxic, as these diffuse easily through cell membranes and deliver the heavy metals to the interior of the cell (Briffa et al., 2020). Once in contact with the content of cells, toxic heavy metals can cause damage in several ways. They disrupt cellular function by competing with essential elements and by binding to negatively charged groups on proteins and nucleic acids. This causes misfolding, aggregation, inappropriate protein activity and interference with DNA repair mechanisms. They also generate reactive oxygen species (ROS) which cause oxidative damage to DNA, proteins and lipid cell membranes (Briffa et al., 2020). As a result many heavy metals are reported to induce toxicity and mutagenicity in most cells, as well as being neurotoxic, carcinogenic and teratogenic to higher organisms (Ali et al., 2019; Briffa et al., 2020).

As heavy metals are elements, they do not degrade and are persistent in the environment after contamination. This causes them to bioaccumulate in the surrounding biota, inducing toxicity in microorganisms, plants and animals. Furthermore, depending on their speciation, the concentrations of heavy metals are biomagnified as they are transferred along the successive trophic levels of food webs (Ali & Khan, 2019). Humans, as well as other organisms on higher trophic levels, are thus at a higher risk of being exposed to heavy metals at concentrations that exceed safe value thresholds.

1.2.2 Heavy metals in water bodies

Heavy metal pollution in water bodies presents significant problems for the health of humans and aquatic organisms present in the environment. Heavy metal pollutants from both natural and anthropogenic sources are dissolved in rainwater and industrial runoff and carried by streams and rivers into larger bodies of water. The bioavailability and toxicity of heavy metals in aquatic environments generally increase in conditions of low alkalinity, pH, water hardness and dissolved oxygen and by elevated temperatures, with ionic and other soluble forms being the most toxic (de Paiva Magalhães et al., 2015). A meta-analysis of global heavy metal concentrations in surface water bodies from 1994 to 2019, conducted by Kumar et al. in 2019, found that global average values of Cr, Mn, Co, Ni, As and Cd exceeded the guidelines established by the World Health Organization (WHO) and United States Environmental Protection Agency (USEPA). The analysis concluded that surface water bodies are generally “highly polluted”, with heavy metal concentrations being disproportionally high in many developing countries in Africa, Asia and South America (Kumar et al., 2019). The concentrations of heavy metals, alongside other pollutants, found in fish and other seafood around the world have also received much attention from both the scientific community and the public, spurring debate regarding the risk-benefit of consuming some foods in large quantities (Castro-González & Méndez-Armenta, 2008).

1.2.3 Zinc; Roles, pollution and human toxicology

Zinc (Zn) is an essential heavy metal and the 30th element on the periodic table. It is the 24th most common metal in the earth’s crust and is commonly used in the production of nanomaterials, brass, paint, cosmetics, textiles and to electroplate metals for rust-prevention (Briffa et al., 2020; Eisler, 1993). Zinc is a ubiquitous trace element in all living organisms, and functions as a cofactor for many enzymes and other biomolecules (Briffa et al., 2020). It plays an important role in gene transcription, development and maintenance of the immune system, cell metabolism and cell signalling modulation (Valko et al., 2005). It is considered relatively nontoxic to humans compared to other heavy metals, and individuals are at a greater risk of suffering from adverse effects associated with zinc deficiency that zinc toxicity (Plum et al., 2010). Therefore the WHO has not established formal guidelines for zinc concentrations in drinking water (WHO, 2017). Major human routes of entry for zinc are through inhalation, oral ingestion and dermal exposure (Plum et al., 2010).

One of the most common symptoms of zinc toxicity in mammals is copper deficiency, which is reported to adversely affect the ratio of low-density-lipoprotein to high-density-lipoprotein and by extension cardiac health in the long term (Fosmire, 1990). Inhalation of zinc-containing smoke can lead to an acute respiratory syndrome referred to as metal fume fever (MFF), a temporary condition that usually affects industrial workers involved in zinc smelting or welding. It is generally not life-threatening, and resolves shortly after onset (Plum et al., 2010). Occupational Safety and Health Administration (OSHA) regulations limits 8-hour workday air concentrations of zinc chloride fumes to 1 mg/m^3 and zinc oxide fumes to 5 mg/m^3 (Roney et al., 2005). Other symptoms of zinc toxicity include impaired immune function, abdominal cramps, vomiting and nausea (Plum et al., 2010; Valko et al., 2005).

1.2.3.1 Zinc in water bodies

In water bodies zinc is found dissolved as its divalent ionic form (Zn^{2+}), as a solid precipitate, adsorbed to particle surfaces and as part of organic and inorganic complexes. Zinc concentrations usually do not exceed 0.01 mg/L in surface waters or 0.05 mg/L in groundwater (WHO, 2017). Excess zinc exposure is comparatively more damaging to aquatic organisms than to humans. In fish it causes gill epithelium dysfunction and subsequent hypoxia, as well as damaging the hearth, kidney and liver tissue (Ibemenuga, 2013). Industrial production of zinc oxide nanoparticles, widely used in cosmetics, rubbers, paints and for biomedical applications (Jiang et al., 2018), represents a particular threat to aquatic environments today. A review article by Kahru & Dubourguier (2010) examining the toxicological effects of nanoparticles on bacteria, algae, crustaceans and fish classified zinc oxide nanoparticles as “extremely toxic” to aquatic environments (LC50 < 0.1 mg/L), although toxicity varies greatly depending on organism and environmental factors (Wu et al., 2019). The toxicity of zinc oxide nanopaticles is attributed in large part to the release of ionic zinc and oxidative damage (Bacchetta et al., 2016; Wu et al., 2019). Modelled estimates of zinc oxide nanoparticle concentrations in surface waters in Europe and the U.S. range between 0.001 to 0.058 µg/L, and are expected to increase in the future (Gottschalk et al., 2009).

1.2.4 Cadmium: Roles, pollution and human toxicology

Cadmium (Cd) is the 48th element on the periodic table, and the 66th most common metal in the earth’s crust (Haynes et al., 2017). It has wide industrial application due to its high corrosion resistance, high ductility, high thermal and electrical conductivity and low melting temperature. This includes the production of batteries, fertilizers and pesticides, electroplating, glassware pigmentation and stabilization in plastic production (Briffa et al., 2020; Rahimzadeh et al., 2017). It is mainly recovered from zinc ores and concentrates, and in 2020 the global cadmium refinery production was about 23.000 metric tonnes (U.S. Geological Survey (USGS), 2021). It is a non-essential heavy metal and considered highly toxic to humans and the environment at low doses, being reported as WHO guidelines limits cadmium intake to 25 µg/kg body weight and concentrations to 0.003 mg/L in drinking water

(WHO, 2017). Although industrial cadmium pollution has decreased in most industrialized nations, it remains a concern for industrial workers and those living in polluted areas. In less developed nations pollution levels can be much higher (Rani et al., 2014). Its major routes of entry for humans are through inhalation and ingestion, and major sources of exposure are cigarette smoking, fuel combustion and food and water contaminated by cadmium from natural leaching or industrial runoff (Rahimzadeh et al., 2017).

Cadmium has a long biological half-life (approximately 20-30 years in humans) and therefore easily bioaccumulates over long periods even from very low chronic exposure doses (Rani et al., 2014). In polluted areas, cadmium has been shown to accumulate in important food sources including rice, leafy vegetables, shellfish and animal offal, posing a health risk to those who consume them (Chunhabundit, 2016). The International Agency for Research on Cancer (IARC) has classified cadmium as a Class 1 carcinogenic (Rani et al., 2014), and other toxic effects include osteoporosis, impaired cellular respiration, neurotoxicity, impaired immune function, infertility, renal dysfunction and cardiovascular disease (Rahimzadeh et al., 2017; Rani et al., 2014; Valko et al., 2005).

1.2.4.1 Cadmium in water bodies

In water bodies cadmium generally exists as its most common divalent ionic form (Cd^{2+}), and concentrations usually do not exceed 0.001 mg/l in unpolluted natural waters (WHO, 2017). It is highly mobile due to its tendency to form soluble complexes with various organic and inorganic compounds, reducing rates of adsorption and precipitation (Kubier et al., 2019). This leads to increased leaching from natural sediments and solid waste material into water bodies, prolonged exposure of aquatic environments and increased possibility of accumulation in both aquatic organisms and humans.

According to the Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR), average cadmium levels have been reported to range between 0.03-1 mg/kg in marine sediments and up to 5 mg/kg in sediments from rivers and lakes. Average cadmium concentrations range between 10-100 ng/L in European rivers, 5-20 ng/L in open seas and 80 - 250 ng/L in French and Norwegian coastal zones (OSPAR, 2002) . The concentrations of cadmium in water bodies will vary greatly and depends on proximity to sources of pollution. The drinking water of mining communities in northern Ghana have been reported to contain mean cadmium concentrations ranging from 0.023-0.534 mg/L, far exceeding the safe levels (0.003 mg/L) established by the WHO (Cobbina et al., 2015). Along the heavily industrialized eastern coastline of Bahrain, sediment concentrations of cadmium have been reported to range between 0.083-19.14 mg/kg dry weight (Naser, 2013).

1.2.5 Remediation technology: Current shortcomings and the prospects of bioremediation

Currently employed remediation technology for the removal of heavy metals from wastewaters consist of classical physicochemical methods (chemical precipitation, ion exchange, adsorption, filtration, membrane separation and solvent extraction) and electrochemical methods (electrocoagulation, electrodeposition and electroflotation) (Qasem et al., 2021). These technologies have varying drawbacks that prevent them from being employed on a large enough scale to properly counteract anthropogenic heavy metal pollution, including high construction and maintenance costs, lower efficiency under industrial conditions, incomplete metal removal and the production of toxic and/or metal containing waste products that need further treatment (Qasem et al., 2021). There is thus a need for the development of cheaper, more efficient and eco-friendly heavy metal remediation methods, and lately the scientific community has shown interest in the potential use of biological methods of detoxification to answer this challenge (Mantzorou et al., 2018).

Bioremediation refers to remediation methods that utilize microorganisms to remove pollutants from soil, water and air, and represents a cost-effective and eco-friendly alternative to conventional methods of heavy metal removal (Mantzorou et al., 2018). By utilizing the detoxifying mechanisms inherent to certain tolerant microorganism species, heavy metals can be transformed into less toxic forms and/or accumulated in/adsorbed to biomass, allowing for easier elimination from the environments and potential further processing (Prabhakaran et al., 2016). Bioremediation can utilize microorganisms indigenous to the site of pollution or isolated elsewhere to treat polluted soil or water in- and ex-situ (Mantzorou et al., 2018). Among the microorganisms of interest microalgal species, particularly diatoms, have shown great promise due to their cost-effectiveness, inherent tolerance to heavy metals and capability to remove them from the aquatic environment (Mantzorou et al., 2018).

1.3 Phycoremediation

Bioremediation using macro- and microalgae is referred to as phycoremediation (Mantzorou et al., 2018). Microalgae are showing great promise for eco-friendly bioremediation, as they have several advantages over other potential organisms (bacteria, higher plants, fungi). They large surface-volume ratios, are adaptable to environmental conditions with different salinity, pH, temperatures and light, do not require agricultural land for cultivation, have high growth rates, have low nutritional requirements, do not produce toxic by-products and their biomass has potential further uses after cultivation (Kumar et al., 2015; Mantzorou et al., 2018). Phycoremediation of heavy metal pollution in particular shows great promise, as microalgae tend to have high tolerances to heavy metals when alive, demonstrate high yields of recovery per unit of mass and are able to remove metals even at low concentrations (Monteiro et al., 2012). Phycoremediation involves the processes of (1) biosorption, where heavy metals bind to extracellular functional groups found on the cell wall, and (2) bioaccumulation, where the algae actively transport heavy metals into the intracellular space. Biosorption alone can be used to conduct phycoremediation using dead microalgal biomass, offering benefits relating to ease of operation, resistance to wastewater toxicity, recovery of adsorbed metals and reuse of biomass in multiple sorption/desorption cycles (Monteiro et al., 2012). On the other hand, some studies have shown that living microalgae preforming bioaccumulation contained more heavy metals intracellularly than adsorbed extracellularly. (Kumar et al., 2015; Monteiro et al., 2012). Living cells may also be more beneficial for the remediation of large bodies of water containing lower concentrations of contaminants (Kaplan, 2013).

1.3.1 Biosorption

Biosorption is a metabolically inactive process, involving the binding (adsorption) of heavy metals cations to functional groups on the cell-surface (Marella et al., 2020). Since it does not require energy or metabolic activity, both living and dead microalgal biomass are capable of this form of phycoremediation (Monteiro et al., 2012). The cell-wall is the microalgae's first defence against damaging compounds. Furthermore, the microalgal cell surfaces contains macromolecules presenting charged binding sites on functional groups, which increase their resistance to heavy metals. Diatoms cell surfaces contain silanol (SI–O–H), amino ($–NH_2$) and carboxyl (–COOH) groups in addition to aldehydes, ketones, esters, extracellular polymeric substances (which can also be excreted for the same purpose), chelating peptides called phytochelatins (PCs) and protein-binding glycoproteins called frustulins, which form three-dimensional porous structures (Figure 1.1) (Marella et al., 2020; Monteiro et al., 2012; Santos et al., 2013; Xiao & Zheng, 2016). These functional groups impart negative charges on the cell wall which can bind heavy metal cations via counterion interactions, preventing them from entering the cell in excessive amounts (Monteiro et al., 2012). The mechanisms of biosorption include ion exchange, physical adsorption, complexation and precipitation (Mantzorou et al., 2018).

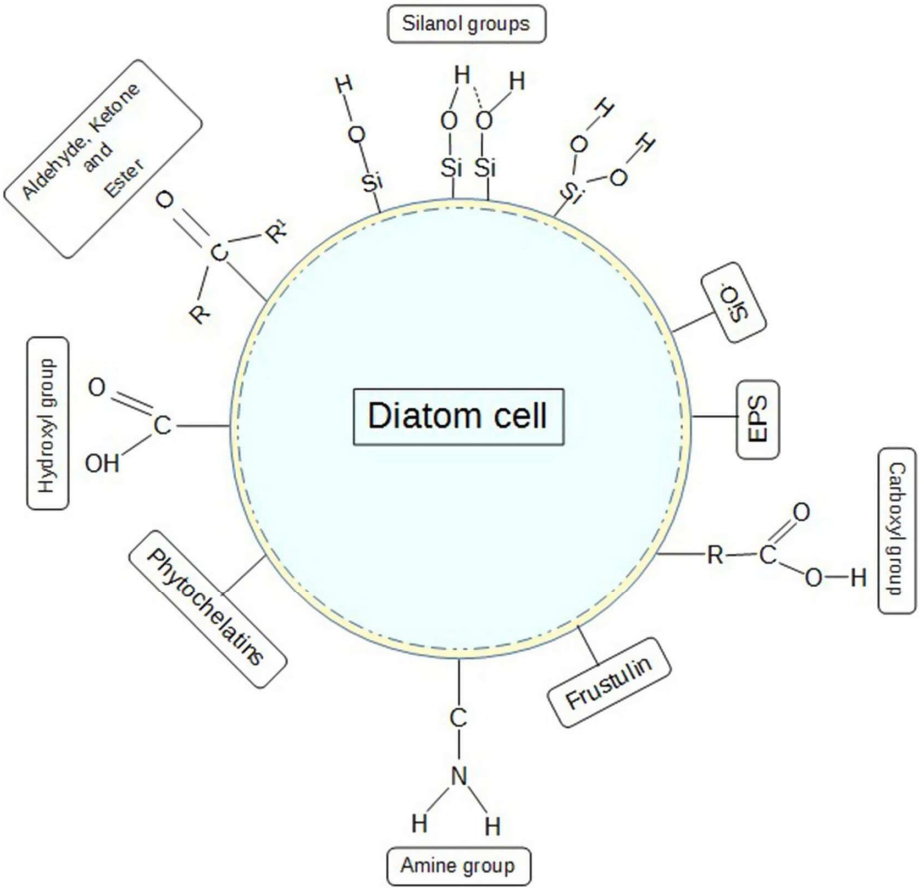


Figure 1.1: Biosorption-facilitating molecules and functional groups that can be found on the surface of diatom cell surfaces. SiO: Silicon monoxide. EPS: extracellular polymeric substances. Figure is reproduced from Marella et al. (2020).

Ion exchange is thought to be the main mechanism of microalgal biosorption (Mantzorou et al., 2018). Normally the functional groups on the cell surface contains light metals ions such as K^+, Na^+, Ca^{2+} and Mg^{2+} or protons at lower pH levels (Brinza et al., 2007; Mantzorou et al., 2018; Monteiro et al., 2012). When the microalgae are exposed to heavy metal cations, these compete with the light metal ions and bind to the functional groups (Brinza et al., 2007). Physical adsorption also occurs via physical forces like Van der Waals and electrostatic interactions between metal ions and extracellular macromolecules, although its contribution to biosorption is minor for most species (Mantzorou et al., 2018). It is possible that some heavy metals form complexes with the functional groups on the cell wall through electrostatic forces and the formation of dipolar bonds (Brinza et al., 2007; Mantzorou et al., 2018). It is thought that the more electropositive a metal ion is, the stronger it will associate with the cell wall (Mantzorou et al., 2018). Precipitation can occur passively due to environmental changes surrounding the microalgal cell surface when heavy metals are bound to functional groups, or actively when microalgae secrete extracellular polymeric substances that bind heavy metals and form precipitates (Mantzorou et al., 2018).

Biosorption also depends on parameters beyond the functional groups present on the cell surface. pH is considered the most significant environmental parameter affecting biosorption because it affects the speciation of the heavy metals, availability of functional groups on the cell surface and competition among metal ions (Mantzorou et al., 2018). The highest biosorption on dead microalgal biomass has been observed to occur in the pH interval between 3 and 6.5, because at such pH values metals are both highly soluble and present in solutions as simple ionic forms (Brinza et al., 2007). This increases both heavy metal adsorption rates and toxicity, a trade-off that must be considered when using live microalgae (Brinza et al., 2007). At pH values lower than 3 protons compete for interaction with functional groups, while pH values above 6.5 tend to facilitate heavy metal precipitation as hydroxides instead of adsorption (Brinza et al., 2007; Monteiro et al., 2012). Very low pH-values can also be harmful to the microalgae (Mantzorou et al., 2018) . Observed effects of temperature on biosorption are weak and sometimes contradictory, and it is generally considered a less important parameter (Mantzorou et al., 2018). Beyond biomass composition, biosorption efficiency can also be affected by biomass density, culture medium composition and the presence of metal cations with similar physico-chemical properties competing for binding to functional groups (Arief et al., 2008; Mantzorou et al., 2018).

1.3.2 Bioaccumulation

Bioaccumulation is an active process, where pollutants are absorbed and accumulated within the cell interior, and can be considered a secondary process to biosorption (Mantzorou et al., 2018). Being an active process means that it requires metabolic activity, and as such can only be performed by living organisms (Mantzorou et al., 2018). The first step of bioaccumulation is the passive sequestration of the compound on cell surface ligands, meaning the biosorption process described in section 1.2.1.1. The second step involves the active transport of the compound across the cell wall (Kumar et al., 2015). Once inside the cell, compounds are bound to intracellular peptides or other molecules and/or moved to intracellular compartments for further detoxification (Danouche et al., 2021). General processes are illustrated in Figure 1.2 (Kaplan, 2013).

1.3.2.1 Heavy metal transport in microalgae

Current understanding of heavy metal transport in microalgae like *P. tricornutum* at the molecular level is largely based on identified similarities with known and putative proteins in plants, fungi, animals and model microalgae like *Chlamydomonas reinhardtii* (Blaby-Haas & Merchant, 2012). Heavy metals are likely transported in and out of algal cells by energy-driven, integral transmembrane metal ion transporter proteins, which control the cellular influx and efflux of metal ions in a concentration dependent manner. (Danouche et al., 2021; Kumar et al., 2015). Several families of heavy metal transporters have been identified in algae, and according to Blaby-Haas and Merchant (2012) these can generally be separated into two groups. Group A refers to metal transporters responsible for moving metal ions into the cytoplasm and include the ZIP (Zrt-, Irt-like Proteins), NRAMP (Natural Resistance-

Associated Macrophage Proteins), CTR (Cu TRansporter) and FTR (Fe TRansporter) families (Blaby-Haas & Merchant, 2012). These transporters increase ion influx into the cytoplasm when the intracellular equilibrium between chelating sites and metal ions are perturbed by metal deficiency. They can be found in the cell membrane and draw metal ions into the cytoplasm from the extracellular space, or in organelle membranes and draw metal ions into the cytoplasm from intracellular storage sites like vacuoles, mitochondria and chloroplasts (Blaby-Haas & Merchant, 2012; Danouche et al., 2021). Group B refers to metal transporters that reduce the concentration of metals in the cytoplasm, and include the P_{1B}-type ATPase, FPN (FerroPortiN), CDF (Cation Diffusion Facilitator) and Ccc1 (Ca(II)-sensitive Cross-Complementer 1)/VIT1 (Vacuolar Iron Transporter 1) families (Blaby-Haas & Merchant, 2012). If cytosolic metal concentrations exceed cellular requirements, or metal-peptide complexes begin affecting cellular metabolism, these efflux transporters discharge metal ions back into the extracellular space or into intracellular storage sites depending on their cellular location (Blaby-Haas & Merchant, 2012; Danouche et al., 2021).

1.3.2.2 Heavy metal detoxification in microalgae

Algae, like other organisms, have developed the ability to synthesize detoxifying metal-binding compounds. Once inside the cell, heavy metals are scavenged by several chelating agents that render the metal ions unavailable for binding to enzymes and other important biomolecules. The most important among these are considered to be the polypeptides metallothioneins (MTs), glutathione (GSH) and PCs (Danouche et al., 2021). Sulfhydryl-groups (-SH) found on cysteine residues present in these polypeptides are thought to play a crucial role in their metal-binding properties, due to carrying a negative charge which increase their affinity towards metal cations (Balzano et al., 2020).

MTs are small (≤300 amino acids), gene-encoded proteins containing high amounts of the amino acid cysteine (15–35% of amino acid content) that are ubiquitous in all living organisms (Balzano et al., 2020). They are be thought to be primarily involved in detoxification during minor heavy metal exposure. (Danouche et al., 2021).

GSH and PCs are enzyme-synthesized polypeptides, as opposed to being gene-encoded like MTs. GSH is a glutamate-cysteine-glycine tripeptide consisting of a glutamate residue bound to the amino group of a cysteine molecule through its carboxyl side chain (γ-glutamylcysteine), which is then bound to a glycine residue by a peptide bond. It is considered the main redox buffer in eukaryotes, and is primarily involved in defence against oxidative and metal stress (Balzano et al., 2020). GSH is also the precursor for the synthesis of PCs, which are made up of multiple units of γ-glutamylcysteine bound to a terminal glycine residue. The general formula for the structure of PCs is (γ-glutamate-cysteine)n-glycine, where n = 1-11 (Balzano et al., 2020; Hirata et al., 2005). The longer the PC chains are, the more stable the complexation with metal ions is (Mantzorou et al., 2018; Perales-Vela et al., 2006). GSH is thought to be the main ligand when intracellular metal concentrations are low, while PCs ensure metal detoxification when concentrations are high (Danouche et al., 2021).

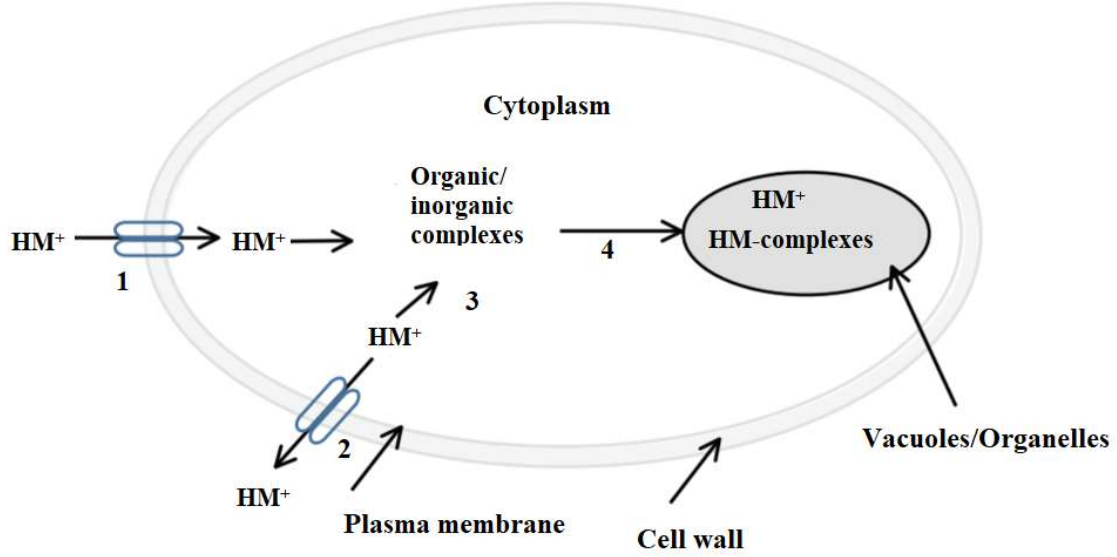


Figure 1.2: Processes of metal transport, detoxification and compartmentalization in microalgae. 1: Active influx of heavy metal ions (HM^+) by transporter proteins. 2: Active efflux of HM^+ by transporter proteins. 3: Detoxifying formation of complexes of heavy metal ions with organic and/or inorganic cell components. 5: Transportation and sequestration of HM^+ or HM-complexes in vacuoles or other organelles. Figure is modified from Kaplan (2013).

1.3.3 The effects of zinc and cadmium on *P. tricornutum*

Cell growth and photosynthetic activity have frequently been used as indicators of cellular stress in autotrophic organisms exposed to heavy metals, as these parameters are highly correlated with overall metabolic health and the proper functioning of many physiological and biochemical processes within the cell (Horvatić & Peršić, 2007; Monteiro et al., 2012; Overnell, 1975; Paunov et al., 2018). Heavy metal toxicity is generally associated with reduced proliferation and growth rates, disturbances of chlorophyll membranes, reduced biosynthesis of chlorophyll and other pigments, interference with the activity of Calvin-cycle enzymes, and disturbances of electron transport chains in Photosystem I and II in both microalgae and plants (Horvatić & Peršić, 2007; Monteiro et al., 2012; Overnell, 1975; Paunov et al., 2018; Sheoran et al., 1990)

1.3.3.1 Cell growth

Horvatić & Peršić (2007) examined the growth rates of *P. tricornutum* in a number of different heavy metal solutions. *P. tricornutum* was grown in medium containing various concentrations of Zn^{2+} and Cd^{2+} and cell density was measured after 72h and 336h of exposure at each concentration. After 72h of exposure, low Cd^{2+} concentrations induced a significant increase in growth rate of *P. tricornutum* compared to the control. Only beginning at a concentration of 5 mg/L and upwards was a reduction in growth rate observed. This is in

accordance with the findings of Torres et al. (1997), where a significant reduction of *P. tricornutum* growth was observed at Cd^{2+} concentrations of 5mg/L and above after four days of exposure. Both studies concluded that *P. tricornutum* was highly resistant to cadmium toxicity (Horvatić & Peršić, 2007; Torres et al., 1997). Zn^{2+} exposure for 72h lead to minor reductions in cell growth at all concentrations, with more notable decline above 10 mg/L (Horvatić & Peršić, 2007). After 336h of exposure, growth rates in moderate concentrations of Zn^{2+} and low concentrations of Cd^{2+} were higher than controls (Horvatić & Peršić, 2007). IC_{50} (concentration at which growth rate of *P. tricornutum* was reduced 50%) was determined to be 5.37 mg/L for Cd^{2+} and 41.85 mg/L for Zn^{2+} after 72h of exposure, and as high as 7560.6 mg/L for Cd^{2+} and 140.63 mg/L for Zn^{2+} after 336 h of exposure (Horvatić & Peršić, 2007). This notable reduction in sensitivity, particularly towards cadmium, show that *P. tricornutum* is strongly capable of adaptation to long term heavy metal exposure in addition to its innate tolerance.

1.3.3.2 Photosynthesis

The effects of heavy metals on photosynthesis can be observed by determining measurements of photosynthetic efficieny like the quantum yields of photosystem I (PSI) and II (PSII), electron transport rates (ETR) and saturating irradiances after exposure to a heavy metal (Maxwell & Johnson, 2000; Ralph & Gademann, 2005). Ji et al. (2018) observed that cadmium inhibited the effective quantum yield and ETR of both PSI and PSII of *P. tricornutum* in a dose-dependent manner. The PSII was shown to be more sensitive to cadmium toxicity than PSI, which is in accordance with previous studies (Ji et al., 2018; Krupa, 1999; Vassilev et al., 2004). Ji et al. (2018) concluded that the reduction in PSI activty was at least in some part caused by the inhibiton of PSII on the donor side and potentially disruption of the Calvin cycle on the acceptor side. They also observed an inhibitory effect of cadmium on the reversible epoxidation of diatoxanthin back to diadinoxanthin, which is an important step of the xanthophyll cycle that confers photoprotection to diatoms (Ji et al., 2018). This has previously been observed in other studies and is thought to contribute to the photoinhibitory effects of cadmium (Bertrand et al., 2001; Ji et al., 2018).

Little seems to be known about the effects of toxic levels of zinc exposure on the photosynthetic activity of *P. tricornutum.* Tripathi & Gaur (2006) exposed the green algae *Scenedesmus* sp. (Chlorophyta) to 5 and 25 µmol/L $ZnCl_2$ (equating to ~ 0.3 and 1.6 mg/L Zn^{2+}, respectively) and measured levels of chlorophyll a (Chl a), chlorophyll b (Chl b) and photosynthetic O_2 evolution at bihourly intervals for 48h. They observed a reduction in cellular levels of both Chl a and Chl b and a reduction in photosynthetic O_2 evolution in the green algae Scenedesmus sp. (Chlorophyta) when exposed to 5 and 25 µmol/L ZnCl2 (equating to ~ 0.3 and 1.6 mg/L Zn^{2+}, respectively). Nguyen-Deroche et al. (2012) examined the effect of 20 µmol/L (equating to ~ 1.3 mg/L) Zn^{2+} on several diatoms and found that it decreased the ETR of the most sensitive species. In higher plants, zinc has also been shown to negatively affect the photosynthetic electron transport processes of the PSII in a similar way to Cd, albeit to a lesser degree (Paunov et al., 2018). It is therefore likely that sufficiently high zinc concentrations also affect the photosynthetic activity of *P. tricornutum.*

1.3.4 Understanding the heavy metal tolerance of *P. tricornutum*

Several studies have attempted to elucidate the molecular mechanisms underlying the impressive heavy metal tolerance, particularly to cadmium, of *P. tricornutum* (Brembu et al., 2011; Ma et al., 2021; Torres et al., 1997). Torres et al. (1997) attributed the cadmium tolerance of *P. tricornutum* to the high levels and long length of synthesized PCs. Ma et al. (2021) also highlighted the presence of high levels of PCs and GSH as an important contributing factor. Recently, efficient genomic analysis has allowed for a deeper understanding of the genetic mechanisms behind the adaptation of *P. tricornutum* to heavy metal exposure. Brembu et al. (2011) conducted a genome-wide profiling of transcriptional changes in *P. tricornutum* when exposed to cadmium at concentrations of 0.123 and 1.230 mg/L. When exposed to 0.123 mg/L of Cd^{2+} only small transcriptional changes were observed, indicating that the cells were able to adapt to the increased Cd^{2+} through non-transcriptional means. At 1.230 mg/L however, a much stronger response was observed. Among the genes that were upregulated by Cd^{2+} exposure, two encode putative metal transporter proteins of the P_{1B}-type ATPase and VIT1/Ccc1 families. These, named *ATPase5-1B* and *VIT1*, were hypothesized to be particularly relevant for the observed cadmium tolerance in *P. tricornutum* by pumping Cd^{2+} out of the cell and sequestering it in the vacuole, respectively (Brembu et al., 2011). This finding was further supported by Ma et al. (2021), who also observed an upregulation of the transcription of *ATPase5-1B* and *VIT1* in response to cadmium exposure, and concluded the encoded proteins play an important role in cadmium homeostasis in *P. tricornutum.*

1.3.4.1 The *VIT1* and *VIT2* genes

VIT1 (NCBI Gene ID: 7197377; https://www.ncbi.nlm.nih.gov/gene) and *VIT2* (NCBI Gene ID: 7197079) genes identified in the sequencing of the *P. tricornutum* genome, encoding proteins whose probable function and designation has been inferred from homology to members of the VIT family in other organisms (Bowler et al., 2008). The vacuolar membrane-localized ion transporter VIT1/Ccc1 has been shown to play an important role in the vacuolar storage of iron in both *Arabidopsis thaliana* seeds and in yeast, and rice homologs of the gene have been shown to encode proteins capable of vacuolar storage of zinc (Kim et al., 2006; Li et al., 2001; Solioz & Vulpe, 1996; Zhang et al., 2012). As both Brembu et al. (2011) and Ma et al. (2021) demonstrated its increased expression when *P. tricornutum* was exposed to cadmium, it was predicted that VIT1 plays an important role in the vacuolar sequestration of cadmium in *P. tricornutum.* VIT2 is predicted to serve a similar purpose, as it is homologous with VIT1 and has also shown increased expression in response to cadmium exposure (Brembu et al., 2011). The presence of multiple VIT-family proteins and/or other genetically dissimilar ion transporters who share similar functions and complement each other have been found in other organisms (Singh et al., 2020; Zhang et al., 2012). This both strengthens the probability of a shared function between VIT1 and VIT2 in *P. tricornutum* and highlights the possibility of genetic and functional redundancy in its metal transporting system.

1.3.4.2 The *ATPase5-1B* gene

The ATPase5-1B gene (NCBI Gene ID: 7203325) encodes a P_{1B}-type ATPase (Brembu et al., 2011). ATPases are another family of membrane-bound metal transporters that are known to be important for the maintenance of ion homeostasis in a broad range of organisms, and particularly for large and toxic heavy metals like cadmium in the case of P-type ATPases (Solioz & Vulpe, 1996). The ATPase5-1B gene shares high similarity with the *A. thaliana* ion transporter HMA4 (Heavy Metal ATPase 4), which is localized in plasma membranes and associated with the cellular efflux of cadmium and zinc in plants (Brembu et al., 2011; Hanikenne et al., 2008). As opposed to being a single copy gene in *A. thaliana*, the closely related *Arabidopsis halleri* encodes three copies of the HMA4 gene, and the hypertolerance of *A. halleri* towards cadmium and zinc has been attributed to the enhanced expression of this gene (Hanikenne et al., 2008). Since *ATPase5-1B* was upregulated in *P. tricornutum* when exposed to cadmium, the encoded protein was predicted to play a similar role in the heavy metal tolerance of *P. tricornutum* by pumping cadmium and zinc out of the cell (Brembu et al., 2011; Ma et al., 2021).

1.4 Genome editing of diatoms using CRISPR/Cas9

Technologies that enable precise gene editing nucleases have improved greatly in recent years, in particular due to the utilization of targeted nucleases like Zinc-Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs), and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) complexes (Yamamoto, 2015). These nucleases can induce DNA double-strand breaks (DSB) at specific sites on a genome, whose induced repair mechanisms can by utilized to enable the creation of targeted nucleotide insertions, deletions and precise sequence changes (Yamamoto, 2015). Cellular repair of induced BSBs occurs by two major pathways: nonhomologous end-joining (NHEJ) and homology-directed repair (HDR) (Lieber, 2010). NHEJ repair happens directly, and without a template, but is error prone and frequently leads to insertion and deletion (indel) mutations of varying length at the site of DSBs (Lieber, 2010; Yamamoto, 2015). As this can lead to shifting of the translational reading frame of the target sequence and cause loss-of-function (or "knockouts") where the activity of its encoded products is disabled, creating DSBs and relying on NHEJ for their repair has become an important technique for functional perturbation studies of genes (Sander & Joung, 2014; Yamamoto, 2015). HDR is a more precise and homology-directed repair pathway that requires a template whose ends are homologous with the ends of the DSB, and can be utilized to introduce desired sequences or induce specific point mutations in a genome (Yamamoto, 2015). Among these technologies, CRISPR has emerged as a remarkably useful tool for precisely and efficiently targeting, editing, modifying, regulating, and marking genomic loci of a variety of cells and organisms in a cost-effective and simplified manner (Doudna & Charpentier, 2014)

1.4.1 CRISPR/Cas9 technology

The CRISPR locus was first described in *Escherichia coli* by Ishino et al. (1987) and later observed in the genome of several bacteria and archaea, where it was identified as part of an adaptive immune system (Doudna & Charpentier, 2014). It is made up of tandem repeats and spacers, where the repeats share the same sequence and the spacers are made up of different sequences derived from invading foreign nucleic acids (Yamamoto, 2015). The CRISPR locus form complexes with RNA-guided CRISPR-associated (Cas) nuclease proteins which recognize and cleave matching sequences of invading nucleic acids (Sander & Joung, 2014). Several bacterial CRISPR systems and Cas proteins have been identified, of which the type II CRISPR system from *Streptococcus pyogenes* has been the most widely studied and utilised for genetic engineering (Doudna & Charpentier, 2014; Sander & Joung, 2014; Yamamoto, 2015). In the naturally occurring type II CRISPR system (Figure 1.4a), foreign nucleotides such as viral DNA or plasmids are incorporated into arrays on the CRISPR locus of the host genome (Sander & Joung, 2014; Yamamoto, 2015). The CRISPR arrays are transcribed and processed into CRISPR RNAs (crRNAs) which contain "protospacer" regions that are complimentary to the foreign nucleotide sequence it was derived from. This crRNA hybridizes to a fixed trans-activating crRNA (tracrRNA), also encoded by the CRISPR system, which can then complex with the Cas9 nuclease to recognize and cleave DNA bearing the protospacer sequences through hybridization (Sander & Joung, 2014).

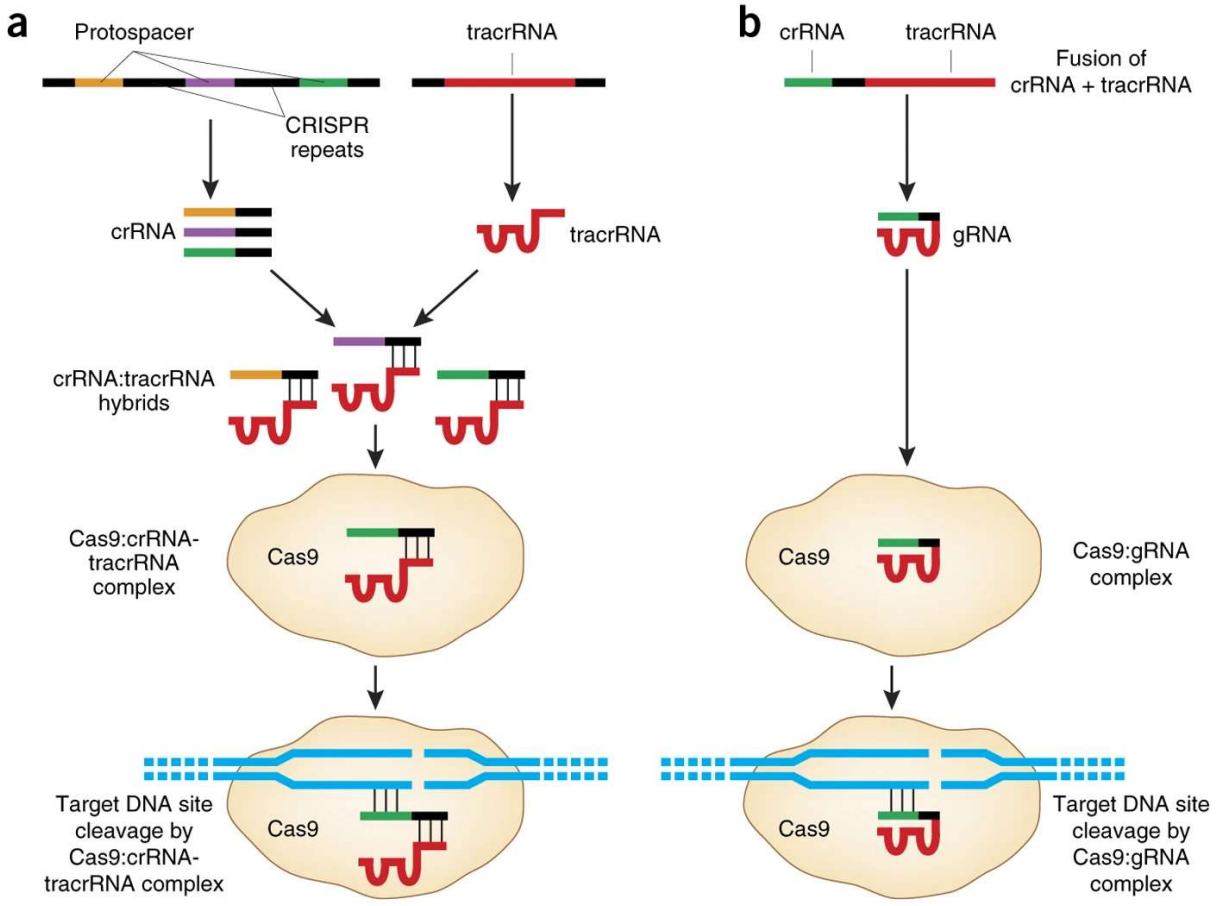


Figure 1.4: The naturally occurring and adapted CRISPR/Cas9 systems. **a)** Naturally occurring CRISPR system. Foreign DNA sequences are incorporated into CRISPR arrays, which produce crRNA with protospacer regions complimentary to the foreign DNA. crRNA hybridizes to fixed tracrRNA also encoded by the CRISPR system, which complexes with the Cas9 nuclease to hybridize to and cleave DNA bearing the same sequence. **b)** Engineered CRISPR system commonly used for gene editing. A fusion of crRNA and tracrRNA called gRNA is utilized, which complexes with Cas9 to cleave DNA sites complementary to the 5′ 20 nt of the gRNA and adjacent to a PAM (protospacer adjacent motif) sequence. Figure is reproduced from Sander & Joung (2014).

The type II CRISPR system from *S. pyogenes* has been adapted for ease-of-use and higher activity when utilized in genetic editing (Sander & Joung, 2014; Yamamoto, 2015). The simplest and most widely used form of the system (Figure 1.4b) is comprised of two components which must be introduced to and/or expressed in target cells or organism to facilitate gene-editing: the Cas9 nuclease and a guide RNA (gRNA) consisting of a fusion between the fixed tracrRNA and a crRNA (Sander & Joung, 2014). The gRNA and the Cas9 nuclease form a complex that recognizes and cleaves at DNA sites that are complementary to the twenty nucleotides at the 5′end of the gRNA, inducing DSB that can be used to form variable indel mutations though the NHEJ pathway or more specific mutations or insertions in conjunction with a single-stranded template through the HDR pathway (Sander & Joung, 2014). For the gRNA/Cas9 complex to successfully bind to and cleave complementary target-DNA sequences, the target site must also lie immediately 5′ of a species-dependent

protospacer-adjacent motif (PAM) sequence, which for the *S. pyogenes* system is 5′-NGG-3′ (Yamamoto, 2015). A good gRNA should not contain poly-T sequences, which can work as a transcriptional terminators, not form secondary structures and avoid being complementary with other sequences on the host genome to avoid off-target editing and subsequent mutation (Yamamoto, 2015). It is also important that the target sequence of the gRNA is positioned on, or upstream of, an important region of the target gene in relation to the function of the translated protein, so that induced mutations result in an inactive gene product. With this system Cas9 nuclease activity can be directed to any DNA sequence adjacent to a PAM site by altering the first 20 nt of the gRNA to correspond to the target sequence (Sander & Joung, 2014).

1.4.2 CRISPR/Cas9 genome editing of the diatom *P. tricornutum*

Several methods have been developed for the transformation of *P. tricornutum* by delivery of vectors that contain all parts of the CRISPR/Cas9 system (Sharma et al., 2018). Biolistic transformation, the delivery of particles coated with vectors directly into host cells using a pressurized gene-gun, has been the most common technique for transformation of diatoms. It creates transgenic lines with stable expression of gRNA and Cas9 and generally result in high mutation efficiencies, but requires high plasmid concentrations, results in relatively low transformation rates and causes cellular damage (Nymark et al., 2016; Sharma et al., 2018). The resulting stable integration of vector DNA components into the host genome may also affect unintended genetic elements, and prolonged high expression of Cas9 has been reported to increase the probability of off-target mutations (Hsu et al., 2013; Sharma et al., 2018). Bacterial conjugation, the transfer of plasmid vectors from specialized bacterial cells to host cells, has recently become a useful alternative tool for the delivery of CRISPR/Cas9 plasmids to *P. tricornutum*, as it allows the plasmid vector to remain as an episome and avoid the permanent and random integration of vector DNA into the host genome (Sharma et al., 2018; Slattery et al., 2018). Through the removal of selection pressure the plasmid can also be eliminated from the host cell when desired, leading to a transient Cas9 expression and reducing the instability of mutant lines (Sharma et al., 2018). Sharma et al. (2018) observed that mutant lines took longer to appear in *P. tricornutum* lines transformed by bacterial conjugation compared to lines transformed by biolistics, and a somewhat reduced mutation efficiency in cells expressing gRNA and Cas9 from an episomal vector.

2 Objectives

The motivation for this project is to better understand the genetic and molecular mechanisms underlying the inherent heavy metal tolerance and accumulation capabilities observed in the diatom *P. tricornutum.* Such knowledge would be important for the continued improvement of microalgal bioremediation technology, which presents itself as one of the most promising solutions to the global challenge of heavy metal polluted waters. Furthermore, identifying genes that enhance resistance to, and accumulation of, heavy metals in diatoms could facilitate the future engineering of hyperaccumulator strains that exhibit remediation capabilities exceeding those found in naturally occurring species.

The aim of this thesis project is to use CRISPR/Cas9 technology to create knockout mutants for the genes encoding the predicted metal transporter proteins VIT1, VIT2 and ATPase5-1B in *P. tricornutum*, believed to confer tolerance to cadmium (and potentially zinc) through vacuolar sequestration (VIT1 and VIT2) and efflux (ATPase5-1B) of metal ions (Brembu et al., 2011). By utilizing traditional cloning techniques, the pPtPuc3m diaCas9_sgRNA plasmid vector (Sharma et al., 2018) will be adapted to contain gRNA sequences that target the genes encoding VIT1, VIT2 and ATPase5-1B. Knockout lines will be created by transforming *P. tricornutum* with these vectors using both biolistic delivery and bacterial conjugation. Mutant strains will be identified by high-resolution melting (HRM) analysis and sanger sequencing. Preliminary characterization of *VIT1*, *VIT2* and *ATPase5-1B* knockout lines will be carried out by comparing cellular growth rate and photosynthetic activity to wild type when exposed to cadmium and zinc. It is hypothesized that knockout mutants will display reduced tolerance to the heavy metals, caused by the inability to effectively remove them from the cytoplasm.

In summary:

1. Create knockout lines of the genes encoding VIT1, VIT2 and ATPase5-1B in *P. tricornutum.*
2. Carry out a characterization experiment on knockout lines to observe potential changes in tolerance towards cadmium and zinc compared to wild type by measuring cell growth and photosynthetic activity.

3 Materials and methods

All experimental work was conducted at the Norwegian University of Science and Technology (NTNU), as a part of the CMBG research group at the Institute of Biology. The experimental design is represented by flowcharts in Appendix A. Details regarding materials, reagents and instruments used over the course of the project are listed in Appendix B.

The first phase of the project was to create CRISPR/Cas9 vectors encoding gRNA that targeted the *VIT1* and *VIT2*, or *ATPase5-1B* genes of interest. Sharma et al. (2018) have developed an efficient plasmid vector for the tranformation of *P. tricornutum* by bacterial conjugation (pPtPuc3m diaCas9_sgRNA, Supplementary Figure D.1) by inserting the Cas9-gRNA cassette from the pKS diaCas9_sgRNA vector developed by Nymark et al (2016) into a modified version of the pPtPuc3 vector (pPtPuc3m) (Karas et al., 2015). The pPtPuc3m backbone contains an origin of transfer (oriT) that facilitates plasmid delivery to *P. tricornutum* cells by bacterial conjugation, and the gRNA cassette contains two *Bsa*I restriction enzyme cutting sites which can be used to make the gRNA target-specific by inserting adapters with target-specific sequences. The plasmid also contains a BleoR protein encoding cassette that confers resistance to the antibiotic Zeocin to be used as a selection marker for transformants (Sharma et al., 2018).

In the second phase *P. tricornutum* was transformed with the targeted vectors using both bacterial conjugation and biolistic bombardment, and CRISPR/Cas9 edited strains with expected loss of function were isolated. The methodology for target specific adapter insertion and subsequent gene editing of *P. tricornutum* was based on pre-existing protocols developed by the CMBG research group (Nymark et al., 2017; Sharma et al., 2018). The final part of the project consisted of characterizing the isolated edited strains of *P. tricornutum* by examining growth and photosynthetic efficiency during heavy metal exposure.

3.1 Creation of gene specific CRISPR/Cas9 vectors

In the first phase of the project, CRISPR/Cas9 vectors were created with gRNA sequences designed to target the *VIT1, VIT2* and *ATPase5-1B* genes., and target specificity was achieved by inserting adapters annealed from oligonucleotides that targeted PAM sites on the genes of interest (Appendix C). A total of four adapters were used targeting two PAM-sites on each gene, with PAM sites on both *VIT* genes being targeted by the same adapter pair. The resulting four plasmid vectors were designated by their target genes and PAM sites: VIT_PAM1, VIT_PAM2, ATPase_PAM1 and ATPase_PAM2

3.1.1 Bacterial amplification of pPtPuc3m_diaCas9_sgRNA plasmid vector

The P. tricornutum specific CRISPR/Cas9 plasmid vector pPtPuc3m_diaCas9_sgRNA (Supplementary figure D.1) described in Sharma et al. (2018) was available in the lab. Bacterial amplification was used to obtain sufficiently high concentrations of the pPtPuc3m_diaCas9_sgRNA vector for the subsequent cloning step. The vector was transformed into competent DH5α-*E.coli* by heat shock treatment. Competent DH5α-*E.coli* (50 µL) were incubated with the pPtPuc3m_diaCas9_sgRNA plasmid (10 ng) on ice for 30 min. The mixture was then incubated in a water bath (42°C) for 45 seconds to achieve heat shock transformation, before being transferred back on ice for another 2 min. Following cooling on ice, 450 µL pre-heated (37°C) LB medium (Appendix B.7) was added to the cells. The mixture was then incubated in a shaker (New Brunswick Scientific) at 220 rpm for 1 hour at 37°C. 250 µL of the mixture was spread over pre-heated (37°C) LB agar plates containing the antibiotic kanamycin (50 µg/mL; Sigma-Aldrich) to select bacterial colonies containing the plasmid, and incubated overnight at 37°C. The following day a single colony from the growth plates was transferred to tubes containing liquid LB medium (3 mL) and kanamycin (50 µg/mL). This culture was then incubated overnight at 37°C with shaking (220 rpm). Plasmids were then isolated from the culture using the E.Z.N.A.® Plasmid Mini Kit I (Omega Bio-tek Inc.) following the manufacturer's instructions. A NanoDrop™ One Spectrophotometer (Thermo Scientific) was used to quantify plasmid DNA by absorbance at 260nm. It was also used to ensure adequate DNA purity after isolation, considering A_{260}/A_{280} absorbance ratios of ≥1.8 as signifying low protein contamination and A_{260}/A_{230} ratios of ≥1.4 as low levels of other organic contaminants (Boesenberg-Smith et al., 2012).

3.1.2 Creation of target specific adapters for *VIT1*, *VIT2* AND *ATPase5-1B*

Adapters were designed by Marianne Nymark at the CMBG group (NTNU Institute of Biology) following the procedures described by Nymark et al. (2016) to target PAM-sites of *VIT1, VIT2* and *ATPase5-1B.* Each adapter was produced by annealing two 24 nt oligonucleotides (Sigma-Aldrich) with 20 nt complementary sequences and 5´ TCGA and AAAC overhangs, as depicted in Figure 3.1 (Sequences in Appendix E.1). The annealing reaction was set up as described in Table 3.1 and incubated on a heating block at 85°C for 10 min. to disrupt possible secondary structures within each oligonucleotide (Nymark et al., 2017). The reaction mixture was then slowly cooled down to room temperature by turning the heating block off while leaving the tubes containing the reaction mixture in the heating block (60-90 min.) to facilitate complementary hybridization (Nymark et al., 2017).

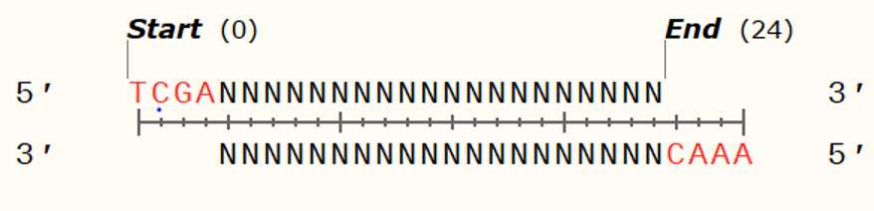


Figure 3.1: General design of adapters. Adapters made by annealing two 24 nt oligonucleotides with 20 nt complementary regions (depicted as black Ns) and resulting 5′ TCGA and AAAC overhangs (red). Figure created in SnapGene® (version 6.0) from GSL Biotech LLC.

Table 3.1: Annealing reaction setup for creation of target specific adapters.

Reagent	Volume (µL)
Forward oligonucleotide (1 µg/µL)	1
Reverse oligonucleotide (1 µg/µL)	1
10x T4 Ligase Buffer	5
ddH_2O	Up to 50

3.1.3 Restriction enzyme digestion of pPtPuc3m_diaCas9_sgRNA plasmid vector

To be ligated with the PAM-site targeting adapters, the pPtPuc3m_diaCas9_sgRNA vector was cut using the restriction enzyme *Bsa*I-HF (NEB) with two *Bsa*I restriction sites placed immediately 5′- to the sgRNA in the pPtPuc3m_diaCas9_sgRNA vector (Supplementary figure D.1). The digestion reaction was set up as described in Table 3.2 and incubated overnight at 37°C to ensure complete digestion of the plasmid. The reaction mixture was then heat inactivated at 65°C for 20 min. before the plasmid was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). Purified plasmid concentrations were determined by absorbance at 260nm using A NanoDrop™ One Spectrophotometer (Thermo Scientific).

Table 3.2: Digestion reaction setup for digestion of pPtPuc3m_diaCas9_sgRNA vector at *Bsa*I restriction sites.

Reagent	Volume (µL)
Vector (82.7 ng/µL)	12.1
*Bsa*I-HF enzyme (20 units/µL)	0.5
10X CutSmart Buffer	2.5
ddH_2O	Up to 25

3.1.4 Ligation of adapters to pPtPuc3m_diaCas9_sgRNA plasmid vector

The target specific adapters were ligated into the *Bsa*I-digested pPtPuc3m_diaCas9_sgRNA plasmid immediately 5´- to the sgRNA using T4 DNA Ligase (NEB). A molar vector to adapter ratio of 1:20 was used, with ligation reaction was set up as described in Table 3.3. The reaction mixture was left at 16°C for about 90 min. before being heat inactivated at 65°C for 10 min. The reaction mixture was then chilled on ice.

Table 3.3: Ligation reaction setup for ligation of pPtPuc3m_diaCas9_sgRNA vector with target specific adapter inserts.

Reagent	Volume (µL)
Vector (82.7 ng/µL)	3.1
Adapter (4 ng/µL)	2.5
T4 DNA Ligase (400 units/µL)	1
10x T4 Ligase Buffer	2
ddH_2O	Up to 20

3.1.5 Transforming *E. coli* with ligated vectors

The ligation mix (Section 3.1.4) was used to transform competent DH5α-*E.coli* cells using the heat shock method as described in Section 3.1.1, using 5 µL of the ligation mix to transform 100 µL of competent cells.

3.1.6 Screening and isolation of plasmids containing inserts

Bacterial colonies that contained a recombinant pPtPuc3m_diaCas9_sgRNA plasmid with target specific adapter inserts were identified by colony PCR screening. Plasmids were subsequently purified from colonies that were identified from the PCR screening and sequenced to confirm adapter insertion.

3.1.6.1 Colony PCR

A selection of 8-16 individual colonies were scraped from the antibiotic selection plates using a pipette tip. Some of each colony was replated, before adding the pipette tip to PCR tubes containing lysis buffer (10 mM Tris-HCl pH 8, 1 mM EDTA, 0.01% Triton X-100) and incubated at 95°C for 10 minutes to lysate the bacterial cells. 5 µL of this lysate was used as template for PCR screening. The PCR reactions were run using the respective forward adapter oligonucleotide and the universal vector specific M13 reverse (Sigma-Aldrich) primer (Supplementary figure D.1), resulting in PCR-products of 484 bp for colonies

containing the pPtPuc3m_diaCas9_sgRNA vector with inserted adapters. The PCR reaction is described in detail in Appendix F.1, and primer sequences are listed in Appendix E.2.

3.1.6.2 Agarose gel electrophoresis

The colony PCR products were analysed by agarose gel electrophoresis to determine which colonies contained vectors with adapter inserts. The electrophoresis was run on a 1% (w/v) agarose gel in 1xTAE buffer 40mM Tris base, 20mM acetic acid, 1mM ethylenediaminetetraacetic acid (EDTA)). The PCR products where mixed with a loading dye (6x: 10mM Tris-HCL (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol and 60 mM EDTA) before being loaded into wells, and GeneRuler^{TM} 1kb DNA ladder Plus (Thermo Scientific) was added to separate wells to determine DNA fragment size. The gel was stained using GelRedTM Nucleic Acid Gel Stain (1:3300 dilution, Biotum, USA) before being analysed using a G BOX Chemi XRQ gel imaging system (Syngene).

3.1.6.3 Plasmid purification

Bacterial colonies that produced PCR products of the expected size for each of the adapters of interest were incubated overnight at 37°C (220 rpm) in 3 mL liquid LB medium (50 µg/mL kanamycin). The following day plasmids were isolated using the E.Z.N.A.® Plasmid Mini Kit I (Omega Bio-tek Inc.) following the manufacturer's instructions. Plasmid concentration and purity was established using NanoDrop^{TM} One Spectrophotometer (Thermo Scientific) as described in Section 3.1.1

3.1.6.4 Sequencing of plasmids

To confirm the gel electrophoresis results, the adapter region of the recombinant pPtPuc3m_diaCas9_sgRNA plasmids were analysed by Sanger sequencing. A sequencing mixture consisting of 400-500 ng plasmid, 2.5 µL M13 reverse primer (µM) and ddH_2O up to 10 µL was set up for each sample and sent to Eurofins Genomics for third party Sanger sequencing using their LightRun^{TM} sequencing service. The sequencing results were then analysed in SnapGene® (version 6.0; GSL Biotech LLC) by multiple alignment with the predicted sequence of the recombinant pPtPuc3m_diaCas9_sgRNA to confirm the presence of the target specific adapter in question.

3.1.7 Verifying plasmid integrity

To confirm the structural integrity of recombinant pPtPuc3m_diaCas9_sgRNA, plasmids were isolated from bacterial cultures as described in Section 3.1.1 and digested with the restriction enzyme *Hind*III (NEB) using the method described in Section 3.1.3 with the NEB buffer (NEB). The digestion product was run on an agarose gel using the method described in Section 3.1.6, and the fragment lengths were compared to the expected lengths based on *Hind*III restriction sites on the intact plasmid (Supplementary Figure D.1; 5874 bp, 3161 bp, 2405 bp, 1178 bp, 190 bp and 11 bp).

3.2 Creating *P. tricornutum* mutant lines

The second phase of the project consisted of transforming *P. tricornutum* with the CRISPR/Cas9 vectors. The diatom cells were transformed following two different bacterial conjugation protocols and by biolistic bombardment. Transformants were isolated on zeocin antibiotic selection plates trough the resistance conferred by the pPtPuc3m_diaCas9_sgRNA vector (Supplementary figure D.1) and screened for lines containing knockout mutations in the target regions of the *VIT1*, *VIT2* and *ATPase5-1B* genes by high-resolution melting (HRM) PCR assays and Sanger sequencing of the region containing the CRISPR/Cas9-sgRNA target site.

3.2.1 Cultivating and counting *P. tricornutum* cells

P. tricornutum were grown in liquid f/2 medium (Guillard's diatom growth medium; Appendix B.6) made with sea salts (20 g/L, Sigma Aldrich S9883), supplemented with f/2 nutrients and vitamins. In subsequent growth on plates, 50% seawater (SW) f/2-Si, 1% (w/v) agar plates were prepared with 0.2 µm-filtered and autoclaved local seawater. Cultures were grown and kept in a white light growth room (65 µmol photon m^{-2} s^{-1}, 16 h photoperiod, 22°C). *P. tricornutum* cells were kept in an exponential growth phase (cell density approx. 0.2-2 x 10^6 cells/mL) through repeated dilutions for approx. 1 week before both conjugation and biolistic transformation was conducted.

P. tricornutum cell counts were performed on culture samples using a NovoCyte™ flow-cytometer (ACEA Biosciences) to determine cellular densities leading up to transformations. The samples were excited by a 488 nm laser to discriminate for cells containing chlorophyll a, and fluorescence emissions were measured by a detector with a 675/30 nm bandpass filter (ACEA Biosciences, 2018).

3.2.2 Bacterial conjugation of *P. tricornutum*

Two different conjugation protocols were followed to transform *P. tricornutum* with the target specific pPtPuc3m_diaCas9_sgRNA vectors. The first protocol had been adapted by the CMBG group from Moosburner et al. (2020) (hereafter referred to as C1), while the second had been developed by the CMBG research group (Sharma et al., 2018) (hereafter referred to as C2).

Both conjugation protocols utilized competent DH10β-*E.coli* cells that had already been transformed with the pTA-MOB plasmid (Appendix D.2). The pTA-MOB plasmid contains all the genes necessary for the conjugative transfer of oriT-containing plasmids, such as pPtPuc3m_diaCas9_sgRNA in this project, and confers resistance to the antibiotic gentamycin as a bacterial selection marker (Strand et al., 2014). The competent DH10β-*E. coli* cells were transformed with the previously isolated target specific plasmids (Section 3.1.6), by the heat shock method (Section 3.1.1), plated on LB-plates containing kanamycin (50 mg/L) and gentamycin (20 mg/L; Sigma Aldrich) to select for bacteria containing both the target specific plasmids and the pTA-MOB mobilization helper plasmid, incubated overnight at 37°C and stored at 4°C for further use.

3.2.2.1 C1 conjugation protocol

In preparation for the day of conjugation 6-well 50% SW f/2-Si 1% (w/v) agar plates containing 5% LB were made, using 8 mL per well. In addition, 1% (w/v) agar f/2-Si selection plates containing zeocin (100 mg/L; Invivogen) were prepared to isolate transformed *P. tricornutum* cultures. 3 zeocin selection plates were made for each transformation. The concentration of *P. tricornutum* was increased leading up to the conjugation to obtain 50 million cells per transformation.

P. tricornutum was grown in an exponential growth phase in liquid f/2 medium for at least a week prior to conjugation. The day before conjugation cell counts of the cultivated *P. tricornutum* were taken using the NovoCyte^{TM} flow-cytometer (ACEA Biosciences), and necessary volumes of culture were harvested to obtain 50 million cells per transformation. The cells were pelleted by centrifugation (first in 50 mL Falcon tubes for 5 min. at 2800g, and then in 2 mL tubes for 5 min. at 5000g) and resuspended in aliquots of 100 µL f/2 for each transformation. The 100 µL concentrated *P. tricornutum* aliquots were added to the centre of separate wells in a 6-well plate, each well containing 8 mL of 50% SW f/2-Si, 5% (v/v) LB and 1% (w/v) agar and spread around by orbital movement. The plates were left to dry completely, before being sealed with parafilm and incubated in the growth room overnight. Colonies of DH10β-*E.coli* containing the recombinant pPtPuc3m_diaCas9_sgRNA plasmid and the pTA-MOB plasmid were transferred from growth plates (stored at 4°C) to 2 mL LB containing kanamycin (50 mg/L) and gentamycin (20 mg/L) and incubated overnight at 37°C (220 rpm).

On the day of conjugation 250 µL of the overnight DH10β-*E.coli* cultures were added to 12.5 mL LB containing kanamycin (50 mg/L) and gentamycin (20 mg/L) in falcon tubes with the lid slightly loose to allow for slight airflow. These cultures were grown until they measured

an optical density (OD_{600}) of 0.8-1 using a SmartSpec™ Plus spectrophotometer (Bio-Rad Laboratories). The cells were then harvested by centrifugation (5 min. at 4000g) and resuspended in 100 µL SOC medium (Appendix B.8), added to a well on the 6-well plates containing *P. tricornutum* from the day before and spread around by orbital motion. The plates were left to dry and resealed with parafilm before being transferred to the growth room for 48 hours.

After 48 hours, 500 µl of f/2 medium was added to each well. The cells were scraped off and the suspension from each conjugation was spread evenly over three 50%SW f/2-Si, 1% (w/v) agar selection plates containing zeocin (100 mg/L; Invivogen). The plates were sealed with parafilm and incubated in the growth room until colonies appeared.

3.2.2.2 C2 conjugation protocol

P. tricornutum cells were grown, counted, and harvested as described in protocol C1 and resuspended in f/2 medium to obtain a concentration of 1.0×10^8 cells/mL. 250 µL of this cell suspension was plated on 50% SW f/2-Si, 1% (w/v) agar plates, which were incubated in the growth room for four days. The cells were scraped off using 500 µL f/2 medium and the concentration was adjusted to 5.0×10^8 cells/mL. DH10β-*E.coli* cells containing a recombinant pPtPuc3m_diaCas9_sgRNA plasmid and the pTA-MOB plasmid were grown as described in protocol C1 until an optical density (OD_{600}) of 0.8-1. The DH10β-*E. coli* cells were spun down (10 min. at 3000g) and resuspended in 500 µL SOC medium. 200 µL of both *P. tricornutum* and *E.coli* suspensions were mixed by pipetting in 1,5 mL Eppendorf tubes, before being plated on 50% SW f/2-Si, 5% (v/v) LB, 1% (w/v) agar plates and incubated in the dark for 90 min at 30 °C. Plates were then incubated in the growth room for two days. After two days the cells were scraped off using both 500 µL f/2 medium and distributed onto three 50% SW f/2-Si,1% (w/v) agar selection plates containing zeocin (100 mg/L). The plates were incubated in the growth room until colonies appeared.

3.2.3 Biolostic bombardment of *P. tricornutum*

P. tricornutum was transformed by biolistic transformation following the protocol described by Nymark et al. (2017).

P. tricornutum was grown in an exponential growth phase in liquid f/2 medium for at least a week prior to transformation. The day before biolistic transformation was conducted, *P. tricornutum* cells were counted using the NovoCyte™ flow-cytometer (ACEA Biosciences), centrifuged for 10 min at 4500 *g*, and resuspended in f/2 medium to a concentration of 1.0×10^8 cells/mL. 500 µL of this cell suspension was plated on 50% SW f/2-Si, 1% (w/v) agar plates and allowed to dry before being transferred to the growth room overnight.

The following day 50 µL aliquots containing 3 mg washed 0.7 µm tungsten microcarriers (Tungsten M10, Bio-Rad Laboratories) were vortexed for 5 min. to disrupt agglomerated particles. While continuing to vortex, 2.5 µg of the target specific vectors, 20 µL spermidine

(0.1 M) and 50 µL $CaCl_2$ (2.5 M) were added to the aliquots before being vortexed for another 3 min. The coated microcarriers were allowed to settle for 1 min. before being spun down briefly in a microfuge, and the supernatant was removed. 140 µL of 70% (v/v) ethanol was added to the microcarriers and removed by pipetting without mixing or centrifuging. The same was repeated with 140 µL of 100% ethanol. The microcarriers were then resuspended in 50 µL 100% of ethanol by carefully tapping the tubes. While vortexing at low speed 10 µL of the coated tungsten particles were removed by pipetting and spread at the centre of a macrocarrier (Bio-Rad Laboratories). This was repeated four more times for each aliquot, until five macrocarriers had been loaded with microcarriers for each plasmid. The macrocarriers were kept at room temperature until the ethanol had evaporated.

The PDS-1000/He biolistic particle delivery system was assembled according to the manufacturer's instructions (PSD-1000/He Biolistic Particle Delivery System manual, Bio-Rad Laboratories). rupture discs (1,550 psi, Bio-Rad Laboratories) were sterilized briefly in 70% (v/v) isopropanol before insertion. The agar plate containing *P. tricornutum* cells was placed in the target shelf 6 cm below the stopping screen (level 2) in the assembled particle delivery system. The plate was bombarded five times, each time moving the plate before the next shot to ensure that subsequent shots hit different parts of the plate. Two plates were used for each of the recombinant pPtPuc3m_diaCas9_sgRNA plasmids. The plates were incubated in the growth room for 24 h.

After 24 h the cells were scraped off each plate using 1 mL f/2 medium, and the suspension was transferred to two 50% SW f/2-Si,1% (w/v) agar selection plates containing zeocin (100 µg/mL). Another 0.5 mL f/2 medium was used to wash off cells from the plates used for biolistic bombardment and transferred to a third zeocin selection plate. The plates were left to dry before being incubated in the growth room until colonies appeared.

3.2.4 Screening for putative gene edited *P. tricornutum* mutants

Transformant colonies that appeared on the zeocin selection plates after conjugation or biolistic transformation were screened for CRISPR/Cas9 induced mutations in the target regions of the *VIT1*, *VIT2* and *ATPase5-1B* genes. High-Resolution Melting (HRM) based PCR assays was used to select transformant colonies with target sequences that differed from the wild type (WT), and Sanger sequencing was used to identify *P. tricornutum* containing knockout mutations in the target region for subsequent isolation and characterization.

Individual primary transformant colonies that appeared on the zeocin selection plates after conjugation or biolistic transformation were transferred to separate wells on 24-well cell culture plates in liquid f/2 medium containing zeocin (50 µg/mL). Plates were sealed with Parafilm and incubated in the growth room for 2 weeks before screening for CRIPSR/Cas9-induced mutations.

3.2.4.1 Preparing algae lysate

The putative mutant colonies' DNA was extracted by lysing algae cells. 100 µL of cell each culture was transferred to a 1.5 mL tube. The tubes were centrifuged for 1 min at 17,000 g, and the supernatant was removed. The pellet was resuspended in 20 µL of lysis buffer (1% triton X-100, 2mM EDTA, 20 mM Tris-HCL, pH 8.0) and vortexed for 30 sec. The tubes were placed on ice for 15 min, before being transferred to a heating block for 10 min at 85°C. The lysate was diluted 1:5 with ddH_2O, and debris was pelleted by brief centrifugation. The lysate was stored at -20°C until further use.

3.2.4.2 PCR amplification of target area

PCR was performed with the algae lysate as template and primers listed in Appendix E.3 to amplify a large amplicon (523 to 800 bp) that contained the target area of interest. The lysate template was run with specific primers designed to amplify the area that contains the target sites of interest (Appendix E) with DreamTaq DNA polymerase (5 U/µL) and 10X buffer (Thermo Scientific). PCR setup is described in Appendix F. An aliquot (5µL) of the PCR product was used for agarose gel electrophoresis (as described in Section 3.1.6) to confirm amplification of the target sequences and to identify large indels.

3.2.4.3 High resolution melting (HRM) analysis of transformants

Specific primers (Appendix E) were used to amplify a short sequence (Approximately 100 bp) containing the target sites for high-resolution melting analysis using the LightCycler 480 High Resolution Melting Master kit (Roche) and a LightCycler 96 instrument (Roche). Another aliquot (5 µL) of the PCR reaction with the large amplicon containing the target area was diluted in a dilution series (up to 1:4,000,000). Reactions for high-resolution melting analysis were set up (Supplementary Table F.5) in 96-well plates by adding 5 µL of the 1:4,000,000 dilution to LightCycler 480 High Resolution Melting Master mix (Roche), $CaCl_2$, and target specific primers. These amplify a nested amplicon (122 to 197 bp) containing the target site and are labelled HRM in Supplementary Table E.3. In addition to the putative gene edited mutants each plate contained non template controls and wild-type (WT) *P. tricornutum* samples. The plates were spun at 1500 g for 2 min before being loaded on a LightCycler 96 instrument (Roche). HRM was performed as indicated in Supplementary Table F.6, and the HRM curves were analysed by comparison to WT using the LightCycler 96 Software (Roche).

3.2.4.4 Sequencing of putative gene edited DNA sequences

P. tricornutum lines that produced different HRM curves than WT were characterized by Sanger sequencing to confirm the presence of indels. 5 µL of the larger PCR products for each putative mutant was purified with the ExS-Pure™ Enzymatic PCR Cleanup Kit (NimaGen) following the manufacturer's instructions, which produced 7 µL of purified PCR-product. Each product was sequenced with a forward and reverse primer (appendix E.4) by combining 3.5 µL samples of the purified PCR product with 2.5 µL primer in a total volume of 10 µL (as described in Section 3.1.6). PCR products that produced multiple bands after gel electrophoresis were instead excised directly from the gel, with each band being purified using the Wizard® SV Gel and PCR Clean-Up System Kit (Promega) according to the manufacturer's instructions. This produced 30 µL of purified PCR product, from which 7.5 µL was added to 2.5 µL of primer and was sent for sequencing (as described in Section 3.1.6). The sequencing data was analysed using the online ICE (Inference of CRISPR Edits) CRISPR analysis tool (https://www.synthego.com/products/bioinformatics/crispr-analysis), to characterize induced mutations, and detect whether potential multiple cell lines where present in a primary transformant colony.

3.2.5 Isolating *VIT2*, *VIT2* and *ATPase5-1B* mutant lines originating from single cells

The *P. tricornutum* lines that were confirmed to contain mutations of interest from Sanger sequencing were heavily diluted (~200 cells/mL). Different volumes (250 µL and 500 µL) of the diluted suspensions were spread on f/2-Si, 1% (w/v) agar selection plates containing zeocin (100 µg/ml) to isolate single cells. The plates were incubated in the growth room until individual colonies appeared, and the screening process (Section 3.2.5) was repeated on these colonies.

3.3 Characterization of mutant lines

The third phase of the project aimed to characterize selected *VIT1*, *VIT2* and *ATPase5-1B* gene edited mutant lines. These lines, along with WT, were grown in medium with and without heavy metals, and the effect on growth kinetics and photosynthetic activity were compared. Two VIT2_PAM1 (V1 and V2) and two ATPase_PAM2 (A1 and A2) mutants were selected for characterization.

3.3.1 Exposure of *P. tricornutum* to zinc and cadmium

A week prior to the heavy metal exposure experiment, mutant lines were transferred from zeocin selection plates to liquid f/2 medium without zeocin for acclimation and kept in an exponential growth phase by regular dilutions. *P. tricornutum* knockout mutant lines (V1, V2, A1 and A2) and WT were then grown in 50 mL flasks, with a starting concentration (Day 0) of ~25,000 cells/mL. Each of the five groups were grown in three different media (treatments): standard f/2 medium (control; CT), f/2 medium containing 5 mg Zn^{2+}/L prepared from a 1000 times concentrated stock solution of $ZnSO_4.7\ H_2O$ and f/2 medium containing 1.25 mg Cd^{2+}/L prepared from a 1000 times concentrated stock solution of $CdSO_4.8/3\ H_2O$. Three replicates of each group and treatment were set up, for a total of 45 flasks. The flasks were incubated in a 24 h photoperiod white light growth room (55-70 µmol photon $m^{-2}\ s^{-1}$, 22°C) for the duration of the experiment. The positioning of the flasks was rotated each day of the experiment to ensure even light exposure. Every 24 hours, a sample of each flask (1 mL) was harvested by pipetting and cells were counted using the NovoCyte™ flow-cytometer (ACEA Biosciences) as described in Section 3.2.1. This was done for 8 consecutive days (Day 1-8), with a single final measurement of each flask taken at Day 17. At Day 3 of the growth experiment a sample from each flask was harvested (2 mL) by pipetting and variable Chl *a* fluorescence was measured using a pulse amplitude modulation (PAM) fluorometer (Multi-Color-PAM, Walz).

3.3.2 Cell growth parameters

When cells are in exponential growth, the rate of cell increase per unit of time is proportional to the number of cells present in the culture at the beginning of any unit of time (Wood et al., 2005). The cell counts obtained from flow cytometry were used to determine population size (N) at each day of measurements and used to calculate the specific growth rate (K') for each time interval (day) using Equation 1. N_0 is the population size at time interval t_0, and N_n is the population size at time interval t_n. The minimum and maximum population sizes from time intervals displaying K' values representing exponential growth phases of the growth curves were used to calculate the maximum specific growth rate K'_{max} using Equation 2. The relative difference in the K'_{max} of mutant lines grown in CT medium and medium containing heavy metals (HM) was compared to those of the WT by calculating the relative K'_{max} using Equation 3.

$$\text{Specific growth rate} = K' = \frac{\ln\left(\frac{N_n}{N_0}\right)}{(t_n - t_0)} \quad (1)$$

$$\text{Maximum specific growth rate} = K'_{max} = \frac{\ln\left(\frac{N_{exp\,max}}{N_{exp\,min}}\right)}{(t_{exp\,max} - t_{exp\,min})} \quad (2)$$

$$\text{Relative } K'_{max} = \frac{K'_{max}Mut(+H\ \ )/K'_{max}Mut(-HM)}{K'_{max}WT(+H\ \ )/K'_{max}WT(-HM)} \quad (3)$$

The effect heavy metal exposure had on the growth of mutant lines compared to WT on the final day of measurement (Day 17) was determined by calculating the relative cell count difference using Equation 4. Relative cell count difference was determined by first calculating the proportional difference in N between a mutant line (Mut) in heavy metal solution (+HM) and the same mutant line in CT medium (-HM), and then calculating the proportional difference between this number and the same number calculated for WT exposed to the same heavy metal and CT.

$$\text{Relative cell count difference} = \frac{N_{D17}\text{Mut(+HM)}/N_{D17}\text{Mut(}-\text{H}\quad)}{N_{D17}\text{WT(+HM)}/N_{D17}\text{WT(}-\text{HM)}} \qquad (4)$$

3.3.3 Variable chlorophyll *a* fluorescence

Variable Chl *a* fluorescence was measured using a pulse amplitude modulation (PAM) fluorometer (Multi-Color-PAM, Walz). The 2 mL samples collected from the growth experiment on day 3 were added to clear-sided cuvettes, and dark-acclimated for 180 seconds. LEDs (light emitting diodes) were used to excite the samples with a weak and nonactinic measuring light and a strong saturation flash (>2000 µmol photons m^{-2} s^{-1}). The samples were then excited by a 440 nm blue light for 30 second intervals with incremental irradiance levels (0-1313 µmol photons m^{-2} s^{-1}; Appendix H). The minimal fluorescence from the dark-adapted samples exposed to the measuring light (F_0) and the maximum fluorescence from dark-adapted exposure to the saturation light (F_M) were obtained from the PAM-measurements using the PamWin-3 Software (Walz). The results of the incremental irradiance exposure were used to create rapid light curves (RLCs; Ralph & Gademann, 2005) with the PamWin-3 Software, which were fitted according to the model described by Eilers & Peeters (1988). F_0 and F_M were used to calculate the maximum quantum yield of PSII ($\Phi_{PSII\ max}$) according to Equation 5, which is a measure of the maximum efficiency of PSII if all its reaction centres were open (Maxwell & Johnson, 2000). The plateau of the RLCs were used to determine the maximum relative electron transport rate $rETR_{max}$, used as a measurement of maximum photosynthetic capacity (Ralph & Gademann, 2005). The intercept of the slope of the RLC in the initial light limiting region (α) and $rETR_{max}$ was used to determine the minimum saturating irradiance (E_k), the irradiance at which photosynthesis is saturated (Ralph & Gademann, 2005).

$$\text{Maximum quantum yield of PSII} = \phi_{PSII\ max} = \frac{F_m - F_o}{F_m} \qquad (5)$$

3.3.4 Statistical analysis

Statistical analysis was conducted in SPSS Statistics 28.0.1 (IBM). One-way ANOVA with Tukey's post hoc test was conducted to identify statistically significant differences in cell growth (K'_{max}) and photosynthetic efficiency ($\Phi_{PSII\ max}$, $rETR_{max}$ and E_k) parameters between all lines from the CT, Zn^{2+} and Cd^{2+} treatment groups, and between each individual line grown in CT compared to when grown in Zn^{2+} and Cd^{2+}. The statistical significance level was set to $P < 0.05$ for all analyses.

4 Results

The aim of this thesis was to investigate if the to investigate if the metal transporters putatively encoded by the *VIT1*, *VIT2* and *ATPase5-1B* genes of the marine diatom *P. tricornutum* play a role in its tolerance and accumulation of the heavy metals zinc and cadmium, by using CRISPR/Cas9 technology to create mutants for these genes. This section describes the results of that experimental work.

4.1 Creation of target specific CRISPR/Cas9 vectors

Marianne Nymark from the CMBG group (NTNU Institute of Biology) designed adapter oligonucleotides that targeted selected PAM-sites on *VIT1, VIT2* and *ATPase5-1B* (Appendix C.1-C.3), using the methods described by Nymark et al. (2016). Four adapters were created to target two PAM-sites on each gene (Appendix E.1) Both *VIT* genes were targeted by the same adapters because, due to their highly similar sequences, they were suspected of having redundant function and it was predicted that knocking out both genes simultaneously might be necessary to observe changes in heavy metal tolerance. Using traditional restriction-ligation cloning, these adapters were inserted into the pPtPuc3m_diaCas9_sgRNA plasmid vector (Supplementary Figure D.1) which has been designed and optimized for CRISPR/Cas9-induced gene editing in P. tricornutum by bacterial conjugation (Nymark et al., 2017; Sharma et al., 2018). After amplification in competent DH5α-*E.coli*, eight colonies were screened by PCR to identify colonies containing pPtPuc3m_diaCas9_sgRNA plasmids that had integrated the respective adapter. The colony PCR screening identified band patterns an amplicon of the expected size of approximately 500 bp for most of the colonies (Appendix G.) from which two colonies were selected for further sequencing.

4.1.1 Plasmid sequencing to confirm target specific adapter insertion

To confirm insertion of adapter in the plasmid indicated by the colony PCR screening, the region covering the gRNA cassette within the pPtPuc3m_diaCas9_sgRNA plasmids of selected colonies was sequenced. Multiple sequence alignment of the sequences that were obtained and the predicted sequence of the pPtPuc3m_diaCas9_sgRNA plasmid containing the desired adapter (Figure 4.1 and Figure 4.2), showed that 24 bp sequences corresponding to the predicted adapter sequences had been inserted into all the plasmids. The resulting plasmid vectors were named by target gene and PAM site: VIT_PAM1, VIT_PAM2, ATPase_PAM1 and ATPase_PAM2.

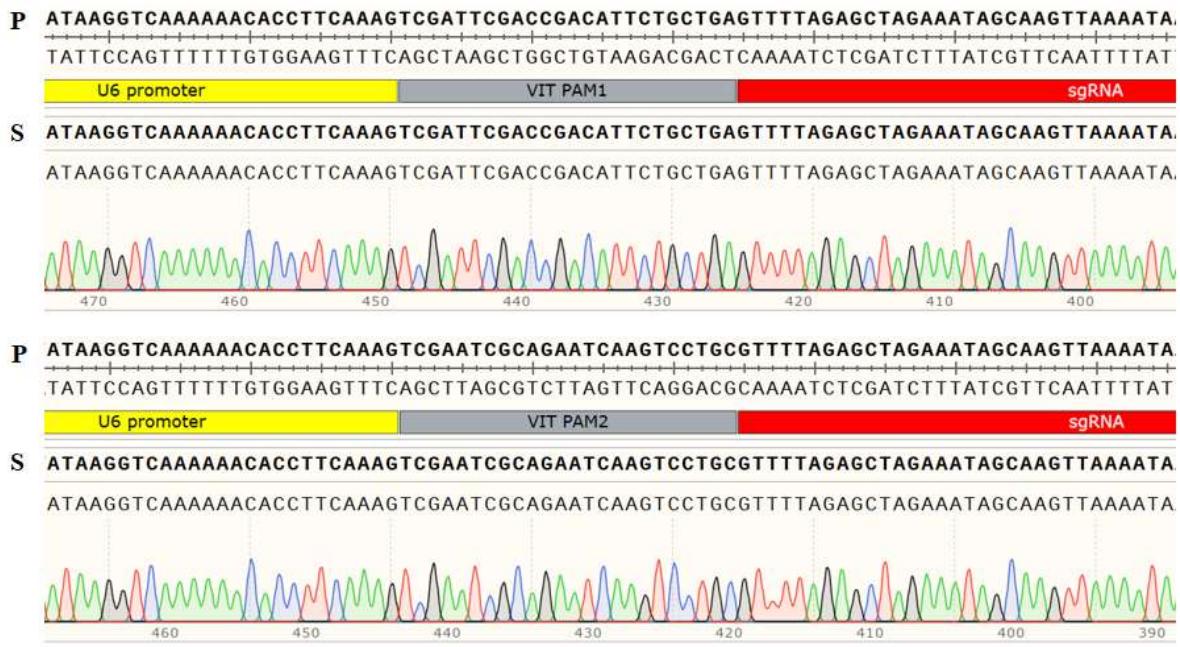


Figure 4.1: Sequencing results for pPtPuc3m_diaCas9_sgRNA plasmid containing inserts targeting PAM-sites on the *VIT1* and *VIT2* genes (VIT_PAM1 and VIT_PAM2 plasmids: S) including the relevant portion of the sequencing chromatogram, aligned with predicted sequence of the pPtPuc3m_diaCas9_sgRNA gRNA cassette containing the expected adapter (P). Alignment and figure produced by SnapGene® software (version 6.0) from GSL Biotech LLC.

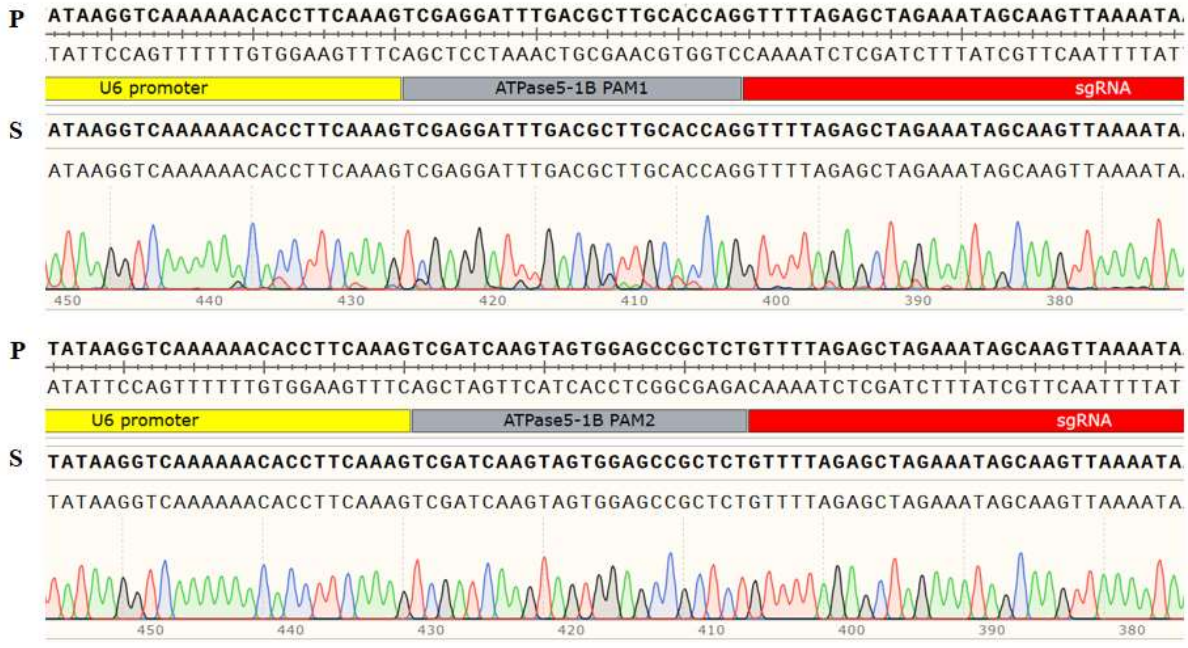


Figure 4.2: Sequencing results for pPtPuc3m_diaCas9_sgRNA plasmid containing inserts targeting PAM-sites on the *ATPase5-1B* gene (ATPase_PAM1 and ATPase_PAM2 plasmid; S) including the relevant portion of the sequencing chromatogram, aligned with predicted sequence of the pPtPuc3m_diaCas9_sgRNA gRNA cassette containing the expected adapter (P). Alignment and figure produced by SnapGene® software (version 6.0) from GSL Biotech LLC.

4.1.2 *Hind*III enzymatic digestion to confirm plasmid stability

As the structural integrity and rearrangement of the pPtPuc3m_diaCas9_sgRNA containing target specific adapters can be an issue (Sharma et al., 2018), they were digested with the restriction enzyme *Hind*III, and the resulting digestion product was run on an agarose gel. The fragment lengths of the digestion product (Figure 4.3) were compared to those predicted from the *Hind*III restriction sites on the intact pPtPuc3m_diaCas9_sgRNA plasmid. Bands were observed that corresponded to the predicted 5874 bp, 3161 bp, 2405 bp, 1178 bp, 190 bp and 11 bp fragments for all target specific plasmid vectors, verifying their structural integrity and the absence of rearrangement.

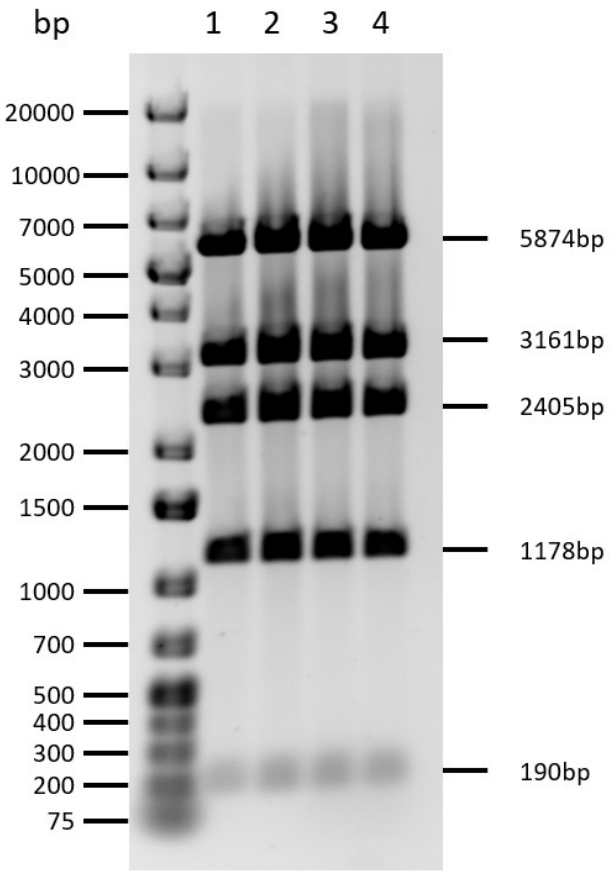


Figure 4.3: Gel picture from *Hind*III enzymatic digestion of target specific plasmid vectors. Bands correspond to 5874 bp, 3161 bp, 2405 bp, 1178 bp and 190 bp (11 bp not visible on gel) fragments expected from *Hind*III restriction sites on the intact original pPtPuc3m_diaCas9_sgRNA plasmid. 1: VIT_PAM1, 2: VIT_PAM2, 3: ATPase_PAM1, 4: ATPase_PAM2.

4.2 Screening, isolation and sequencing of *P. tricornutum* mutant lines

The isolated pPtPuc3m_diaCas9_sgRNA containing target specific adapters (VIT_PAM1, VIT_PAM2, ATPase_PAM1 and ATPase_PAM2) were used to transform *P. tricornutum* cells. The plasmids encoded the necessary products to induce CRISPR/Cas9 gene editing at the target sites of the *VIT1*, *VIT2* and *ATPase5-1B* genes of interest on the diatom genome and create mutant lines. *P. tricornutum* cells were initially transformed by bacterial conjugation following the C1 protocol (Section 3.2.2.1), before being grown on selection plates containing Zeocin to select for colonies expressing the *BleoR* gene present on the pPtPuc3m_diaCas9_sgRNA plasmid (Supplementary Figure D.1). After approximately 3-4 weeks, primary colonies were large enough to be transferred to liquid medium for further cultivation, propagation and characterization. The number of obtained colonies remained low despite several repeated conjugations and growth periods utilising fresh selection plates, causing significant time delays. The decision was therefore made to transform *P. tricornutum* with the target specific plasmids utilizing a different conjugation protocol established in the lab (Sharma et al., 2018) (C2; Section 3.2.2.2). Finally, biolistic bombardment (3.2.3) was performed with the same plasmids to ensure sufficient colonies. These primary transformed colonies were similarly grown on zeocin selection plates and transferred to liquid medium after approximately 3-4 weeks.

4.2.1 HRM screening and sequencing of primary *P. tricornutum* colonies

These colonies of *P. tricornutum* cells were then screened for mutations at their target sites by HRM analysis. Therefore, genomic DNA was extracted and used as a template for PCR amplification of large amplicons (523 to 800 bp) containing the target sites of interest. These PCR products served then as templates for HRM analyses of shorter DNA sequences (122 to 197 bp) containing the target sites of interest. The number of primary colonies screened for the different transformation methods and the apparent mutation efficiency based on the number of colonies displaying a melting behaviour differing from WT during HRM are presented in Table 4.1.The C1 protocol produced only 0-5 colonies per plate, while the C2 protocol produced 20-50 colonies per plate. Biolistics produced inconsistent numbers of colonies for each plasmid (2-<50 per plate). Overall the mutation efficiency was low for colonies transformed through bacterial conjugation and varied between 2.7 and 20%. This led to several rounds of HRM analyses being required to identify enough putative mutants. The decision to use biolistics was made due to the low number of putative mutants identified for the ATPase-PAM1 (1) and ATPase-PAM2 (2) plasmids. When using biolistics the mutation efficiency appeared higher for colonies transformed with the VIT_PAM1 and ATPase_PAM1 plasmids, while no colonies were mutated by the VIT_PAM2 or ATPase_PAM2 displayed divergent melting behaviour from WT. Larger PCR amplicons pertaining to the colonies displaying divergent melting behaviour, in addition to some from randomly chosen colonies which did not display divergent melting behaviour, were sequenced, and the Sanger chromatogram trace files were analysed using the Synthego ICE CRISPR analysis tool which deconvolutes traces and infers indels in the presence of sequence mixtures (Conant et al., 2022). Only some of the putative mutant colonies did in fact contain indel edits in their target

region. Sequencing of some colonies that did not display divergent melting behaviour during HRM analysis also showed the presence of apparent indel mutations. Furthermore, the sequencing chromatograms of all the colonies containing edits displayed moderate-to-high background levels, and the Synthego ICE CRISPR analysis tool identified them as containing mixtures of different indel mutations in the target region as well as WT. In total 9 VIT_PAM1 and 3 ATPase-PAM1 colonies from biolistic transformation and 2 APTase-PAM2 colonies from conjugation were selected for dilution and further isolation of individual cell lines due to sequencing revealing the presence of frameshifting indel mutations.

Table 4.1: Mutation efficiency as indicated by HRM analysis of primary colonies of *P. tricornutum* colonies transformed by bacterial conjugation and biolistics.

Transformation method	Plasmid used	Number of primary colonies screened	Number of putative mutants identified	Apparent mutation efficiency (%)
Conjugation	VIT_PAM1	34	3	8.8
	VIT_PAM2	25	5	20
	ATPase_PAM1	37	1	2.7
	ATPase_PAM2	21	2	9.5
Biolistics	VIT_PAM1	25	7	28
	VIT_PAM2	2	0	0
	ATPase_PAM1	7	3	42.9
	ATPase_PAM2	25	0	0

4.2.2 Isolation of individual gene-edited *P. tricornutum* colonies and further characterization of indels by HRM and sequencing

To separate gene-edited lines from each other and any WT cells in the primary colonies and isolate individual lines with bi-allelic mutations, the colony cultures were highly diluted and replated on selection medium. A selection of 4 VIT_PAM1, 17 ATPase-PAM1 and 2 ATPase-PAM2 individual colonies appeared after 3-4 weeks and were than submitted to HRM analysis and sequencing as before. In the end four mutant lines were established: The mutant lines called A1 and A2, both of which contained indels on the *ATPase5-1B* gene and were edited by CRISPR/Cas9 activity conferred from the ATPase_PAM2 episomal plasmid delivered through bacterial conjugation. The lines called V1 and V2, both of which contained indels in the *VIT1* and the *VIT2* gene induced by the VIT_PAM1 plasmid delivered through biolistics. All these mutant lines displayed a divergent melting behaviour during HRM analysis when compared to the DNA fragments of the target area in WT. The A1 mutant DNA fragment displayed a melting curve containing two peaks, both with lower melting temperatures than WT (Figure 4.4). The A2 mutant displayed a slightly higher melting point than WT (Figure 4.5). The V1 mutant displayed a melting curve with two peaks for the DNA fragment containing the target area of the *VIT1* gene: One right shifted peak and a small left shifted peak with compared to WT. For the DNA fragment containing the *VIT2* gene, the melting curve was right shifted compared to WT (Figure 4.6). The V2 mutant displayed a melting curve with two peaks, one with a lower and one with a higher melting temperature

than WT for the *VIT1*-fragment, and a melting curve comparable to that of the V1 mutant for the *VIT2*-fragment (Figure 4.7).

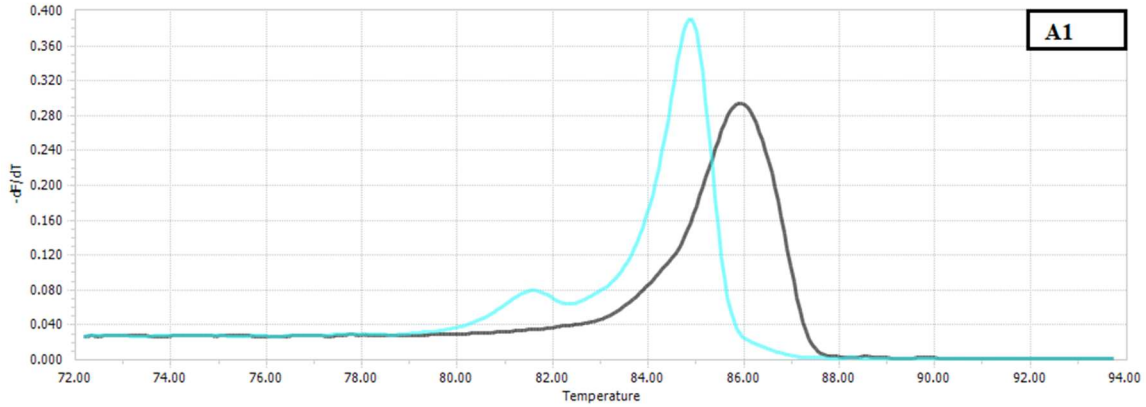


Figure 4.4: Normalized melting curves generated from high resolution melting (HRM) analysis of PCR products containing the target area of the ATPase-PAM2 plasmid on the *ATPase5-1B* gene, amplified from DNA extracted from *P. tricornutum* cell lines. The light blue line represents the melting curve of the A1 mutant line (edited by the ATPase-PAM2 plasmid). The black line represents the melting curve of the same PCR product from WT for comparison. Generated with LightCycler 96 Software (Roche).

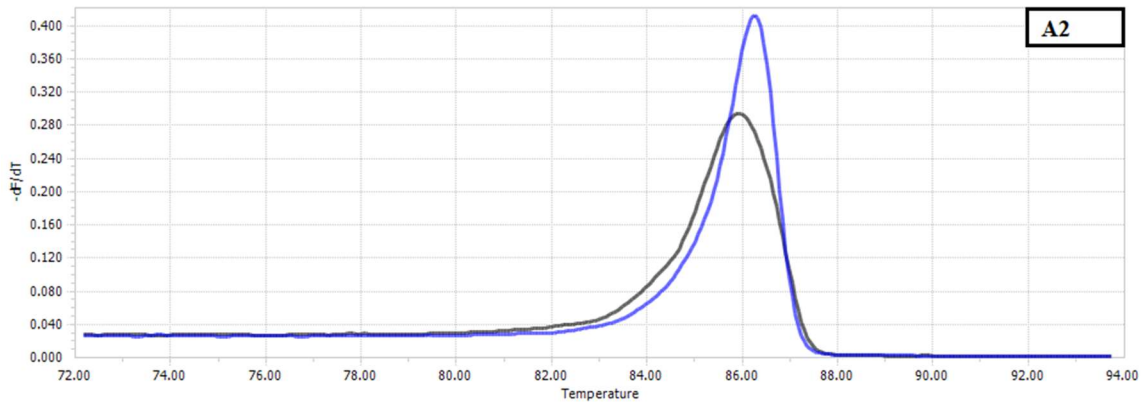


Figure 4.5: Normalized melting curves generated from high resolution melting (HRM) analysis of PCR products containing the target area of the ATPase-PAM2 plasmid on the *ATPase5-1B* gene, amplified from DNA extracted from *P. tricornutum* cell lines. The blue line represents the melting curve of the A2 mutant line (edited by the ATPase-PAM2 plasmid). The black line represents the melting curve of the same PCR product from WT for comparison. Generated with LightCycler 96 Software (Roche).

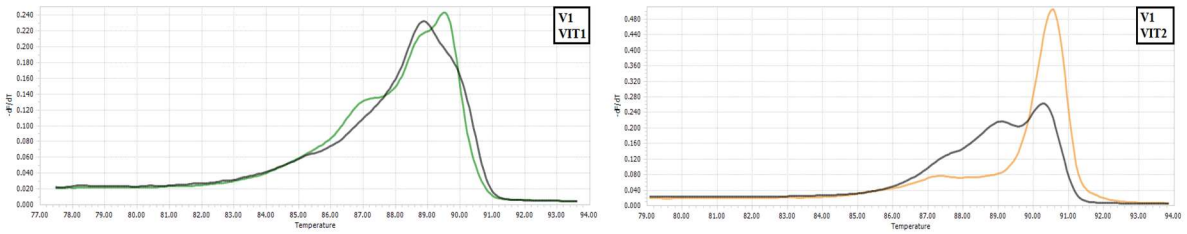


Figure 4.6: Normalized melting curves generated from high resolution melting (HRM) analysis of PCR products containing the target area of the VIT-PAM1 plasmid on the *VIT1* and *VIT2* genes, amplified from DNA extracted from *P. tricornutum* cell lines. The green line represents the melting curve of the V1 mutant line (edited by the VIT-PAM1 plasmid) for the *VIT1* gene, while the orange line represents the V1 melting curve for the *VIT2* gene. The black lines represent the melting curves of the corresponding PCR products from WT for comparison. Generated with LightCycler 96 Software (Roche).

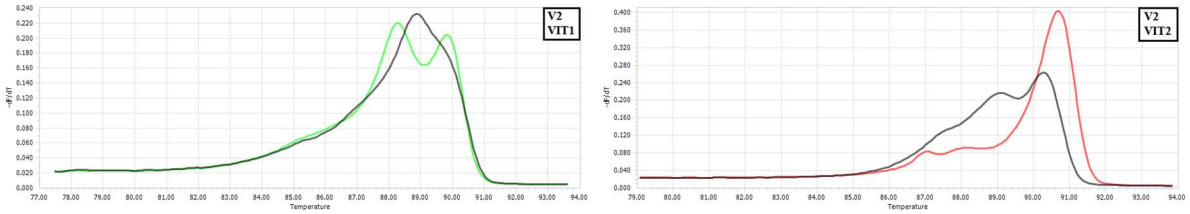


Figure 4.7: Normalized melting curves generated from high resolution melting (HRM) analysis of PCR products containing the target area of the VIT-PAM1 plasmid on the *VIT1* and *VIT2* genes, amplified from DNA extracted from *P. tricornutum* cell lines. The light green line represents the melting curve of the V2 mutant line (edited by the VIT-PAM1 plasmid) for the *VIT1* gene, while the red line represents the V2 melting curve for the *VIT2* gene. The black lines represent the melting curves of the corresponding PCR products from WT for comparison. Generated with LightCycler 96 Software (Roche).

The four mutant lines were sequenced by Sanger sequencing of their larger PCR products amplified preceding HRM analysis, and the sequencing data was analyzed using the Synthego ICE CRISPR analysis tool. Sequencing analysis indicated that some of the mutant lines were not yet completely isolated from other cell lines in the selected colonies, but due to time constraints the decision was made to move forward with the project without further repeating the isolation protocol. The A1 (ATPase_PAM2) mutant line produced a chromatogram with some background noise showing a 13 bp deletion(Figure 4.8). The ICE CRISPR analysis reported that 96 % of the sequencing sample was made up of amplicons containing the same 13 bp deletion, while the remaining 4 % were made up of amplicons containing various smaller deletions (data not shown). The A2 (ATPase_PAM2) mutant line produced a chromatogram showing a 2 bp insertion, with smaller levels of background noise following the site of the insertion (Figure 4.9). ICE CRISPR analysis reported that 94% of the sample was made up of amplicons containing the 2 bp insertion, and the remaining 6% were comprised of amplicons containing a 7 bp deletion (data not shown). The V1 (VIT_PAM1) mutant produced a clear chromatogram containing a 216 bp insertion from the amplicon sample containing the *VIT2* target site, which was confirmed to be the only sequence present in the sample by the ICE CRISPR analysis (Figure 4.10). A NCBI BLAST Analysis of the 216 bp sequence showed it to be a partial sequence from a predicted voltage-gated chloride channel in the *P. tricornutum* genome (NCBI Gene ID: 7195144; https://blast.ncbi.nlm.nih.gov/Blast.cgi) with a 100 % nucleotide match. The amplicon sample containing the *VIT1* target site, however, displayed a chromatogram with high background noise (not depicted) and was inferred to be made up of amplicons containing a 1 bp insertion (70%), a 3 bp deletion (22%) and without edits (WT)(data not shown). The V2 (VIT_PAM1) mutant produced a clear chromatogram showing a 1 bp insertion from the PCR product containing the *VIT2* target site, confirmed to be the only sequence present by the ICE CRISPR analysis (Figure 4.11). Similarly to the V1 mutant, the V2 sample of the amplicon containing the *VIT1* target site produced a chromatogram with high background noise (data not shown) and ICE CRISPR analysis predicted it to be made up of a mixture of amplicons containing two different 12 bp deletions (32 % and 13 %), 3 bp deletions (18 %), 1 bp insertions (23%) and without edits (WT; 7 %). The observed insertions and deletions in all the mutant lines occurred as expected in or close to their respective target site.

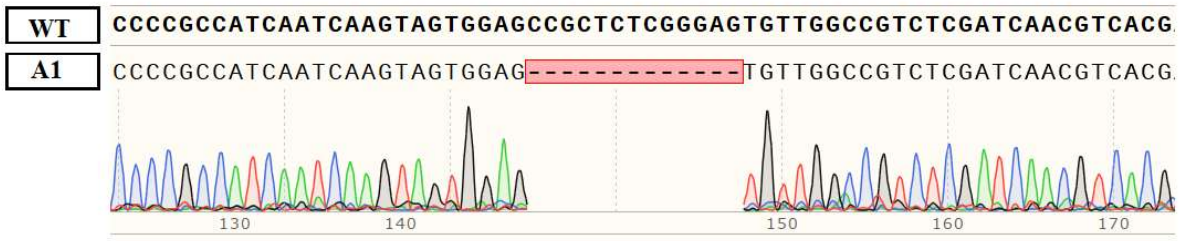


Figure 4.8: Sequencing alignment of the target region of the *ATPase5-1B* gene for the *P. tricornutum* A1 (ATPase_PAM2) mutant and WT. Chromatogram and sequence alignment generated by SnapGene® software (version 6.0) from GSL Biotech LLC.

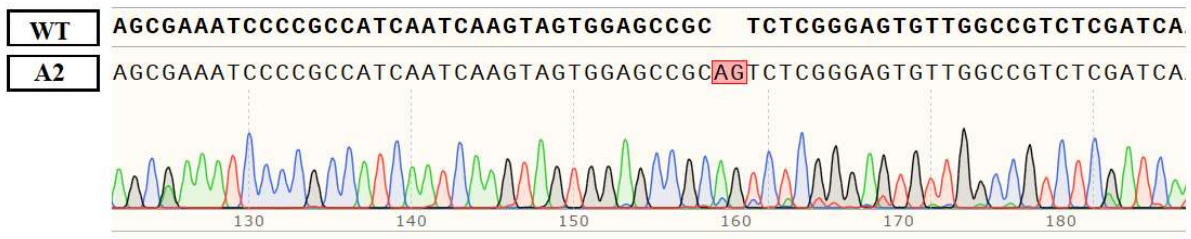


Figure 4.9: Sequencing alignment of the target region of the *ATPase5-1B* gene for the *P. tricornutum* A2 (ATPase_PAM2) mutant and WT. Chromatogram and sequence alignment generated by SnapGene® software (version 6.0) from GSL Biotech LLC.

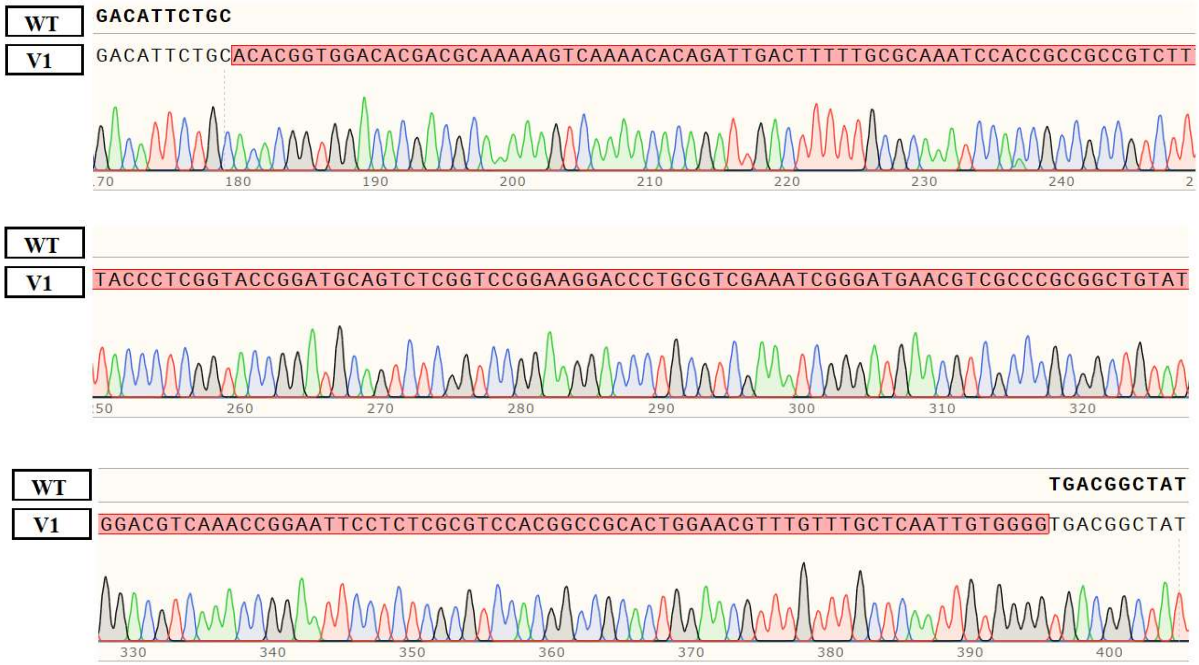


Figure 4.10: Sequencing alignment of the target region of the *VIT2* gene for the *P. tricornutum* V1 (VIT_PAM1) mutant and WT. Chromatogram and sequence alignment generated by SnapGene® software (version 6.0) from GSL Biotech LLC.

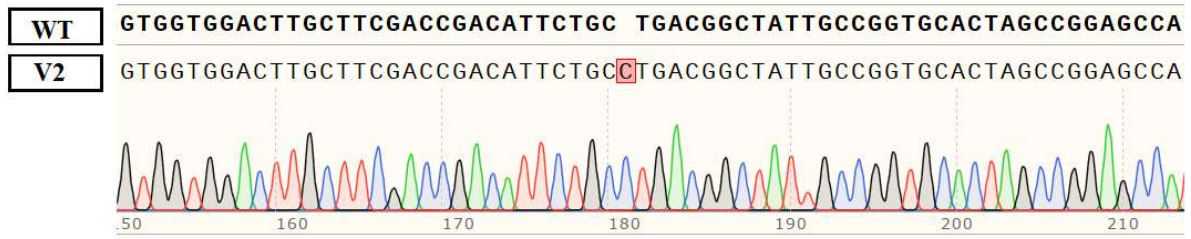


Figure 4.11: Sequencing alignment of the target region of the *VIT2* gene for the *P. tricornutum* V2 (VIT_PAM1) mutant and WT. Chromatogram and sequence alignment generated by SnapGene® software (version 6.0) from GSL Biotech LLC.

4.3 Characterization of *P. tricornutum* knockout lines

To observe whether the above mentioned mutations of the *ATPase5-1B*, *VIT1* and *VIT2* genes affected the heavy metal tolerance of *P. tricornutum*, an experiment was conducted to investigate any differences in cell growth parameters and photosynthetic efficiency between the A1, A2, V1, V2 mutant lines and WT. Cultures from each cell line were cultivated in f/2 medium and acclimated to 55-70 μmol photon m^{-2} s^{-1} continuous light conditions at 22°C, before subsequently being grown in three different media for the duration of the experiment: f/2 medium containing 5 mg/L Zn^{2+}, f/2 medium containing 1.25 mg/L Cd^{2+} and standard f/2 medium as a control. The heavy metal concentrations were determined from concentrations shown to induce mild growth inhibition by Horvatić & Peršić (2007), as this was predicted to be the concentrations where changes in tolerance would be most noticeable in mutant lines. Three replicate cultures were grown and measured for each cell line in each medium.

4.3.1 Growth inhibition of *P. tricornutum* mutant lines compared to wild type when exposed to zinc and cadmium

The growth inhibition experiment began with the acclimated *P. tricornutum* cells lines in exponential growth (initial concentration ~25,000 cells/mL). The number of cells were measured every 24 hours on 8 consecutive days, and with a final measurement also conducted at Day 17. The cell count data was used to create growth curves for each cell line in each growth medium, as well as calculating their specific growth rates (K') for each time point using Equation 1 (Section 3.3.2) to identify their exponential growth phase. All mutant lines displayed exponential growth phases between measurements on Day 1 and Day 3-4 irrespective of heavy metal exposure. Maximum specific growth rates (K'_{max}) were then calculated for the exponential growth phases using Equation 2. The relative K'_{max} was calculated using Equation 3 to measure whether heavy metals affected K'_{max} differently in the mutant lines compared to WT. Finally the total cell counts for each line in each growth medium measured at the end of the experiment (Day 17) was used to calculate the relative cell count difference, to discover if the final cell concentrations were differently affected by heavy metals for the mutant lines than for WT.

4.3.1.1 Growth of *P. tricornutum* mutant lines in control medium

The growth curves of cell lines grown in CT medium show that WT grows more rapidly than the four mutant lines, while V1, V2 and A1 grew more rapidly than A2 (Figure 4.12). All mutant lines displayed significantly lower K'_{max} values (Table 4.2; Supplementary Table I.1) than WT at the $P < 0.05$ level when grown in CT medium, while differences observed between mutant lines were not significant (Appendix J.1). WT reached the highest final cell concentration, with the V1, V2 and A1 mutant lines reaching similar final concentrations and the A2 mutant line reached the lowest final concentration by the end of the experiment. The mean final cell concentrations (n = 3) at Day 17 were 13.2 x 10^6 cells/mL (SD = 1) for WT, 10.8 x 10^6 cells/mL (SD = 0.5) for V1, 11 x 10^6 cells/mL (SD = 0.5) for V2, 10.5 x 10^6 cells/mL (SD = 0.6) for A1 and 7.8 x 10^6 cells/mL (SD = 0.3) for A2. All mutant lines had significantly lower final cell counts than WT (Appendix J.4).

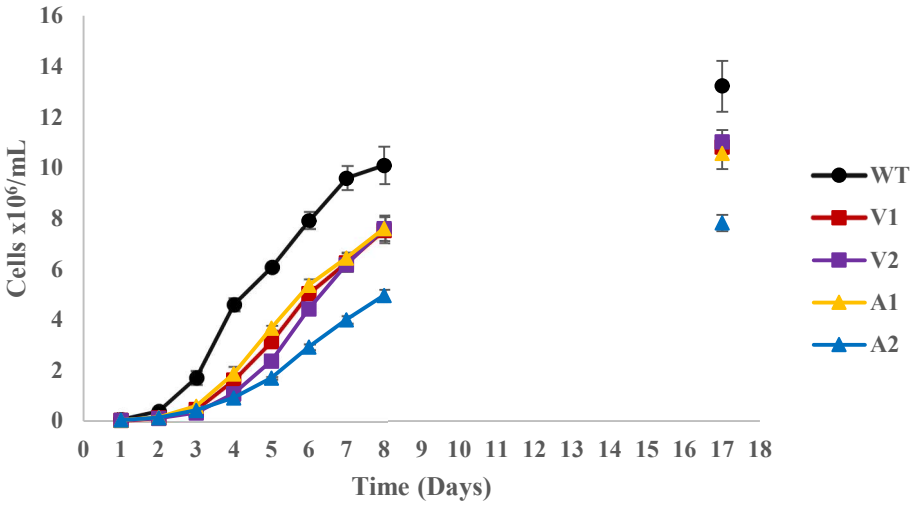


Figure 4.12: Cell counts from A1 (ATPase_PAM2; Yellow), A2 (ATPase_PAM2; Blue), V1 (VIT_PAM1; Red), V2 (VIT_PAM1; Purple) mutant lines and WT (Black) cultivated in standard f/2 medium (24 h photoperiod 55-70 µmol photon m^{-2} s^{-1}, 22°C). Measurements taken every 24 h for the first 8 days, with a final measurement taken at day 17. Mean cell counts and standard deviations calculated from n = 3 replicas of each cell line and measurement. Created with Microsoft Excel.

4.3.1.2 Effect of zinc on cell growth of *P. tricornutum* mutant lines

The growth curves created for cell lines grown in f/2 medium containing 5 mg/L Zn^{2+} (Figure 4.13) have a similar shape to those of the cell lines grown in CT medium: WT displayed the highest growth rate and final cell concentration, the V1, V2 and A1 mutant lines displayed similar final concentrations and the A2 mutant line had the lowest final concentration. Again, WT had a significantly higher K'_{max} than the mutant lines at the $P < 0.05$ level (Table 4.2; Supplementary Table I.1). Significant differences in K'_{max} were observed between the V1 and

A2 mutant lines, as well as the A1 and A2 mutant lines in the presence of Zn^{2+} (Appendix J.2). Among all cell lines, only the A2 mutant line displayed a significantly different K'_{max} when grown in the 5 mg/L Zn^{2+} medium compared to the same line grown in CT medium. This was emphasized by the relative K'_{max} values, which were close to 1 (i.e. equal to WT) for the V1 (0.984), V2 (0.978) and A1 (0.947) mutant lines while lower for the A2 mutant line (0.878)(Table 4.2). The mean final cell concentration (n = 3) at Day 17 was 14.7 x 10^6 cells/mL (SD = 0.2) for WT, which is significantly higher than the final cell concentration of WT when grown in CT medium (13.2 x 10^6 cells/mL; SD = 1).

Mean final cell concentrations were 10.7 x 10^6 cells/mL (SD = 0.2) for V1, 11.5 x 10^6 cells/mL (SD = 0.2) for V2, 10.1 x 10^6 cells/mL (SD = 0.1) for A1 and 8.1 x 10^6 cells/mL (SD = 0.3) for A2. All mutant lines had significantly lower final cell concentrations than WT when exposed to zn^{2+}, but none had significantly different final concentrations from when grown in CT medium (Appendix J.4). The relative cell count differences were close to 1 for all the mutant lines (0.892 for V1, 0.940 for V2, 0.859 for A1 and 0.934 for A2) (Table 4.3), indicating little change in the relative effect of zinc on the final cell concentration on the mutant lines compared to that on WT.

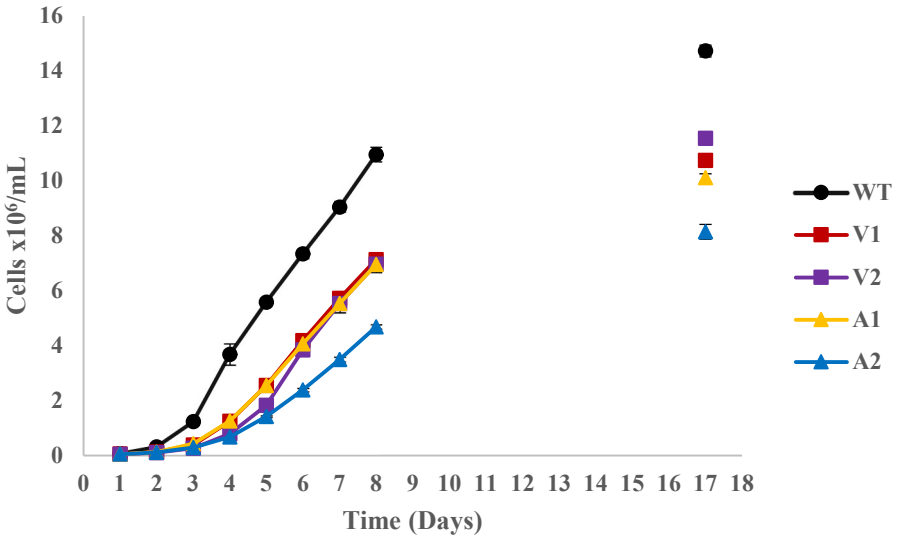


Figure 4.13: Cell counts from A1 (ATPase_PAM2; Yellow), A2 (ATPase_PAM2; Blue), V1 (VIT_PAM1; Red), V2 (VIT_PAM1; Purple) mutant lines and WT (Black) cultivated in f/2 medium containing 5 mg/L Zn^{2+} (24 h photoperiod 55-70 µmol photon m^{-2} s^{-1}, 22°C). Measurements taken every 24 h for the first 8 days, with a final measurement taken at day 17. Mean cell counts and standard deviations calculated from n = 3 replicas of each cell line and measurement. Created with Microsoft Excel.

Table 4.2: Maximum specific growth rate (K'_{max}) and relative K'_{max} calculated for A1, A2, V1 and V2 mutant lines and WT when grown in f/2 medium containing 5 mg/L Zn^{2+} compared to control. Mean K'_{max} and standard deviation calculated from n = 3 replicas. Cell lines denoted by statistical significance as determined by ANOVA analysis, with a: statistically significant difference ($p < 0.05$) between denoted line and WT in the same treatment group (with or without Zn^{2+}), b: statistically significant difference between denoted line in Zn^{2+} medium compared to the same line in CT medium.

5 mg/L Zn^{2+}			
Line	**K'_{max}**	**SD (n = 3)**	**relative K'_{max}**
WT (CT)	1.59	0.036	
WT (Zn^{2+})	1.48	0.042	
V1 (CT)	1.22[a]	0.059	0.984
V1 (Zn^{2+})	1.12[a]	0.034	
V2 (CT)	1.13[a]	0.085	0.978
V2 (Zn^{2+})	1.02[a]	0.103	
A1 (CT)	1.28[a]	0.048	0.947
A1 (Zn^{2+})	1.12[a]	0.059	
A2 (CT)	1.14[a]	0.021	0.878
A2 (Zn^{2+})	0.93[ab]	0.026	

a: Significant difference ($p < 0.05$) between a mutant line (CT) and WT (CT) or a mutant line (Zn^{2+}) and WT (Zn^{2+})

b: Significant difference ($p < 0.05$) between a line (CT) and the same line (Zn^{2+})

Table 4.3: Relative cell count differences calculated for A1, A2, V1 and V2 mutant lines and WT when grown in f/2 medium containing 5 mg/L Zn^{2+} compared to control. Calculated from the relative effect of Zn^{2+} on the final cell concentration (day 17) of each mutant line compared to that on WT.

Line	**Relative cell count difference**
V1	0.892
V2	0.940
A1	0.859
A2	0.934

4.3.1.3 Effect of cadmium on cell growth of *P. tricornutum* mutant lines

Growth curves for cell lines cultivated in 1.25 mg/L Cd^{2+} medium (Figure 4.13) display larger variations in cell growth compared to those cultivated in Zn^{2+}. WT still grew faster than the mutant lines, followed by the V1 and V2 mutant lines. The final concentration of the A1 mutant was closer to that of the A2 mutant than when cultivated in CT or Zn^{2+} medium, with a growth curve that seemingly flattens out between measurement on Day 8 and the final measurement on Day 17. K'_{max} of WT remained statistically higher that the other mutant lines at the $P < 0.05$ level (Table 4.4; Supplementary Table I.1), while K'_{max} of the A2 mutant line remained statistically different from the V2 and A1 lines (Appendix J.3). No mutant lines displayed a significantly different K'_{max} when cultivated in 1.25 mg/L Cd^{2+} medium compared to CT. Relative K'_{max} values were close to 1 for all mutant lines, indicating little effect of cadmium on their maximum growth rate compared to that of WT.

Mean final cell concentrations were 11 x 10^6 cells/mL (SD = 0.4) for WT, 7.4 x 10^6 cells/mL (SD = 0.5) for V1, 9.2 x 10^6 cells/mL (SD = 0.4) for V2, 5.3 x 10^6 cells/mL (SD = 0.5) for A1 and 4.6 x 10^6 cells/mL (SD = 0.03) for A2. All mutant lines had significantly lower final cell concentrations than WT, and all lines had significantly lower final cell concentrations than when grown in CT medium (Appendix J.4). Relative cell count differences (Table 4.5) for the V1 and V2 mutants remained close to 1 (0.825 for V1 and 0.997 for V2). Those for the A1 and A2 mutants were lower (0.600 for A1 and 0.707 for A2), indicating a stronger inhibitory effect of Cd^{2+} on the final cell concentration of these mutant lines compared to that of Cd^{2+} on WT.

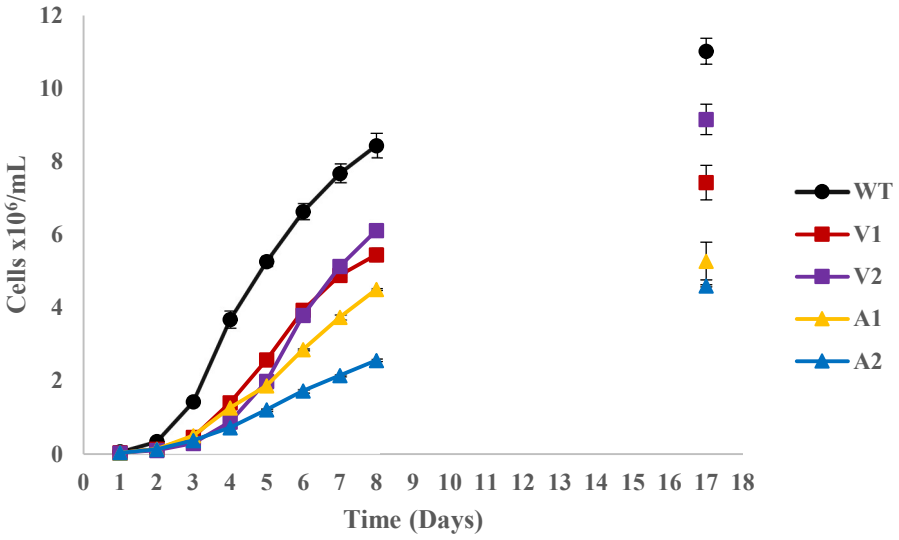


Figure 4.14: Cell counts from A1 (ATPase_PAM2; Yellow), A2 (ATPase_PAM2; Blue), V1 (VIT_PAM1; Red), V2 (VIT_PAM1; Purple) mutant lines and WT (Black) cultivated in f/2 medium containing 1.25 mg/L Cd^{2+} (24 h photoperiod 55-70 µmol photon m^{-2} s^{-1}, 22°C). Measurements taken every 24 h for the first 8 days, with a final measurement taken at day 17. Mean cell counts and standard deviations calculated from n = 3 replicas of each cell line and measurement. Created with Microsoft Excel.

Table 4.4: Maximum specific growth rate (K'_{max}) and relative K'_{max} calculated for A1, A2, V1 and V2 mutant lines and WT when grown in f/2 medium containing 1.25 mg/L Cd^{2+} compared to control. Mean K'_{max} and standard deviation calculated from n = 3 replicas. Cell lines denoted by statistical significance as determined by ANOVA analysis, with a: statistically significant difference ($p < 0.05$) between denoted line and WT in the same treatment group (with or without Cd^{2+}).

1.25 mg/L Cd^{2+}			
Line	**K'_{max}**	**SD (n = 3)**	**relative K'_{max}**
WT (CT)	1.59	0.036	
WT (Cd^{2+})	1.52	0.023	
V1 (CT)	1.22[a]	0.059	0.999
V1 (Cd^{2+})	1.17[a]	0.030	
V2 (CT)	1.13[a]	0.085	0.975
V2 (Cd^{2+})	1.05[a]	0.069	
A1 (CT)	1.28[a]	0.048	1.061
A1 (Cd^{2+})	1.30[a]	0.044	
A2 (CT)	1.14[a]	0.021	0.984
A2 (Cd^{2+})	1.08[a]	0.026	

a: Significant difference ($p < 0.05$) between a mutant line (CT) and WT (CT) or a mutant line (Cd^{2+}) and WT (Cd^{2+})

Table 4.5: Relative cell count differences calculated for A1, A2, V1 and V2 mutant lines and WT when grown in f/2 medium containing 1.25 mg/L Cd^{2+} compared to control. Calculated from the relative effect of Cd^{2+} on the final cell concentration (day 17) of each mutant line compared to that on WT.

Line	**Relative cell count difference**
V1	0.825
V2	0.997
A1	0.600
A2	0.707

4.3.2 Variable chlorophyll *a* fluorescence of *P. tricornutum* knockout lines when exposed to zinc and cadmium

In addition to investigating the effects 5 mg/L Zn^{2+} and 1.25 mg/L Cd^{2+} had on cell growth in the mutant lines, variable chlorophyll *a* fluorescence was conducted to observe if the photosynthetic activity of the mutant lines were affected differently by the heavy metals than that of WT. Samples were extracted for each mutant line and WT in each treatment group (ZN, CD and CT) on Day 3 of the growth experiment. Three replicas of each cell line and treatment group where measured.

The minimal fluorescence after exposure to the measuring light (F_0) and the maximal fluorescence after exposure to the saturation light (F_M) were used to calculate the maximum quantum yield of photosystem II (PSII; $\Phi_{PSII\ max}$) according to Equation 5 (Section 3.3.3). $\Phi_{PSII\ max}$ indicates the maximum amount of energy used for photochemistry by PSII (Ralph & Gademann, 2005) and was used as a measure of photosynthetic efficiency and thereby general cellular fitness in response to heavy metal exposure.

The fluorescence data from measurements done at incremental irradiance levels was used to generate rapid light curves (RLC) with the PamWin-3 Software (Walz). The RLCs were used to determine the maximum relative electron transport rate ($rETR_{max}$) and the minimum saturating irradiance (E_k) of the samples to better understand the effects of heavy metal exposure on the photosynthetic apparatus of the mutant lines compared to that of WT.

4.3.2.1 Effect of zinc on photosynthesis-related parameters of *P. tricornutum* mutant lines

$\Phi_{PSII\ max}$

The concentration of Zn^{2+} used in our assay had seemingly no effect on the $\Phi_{PSII\ max}$ of WT compared to CT (Figure 4.15A; Appendix J.2). All mutant lines displayed a lower $\Phi_{PSII\ max}$ than WT when cultivated in CT medium, but the differences were only statistically significant at the $P < 0.05$ level for the A1 and A2 mutants (Figure 4.15A; Appendix J.1). Similarly, all mutant lines displayed lower $\Phi_{PSII\ max}$ than WT when exposed to Zn^{2+}, but the differences were only significant for the A1 and A2 mutants. The $\Phi_{PSII\ max}$ was slightly lower for the V1 mutant in the presence of Zn^{2+} than in CT medium, and slightly higher for A1 and A2. None of these differences were statistically significant, however (Appendix J.2).

$rETR_{max}$

No major differences in $rETR_{max}$ were observed between WT in the CT and Zn^{2+} treatment group (Figure 4.15B; Appendix J.2). All mutant lines displayed a significantly lower $rETR_{max}$ than WT when cultivated in CT medium (Figure 4.15B; Appendix J.1). The V1, A1 and A2 mutants also displayed significantly lower $rETR_{max}$ than WT in the presence of Zn^{2+}. $rETR_{max}$ was higher for V1, lower for V2, lower for A1 and equal for A2 in the Zn^{2+} treatment group compared to CT, but none of the differences were statistically significant (Appendix J.2).

E_k

E_k was similar for WT in both the CT and Zn^{2+} treatment groups (Figure 4.15C; Appendix J.2). All mutant lines displayed lower E_k than WT in both treatment groups, but the differences were only statistically different for the A2 mutant (Figure 4.15C). E_k was higher in the presence of Zn^{2+} than in the CT medium for the V1 and V2 mutants and lower in Zn^{2+} for the A1 and A2 mutants, but again none of the differences between the same lines in different treatment groups were statistically significant (Appendix J.2).

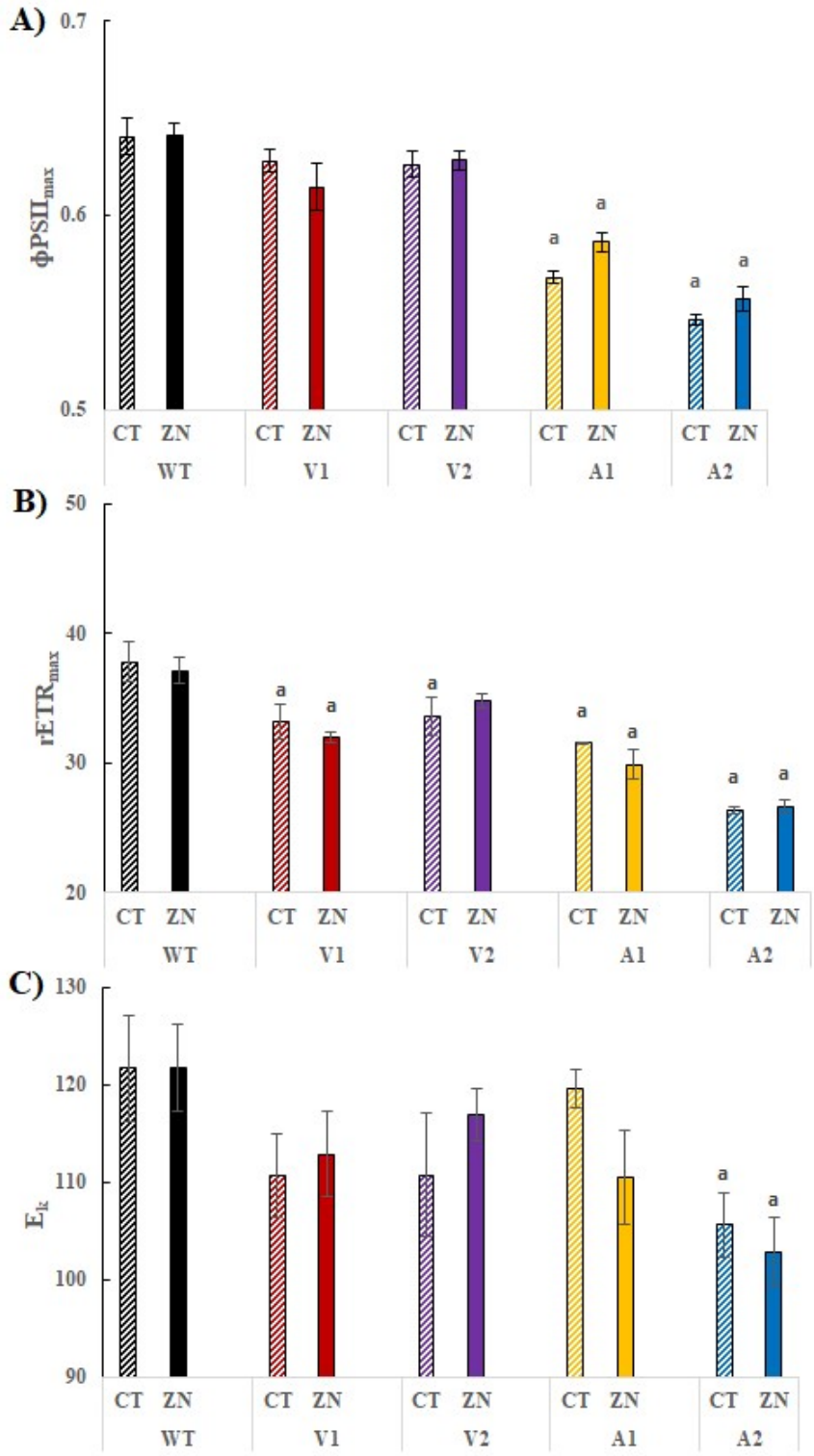


Figure 4.15: A) Maximum quantum yield of PSII ($\Phi_{PSII\ max}$), B) maximum relative electron transport rate ($rETR_{max}$) and C) minimum saturating irradiance (E_k) for the A1 (Yellow), A2 (Blue) V1 (Red) and V2 (Purple) mutant lines and WT (Black) cell lines cultivated in f/2 medium containing 5 mg/L Zn^{2+} and standard f/2 medium (CT). Means and standard deviations calculated from n = 3 replicas. Filled columns: Zn^{2+} treatment group, striped columns: CT treatment group, a: statistically significant difference ($p < 0.05$) between denoted line and WT in the same treatment group (with or without Zn^{2+}) Created with Microsoft Excel.

4.3.2.2 Effect of cadmium on photosynthesis-related parameters of *P. tricornutum* mutant lines

$\Phi_{PSII\ max}$

$\Phi_{PSII\ max}$ was lower for WT in the Cd^{2+} treatment group than in CT, but the difference was not statistically significant (Figure 4.16A; Appendix J.3). $\Phi_{PSII\ max}$ values for V1 and V2 where not significantly different from WT when exposed to Cd^{2+}, but they were significantly lower for A1 and A2. There were no major differences between the Cd^{2+} and CT treatment groups for either V1 or V2. $\Phi_{PSII\ max}$ was slightly higher in the presence of Cd^{2+} for both A1 and A2, but the difference was not statistically significant either.

$rETR_{max}$

$rETR_{max}$ remained unchanged in the Cd^{2+} treatment group for WT compared to CT (Figure 4.16B; Appendix J.3). All mutant lines displayed significantly lower $rETR_{max}$ values compared to WT when exposed to Cd^{2+}. $rETR_{max}$ values were slightly higher in the Cd^{2+} treatment group than in CT medium for V1 and V2, and slightly lower for A1 and A2, but none of these differences were statistically significant.

E_k

E_k was slightly lower for WT in the Cd^{2+} treatment group, but the difference was not significant (Figure 4.16C; Appendix J.3). All mutant lines displayed a lower E_k than WT when exposed to Cd^{2+}, but the difference was only statistically significant for A2. E_k was higher in the Cd^{2+} treatment group than in CT medium for V1, and V2 and lower for A1 and A2, but none of these differences were statistically significant.

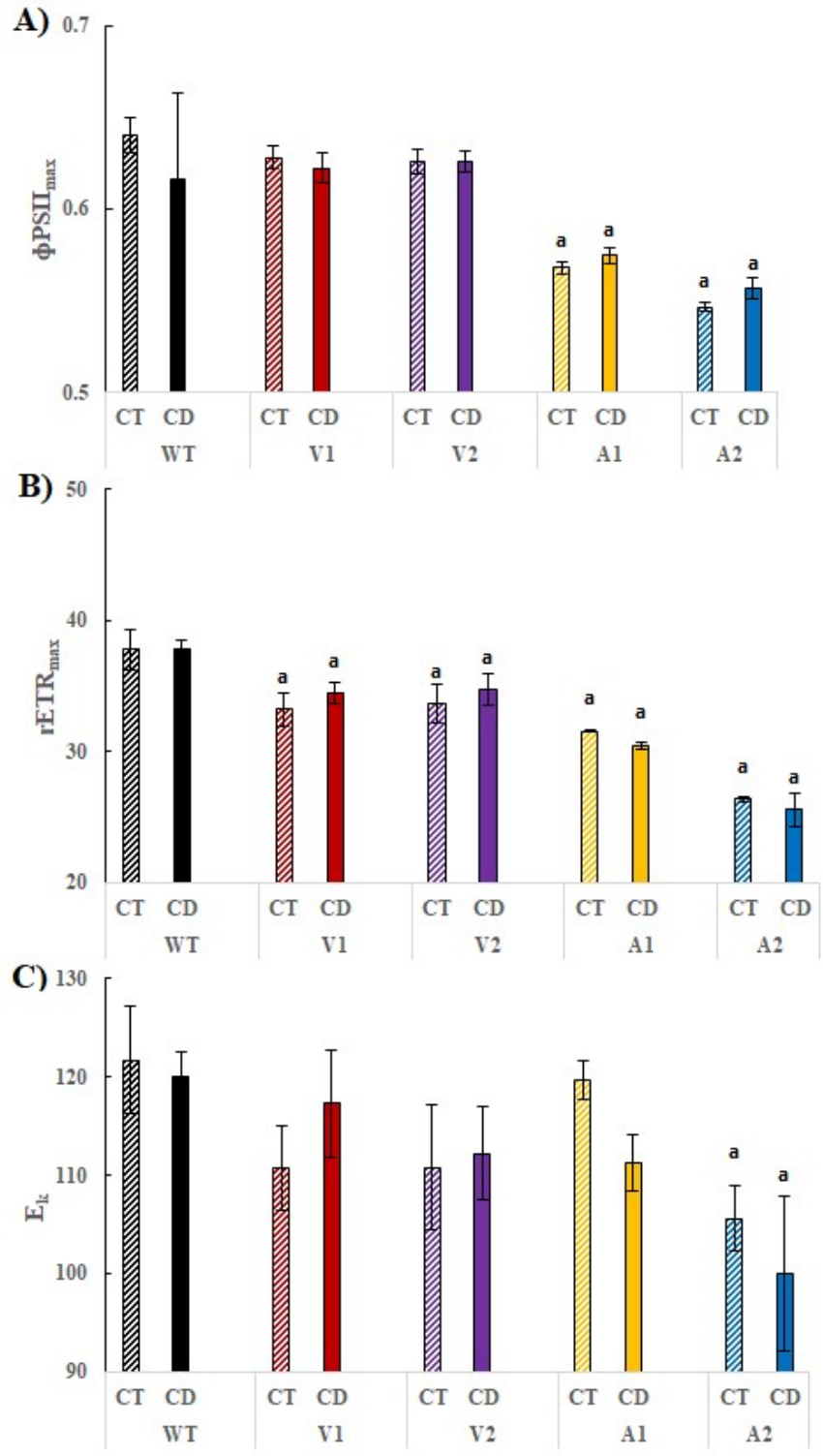


Figure 4.16: A) Maximum quantum yield of PSII ($\Phi_{PSII\ max}$), B) maximum relative electron transport rate ($rETR_{max}$) and C) minimum saturating irradiance (E_k) for the A1 (Yellow), A2 (Blue) V1 (Red) and V2 (Purple) mutant lines and WT (Black) cell lines cultivated in f/2 medium containing 1.25 mg/L Cd^{2+} and standard f/2 medium (CT). Means and standard deviations calculated from n = 3 replicas. Filled columns: Cd^{2+} treatment group, striped columns: CT treatment group, a: statistically significant difference ($p < 0.05$) between denoted line and WT in the same treatment group (with or without Cd^{2+}) Created with Microsoft Excel.

5 Discussion

The aim of this project was to explore the roles played by three predicted metal transporter proteins (Brembu et al., 2011) in the observed heavy metal tolerance of *P. tricornutum* utilizing CRISPR/Cas9 gene editing technology to create knockout lines. Microalgae have recently received much attention for their potential use in sustainable bioremediation of heavy metal-containing wastewater, due to species displaying high innate tolerances and the ability to effectively accumulate heavy metals. Advancements in genetics and molecular biology are currently helping scientists to better understand the underlying mechanisms responsible for these traits, and many see this knowledge being used to engineer microalgal strains with increased tolerances and higher accumulation capabilities (Ranjbar & Malcata, 2022). Such microalgal strains might be of high commercial interest as part of developing a more sustainable and cost-effective next generation of remediation technology. For this reason, *P. tricornutum,* a marine diatom already noted for its heavy metal resistance (Horvatić & Peršić, 2007; Torres et al., 1997), was selected for this project, with the goal of contributing to the existing body of knowledge on its heavy metal metabolism and perhaps identify targets for future work on commercial bioremediation technology. Compatible plasmids were designed to encode CRISPR/Cas9 systems targeting the predicted metal transporter genes *VIT1*, *VIT2* and *ATPase5-1B* (Brembu et al., 2011), and were delivered to *P. tricornutum* cell both by bacterial conjugation and biolistics to create knockout-mutants. Cell lines thought to contain knockout mutations in the *VIT1*/*VIT2* (V1 and V2) and *ATPase5-1B* (A1 and A2) genes were identified. The mutant lines were subjected to zinc and cadmium, two heavy metals whose cellular concentrations were believed to be modulated by the encoded metal transporters of interest, to characterize their effects on cell growth and photosynthetic efficiency. Time available for the experimental work was impacted at a critical time by restrictions related to the onset of the Covid-19 pandemic, which influenced the remaining course of the project.

5.1 Creation of target specific CRISPR/Cas9 plasmid vectors

Plasmid CRISPR/Cas9 vectors targeting the *VIT1*, *VIT2* and *ATPase5-1B* genes were created using traditional restriction-ligation cloning technology. Pairs of plasmids were designed to target two different PAM-sites on each gene and increase the likelihood of obtaining edited mutants. Due to their high degree of sequence similarity, *VIT1* and *VIT2* were suspected of having similar, and perhaps redundant, cellular functions. For this reason, plasmids were designed to target both genes simultaneously and possibly obtain mutants where both genes had been knocked out. Colony PCR screening followed by Sanger sequencing using vector specific primers was used to identify plasmids containing the correct target specific adapter inserts at the 5′-end of the gRNA. Alignment of the Sanger sequencing data verified the successful insertion of the adapters at their expected location on the gRNA cassette (Figure 4.1 and Figure 4.2). *Hind*III enzymatic digestion was conducted on each plasmid to verify plasmid integrity, and the resulting fragment lengths indicated that the target specific plasmids remained stable after adapter insertion and amplification in competent *E. coli*. With

all plasmids encoding gRNA adapter sequences complementary to parts of the target genes immediately 5′ of a PAM sequence, they are expected to facilitate hybridization and subsequent cleavage by the CRISPR/Cas9 complex (Yamamoto, 2015). In the absence of an added template, the DSB induced by the CRISPR/Cas9 complex will be repaired through the NHEJ pathway, leading to the possible formation of indel mutations (Sander & Joung, 2014).

5.2 Creation of *P. tricornutum* mutants

5.2.1 Transformation and mutation efficiency

Initially, the plan was to transform *P. tricornutum* cells with the target specific plasmid through bacterial conjugation with competent *E. coli*, following the C1 protocol (Section 2) adapted from Moosburner et al. (2020) by the CMBG group to be a simple method yielding sufficient transformants to identify knockout mutants. Conjugation was preferred because it only introduces vectors as episomes, allowing transient expression of Cas9 and gRNA until knockout-mutations are identified and reducing the chances of off-target mutations or re-editing at the target site (Sharma et al., 2018). This, however, resulted in few primary colonies appearing on zeocin-containing selection plates for the different plasmid treatments (Table 3.1). The conjugation protocol was repeated twice, but colony numbers remained insufficiently low. With each conjugation requiring 3-4 weeks for colonies to appear, this caused significant time delays despite repeated transformations being conducted concurrently. Finally, the choice was made to transform *P. tricornutum* cells utilizing the C2 protocol developed by the CMBG research group (Sharma et al., 2018) instead. The C2 protocol utilizes liquid suspensions of *P. tricornutum* and *E. coli* before plating, as opposed to the simplified C1 protocol which adds competent *E. coli* to *P. tricornutum* already cultivated on plates, and is reported to have a high transformation efficiency of 500-1000 transformants/10^8 cells (Sharma et al., 2018). The C2 protocol did produce a higher number of colonies than C1, although the transformation efficiency only equated to 20-50 transformants/10^8 cells. Transformation efficiencies were thus lower than with both protocols. Since repeated transformations were carried out with newly made growth mediums and transformed competent *E. coli* cells without changing the outcome, it seems unlikely that the cause was human error. The plasmids were considered unlikely to be toxic for *P. tricornutum*, as the same plasmid backbone had been previously used Also, the *Hin*dIII digestion did not reveal the plasmids to be unstable or to have rearranged in *E. coli*, something which could have explained the low efficiency (Karas et al., 2015). The *E. coli* cells containing the pTA-MOB helper plasmid were used successfully for similar work within the CMBG research group during the same timeframe (unpublished data), making it unlikely that they were responsible. It is therefore difficult to properly explain what caused low transformation efficiencies for bacterial conjugation with the available data. In addition, *P. tricornutum* was also transformed through biolistic bombardment with tungsten particles coated with the target specific vectors in case conjugation did not produce enough transformants or subsequent mutants. Biolistics produced a satisfactory number of colonies for the VIT_PAM1 and ATPase_PAM2 plasmids, but low numbers for VIT_PAM2 and ATPase_PAM1. Such variety is not unexpected, as low consistency and reproducibility are

known drawbacks of the particle delivery of genetic material (Miller et al., 2021). The method leads to random DNA fragment integration in the host genome and can cause damage to cells, both of which can affect cell growth and expression of delivered genetic material (Liu et al., 2019). Notably, any cell that did not integrate a plasmid fragment containing the *BleoR* gene encoding zeocin-resistance would not survive to form a colony on the selection plates despite functioning CRISPR/Cas9 expression.

Based on which colony samples displayed divergent melting behaviour compared to WT during HRM analysis, mutation efficiency in colonies expressing CRISPR/Cas9 on episomal plasmids delivered by bacterial conjugation ranged from 2.7-20 %. This is lower than the reported efficiency of 25–33 % for the similar conjugation protocol described in Sharma et al. (2018). One of the things that can cause a low rate of CRISPR/Cas9-induced mutations is the presence of polymorphisms on the recognition site of the target gene. The PAM sequence (5′-NGG-3′ for the commonly used *S. pyogenes* system) must be conserved in the host cell for the CRISPR/Cas9 complex to recognize the target site. The target gene sequence must also match the gRNA sequence for hybridization and cleave to occur. The specificity of gRNA increases with increasing proximity to the PAM site, with the closest 12 bp having the highest targeting specificity (Yamamoto, 2015). Accumulated random polymorphisms in the target sequence of the *P. tricornutum* cells transformed, compared to the known sequence used to design the target specific adapters, would result in a lower binding affinity of the CRISPR/Cas9 complex and subsequently lower mutation efficiency. However, no polymorphisms were found when aligning sequencing data from WT colonies with the predicted target sequences of each plasmid on their respective target genes, and therefore do not explain these observations. Mutation efficiency can also appear lower if the method of assessment underrepresents the actual number of mutants present during screening. This appeared to be a posgsibility, since Sanger sequencing later identified mutations in samples that displayed no divergent melting behaviour from WT during HRM analysis. The actual number of mutants might therefore have been higher than what the HRM analysis indicated. Some colonies also displayed divergent melting behaviour despite not containing being edited, however, making assessments of whether HRM analysis under- or overestimated the true mutation efficiency difficult. Possible limitations of the HRM assessment approach are discussed in more detail below. Another factor which could influence the observed mutation efficiency is the amount of time primary mutant colonies were cultivated on selection plates before being screened by HRM and sequenced. Sharma et al. (2018) observed that mutations took longer to appear for *P. tricornutum* cells transformed through bacterial conjugation than through biolistics using a similar plasmid vector, taking three months to reach a mutation efficiency of 25–33 %. Some of the earlier conjugation plates had been cultivated for approximately this long before being assessed by HRM, but most primary colonies from the repeated transformation attempts were assessed much earlier due to time constraints. Sharma et al. (2018) observed that the numbers of episomal CRISPR/Cas9 copies in transformed cells were less than one third of CRISPR/Cas9 copies integrated in the genome of cells transformed by biolistics on average, and were generally correlated with lower expression of Cas9 mRNA. A similarly low expression of Cas9 in the present study might have led to lower levels of nuclease activity and therefore a lower mutation efficiency in the colonies transformed by conjugation than by biolistics. qPCR could potentially have been used to investigate the expression of Cas9 at the transcriptional level to better assess this possibility. Furthermore, Sharma et al. (2018) hypothesize that since cells transformed by conjugation

have high cell divisions with minimal lag phases after transformation, so that they could potentially undergo multiple cell divisions before Cas9 induces mutagenesis. In such cases the primary colony might contain only a small fraction of mutated cells, below the detection limit of the HRM analysis. As a similar protocol was followed when attempting to create mutant lines for this project, it is likely that similar factors affected the observed mutation efficiency. There is also a possibility that the episomes underwent rearrangements during or after conjugative plasmid transfer to *P. tricornutum* cells, and thereby lost mutagenic function. Although this is less likely since plasmid stability was observed after bacterial transformation, we did not verify the absence of rearrangement of the plasmids during transfer to *P. tricornutum* and their stability as episomes in the diatom

Mutation efficiencies from biolistic transformation for the VIT_PAM1 (28%) and ATPase_PAM1 (42.9%) plasmids were consistent with previous observations made by Sharma et al. (33-50%; 2018), while the VIT_PAM2 and ATPase_PAM2 vectors produced no apparent putative mutants. In the case of VIT_PAM2 this is likely due to a low number of primary colonies (2) which reduce the likelihood of identifying any putative mutants. Why the ATPase_PAM2 vector produced no apparent mutations might be explained by the previously discussed drawbacks of biolistic transformation. Identifying mutants required the stable integration of the genes encoding both zeocin-resistance and functional target specific CRISPR/Cas9 nuclease. If only the latter were integrated into the host genome, or the transformed cells were lethally damaged by the particle bombardment itself, no mutant colonies would appear on the selection plates. It is important to note that biolistics was conducted using plasmids designed for conjugation due to time constraints. There is approximately 2000 bp between the *BleoR* gene encoding zeocin-resistance and the sequence encoding Cas9 and the gRNA (Supplementary Figure D.1), which increases the likelihood of only one of the two being integrated into the *P. tricornutum* genome. Since the plasmids produced mutants after transformation by conjugation, the lack of mutants was likely not caused by a lack of targeting specificity.

5.2.2 HRM screening and sequencing of isolated mutant lines

The amount of time spent on obtaining putative mutants, in addition to restrictions imposed on laboratory work in relation to the outbreak of the Covid-19 pandemic, caused significant time constraints on the remaining part of the project. For that reason, the screening and isolation of pure single cell-derived mutant lines had to be cut short. This led to varying purity of the final mutant lines utilized moving forward. This is also the reason for using the Synthego ICE data analysis tool on sequencing results from PCR products to ascertain whether mutations were mono- or bi-allelic (carrying a mutation on one or both alleles, respectively) by deconvoluting the different sequence traces, instead of cloning the PCR product into a plasmid to be able to isolate and sequence individual amplicons as originally planned.

HRM allows detection of possible mutants by analyzing the denaturing melting behavior of a PCR product containing the CRISPR/Cas9 target sequence. The melting behavior will depend on the length, GC content and nucleotide sequence variation of the PCR product (Herrmann et al., 2006), and can therefore be used to identify mutations when comparing a potential

mutant to WT. Deletion of nucleotides will generally result in a lower melting temperature for the PCR product, while insertions result in higher melting temperatures (Wittwer et al., 2003). This can then be visualized in normalized melting curves as left shifted peaks for deletions and right shifted curves for insertions when compared to the melting behavior of the WT PCR product. If the DNA sequence does not contain any mutations, the melting behavior and subsequent peaks will be highly similar to WT. HRM analysis of PCR products from the mutant lines (Figures 4.4-4.7) displayed melting behavior that differed from the WT control samples in a variety of ways, which is representative of the variety of possible indel mutations that can occur when DSBs caused by the CRISPR/Cas9 complex are repaired through the NHEJ pathway (Sander & Joung, 2014).

5.2.2.1 A1 mutant line

The A1 mutant line (ATPase_PAM2; Figure 4.4) displayed a left shifted curve compared to the WT control, indicating a nucleotide deletion. It also displayed a smaller second peak shifted further left. Multiple peaks are caused by the melting of heteroduplex species made up of different PRC products with base pair mismatches (Herrmann et al., 2006). This indicated that the sample contained a mutant with a heterozygous genotype (carrying two different mutations on each allele) or a mixture of cell lines with different genotypes. The Sanger sequencing data and subsequent ICE analysis further supported these observations, showing that the sample contained mostly PCR products with a 13 bp deletion (96 %) and smaller amounts of various deletions. The chromatogram from Sanger sequencing (Figure 4.8) displayed a background noise that might have been interpreted as other deletions by the ICE analysis. Since the HRM analysis displayed multiple peaks however, it is probable that different PCR products were present in A1 samples. If the 13 bp deletion were present in a bi-allelic heterozygous cells, the other allele would have made up a similarly high fraction of the sequences identified from the ICE analysis. It is therefore likely that the A1 mutant line was comprised mainly of a homozygote bi-allelic mutant with a 13 bp deletion in *ATPase5-1B* and possibly contaminated by small amounts of other mutants in the same gene.

5.2.2.2 A2 mutant line

The A2 mutant line (ATPase_PAM2; Figure 4.5) displayed a clear, slightly right shifted curve compared to WT controls, which indicated an insertion. This was reinforced by the sequencing data, with most PCR products containing a 2 bp insertion (94 %) in *ATPase5-1B*. ICE analysis also identified the presence of a 7 bp deletion (6 %). The chromatogram (Figure 4.9) showed divergent sequence traces only after the target sequence, indicating that several mutant sequences were present in the sample. The different prevalence between the two sequences identified indicated that they were not carried by the same heterozygous cells, and thus belonged to different mutants. Since no other mutations or WT sequence were detected, it is possible that both the 2 bp insertion and 7 bp deletion mutants present in the A2 line were bi-allelic. Another possibility is that the continuous expression of CRISPR/Cas9 by the episomal vector, maintained by the selection pressure of the zeocin selection plates, led to re-

editing of cells present in the colony. A small subsection of pure 2 bp insertion mutants might have undergone re-editing of *ATPase5-1B* by CRISPR/Cas9 towards the end of the cultivation phase, generating a 7 bp deletion on one allele.

5.2.2.3 V1 mutant line

Since *VIT1* and *VIT2* share the target sequence of the gRNA encoded by the VIT_PAM1 plasmid, both genes could potentially have been edited in the same cell line. The HRM curve for the *VIT1* PCR product (Figure 4.6) displayed two peaks, indicating that heteroduplex species were formed during annealing. Sequencing analysis supported this, indicating a mixture of products with a 1 bp insertion (70%), a 3 bp deletion (22%) and the WT sequence. Although HRM analysis of the PCR product containing the *VIT2* target site (Figure 4.6) produced only a slight right shift of the melting peak compared to WT control, sequencing identified a 216 bp insertion in *VIT2* (Figure 4.10). NCBI BLAST analysis of the sequence identified it as a fragment of a predicted voltage-gated chloride channel gene in the *P. tricornutum* genome (NCBI Gene ID: 7195144).

The insertion of such a product into the target sequence is difficult to explain by NHEJ, which generally produces small indels (Yamamoto, 2015). As this fragment does not present substantial sequence homology to VIT2 it is also difficult to assess whether the insertion was caused by HDR (Yamamoto, 2015). Microhomology-mediated end joining, which relies on small stretches of homology on either side of the DSB, is likely not the cause of the insertion either, as it causes deletions flanking break sites which were not present (Wang & Xu, 2017). This mutant would not have been selected for further investigation if more mutants likely to carry knockout mutations were identified. Since the sequence trace (Figure 4.10) indicated that all PCR products contained the 216 bp insert in the *VIT2* gene, V1 is considered a pure mutant line that is homozygous bi-allelic for *VIT2*. It is therefore possible that this mutation was present before further isolation of primary transformants, while mutations on the *VIT1* continued during cultivation of the isolated colony. Due to having the highest prevalence among the *VIT1* PCR products, it is possible some cells are bi-allelic for a 1bp insertion. It is difficult to determine if the 3bp deletion was bi-allelic, or a mono-allelic mutation due to a re-edition of the 1 bp insert mutant. Unedited PCR products were likely from homozygous unedited WT transformants, but this also difficult to confirm from the available screening data.

5.2.2.4 V2 mutant line

The V2 mutant line (VIT_PAM1) was screened and sequenced in a similar was to V1. HRM curves displayed twin peaks indicating heteroduplex species formation for the PCR products containing the *VIT1* target site, and a clear right shifted peak indicating an insert in *VIT2* (Figure 4.7). Sequencing and ICE analysis showed that the *VIT1* PCR products indeed contained a mixture of deletions, insertions, and unedited sequences in the target site of *VIT1*. Sequencing analysis of *VIT2* PCR products (Figure 4.11) identified a 1 bp insertion mutation as the only sequence present, which is consistent with the observed melting behavior. V2 is thus a pure line with the same 1bp insertion in both alleles of the *VIT2* gene. It is likely that

the 1bp insertion in *VIT2* occurred very early in the V2 mutant line, with mutations on *VIT1* occurring in subsections of the colony later. Since many of the predicted mutations in *VIT1* have a similar relative prevalence, it is possible that they represent different mono-allelic mutations in heterozygous cells, but this is difficult to determine from the available data. Whether unedited sequences belonged to homozygous WT cells or heterozygous mono-allelic mutants is also unknown, although their relatively lower prevalence might indicate the former.

5.2.3 Challenges when screening for mutants

Several factors can affect the output of HRM analysis and the interpretation of sequencing data for CRIPSR/Cas9 mutants. HRM analysis was conducted on short amplicons of 122 bp to 197 bp, which means that a sufficiently large deletion at the target site could remove the binding site of one or both HRM primers. In such a case no product would be amplified for the deletion mutant, and the HRM analysis could produce a false negative if only unedited WT sequences were still amplified. Since the larger PCR products (approximately 500-800 bp) amplified before HRM were also analyzed by gel electrophoresis however, such large deletions would have been detected. If deletions that were several hundred base-pairs in size occurred, they could remove primer binding sites for the larger PCR products as well. In this case no PCR products would be produced to begin with, and the mutant would not be identified. As mentioned before, surprisingly some samples that displayed a WT melting curve during HRM analysis were shown to contain mutations in the target sequence by Sanger sequencing. Hence there was not always a strict concordance in the results from these two analyses.

P. tricornutum is a diploid organism, and thus carries two alleles of each gene. A mutant cell can therefore carry a mutation on one or both alleles or carry different mutations on each allele. The sequencing of individual PCR amplicons after having cloned PCR products into plasmids was planned to properly characterize both alleles of any isolated heterozygous mutants but, due to time constraints this was not possible. Hence, genotyping of individual cells from the mixed colonies producing heterozygous PCR products were indirectly inferred from HRM and ICE output instead. The screening challenges were further exacerbated by the decision to design CRISPR/Cas9 plasmids that targeted both the *VIT1* and *VIT2* genes simultaneously, since a more extensive screening process becomes necessary to obtain mutants with bi-allelic mutations in both genes. This resulted in both the V1 and V2 mutant lines only being pure mutants for the *VIT2* gene, while no pure mutants were isolated for *VIT1*.

5.2.4 Were knockout mutations for target genes created?

When CRISPR/Cas9 gene editing is used to generate DSBs without a template sequence added, DSB repair will usually be made through the error prone NHEJ pathway. This will result in the creation of random indel mutations near the target sequence of the CRISPR/Cas9 nuclease complex, which can lead to the inactivation of the protein encoded by the target gene and thereby create a knockout (loss-of-function) mutant. Indel mutations that disrupt the 3 bp codon based translational reading frame of the gene sequence will change the resulting downstream polypeptide sequence during translation, which will often completely change the final structure of the encoded protein and lead to inactivation. Frameshift mutations might also produce early stop codons on the transcribed mRNA sequence, leading to premature termination of protein translation.

Among multiple colonies screened and sequenced after the isolation step, the A1, A2, V1 and V2 mutant lines were selected due to being the most purified and because they carried the mutations predicted to most likely result in loss of function for the target genes. If it had been possible, only pure homozygous bi-allelic mutants carrying indels that would result in frameshifts would have been selected.

5.2.4.1 Do the mutations in *ATPase5-1B* lead to non-functional transporters?

The effects of a mutation on the activity of *P. tricornutum ATPase5-1B* can similarly be interpreted by understanding the documented general structure and function of P_{1B}-type ATPases. This protein sub-family of ATPases generally consists of a varying number of transmembrane domains (TMDs) and transport metals across membranes through a multi-step catalytic mechanism requiring ATP (Argüello et al., 2007). ATP binding to an ATP binding domain on the cytosolic side phosphorylates the protein, allowing metal ion occlusion within funnel structures present on the cytosolic side of metal binding TMDs. This also causes the rotation of a cytosolic energy transducing domain, leading to conformational changes that allow for extracellular release of the metal ion, followed by subsequent dephosphorylation and a return to the initial conformation. Amino acid residues allowing metal binding are usually located on one or more C-terminal TMDs. The ATP binding domain is located next to metal-binding TMDs, while the energy transducing domain is located between the adjacent TMD pair towards the N-terminus. The TMDs of P_{1B}-type ATPases are also often flanked on either side by cytosolic N-terminal and C-terminal metal binding domains that are thought to help regulate the transport of metal ions to the metal binding TMDs (Argüello et al., 2007).

Although the *P. tricornutum* ATPase5-1B protein is described as belonging to the P_{1B}-type ATPases (Brembu et al., 2011) its protein entry in UniProt (UniProt ID: B7G6B2) shows only a predicted N-terminal heavy metal associated domain in position 36-102 and four transmembrane domains further towards the C-terminus. The ATPase_PAM2 target site begins only 161 bp into the protein coding sequence of *ATPase5-1B* (Appendix C.3). The A1 mutant line contained a 13 bp deletion in the target site, which would cause a frameshift at

the amino acid residue in position 59. In addition, this would lead to a premature stop codon after another 30 amino acids, which would prevent the translation of all TMDs, the ATP binding domain and the energy transducing domain, clearly preventing the formation of a functional ATPase5-1B. The A1 line is therefore considered a knockout mutant for ATPase5-1B with possible contamination since sequencing analysis indicated the presence of other genotypes. The A2 line carried a majority of 2 bp insertions on the *ATPase5-1B* gene, leading to a frameshift beginning at amino acid residue position 60 and a premature stop codon at position 93. Similarly, the 7 bp deletion leads to a frameshift and stop codon, which would disrupt the function of critical protein domains as discussed above. The A2 line is therefore also considered a knockout mutant for ATPase5-1B.

5.2.4.2 Do the mutations in *VIT1* and *VIT2* lead to non-functional transporters?

The structure of the metal transporters encoded by both the *P. tricornutum VIT1* and *VIT2* genes have not been well studied. This makes it a challenge to identify and assess the sections and possible domains that are essential to their function, and thus which mutations would result in a functional knockout. More literature is available for similar proteins present in other organisms, however, from which better understanding of their *P. tricornutum* counterparts might be inferred. The recently described crystal structure and functional analysis of the *Eucalyptus grandis Eg*VIT1 protein has led to the identification of conserved domains and residues within the VIT1/Ccc1 family essential to metal transport (Kato et al., 2019). VIT1/Ccc1 family proteins, including those of *P. tricornutum*, possess five transmembrane domains (TMD1-5 from the N-terminus to C-terminus) which span the vacuolar membrane (Sorribes-Dauden et al., 2020). According to the UniProt protein database (https://www.uniprot.org/) prediction the *P. tricornutum* VIT1 (UniProt ID: B7FRU5) and VIT2 (UniProt ID: B7FRU4) proteins also contain five transmembrane domains. The *Eg*VIT1 protein has been shown to form dimers where the interface between the two protomers form a metal translocation channel in the membrane (Kato et al., 2019), which might also be the case for VIT1 and VIT2 in *P. tricornutum*. In this channel TMD1 and TMD2 create a hydrophilic environment suitable for coordinated metal transport from the cytoplasm into the vacuole. TMD1 and TMD2 contain amino acids that are responsible for coordinating metal transport and selectivity and participate in preventing proton leakage into the cytosol (Kato et al., 2019; Sorribes-Dauden et al., 2020). Some of these are conserved in VIT1 and VIT2 of *P. tricornutum*, likely serving the same function(Sorribes-Dauden et al., 2020), such as the glycine residue on TMD2 that has been shown to be critical for VIT1-mediated iron transport in A. thaliana (Mary et al., 2015). The VIT_PAM1 target site on *VIT1* begins at 347 bp of the open reading frame (Appendix C.1), equating to approximately 116 amino acid residues into the translated polypeptide sequence. The target site on *VIT2* begins at 323 bp Appendix C.2), equating to approximately amino acid residue 108 after translation. This places the target site roughly between the positions of the critical TMD1 and TMD2 domains of both *VIT1* (TMD1: 92-114, TMD2: 120-142) and *VIT2* (TMD1: 84-106, TMD2: 112-134).

The V1 mutant had a large 216 bp insert in the target area of *VIT2*. This will lead to the translation of 72 additional amino acid residues between TMD1 and TMD2, likely causing significant conformational changes. Since the relative position of these TMDs is essential for the correct dimerization and channel formation of the transporter complex, it is probable that this mutation would lead to a non-functional VIT2 in the V1 mutant. The V2 mutant line had a 1 bp insertion in the target area of *VIT2*, leading to a frameshift mutation beginning at the corresponding 114 amino acid residue position. This frameshift would lead to conserved glycine and methionine residues on TMD2 being changed to arginine and histidine, respectively. Without these critical residues metal transport would be disrupted. In addition, a premature stop codon will occur after the amino acid in position 130, leading to the loss of TMD3-5 during translation. This would likely severely affect protein structure and dimerization and thereby inactivating functional metal transport mediated by VIT2 in the V2 mutant. Both V1 and V2 mutant lines contained mixtures of different mutations on *VIT1*. V1 contained mostly a 1 bp insertion (70 %), which is likely to inhibit VIT1 function as previously discussed, but also in in-frame 3 bp deletion and WT, which would likely not severely impact protein function. V2 contained several in-frame deletions for VIT1 in addition to the WT sequence, with only 23 % of the detected PCR products carried a frame-shifting 1 bp insertion. Both V1 and V2 might therefore have reduced VIT1 activity, with likely a bigger effect for V1, although this is speculative.

The V1 mutant line is therefore though to be a probable knockout of VIT2 and a possible partial reduction/knock-down of VIT1. V2 is considered a knockout mutant for VIT2, with a possible partial reduction/knock-down of VIT1. This presents another problem with gRNAs targeting both *VIT1* and *VIT2*. It is not certain that the predicted metal transporters encoded by the genes are exactly functionally redundant, and by only creating mutants with both genes disrupted, or partially disrupted in this case, it is not possible to properly assess the individual roles each gene plays in heavy metal metabolism during later characterization of mutants. If it is possible to design gRNAs that are only specific for either *VIT1* or *VIT2*, something which would require sufficient polymorphisms adjacent to a suitable PAM site on each gene, their individual and combined effects on heavy metal tolerance could be compared. The genotyping and predicted protein function of each mutant line is summarized in Table 5.1 for clarity.

Table 5.1: The final mutant lines used for the heavy metal exposure experiment, including the CRISPR/Cas9 plasmids used to induce mutations, genes edited, genotyping results, and the predicted effects of mutations on target protein function.

Mutant line	Plasmid used	Edited gene	Indel (% of sequence traces)	Predicted effect on protein function
A1	ATPase_PAM2	*ATPase5-1B*	-13 (96%), various (4%)	Knockout (Possible contamination)
A2	ATPase_PAM2	*ATPase5-1B*	+2 (94%), -7 (6%)	Knockout
V1	VIT_PAM1	*VIT1*	+1 (70%), -3 (22%), WT (8%)	Downregulation/knockdown
		VIT2	+216	Probable knockout
V2	VIT_PAM1	*VIT1*	-12 (32%), -3 (18%), +1 (23%), -12 (13%), WT (7%)	Downregulation/knockdown
		VIT2	+1	Knockout

5.3 Characterization of growth and photosynthetic parameters of *P. tricornutum* mutants when exposed to zinc and cadmium

To investigate the roles *VIT1*, *VIT2* and *ATPase5-1B* play in the heavy metal tolerance of *P. tricornutum,* an experiment was designed to investigate whether the CRIPR/Cas9-induced mutations of these genes led to changes in the heavy metal tolerance of the mutated *P. tricornutum* lines. Heavy metals used for the experiment were selected based to the available literature. Expressions of *VIT1*, *VIT2* and *ATPase5-1B* are all upregulated when *P. tricornutum* is exposed to cadmium, as demonstrated by both Brembu et al. (2011) and Ma et al. (2021), and their encoded metal transporters are therefore predicted to play important roles in adaptive heavy metal tolerance. Furthermore, proteins similar to ATPase5-1B in other species are associated with hypertolerance and hyperaccumulation of both cadmium and zinc (Hanikenne et al., 2008). VIT1 homologs are also know to bind zinc (Zhang et al., 2012). There is therefore reason to believe that *VIT1*, *VIT2* and *ATPase5-1B* are important for the transport of both zinc and cadmium. The concentrations selected for the present study were based on available literature on the tolerance of WT *P. tricornutum* to each heavy metal, using concentrations slightly lower than those known to induce high toxicity (Horvatić & Peršić, 2007). The goal was to use a concentration where an increased sensitivity to heavy metals in mutants would result in higher toxicity without severely affecting WT, and avoid an outcome where high general toxicity in all cell lines made lower mutant tolerances indistinguishable from WT.

The VIT1 and VIT2 proteins are predicted to be involved in vacuolar storage of metals (Brembu et al., 2011), making it likely that loss of their activity would lead to mutants having a reduced ability to effectively move excess zinc and cadmium ions from the cytosol into storage vacuoles for detoxification. This would likely lead to higher cytosolic heavy metal concentrations during exposure. The ATPase5-1B protein is likely responsible for cellular efflux of zinc and cadmium, so knockout mutants will likely have a reduced ability to remove excess heavy metal ions from the cell during exposure. Cell growth and photosynthetic efficiency were chosen as indicators of high cytosolic heavy metal concentrations, as these are known to be impacted by heavy metal toxicity.

5.3.1 Effects of zinc and cadmium on cell growth of mutant lines

Neither medium containing zinc nor cadmium caused a significant change in the maximum specific growth (K'_{max}) rate of any *P. tricornutum* cell lines compared to control medium (CT), with the only exception being A2 when grown in medium containing Zn^{2+}. This could mean that the exponential growth of the A2 mutant line was inhibited by zinc, but it is more likely to be an experimental error since the same line displayed no difference in K'_{max} from CT when grown in the comparatively much more toxic Cd^{2+} medium. The lack of significant effects on K'_{max} could mean that the concentrations used were too low to induce toxicity in the mutant lines. In fact, the zinc medium caused a higher final cell concentration for WT than the CT medium. This growth stimulating effect of long term zinc exposure at lower concentrations is consistent with the literature (Horvatić & Peršić, 2007), but indicates that

the intended critical concentrations of each heavy metal might have been underestimated under our experimental conditions. Since K'_{max} was calculated based on growth data from the first days of exponential growth, this could also indicate that the effects of disrupting the target genes would not be visible after such a short period. This assumption might be further investigated when observing the growth of cell lines later in the experiment. The V1 and V2 mutant lines consistently grew at similar rates. There was no notable difference in the relative final concentration for V1 or V2 when exposed to zinc or cadmium compared to WT, however, indicating that the mutants did not have lower long-term tolerance for the heavy metals either. This could mean that the concentrations used were simply too low for any difference in heavy metal tolerance to be observed between mutants and WT or indicate that the mutations carried by the V1 and V2 mutants did not result in higher toxicity when exposed to cadmium or zinc. The A2 mutant line grew consistently slower than the V1 and V2 lines. The A1 line, however, grew very similarly to V1 and V2 up until Day 8. Between In the presence of cadmium the growth of A1 had apparently slowed down between Day 8 and Day 17, and both A1 and A2 ended up with similar final concentrations. The relative final concentrations of A1 and A2 were not notably different from WT between the zinc and CT treatments but were lower than for WT between the cadmium and CT treatments. This means that the A1 and A2 mutants were more sensitive to long term cadmium exposure than WT, possibly due to toxicity caused by higher cytosolic cadmium concentrations, indicating a reduction in adaptive tolerance.

5.3.2 Effects of zinc and cadmium on photosynthetic efficiency of mutant lines

Variable chlorophyll *a* fluorescence showed that the mutant lines had small but significant reductions in various photosynthetic efficiency parameters compared to WT. These reductions were present even in the CT treatment, however, and no significant differences were observed between treatments for any mutant line. If loss of function of VIT1, VIT2 and ATPase5-1B had led to a reduced ability to remove heavy metals from the cytoplasm, it would be expected that heavy metal toxicity would disproportionally reduce the photosynthetic efficiency of the mutant lines. The photosynthetic efficiency of the mutant lines, particularly A1 and A2, was somewhat reduced compared to WT irrespective of heavy metal exposure, but no indication of higher heavy metal sensitivity was observed for any of the mutant lines. It is again possible that differences between mutant lines and WT might have been observed at higher heavy metal concentrations, and it would therefore have been interesting to repeat the experiment with different concentrations of zinc and cadmium. Since the measurement of variable chlorophyll *a* fluorescence was conducted on Day 3 of the experiment, the calculated parameters would only be able to describe differences in tolerance to short term heavy metal exposure. As the growth data indicated, however, it seems like the A1 and A2 mutants displayed higher sensitivity to cadmium exposure in the long term. Fluorescence measurement was conducted early to prevent buildup of dead cells present in the liquid cultures from affecting the measurements, resulting in this test not being useful for identifying changes in long term adaptability to heavy metal exposure. Nevertheless, for the V1 and V2 lines this is an additional indication that their mutations did not result in higher cellular toxicity when exposed to zinc or cadmium.

5.3.3 Can the mutations generated in *VIT1*, *VIT2* and *ATPase5-1B* explain the phenotypes observed for the mutant lines?

5.3.3.1 Mutant lines displayed lower cell growth and photoinhibition irrespective of heavy metal exposure

All mutant lines displayed lower cell growth and varying degrees of photoinhibition compared to WT irrespective of heavy metal exposure. This could be because of off-target mutations caused by the Cas9 nuclease. The *S. pyogenes* Cas9 system utilized in the project has been shown to bind to 5′-NGA-3′ and 5′-NAG-3′ PAM sequences in addition to 5′-NGG-3′, and the specificity of the gRNA target sequence decreases with distance from the PAM site (Yamamoto, 2015). This means that the nuclease can recognize and cleave unintended targets on the host genome that display sufficient similarity to the intended target sequence, potentially leading to disruption of genes that are important for general cellular fitness. The V1 and V2 lines displayed similar fitness, which could be explained by the VIT_PAM1 causing similar off-target mutations in both cell lines. If the A1 and A2 mutant lines carried different degrees of off-target mutations, this could explain why divergent cell growth was observed between the two in the beginning of the experiment. V1 and V2 might also display reduced cell fitness due to genome damage commonly sustained when cells are being transformed through biolistics (Liu et al., 2019). Genome-wide sequencing would be necessary to properly assess the extent of off-target mutations or other random genome damage, however. Another possibility is that the target genes are involved in the general metal homeostasis of *P. tricornutum*, rather than (or in addition to) heavy metal tolerance. For example, *VIT1*, *VIT2* and *ATPase5-1B* homologs have all been shown to transport several different metal ions in other organisms, many of which are parts of regular metabolism (Argüello et al., 2007; Zhang et al., 2012). VIT1-like transporters might also be implicated in maintaining the proton-gradient between the vacuolar space and cytosol (Kato et al., 2019). If this is the case, disruption of these genes could lead to mutants with broad reductions in cellular fitness despite no reductions in zinc or cadmium tolerance being observed.

5.3.3.2 Loss of ATPase5-1B appears to inhibit long-term cadmium tolerance

It was predicted that the knockout lines obtained for the *VIT2* (V2 and probably V1) and *ATPase5-1B* (A1 and A2) genes would have a reduced ability to remove heavy metals from the cytosol by vacuolar storage and efflux, respectively, leading to increased sensitivity to zinc and cadmium exposure. The V1 and V2 lines were also considered to have reduced VIT1 activity, due to a subset of each cell line carrying frameshift mutations on this gene as well. None of the mutant lines displayed reduced tolerance to heavy metals when observing exponential cell growth or photosynthetic efficiency, however. This could be because these measurements were conducted after short-term exposure. Studies have shown that the tolerances of *P. tricornutum* to cadmium and possibly zinc are higher after long-term exposure compared to short-term exposure (Horvatić & Peršić, 2007), indicating that optimal heavy metal tolerance requires adaptation. This long-term adaptation likely involves greater changes in gene expression of proteins like the *VIT1/2* and *ATPase5-1B* metal transporters.

The effect of disrupting such adaptive responses might not be clearly distinguishable as early as these parameters were determined. The A1 and A2 *ATPase5-1B* knockout lines did show a reduced growth, which indicates that the *ATPase5-1B* is indeed involved in long term adaptive tolerance to the metal.

The A1 and A2 mutants were cultivated in medium containing zeocin up to the acclimation period prior to heavy metal exposure, selection pressure would conserve the CRISPR/Cas9 episome due to the zeocin-resistance it conferred. This means that there was a period after genotyping of these mutant lines by HRM and sequencing where re-editing could have occurred (Sharma et al., 2018), potentially changing the mutations identified in the *ATPase5-1B* gene. This is considered unlikely for the 13 bp deletion, however. Since V1 and V2 were transformed by biolistics Cas9 and gRNA were integrated into the *P. tricornutum* genome and continually expressed, potentially also leading to re-editing at the target site of *VIT1* and/or *VIT2*. While this is unlikely occur in the case of large indels in the target sequence, it might cause mutations that re-establish the original reading frame for mutants with small 1 or 2 bp insertions and deletions. Such re-edits could largely restore the function of a previously presumed knockout, and thereby result in cells behaving as a WT rather than the presumed mutant line in the heavy metal tolerance experiment. Since both *VIT1/2* mutants behaved similarly throughout the experiment, and the *ATPase5-1B* mutants displayed similar long-term sensitivity to cadmium, it seems unlikely that this occurred. Characterizing each cell line immediately before or after the heavy metal exposure experiment would identify if this occurred. It is also possible that the reported upregulations of *VIT1*, *VIT2* and *ATPase5-1B* by cadmium (Brembu et al., 2011; Ma et al., 2021) are part of a nonspecific heavy metal stress response without them transporting neither zinc nor cadmium specifically, which could explain why the exponential growth and photosynthetic efficiency of mutant lines were not disproportionally more inhibited by zinc or cadmium than WT. The reduced long-term tolerance of the A1 and A2 lines makes this unlikely in the case of *ATPase5-1B* however, and there are other explanations for the lack of an increased toxicity observed in the *VIT1/2* mutants.

5.3.3.3 The possibility of functional redundancy in of *P. tricornutum* heavy metal tolerance compensating for loss of VIT1, VIT2 and ATPase5-1B

Functional redundancy in the heavy metal tolerance system of *P. tricornutum* might explain why little indication of higher toxicity was observed in mutant lines during heavy metal exposure. Organisms that are highly tolerant towards different environmental stressors are known to utilize numerous defensive mechanisms to maintain normal cell function (Coker, 2019). This is also likely the case for *P. tricornutum*. Diatoms are known to protect themselves from heavy metals through extracellular biosorption macromolecules, metal transporters for removal and compartmentalization and with detoxifying compounds like phytochelatins (PCs) (Danouche et al., 2021; Kumar et al., 2015). Exposure to cadmium has already been shown to increase the expression of many different genes in *P. tricornutum*, including genes coding for several metal transporters that were not targeted during this project (Brembu et al., 2011; Ma et al., 2021), and the post-transcriptional production of PCs (Brembu et al., 2011; Morelli & Fantozzi, 2008). In addition to the cadmium inducible *VIT1*

and *VIT2*, the *P. tricornutum* genome encodes a third VIT1/Ccc1 homolog (Brembu et al., 2011), which further indicates the presence of a diversified and multi-layered approach to dealing with heavy metals. Since several of these mechanisms and genes likely function independently of VIT1, VIT2 and ATPase5-1B proteins, it is quite possible that they would be able to compensate for a lack of their activity at the heavy metal concentrations used in our experiment. For example, reduction in cellular efflux of cadmium caused by ATPase5-1B disruption might lead to an upregulation of other transporters able to pump cadmium out of the cell or the increased production of PCs, thereby still maintaining non-toxic cytosolic concentrations. Such compensatory adaptation could cause mutant lines to display a sensitivity to cadmium similar to the WT despite a part of their heavy metal defences being disabled. This could also explain why the V1 and V2 lines grew better than A1 and A2, as they only carried knockouts of the *VIT2* gene and a knockdown of *VIT1* at best. If these genes are redundant, the expression of functional VIT1 proteins could compensate for a lack of VIT2 transporters. It would be interesting to measure differences in gene expression, and other detoxification mechanisms such as the production of PCs, between these mutant lines and WT when exposed to heavy metals to identify if such compensation takes place.

5.3.3.4 Heavy metal exposure might have exerted selection pressure on mixed mutant lines

The observed results might also have been affected by the presence of multiple genotypes in some mutant lines because of the incomplete screening process. If some of these cells still expressed functional target proteins, either by being unedited or by carrying a mutation that does not affect transport activity, they might have outcompeted cells carrying knockout mutations under the selection pressure of heavy metal exposure due to retaining higher tolerances. Since both the V1 and V2 lines contained cells carrying unedited and in/frame edited *VIT1* genes, it is possible that the relative prevalence of these genotypes dominated during the growth experiment due to their functioning *VIT1* gene providing higher tolerances. This might have contributed to why no reduction in heavy metal tolerance was observed in the V1 and V2 lines, since both lines would have even higher compensatory VIT1 activity than predicted based on the sequencing data. Due to A1 and A2 displaying similar long-term tolerance to cadmium, it seems unlikely that the suspected contamination of A1 with cells presenting other mutations in *ATPase5-1B* than the 13 bp deletion had a similar effect. Sequencing the target genes for each line again at the end of the heavy metal exposure experiment could have been used to reveal any significant shifts in genotype composition.

5.3.3.5 The potential significance of measuring cellular contents of zinc and cadmium

It is important to point out that the selected test parameters only indirectly measure the effects of knocking out the target genes on *P. tricornutum* by the effects of heavy metal toxicity on growth and photosynthesis. But, as discussed, aspects of heavy metal tolerance might be disrupted without causing an outright decrease in cellular fitness if other parts of the tolerance system compensate for the disabled proteins. An important part of the heavy metal metabolism of *P. tricornutum* that was not investigated during the characterization of mutants, and which might have great commercial significance, is their ability to sequester excess heavy metal into vacuoles. The plan was originally to compare the cellular contents of zinc and cadmium of mutant lines to WT when exposed to these heavy metals using inductively coupled plasma mass spectrometry (ICP-MS) analysis. Obtaining such data would have been a very interesting endpoint for this project, but there was not enough time for this analysis. If a *VIT1/2* knockout line had a lower cellular heavy metal content after exposure compared to WT, despite displaying similar sensitivity, it could be inferred that *VIT1* and/or *VIT2* indeed were involved in vacuolar storage of the metal and cellular efflux was likely upregulated to compensate for its loss of function. It could also mean that the loss of VIT1 and VIT2 function led to decreased cellular influx, due to downregulation of other metal transporters or upregulation of extracellular molecules that inhibit the uptake of heavy metals. Conversely, if higher intracellular heavy metal concentrations were observed for *ATPase5-1B* knockouts with similar apparent tolerance to WT, it is likely that vacuolar storage mechanisms where upregulated to compensate for a reduction in cellular efflux. It is also possible that VIT1, VIT2 and ATPase5-1B are located differently than predicted. A study of the cellular localization of these transporters would help interpret such results better. Since the interest in *P. tricornutum* heavy metal tolerance is largely related to its potential use for bioremediation of wastewater, better understanding of the molecular factors influencing accumulation efficiencies are of high importance.

6 Conclusion

The aim of this project was to create knockouts of the predicted metal transporter encoding genes *VIT1*, *VIT2* and *ATPase5-1B* using the CRISPR/Cas9-based gene editing tool, to investigate the roles these proteins play in the reported heavy metal tolerance of the marine diatom *P. tricornutum*. Plasmids designed to express Cas9 nuclease and one gRNA targeting either *VIT1* and *VIT2* or *ATPase5-1B* were successfully created using traditional restriction-ligation cloning. These plasmids were subsequently used to transform *P. tricornutum* by bacterial conjugation and biolistics, and transformants were screened for indels by HRM analysis and Sanger sequencing. Both transformation and mutation efficiencies were low for cells transformed through conjugation, and inconsistent for cells transformed through biolistics. Transformed cells successfully underwent CRISPR/Cas9 mediated mutation, but due to time constraints the screening and isolation of single cell-derived pure mutant lines had to be concluded prematurely. Four mutant lines were chosen to proceed with the project: two lines with mutations that would generate non-functional ATPase5-1B, two lines with mutations that would lead to a decrease in VIT1 activity and to non-functional VIT2. The heavy metal tolerances of these four mutant lines were compared to wild type *P. tricornutum* when exposed to zinc and cadmium, two metals predicted to be transported by the targeted metal transporters based on the available literature. None of the mutant lines displayed statistically significant changes in exponential growth or photosynthetic efficiency when exposed to 5 mg/L Zn^{2+} or 1.25 mg/L Cd^{2+}, indicating that the mutant lines did not experience higher short-term heavy metal toxicity at these concentrations. The *ATPase5-1B* mutants displayed higher relative long term growth inhibition than WT when exposed to 1.25 mg/L Cd^{2+}, indicating that the encoded metal transporter might be involved in adaptive heavy metal tolerance in *P. tricornutum*. It is hypothesized that compensatory upregulation of other transporters and higher production of detoxifying molecules like PCs could explain the lack of increased toxicity observed in the *VIT2* and *ATPase5-1B* knockouts. A WT-like heavy metal tolerance for the *VIT2* knockouts might also be explained by functionalVIT1 still being produced in these lines. Further studies are required, particularly obtaining pure bi-allelic knockout mutants for the targeted genes and investigating their cellular content of heavy metals upon exposure. These findings might be of great interest for future development of microalgal bioremediation technologies

7 Future work

The insights gained from this project can be used to guide further investigations of metal transporters in *P. tricornutum*. To verify the observations made in this thesis, complete purification of mutants would be necessary to remove the possible error-source that a contamination with cells encoding functional gene products represent when measuring the effect of heavy metals on mutant cultures. It would also be an advantage to utilize more than two knockout mutants for each gene to better account for the effects of potential off-target mutations. Using multiple heavy metal concentrations, and perhaps other heavy metals than zinc and cadmium, might increase the likelihood of observing differences between mutant lines and WT. Generating mutants carrying knockouts of *VIT1* and *VIT2* individually and simultaneously would help identify to which degree they share function and whether they are able to compensate for a loss of function in the other. The function of each gene might also be investigated more easily by expressing them in alternative organisms like yeast cells. Similar experiments could furthermore be used to investigate the function of other putative *P. tricornutum* metal transporters (Brembu et al., 2011; Ma et al., 2021).

To assess whether the disruption of *VIT1*, *VIT2* or *ATPase5-1B* leads to compensatory upregulation of other genes involved in heavy metal tolerance, broad RNA sequencing or qPCR of candidate genes could be conducted on knockout lines after exposure to zinc or cadmium. Proteomic analysis could also be conducted to identify a possible compensatory overproduction of other transporter proteins and PCs. Proteomics would also determine if target proteins were knocked down/out.

ICP-MS or methods of intracellular metal localization could help identify changes in heavy metal metabolism in knockout lines that do not affect indicators of general cellular toxicity such as reduced growth and photosynthetic efficiency. Determination of changes in cellular heavy metal contents would also give clearer indications of whether the functions of *VIT1*, *VIT2* and *ATPase5-1B* are consistent with predictions made from homologous genes.

The generation of *P. tricornutum* mutants carrying duplicates of target genes would provide an opportunity to further investigate their function and possible applications for improved bioremediation technology. If the metal transporter proteins investigated here are indeed involved in adaptive heavy metal tolerance as predicted, a mutant carrying multiple copies of its encoding gene might display hypertolerance as observed in other organisms carrying multiple copies of homologous genes (Hanikenne et al., 2008). This would also present a possible method of generating hypertolerant species that could be used for more efficient bioremediation in the future.

8 References

ACEA Biosciences. (2018). *NovoCyte ® Flow cytometer operator's guide*.

Ali, H., & Khan, E. (2018). What are heavy metals? Long-standing controversy over the scientific use of the term 'heavy metals'–proposal of a comprehensive definition. *Toxicological & Environmental Chemistry*, *100*(1), 6-19.

Ali, H., & Khan, E. (2019). Trophic transfer, bioaccumulation, and biomagnification of non-essential hazardous heavy metals and metalloids in food chains/webs—Concepts and implications for wildlife and human health. *Human and Ecological Risk Assessment: An International Journal*, *25*(6), 1353-1376.

Ali, H., Khan, E., & Ilahi, I. (2019). Environmental chemistry and ecotoxicology of hazardous heavy metals: environmental persistence, toxicity, and bioaccumulation. *Journal of Chemistry*, *2019*, 6730305.

Allen, A. E., Dupont, C. L., Oborník, M., Horák, A., Nunes-Nesi, A., McCrow, J. P., Zheng, H., Johnson, D. A., Hu, H., & Fernie, A. R. (2011). Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature*, *473*(7346), 203-207.

Argüello, J. M., Eren, E., & González-Guerrero, M. (2007). The structure and function of heavy metal transport P1B-ATPases. *Biometals*, *20*(3), 233-248.

Arief, V. O., Trilestari, K., Sunarso, J., Indraswati, N., & Ismadji, S. (2008). Recent progress on biosorption of heavy metals from liquids using low cost biosorbents: characterization, biosorption parameters and mechanism studies. *Clean – Soil, Air, Water*, *36*(12), 937-962.

Armbrust, E. V. (2009). The life of diatoms in the world's oceans. *Nature*, *459*(7244), 185-192.

Bacchetta, R., Maran, B., Marelli, M., Santo, N., & Tremolada, P. (2016). Role of soluble zinc in ZnO nanoparticle cytotoxicity in *Daphnia magna*: A morphological approach. *Environmental Research*, *148*, 376-385.

Balzano, S., Sardo, A., Blasio, M., Chahine, T. B., Dell'Anno, F., Sansone, C., & Brunet, C. (2020). Microalgal metallothioneins and phytochelatins and their potential use in bioremediation. *Frontiers in Microbiology*, *11*, 517.

Benoiston, A.-S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., & Bowler, C. (2017). The evolution of diatoms and their biogeochemical functions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1728), 20160397.

Bertrand, M., Schoefs, B., Siffel, P., Rohacek, K., & Molnar, I. (2001). Cadmium inhibits epoxidation of diatoxanthin to diadinoxanthin in the xanthophyll cycle of the marine diatom *Phaeodactylum tricornutum*. *FEBS Letters*, *508*(1), 153-156.

Blaby-Haas, C. E., & Merchant, S. S. (2012). The ins and outs of algal metal transport. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1823*(9), 1531-1552.

Boesenberg-Smith, K. A., Pessarakli, M. M., & Wolk, D. M. (2012). Assessment of DNA yield and purity: an overlooked detail of PCR troubleshooting. *Clinical Microbiology Newsletter*, *34*(1), 1-6.

Bohlin, K. (1897). Zur morphologie und biologie einzelliger algen (vorläufige mitteilung). *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, *54*(9), 507-529.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R. P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., . . . Grigoriev, I. V. (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, *456*(7219), 239-244.

Bozarth, A., Maier, U.-G., & Zauner, S. (2009). Diatoms in biotechnology: modern tools and applications. *Applied Microbiology and Biotechnology*, *82*(2), 195-201.

Brembu, T., Jørstad, M., Winge, P., Valle, K. C., & Bones, A. M. (2011). Genome-wide profiling of responses to cadmium in the diatom *Phaeodactylum tricornutum*. *Environmental Science & Technology*, *45*(18), 7640-7647.

Briffa, J., Sinagra, E., & Blundell, R. (2020). Heavy metal pollution in the environment and their toxicological effects on humans. *Heliyon*, *6*(9), e04691.

Brinza, L., Dring, M. J., & Gavrilescu, M. (2007). Marine micro and macro algal species as biosorbents for heavy metals. *Environmental Engineering & Management Journal 6*(3), 237-251.

Castro-González, M., & Méndez-Armenta, M. (2008). Heavy metals: Implications associated to fish consumption. *Environmental Toxicology and Pharmacology*, *26*(3), 263-271.

Chunhabundit, R. (2016). Cadmium exposure and potential health risk from foods in contaminated area, Thailand. *Toxicological Research*, *32*(1), 65-72.

Cobbina, S. J., Duwiejuah, A. B., Quansah, R., Obiri, S., & Bakobie, N. (2015). Comparative assessment of heavy metals in drinking water sources in two small-scale mining communities in northern Ghana. *International Journal of Environmental Research and Public Health*, *12*(9), 10620-10634.

Coker, J. A. (2019). Recent advances in understanding extremophiles. *F1000Research*, *8*, F1000 Faculty Rev-1917.

Conant, D., Hsiau, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., & Holden, K. (2022). Inference of CRISPR edits from sanger trace data. *The CRISPR Journal*, *5*(1), 123-130.

Daboussi, F., Leduc, S., Maréchal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D. F., Cavarec, L., & Duchateau, P. (2014). Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nature Communications*, *5*, 3831.

Danouche, M., El Ghachtouli, N., & El Arroussi, H. (2021). Phycoremediation mechanisms of heavy metals using living green microalgae: physicochemical and molecular approaches for enhancing selectivity and removal capacity. *Heliyon*, 7(7), e07609.

de Paiva Magalhães, D., da Costa Marques, M. R., Baptista, D. F., & Buss, D. F. (2015). Metal bioavailability and toxicity in freshwaters. *Environmental Chemistry Letters*, *13*(1), 69-87.

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., & Falciatore, A. (2009). Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Research*, *37*(14), e96.

De Tommasi, E., Gielis, J., & Rogato, A. (2017). Diatom frustule morphogenesis and function: a multidisciplinary survey. *Marine Genomics*, *35*, 1-18.

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096.

Duffus, J. H. (2002). "Heavy metals" a meaningless term? . *Pure and Applied Chemistry*, *74*(5), 793-807.

Eilers, P., & Peeters, J. (1988). A model for the relationship between light intensity and the rate of photosynthesis in phytoplankton. *Ecological Modelling*, *42*(3-4), 199-215.

Eisler, R. (1993). Zinc hazards to fish, wildlife, and invertebrates: a synoptic review. *Contaminant Hazard Reviews Report 26,* US Department of the Interior, Fish and Wildlife Service, 106

Falciatore, A., & Bowler, C. (2002). Revealing the molecular secrets of marine diatoms. *Annual Review of Plant Biology*, *53*(1), 109-130.

Fiore, C. L., Jarett, J. K., Olson, N. D., & Lesser, M. P. (2010). Nitrogen fixation and nitrogen transformations in marine symbioses. *Trends in Microbiology*, *18*(10), 455-463.

Fosmire, G. J. (1990). Zinc toxicity. *The American Journal of Clinical Nutrition*, *51*(2), 225-227.

Galas, L., Burel, C., Schapman, D., Ropitaux, M., Bernard, S., Bénard, M., & Bardor, M. (2021). Comparative structural and functional analyses of the fusiform, oval, and triradiate morphotypes of *Phaeodactylum tricornutum* Pt3 strain. *Frontiers in Plant Science*, *12*, 584.

Gottschalk, F., Sonderer, T., Scholz, R. W., & Nowack, B. (2009). Modeled environmental concentrations of engineered nanomaterials (TiO2, ZnO, Ag, CNT, fullerenes) for different regions. *Environmental Science & Technology*, *43*(24), 9216-9222.

Guiry, M. D. (2012). How many species of algae are there? *Journal of Phycology*, *48*(5), 1057-1063.

Hanikenne, M., Talke, I. N., Haydon, M. J., Lanz, C., Nolte, A., Motte, P., Kroymann, J., Weigel, D., & Krämer, U. (2008). Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of *HMA4*. *Nature*, *453*(7193), 391-395.

Haynes, W. M., Lide, D. R., & Bruno, T. J. (2017). Abundance of elements in the Earth's crust and in the sea. In *Handbook of Chemistry and Physics* (97 ed., pp. 2380). CRC Press.

He, L., Han, X., & Yu, Z. (2014). A rare *Phaeodactylum tricornutum* cruciform morphotype: culture conditions, transformation and unique fatty acid characteristics. *PLoS One*, *9*(4), e93922.

Herrmann, M. G., Durtschi, J. D., Bromley, L. K., Wittwer, C. T., & Voelkerding, K. V. (2006). Amplicon DNA melting analysis for mutation scanning and genotyping: cross-platform comparison of instruments and dyes. *Clinical Chemistry*, *52*(3), 494-503.

Hirata, K., Tsuji, N., & Miyamoto, K. (2005). Biosynthetic regulation of phytochelatins, heavy metal-binding peptides. *Journal of Bioscience and Bioengineering*, *100*(6), 593-599.

Horvatić, J., & Peršić, V. (2007). The Effect of Ni2+, Co2+, Zn2+, Cd2+ and Hg2+ on the Growth Rate of Marine Diatom *Phaeodactylum tricornutum* Bohlin: Microplate Growth Inhibition Test. *Bulletin of Environmental Contamination and Toxicology*, *79*(5), 494-498.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., & Shalem, O. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*(9), 827-832.

Ibemenuga, K. N. (2013). Bioaccumulation and toxic effects of some heavy metals in freshwater fishes. *Animal Research International*, *10*(3), 1792–1798.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the *IAP* gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429-5433.

Ji, Y., Xie, X., & Wang, G. (2018). Effects of the heavy metal cadmium on photosynthetic activity and the xanthophyll cycle in *Phaeodactylum tricornutum*. *Journal of Oceanology and Limnology*, *36*(6), 2194-2201.

Jiang, J., Pi, J., & Cai, J. (2018). The advancing of zinc oxide nanoparticles for biomedical applications. *Bioinorganic Chemistry and Applications*, *2018*, 1062562

Kahru, A., & Dubourguier, H.-C. (2010). From ecotoxicology to nanoecotoxicology. *Toxicology*, *269*(2-3), 105-119.

Kaplan, D. (2013). Absorption and adsorption of heavy metals by microalgae. In A. Richmond & Q. Hu (Eds.), *Handbook of Microalgal Culture: Applied Phycology and Biotechnology* (Vol. 2, pp. 602-611). Wiley-Blackwell.

Karas, B. J., Diner, R. E., Lefebvre, S. C., McQuaid, J., Phillips, A. P., Noddings, C. M., Brunson, J. K., Valas, R. E., Deerinck, T. J., & Jablanovic, J. (2015). Designer diatom episomes delivered by bacterial conjugation. *Nature Communications*, *6*(1), 1-10.

Kato, T., Kumazaki, K., Wada, M., Taniguchi, R., Nakane, T., Yamashita, K., Hirata, K., Ishitani, R., Ito, K., & Nishizawa, T. (2019). Crystal structure of plant vacuolar iron transporter VIT1. *Nature Plants*, *5*(3), 308-315.

Kim, S. A., Punshon, T., Lanzirotti, A., Li, L., Alonso, J. M., Ecker, J. R., Kaplan, J., & Guerinot, M. L. (2006). Localization of iron in Arabidopsis seed requires the vacuolar membrane transporter VIT1. *Science*, *314*(5803), 1295-1298.

Krupa, Z. (1999). Cadmium against higher plant photosynthesis-a variety of effects and where do they possibly come from? *Zeitschrift für Naturforschung - Section C Journal of Biosciences*, *54*(9-10), 723-729.

Kubier, A., Wilkin, R. T., & Pichler, T. (2019). Cadmium in soils and groundwater: a review. *Applied Geochemistry*, *108*, 104388.

Kumar, K. S., Dahms, H.-U., Won, E.-J., Lee, J.-S., & Shin, K.-H. (2015). Microalgae–a promising tool for heavy metal remediation. *Ecotoxicology and environmental safety*, *113*, 329-352.

Kumar, V., Parihar, R. D., Sharma, A., Bakshi, P., Sidhu, G. P. S., Bali, A. S., Karaouzas, I., Bhardwaj, R., Thukral, A. K., & Gyasi-Agyei, Y. (2019). Global evaluation of heavy metal content in surface water bodies: A meta-analysis using heavy metal pollution indices and multivariate statistical analyses. *Chemosphere*, *236*, 124364.

Lewin, J. C. (1958). The taxonomic position of *Phaeodactylum tricornutum*. *Microbiology*, *18*(2), 427-432.

Li, L., Chen, O. S., Ward, D. M., & Kaplan, J. (2001). CCC1 is a transporter that mediates vacuolar iron storage in yeast. *Journal of Biological Chemistry*, *276*(31), 29515-29519.

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry*, *79*, 181-211.

Liu, J., Nannas, N. J., Fu, F.-f., Shi, J., Aspinwall, B., Parrott, W. A., & Dawe, R. K. (2019). Genome-scale sequence disruption following biolistic transformation in rice and maize. *The Plant Cell*, *31*(2), 368-383.

Ma, J., Zhou, B., Chen, F., & Pan, K. (2021). How marine diatoms cope with metal challenge: Insights from the morphotype-dependent metal tolerance in *Phaeodactylum tricornutum*. *Ecotoxicology and Environmental Safety*, *208*, 111715.

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., Wincker, P., Iudicone, D., de Vargas, C., & Bittner, L. (2016). Insights into global diatom distribution and diversity in the world's ocean. *Proceedings of the National Academy of Sciences USA*, *113*(11), E1516-E1525.

Mann, D. G., & Vanormelingen, P. (2013). An inordinate fondness? The number, distributions, and origins of diatom species. *Journal of Eukaryotic Microbiology*, *60*(4), 414-420.

Mantzorou, A., Navakoudis, E., Paschalidis, K., & Ververidis, F. (2018). Microalgae: a potential tool for remediating aquatic environments from toxic metals. *International Journal of Environmental Science and Technology*, *15*(8), 1815-1830.

Marella, T. K., Saxena, A., & Tiwari, A. (2020). Diatom mediated heavy metal remediation: a review. *Bioresource Technology*, *305*, 123068.

Martin-Jézéquel, V., & Tesson, B. (2012). *Phaeodactylum tricornutum* polymorphism: an overwiev. In K. Heimann & C. Katsaros (Eds.), *Advances in Algal Cell Biology* (pp. 43-80). De Gruyter.

Martino, A. D., Meichenin, A., Shi, J., Pan, K., & Bowler, C. (2007). Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions. *Journal of Phycology*, *43*(5), 992-1009.

Mary, V., Schnell Ramos, M., Gillet, C., Socha, A. L., Giraudat, J., Agorio, A., Merlot, S., Clairet, C., Kim, S. A., & Punshon, T. (2015). Bypassing iron storage in endodermal vacuoles rescues the iron mobilization defect in the natural resistance associated-macrophage protein3natural resistance associated-macrophage protein4 double mutant. *Plant Physiology*, *169*(1), 748-759.

Maxwell, K., & Johnson, G. N. (2000). Chlorophyll fluorescence - a practical guide. *Journal of Experimental Botany*, *51*(345), 659-668.

Miller, K., Eggenberger, A. L., Lee, K., Liu, F., Kang, M., Drent, M., Ruba, A., Kirscht, T., Wang, K., & Jiang, S. (2021). An improved biolistic delivery and analysis method for evaluation of DNA and CRISPR-Cas delivery efficacy in plant tissue. *Scientific Reports*, *11*, 7695

Monteiro, C. M., Castro, P. M., & Malcata, F. X. (2012). Metal uptake by microalgae: underlying mechanisms and practical applications. *Biotechnology Progress*, *28*(2), 299-311.

Moosburner, M. A., Gholami, P., McCarthy, J. K., Tan, M., Bielinski, V. A., & Allen, A. E. (2020). Multiplexed knockouts in the model diatom *Phaeodactylum* by episomal delivery of a selectable cas9. *Frontiers in Microbiology*, *11*, 5

Morelli, E., & Fantozzi, L. (2008). Phytochelatins in the diatom *Phaeodactylum tricornutum* Bohlin: an evaluation of their use as biomarkers of metal exposure in marine waters. *Bulletin of Environmental Contamination and Toxicology*, *81*(3), 236-241.

Naser, H. A. (2013). Assessment and management of heavy metal pollution in the marine environment of the Arabian Gulf: a review. *Marine Pollution Bulletin*, *72*(1), 6-13.

Nguyen-Deroche, T. L. N., Caruso, A., Le, T. T., Bui, T. V., Schoefs, B., Tremblin, G., & Morant-Manceau, A. (2012). Zinc affects differently growth, photosynthesis, antioxidant enzyme activities and phytochelatin synthase expression of four marine diatoms. *The Scientific World Journal*, *2012*, 982957

Nymark, M., Sharma, A. K., Hafskjold, M. C., Sparstad, T., Bones, A. M., & Winge, P. (2017). CRISPR/Cas9 gene editing in the marine diatom *Phaeodactylum tricornutum*. *Bio-protocol*, *7*(15), e2442.

Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M., & Winge, P. (2016). A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific Reports*, *6*, 24951.

OSPAR. (2002). OSPAR background document on cadmium. *Hazardous Substances Series,* OSPAR Commission, 58 p

Overnell, J. (1975). The effect of heavy metals on photosynthesis and loss of cell potassium in two species of marine algae, *Dunaliella tertiolecta* and *Phaeodactylum tricornutum*. *Marine Biology*, *29*(1), 99-103.

Paunov, M., Koleva, L., Vassilev, A., Vangronsveld, J., & Goltsev, V. (2018). Effects of different metals on photosynthesis: cadmium and zinc affect chlorophyll fluorescence in durum wheat. *International Journal of Molecular Sciences*, *19*(3), 787.

Perales-Vela, H. V., Peña-Castro, J. M., & Canizares-Villanueva, R. O. (2006). Heavy metal detoxification in eukaryotic microalgae. *Chemosphere*, *64*(1), 1-10.

Pfister, L., Wetzel, C. E., Klaus, J., Martínez-Carreras, N., Antonelli, M., Teuling, A. J., & McDonnell, J. J. (2017). Terrestrial diatoms as tracers in catchment hydrology: a review. *Wiley Interdisciplinary Reviews: Water*, *4*(6), e1241.

Plum, L. M., Rink, L., & Haase, H. (2010). The essential toxin: impact of zinc on human health. *International Journal of Environmental Research and Public Health*, *7*(4), 1342-1365.

Prabhakaran, P., Ashraf, M. A., & Aqma, W. S. (2016). Microbial stress response to heavy metals in the environment. *RSC Advances*, *6*(111), 109862-109877.

Qasem, N. A., Mohammed, R. H., & Lawal, D. U. (2021). Removal of heavy metal ions from wastewater: A comprehensive and critical review. *NPJ Clean Water*, *4*(1), 1-15.

Rafati-Rahimzadeh, M., Rafati-Rahimzadeh, M., Kazemi, S., & Moghadamnia, A. (2017). Cadmium toxicity and treatment: An update. *Caspian Journal of Internal Medicine*, *8*(3), 135-145.

Ralph, P. J., & Gademann, R. (2005). Rapid light curves: a powerful tool to assess photosynthetic activity. *Aquatic Botany*, *82*(3), 222-237.

Rani, A., Kumar, A., Lal, A., & Pant, M. (2014). Cellular mechanisms of cadmium-induced toxicity: a review. *International Journal of Environmental Health Research*, *24*(4), 378-399.

Ranjbar, S., & Malcata, F. X. (2022). Is genetic engineering a route to enhance microalgae-mediated bioremediation of heavy metal-containing effluents? *Molecules*, *27*(5), 1473.

Roney, N., Smith, C. V., Williams, M., Osier, M., & Paikoff, S. J. (2005). Toxicological profile for zinc. *Agency for Toxic Substances Disease Registry (ATSDR)*, 8.

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, *32*(4), 347-355.

Santos, J., Almeida, S. F., & Figueira, E. (2013). Cadmium chelation by frustulins: a novel metal tolerance mechanism in *Nitzschia palea* (Kützing) W. Smith. *Ecotoxicology*, *22*(1), 166-173.

Serif, M., Dubois, G., Finoux, A.-L., Teste, M.-A., Jallet, D., & Daboussi, F. (2018). One-step generation of multiple gene knock-outs in the diatom *Phaeodactylum tricornutum* by DNA-free genome editing. *Nature Communications*, *9*(1), 3924.

Sharma, A. K., Nymark, M., Sparstad, T., Bones, A. M., & Winge, P. (2018). Transgene-free genome editing in marine algae by bacterial conjugation–comparison with biolistic CRISPR/Cas9 transformation. *Scientific Reports*, *8*, 14401.

Sheoran, I., Singal, H., & Singh, R. (1990). Effect of cadmium and nickel on photosynthesis and the enzymes of the photosynthetic carbon reduction cycle in pigeonpea (*Cajanus cajan* L.). *Photosynthesis Research*, *23*(3), 345-351.

Siaut, M., Heijde, M., Mangogna, M., Montsant, A., Coesel, S., Allen, A., Manfredonia, A., Falciatore, A., & Bowler, C. (2007). Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*. *Gene*, *406*(1-2), 23-35.

Singh, R., Ranaivoarisoa, T. O., Gupta, D., Bai, W., & Bose, A. (2020). Genetic redundancy in iron and manganese transport in the metabolically versatile bacterium *Rhodopseudomonas palustris* TIE-1. *Applied and Environmental Microbiology*, *86*(16), e01057-20.

Slattery, S. S., Diamond, A., Wang, H., Therrien, J. A., Lant, J. T., Jazey, T., Lee, K., Klassen, Z., Desgagné-Penix, I., & Karas, B. J. (2018). An expanded plasmid-based genetic toolbox enables Cas9 genome editing and stable maintenance of synthetic pathways in *Phaeodactylum tricornutum*. *ACS Synthetic Biology*, *7*(2), 328-338.

Solioz, M., & Vulpe, C. (1996). CPx-type ATPases: a class of P-type ATPases that pump heavy metals. *Trends in Biochemical Sciences*, *21*(7), 237-241.

Sorribes-Dauden, R., Peris, D., Martínez-Pastor, M. T., & Puig, S. (2020). Structure and function of the vacuolar Ccc1/VIT1 family of iron transporters and its regulation in fungi. *Computational and Structural Biotechnology Journal*, *18*, 3712-3722.

Strand, T. A., Lale, R., Degnes, K. F., Lando, M., & Valla, S. (2014). A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PLoS One*, *9*(3), e90372.

Tchounwou, P. B., Yedjou, C. G., Patlolla, A. K., & Sutton, D. J. (2012). Heavy metal toxicity and the environment. *Molecular, Clinical and Environmental Toxicology*, 133-164.

Torres, E., Cid, A., Fidalgo, P., Herrero, C., & Abalde, J. (1997). Long-chain class III metallothioneins as a mechanism of cadmium tolerance in the marine diatom *Phaeodactylum tricornutum* Bohlin. *Aquatic Toxicology*, *39*(3), 231-246.

Tripathi, B., & Gaur, J. (2006). Physiological behavior of *Scenedesmus* sp. during exposure to elevated levels of Cu and Zn and after withdrawal of metal stress. *Protoplasma*, *229*(1), 1-9.

U.S. Geological Survey (USGS). (2021). *Mineral commodity summaries 2021*. U.S. Geological Survey, 200 p

Valko, M., Morris, H., & Cronin, M. (2005). Metals, toxicity and oxidative stress. *Current Medicinal Chemistry*, *12*(10), 1161-1208.

Vardi, A., Thamatrakoln, K., Bidle, K. D., & Falkowski, P. G. (2008). Diatom genomes come of age. *Genome Biology*, *9*(12), 245.

Vareda, J. P., Valente, A. J., & Durães, L. (2019). Assessment of heavy metal pollution from anthropogenic activities and remediation strategies: A review. *Journal of Environmental Management*, *246*, 101-118.

Vassilev, A., Lidon, F. C., Ramalho, J. C., Matos, M. d. C., & Bareiro, M. G. (2004). Shoot cadmium accumulation and photosynthetic performance of barley plants exposed to high cadmium treatments. *Journal of Plant Nutrition*, *27*(5), 775-795.

Wang, H., & Xu, X. (2017). Microhomology-mediated end joining: new players join the team. *Cell & bioscience*, *7*(1), 1-6.

Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G., & Pryor, R. J. (2003). High-resolution genotyping by amplicon melting analysis using LCGreen. *Clinical Chemistry*, *49*(6), 853-860.

Wood, A. M., Everroad, R., & Wingard, L. (2005). Measuring growth rates in microalgal cultures. *Algal Culturing Techniques*, *18*, 269-288.

WHO. (2017). *Guidelines for drinking-water quality: fourth edition incorporating first addendum* (4 ed.). ISBN: 978-92-4-154995-0.

Wu, F., Harper, B. J., & Harper, S. L. (2019). Comparative dissolution, uptake, and toxicity of zinc oxide particles in individual aquatic species and mixed populations. *Environmental Toxicology and Chemistry*, *38*(3), 591-602.

Xiao, R., & Zheng, Y. (2016). Overview of microalgal extracellular polymeric substances (EPS) and their applications. *Biotechnology Advances*, *34*(7), 1225-1244.

Yamamoto, T. (2015). *Targeted genome editing using site-specific nucleases*. Springer.

Zhang, Y., Xu, Y. H., Yi, H. Y., & Gong, J. M. (2012). Vacuolar membrane transporters OsVIT1 and OsVIT2 modulate iron translocation between flag leaves and seeds in rice. *The Plant Journal*, *72*(3), 400-410.

8 Appendix

A: Experimental workflow charts

B: Materials

C: Annotated gene sequences

D: Plasmid maps

E: Primers and adapter oligos

F: PCR reactions and thermocycler programs

G: Gel images from PCR colony screening

H: PAM protocol

I: Calculated parameters from characterization of mutant lines

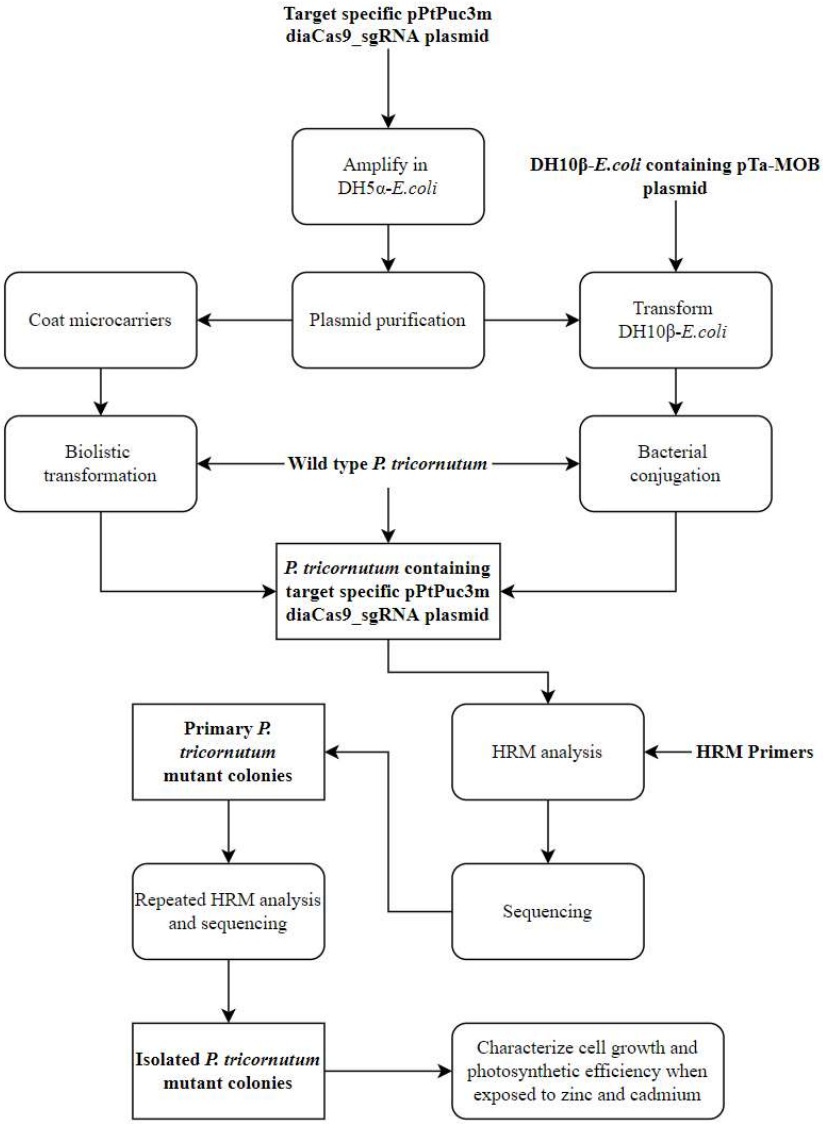
J: Statistical analysis

Appendix A: Experimental workflow charts

A.1 Creating CRISPR/Cas9 plasmid vectors with target specific gRNA

Supplementary Figure A.1: Chart showing the general workflow of the creation of target specific CRISPR/Cas9 vectors. The pPtPuc3m diaCas9_sgRNA plasmid vector was amplified in E. *coli.* Two oligonucleotides were annealed into an adapter designed to target the *VIT1* and *VIT2* or *ATPase5-1B* genes of interest and inserted into the pPtPuc3m diaCas9_sgRNA plasmid by restriction-ligation cloning. E. *coli* was then transformed with the target specific plasmid vectors. Colony PCR was used to screen for colonies potentially containing the recombinant plasmid with adapter insert, from which plasmids were purified and sequenced to confirm adapter insertion.

A.2 Creation and characterization of knockout lines of *P. tricornutum* by CRISPR/Cas9 gene editing



Supplementary Figure A.2: Chart showing the general workflow of the creation and characterization of knockout lines of *P. tricornutum* by CRISPR/Cas9 gene editing. Plasmids with confirmed target specific adapter inserts were amplified in E. *coli* and purified. The purified plasmids were then used to transform wild type *P. tricornutum* by biolistics and by bacterial conjugation. Primary colonies containing mutant lines for the *VIT1*, *VIT2* and *ATPase5-1B* genes were identified and isolated by HRM analysis and subsequent sequencing, which was then repeated in an attempt to isolate purified mutants. Lines shown to contain knockout mutations were cultivated and used in a preliminary characterization experiment where growth inhibition and variable Chl a fluorescence was measured during exposure to zinc and cadmium and compared to wild type and to lines grown in control medium.

Appendix B: Materials

This appendix contains the materials used throughout this thesis project.

B.1 Instruments

Supplementary Table B.1: Instruments (and manufacturers) used for the thesis project.

Instrument	Manufacturer
NanoDrop™ One	Thermo Scientific
Innova 44 Incubator Shaker	Eppendorf AG
T100™ PCR Thermal Cycler	Bio-Rad Laboratories
ACEA NovoCyte™ Flow Cytometer	ACEA Biosciences, Inc.
G BOX Chemi XRQ Gel Imaging System	Syngene
LightCycler 96 RT PCR System	Roche
Biolistic PDS-1000/He Particle Delivery System	Bio-Rad Laboratories
SmartSpec™ Plus Spectrophotometer	Bio-Rad Laboratories
Multi-Color-PAM Fluoresence Analyser	Walz
Heating block	Thermo Scientific
AccuSpin Micro 17R Centrifuge	Fisher Scientific
Megafuge 1.0R Centrifuge	Heraeus
Water bath	Grant
Autoclaver	
Vortex	
Refrigerator	
Freezer	
Sterile bench	

B.2 Kits

Supplementary Table B.2: Kits used for the thesis project.

Kit	Manufacturer
E.Z.N.A.® Plasmid Mini Kit I #D6942-02	Omega Bio-tek Inc.
LightCycler 480 High Resolution Melting Master kit #4909631001	Roche
ExS-Pure™ Enzymatic PCR Cleanup Kit #EXS-500	NimaGen
Wizard® SV Gel and PCR Clean-Up System #A9282	Promega

B.3 Enzymes

Supplementary Table B.3: Enzymes and associated buffers (and producers) used for the thesis project.

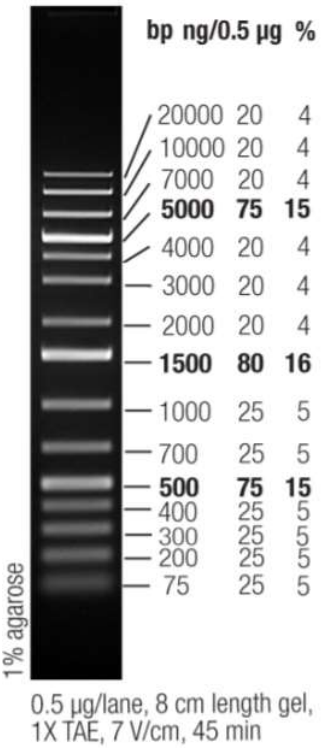
Enzyme	Producer	Buffer
DreamTaq DNA Polymerase #EP0701	Thermo Scientific	10X DreamTaq Buffer
*Bsa*I-HF #R3733	New England biolabs (NEB)	10X CutSmart® Buffer
T4 DNA Ligase #M0202S	New England biolabs (NEB)	10X T4 DNA Ligase Buffer
*Hind*III #R0104S	New England biolabs (NEB)	NEB Buffer 2.1

B.4 Bacteria

The bacteria used for this project were chemically competent DH5α-*E.coli* and DH10β-*E.coli* from laboratory stock. DH5α-*E.coli* was used for plasmid amplification and DH10β-*E.coli* (containing the pTa-MOB plasmid) was used for bacterial conjugation of *P. tricornutum.*

B.5 Gel electrophoresis DNA ladder and loading dye

GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific™; #SM1331): The DNA ladder contains 15 chromatography-purified DNA fragment of incremental length (bp): 75, 200, 300, 400, 500, 700, 1000, 1500, 2000, 3000, 4000, 5000, 7000, 10000 and 20000. From these, the 500, 1500 and 5000 bp bands are reference bands for orientation. The DNA ladder is dissolved in 1X Blue Gel Loading Dye. Band formation is depicted in Supplementary Figure B.1.



Supplementary Figure B.1: GeneRuler™ 1 kb Plus DNA Ladder (Thermo Scientific™; #SM1331) band formation. Source: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0013047_GeneRuler_1kb_Plus_DNALadder_250ug_UG.pdf

Supplementary Table B.6: 1X Blue Gel Loading Dye (NEB; #B7021s); ingredients and concentrations.

Ingredients	Concentrations
Ficoll®-400	2.5%
EDTA	11 mM
Tris-HCl (pH 8.0 @ 25°C)	3.3 mM
SDS	0.017%
Bromophenol blue.	0.015%

B.6 Guillard's (f/2) medium for diatoms

Guillard's (f/2) growth medium for diatoms was made by dissolving 20 g Sea Salt (S9883, Sigma-Aldrich) in 1 L of ddH_2O. The salt solution was autoclaved for 20 minutes at 120˚C, and the stock ingredients listed in Supplementary Table B.7 were added to the cold solution under sterile conditions through a 0.2 µm sterile filter. 50% (v/v) SW, f/2-Si, 1% (w/v) agar plates were made using 50% (v/v) filtered seawater from the Trondheim fjord, ddH_2O and 1% (w/v) agar, which was similarly autoclaved for 20 minutes at 120˚C. The remaining ingredients were added through a 0.2 µm sterile filter once the solution had cooled to < 55 ˚C under constant stirring, except sodium metasilicate which was not added (Supplementary Table B.8). The agar solution was plated on a sterile bench when sufficiently cool, left to solidify and stored in the dark at 4 ˚C. Zeocin was added to the medium when selecting for transformed *P. tricornutum*, at a concentration of 50 µg/mL in the liquid medium and 100 µg/mL for the growth plates.

Supplementary Table B.7: Ingredients and quantities of stock solutions (Supplementary Table B.9) used to make Guillard's (f/2) medium.

Ingredients	Volume (mL/L seawater)
$NaNO_3$	1.0
$NaH_2PO_4.2H_2O$	1.0
Trace elements stock solution (1)	1.0
Vitamin mix stock solution (2)	0.5
Sodium metasilicate stock solution (3)	1.0

Supplementary Table B.8: Ingredients and quantities used to make 50% (v/v) SW f/2-Si, 1% (w/v) agar plates.

Ingredients	Quantity per L
$NaNO_3$	1.0 mL
$NaH_2PO_4.2H_2O$	1.0 mL
Trace elements stock solution (1)	1.0 mL
Vitamin mix stock solution (2)	0.5 mL
BactoAgar	10 g
Filtered seawater (0.2µL sterile filter)	500 mL
ddH_2O	490 mL

Supplementary Table B.9: Trace elements, vitamins and inorganic nutrients in stock solutions used to make Guillard's (f/2) medium.

Ingredients	Concentration (g/L)
(1) Trace elements (chelated)	
NA_2EDTA	4.36
FeCl3.6H2O	3.15
CuSO4.5H2O	0.01
ZnSO4.7H2O	0.022
CoCl2.6H2O	0.01
MnCl2.4H2O	0.18
Na2MoO4.2H2O	0.006
(2) Vitamin mix	
Cyanocobalamin (Vitamin B_{12})	0.0005
Thiamine HCL (Vitamin B_1)	0.1
Biotin	0.0005
(3) Sodium metasilicate	
$Na_2SiO_3.9H_2O$	30.0
$NaNO_3$	75
$NaH_2PO_4.2H_2O$	5.65

B.7 LB medium and agar plates

Luria-Bertani (LB) medium was made by adding the ingredients listed in Supplementary Table B.10 to ddH_2O. The solution was autoclaved for 20 minutes at 120˚C, and antibiotics (Gentamycin final concentration: 20 µg/mL and/or Kanamycin final concentration: 50 µg/mL) were added when the solution had cooled to < 55 ˚C. LB-agar plates were made and stored in the same way as f/2 plates (Appendix B.6).

Supplementary Table B.10: Ingredients and quantities used to make LB medium

Reagent	Concentration (g/L)
NaCl	5
Tryptone	10
Yeast Extract	5
Agar (only for agar plates)	15

B.8 SOC medium

Supplementary Table B.11: Ingredients and concentrations used to make SOC medium.

Reagent	Concentration
Tryptone	2%
Yeast extract	0.5%
NaCl	10 mM
KCl	2.5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	20 mM

B.9 Computer software

Supplementary Table B.12: Computer software (and manufacturers) used for the thesis project.

Software	Manufacturer
SnapGene (V. 6.0)	GSL Biotech LLC
ACEA NovoExpressTM software	ACEA Biosciences, Inc.
Ligh cycler 96 software	Roche
PamWin-3 Software	Walz
SPSS Statistics (V. 28.0.1)	IBM
ICE CRISPR Analysis Tool	Synthego
Excel	Microsoft

Appendix C: Annotated gene sequences

Full length coding sequences of the *P. tricornutum* target genes selected for this project. All sequences are shown in 5' → 3' direction. Primer sequences are highlighted in colours. The two 20bp Cas9-target sites (followed by NGG PAM sites) are indicated in red and green, respectively, and underlined.

C.1 The *VIT1* (NCBI Gene ID: 7197377) sequence

ATGAACGCGCCAGTCGACCCAGCACAGAGCCCGGCGACGGACCAAACTGAGAAAGATATGGAATCGGGACC
AGTCTTAAAGGAATCCATGGGTGATGAACCACCACTGGCCTACGTCGACACGAGTATCGAAAAGGGAAAATT
GAAGATTCGGAACGATAGCGATGCTACAACAGCATCCGATCCTACGAAAACGACTGTTGTTGCGACTCCGCTA
ACCTCTTCCGACGAACTCGCGGACCACTTGGGAGGATCACGTCAATACTGGCGGGATATTATCCTCGGAGTGA
ACGACGGTCTCGTCTCCACATTTCTGCTCGTAGCCGGCGTCGCCGGTGGTGGACTTGCTTCGACCGACATTCTG
CTGACGGCTATTGCCGGTGCACTAGCCGGAGCCATCAGCATGTGCGCTGGTGAATACGTCGCGACGAAATCGC
AGAATCAAGTCCTGCAGGGAGAAATCAATCTCGAAAAGCAACACGTGCAGTTTCGCTTTGAAGAGGAGATCG
CTGAAATGGAAGCGGAACTACTACCAATCATTGGCATTGGTGAAGATCAGCCGGAACTGCGCGAGAAGCTGT
TGGAGTACTATACGGCCCATCCCGATGCCTTGCTCAAGATTATGATTGCACTCGAGTTTGGTGTTGTGGATCAC
GAGCAACGATCAGCCGTATGTGCAGGTCTGACTTCGTGTGGCACTTTTCTTCTGGGAACTCTTCCGTCCGTGTT
GCCCTTTGTGTTCTTGGACGATCCAACAGTGGGTCTGATTGTGGCGGCGATCGTTGCGACCCTGACGTTGTTGT
TGGTGGGGGCGATCAAAACGTGGGCCACGCGCGGCAACGCTTGGACGGCTGCCTTGGAGAATCTGGTCATTG
CCGGTGTCGGTGGTGGATTTGCCTACGGTATTGGTCTTTTGTTCGATTCAACCTTGCATGGATAA

VIT1_F

VIT1_R

VIT1_HRM_F

VIT1_HRM_R

VIT_PAM1

VIT_PAM2

C.2 The *VIT2* (NCBI Gene ID: 7197079) sequence

ATGAGCTCAGAAGTCAATTCAGACCGGGCTTTGGTGGCATATCGTGCCGTGGAAGCCATGGAATTTGAGGAAA
ATTCGCTTACAACGGAACCGGTCCTTGCGAAATCCCATAAGTCAGCTACATTGAGTCGGTGTGATCCAGATAC
CACACCGTCTTTCAGTCTGAACTCTGACACAAGTTGCTTGGCCCCTGCGTACTCACCCAAACTGGCGGATCATT
TGGGAGGATCACGTCAATACTGGCGCGACATGATCCTCGGAGTGAATGACGGTCTCGTCTCCACATTTCTGCT
CGTGGCCGGCGTCGCCGGTGGTGGACTTGCTTCGACCGACATTCTGCTGACGGCTATTGCCGGTGCACTAGCC
GGAGCCATCAGCATGTGCGCTGGTGAATACGTCGCGACGAAATCGCAGAATCAAGTCCTGCAGGGAGAAATC
GACCTCGAAAAGCTGCACGTTCGGTTTCGACCCAAGGAGGAAATCGCCGAGGTTGAAGCGCAACTGCTGACG
GCAATTGGCATTGGCGAAGATCAGCCGGAATTGCGGGAAACTTTGTTACAATATTACACTGCCCATCCAGAAT
CTTTACTCAAGATTATGATTGCACTCGAGTTTGGTGTTGTGGACCACGAGCAACGATCTCCTCTGTGTGCAGGT
CTGACTTCGTGTGGCACTTTTCTCCTGGGAACTCTTCCGTCCGTGTTGCCTTTTGTGTTCTTGGACGATCCGACA
GTGGGTCTGATTGTGGCGGCGGTCATTGCGACCCTGACGTTGTTGTTAGTGGGGGCAATAAAAACATGGGCCA
CGCGTGGCAGCGCTTGGATGGCTGCATTGGATAATCTGGTCATTACCGGTTGTGGGGGTACCTTTGCCTACGGT
ATTGGTCTATTCTTCGATGCGGTTGTGCAATGA

VIT2_F

VIT2_R

VIT2_HRM_F

VIT2_HRM_R

VIT_PAM1

VIT_PAM2

C.3 The ATPase5-1B (NCBI Gene ID: 7203325) sequence

TTGTGACAACCTTACCCCAGGCAACTTCTTTGGCGATGGCACGGAGTGTGATAGCGTGTAGTCAAATTTGCTGT
GCTTCCGAAATTCCCATCATCAATAACATTTTGGAACCAATGGCGGGTGTGACCCTGGTCCGCGTTAATGTTCC
ACTCAAGCAAGTATTGATCGATCACGATGCAAAGATTGTTCCTGCCGCTACAATGGAGCGTGCCTTGAACAAG
CAAGGCCTTGGTGCAAGCATACTTCGGGATGGGACTTTCACTACCGCTGCAACGTCGAAACCAAAAACCACTG
GTGCAAGCGTCAAATCCACCTTTTATGTTGCCCACATTTGTTGTGCCAGCGAAATCCCCGCCATCAATCAAGTA
GTGGAGCCGCTCTCGGGAGTGTTGGCCGTCTCGATCAACGTCACGAGTAAAATGGTGTACGTCACGCACGATA
CAACTAGCATTGCTGCGTCCCAAATTTGTGCTGTTTTGAACGAACAAGGGTTTGGTGCCGAAGTCAAGTCAGA
CGGGGGAGAAAATTTCATTGGTTCGCTTTTTTGCGAGTCCACGCTCGCTTTCCAGATGGACGACAGCGGCCCG
AAAGATTACACGACCATCCTGACTTCCTTCTTAAGGGAATACTTCGATCCATCCCAGATTCAAATGGAAAGTTT
CGTTTTGAACGTTCCGCAACAGCAAATCAAGGTCGTTCACAATCCATTTTCTGTTTCGGCGTCTTCGATTGCCG
AGACCTTGCAAGCAAAAACTGGTATTCGAGCTACCGTCCTGCAGGATGGAGCCGCACCATCCTCATTCGACTG
GTCGACTTTAGACTTTTTGGACAACGTCAGCGTCGACAATGCCGAAGAAGAATCAGTCAACTATCCCAAACCG
GCAATTACTCTTTCAGGTATTTGCTGGTTTTTATCCATGTTGAGTTTACTAGGTGGTCAGCTAGATATCCTCAAA
TATGTCGGTCTTGCCGCCGTTGCCTTTGGTTTGCCCACAATTTCGGGCAAAGCATTTCGTACGCTTTCACGCTG
GCAATTTGACACAAACTGTCTAATGTTCTTAGCCGCCATGGGAGCTGTGGTGTTACAAGAGTATACCGAGGCG
GCAGCGGTCGTCTTTTTGTTCGCCATTTCAGAGTGGCTGGAAGTACGCGCAACGGCTCGTGCCCGACATGCCCT
ATCGGCAATTGTACACTTACGACCCGAAAAGGCCAATCTGGTGCATCCACAGACACGACAATTGGTGGTGATT
CCGGCGTCGGCGGTTCCTGTCGGGGCGCTAGTGTCGGTGAAGACCGGCGACAAGATCCCGTGCGATGGTATTG
TCGTGGAAGGCACAACGACTGTAGACGAATCGTCCTTGACAGGAGAGGCCCGCCCCATTCGAAAAGGTTTAC
ATGACGTTGTTTCGGGTGGCACGGTGAACTCTGGAATGGCCCAGATCATGGTGCGAACGACATCGACCTCGGA
AAACTCGGCCGTGTCTCGTTTGATTCGGCTGGTCGAGGAAGCTCAAGCCAACCGATCGGACACGGAAAAGCTT
GTGGACGAGTTTGCCAAGATATACACACCGTTCGTCGTGCTCGCGGCCGTTCTCATGTGCAGTGTACCGTGGG
CTTTTGGACAGGAAACCGGTAGAACGTGGACTGAGAACGGGCTGATTCTGATCGTTGTGGCCTGTCCGTGTGC
CATGATCATTTCGACGCCCGTGTCTTACGTTGCTGGTCTTGCCGCGACGGCTCAAAATGGAATTCTTATCAAAG
GTGGTGCACACTTGGAGGCGCTCAGCCTCGTGAAGCATATTTGCTTTGACAAGACGGGTACGCTGACGAACGG
GGAGTTTGCGTTGCTCAGCCTGGAAGACTTTGCGAAGAATATGTCGCGTAAGGAGGTATTTGAACACCTGGCC
TTGATGGAAGAGCGCGCGTCGCATCCTGTGGCGCAAGCCATCTTGACCGGTGCCCGCAACGAGAAAGTTAGTA
TACCGAAAGATACGGAACTCGAGAGGCACACCATCATTGCTGGCGAGGGCGTTTTGGGCATTATCAACGGGC
GCGAGGTGCACGTTGGTAACGAGAGGATGTTTGGTCGTATCGGTTTATTGAAGGACGTACCCGAGACGGTACG
GGCCAAGGTAGAATCGTGGAAGGGTCTTGGAGGGACGATTGGGTACATGAGTATAGAAGGGGAGGGAATTGT
GTGTGCGTACTGCGCGGCCGATGGTGTCCGAGCCGAGTCTGCTAGTGTTGTGAGTCGTTTGAGGAAGTGCGGT
ATTGAAGTGACCATGTTGACCGGGGACAATCGCGACGCAGCGTTGGCAGTTGGAGCGCAAGTGGGTTTGTCTG
AGCAGGAGATCCAGTCCAAGCTACTGCCCGAAGAGAAGCTGGCGTTTGTTGAATCTTTGAGCTCTGGCCGTAC
TGGAGGTTCAATTTTGTCGAATCCATGTGGTCAGCGGCGCTTGGTGATGATGTGTGGGGACGGGGTGAACGAC
GCGCCAGCTTTGGCAGCTGCGGACATTGGTGTAGCAATGGGTGCTGGGGCGGCGCTTGCAATGGAAACTGCG
GACATCGCACTACTGGATTCAAATCTAGAGAAGCTCGAATATAGCATCAGGATGGGCCAAAGGGTTACGCGT
AAAATCAAAGAAAACGTGGCTTTTAGTTTAACCGTGAAATTTGTGGTTCTTGGCTTCGCACTTGCCGGGCTGAC
TCATCTTTGGGCAGCTATAGCAAGCGACGTCGGAGCCATGATTTTAGTTACCTTAAATGCAATGTTGCTCTTGC
CGAAACGCCAGAG

Non-coding

ATPase5-1B_F

ATPase5-1B_R

ATPase5-1B_P1_HRM_F

ATPase5-1B_P1_HRM_R

ATPase5-1B_P2_HRM_F

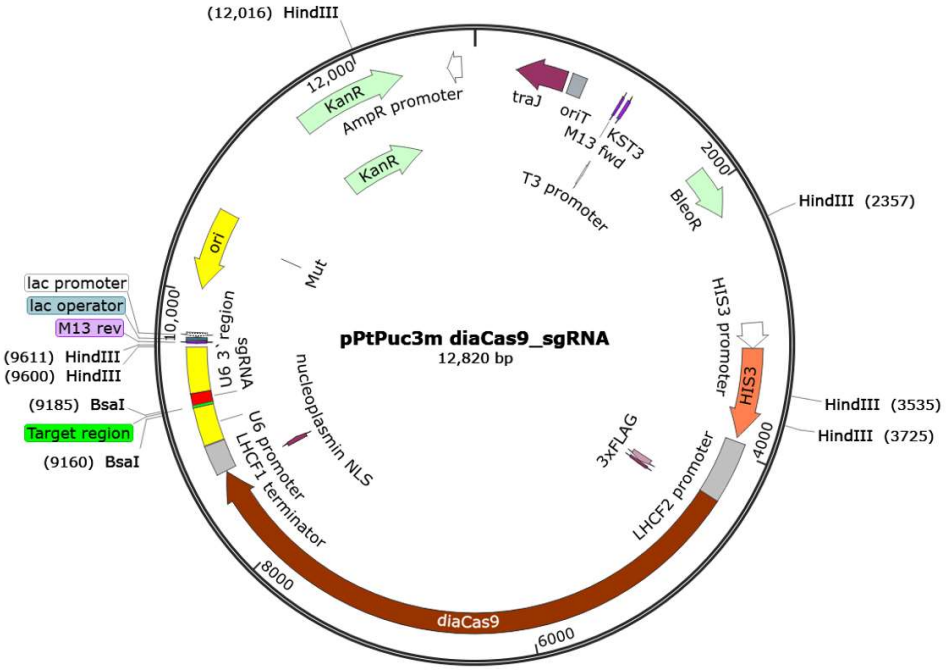
ATPase5-1B_P2_HRM_R

ATPase5-1B_PAM1

ATPase5-1B_PAM2

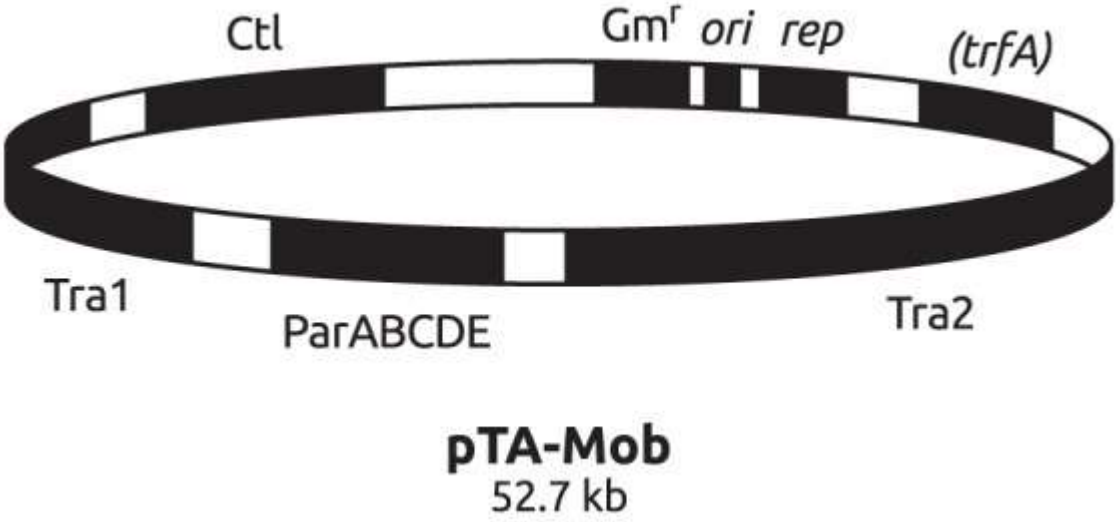
Appendix D: Plasmid maps

D.1 pPtPuc3m diaCas9_sgRNA plasmid map



Supplementary Figure D.1: Map of the pPtPuc3m diaCas9_sgRNA plasmid with important features being: diaCas9: Cas9 codon optimized for *P. tricornutum*; sgRNA: single guide RNA with *P. tricornutum* U6 promoter and 3' region; NLS: bipartite nuclear localization signal from nucleoplasmin; KanR: confers resistance to kanamycin in bacteria; oriT: origin of transfer; BleoR: confers resistance to Zeocin in *P. tricornutum*; M13 fwd and M13 rev: binding sequences for standard M13 fwd and rev primers; BsaI: restriction site for *Bsa*I restriction enzyme; HindIII: restriction site for *Hind*III restriction enzyme. Figure created in SnapGene® (version 6.0) from GSL Biotech LLC

D.2 pTa-MOB plasmid map



Supplementary Figure D.2: Map of mobilization helper plasmid pTA-Mob with relevant regions shown. Gm^r: confers resistance to Gentamycin; *rep*: pBBR1 replication protein; *ori*: pBBR1 replication origin; *trfA*: replication initiation protein gene from the RK2 replicon (not active due to lack of RK2 replication origin); Tra1 and Tra2: contains *tra* genes necessary for conjugative transfer of oriT containing plasmids; ParABCDE: stabilization region encoding the gene products ParA, B, C, D and E; Ctl: central control operon of RK2. Image is reproduced from Strand et al. (2014).

Appendix E: Primers and adapter oligos

All sequences are shown in 5' → 3' direction.

All nucleotides were ordered from Sigma-Aldrich.

E.1 Adapters

Table E.1: Sequence of the forward and reverse complementary oligos annealed to make target specific adapters for CRISPR/Cas9 editing of *P. tricornutum* genes *VIT1, VIT2* and *ATPase5-1B.*

Adapter	PAM F oligo sequence	PAM R oligo sequence
VIT_PAM1	TCGATTCGACCGACATTCTGCTGA	AAACTCAGCAGAATGTCGGTCGAA
VIT_PAM2	TCGAATCGCAGAATCAAGTCCTGC	AAACGCAGGACTTGATTCTGCGAT
ATPase5-1B_PAM1	TCGAGGATTTGACGCTTGCACCAG	AAACCTGGTGCAAGCGTCAAATCC
ATPase5-1B_PAM2	TCGATCAAGTAGTGGAGCCGCTCT	AAACAGAGCGGCTCCACTACTTGA

E.2 PCR primers used in colony screening

Table E.2: Sequence of primers used in colony screening to identify colonies containing plasmids with target specific adapter inserts. Forward primers are marked FP and reverse primers are marked RP.

Primer	Orientation	Sequence
VIT_PAM1 F	FP	TCGATTCGACCGACATTCTGCTGA
VIT_PAM2 F	FP	TCGAATCGCAGAATCAAGTCCTGC
ATPase5-1B_PAM1 F	FP	TCGAGGATTTGACGCTTGCACCAG
ATPase5-1B_PAM2 F	FP	TCGATCAAGTAGTGGAGCCGCTCT
M13 Rev	RP	CAGGAAACAGCTATGAC

E.3 PCR primers used in screening for targeted gene mutations

Table E.3: Sequence of primers used to screen for gene targeted mutations and expected amplicon size. Includes primers used for PCR amplification of the larger target area and those used to amplify smaller fragments for HRM. Forward primers are marked FP and reverse primers are marked RP.

Primer	Orientation	Sequence	Amplicon size (bp)
VIT1_F	FP	CTACGTCGACACGAGTATCGAA	800
VIT1_R	RP	GACCAATACCGTAGGCAAATCC	
VIT1_HRM_F	FP	CGTCTCCACATTTCTGCTCGTA	197
VIT1_HRM_R	RP	CAAAGCGAAACTGCACGTGTT	
VIT2_F	FP	AGTCAGCTACATTGAGTCGGTG	774
VIT2_R	RP	GACCAATACCGTAGGCAAAGGT	
VIT2_HRM_F	FP	CGTCTCCACATTTCTGCTCGTG	195
VIT2_HRM_R	RP	GGTCGAAACCGAACGTGCA	
ATPase5-1B_F	FP	GATTGTTCCTGCCGCTACAATG	523
ATPase5-1B_R	RP	ATTGTGAACGACCTTGATTTGC	
ATPase5-1B_P1_HRM_F	FP	AGCATACTTCGGGATGGGACTT	139
ATPase5-1B_P1_HRM_R	RP	GCTCCACTACTTGATTGATGGC	
ATPase5-1B_P2_HRM_F	FP	ATTTGTTGTGCCAGCGAAATCC	122
ATPase5-1B_P2_HRM_R	RP	ATGCTAGTTGTATCGTGCGTGA	

E.4 Primers used for sequencing putative gene edited DNA products

Table E.4: The sequences of primers used for Sanger sequencing.

Sequencing described in section	Primer	Sequence
3.1.6	M13 Rev	CAGGAAACAGCTATGAC
3.2.5	VIT1_F	CTACGTCGACACGAGTATCGAA
	VIT1_R	GACCAATACCGTAGGCAAATCC
	VIT2_F	AGTCAGCTACATTGAGTCGGTG
	VIT2_R	GACCAATACCGTAGGCAAAGGT
	ATPase5-1B_F	GATTGTTCCTGCCGCTACAATG
	ATPase5-1B_R	ATTGTGAACGACCTTGATTTGC

Appendix F: PCR reactions and thermocycler programs

F.1 PCR setup for colony screening

Supplementary Table F.1: PCR reaction mixture used for colony PCR.

Reagent	Volume (µL)
Template	5
DreamTaq Buffer 10X	2.5
dNTP mix (2 mM)	2.5
Forward primer (10 µM)	1.25
M13 reverse primer (10 µM)	1.25
DreamTaq DNA Polymerase (5 U/µL)	0.125
H_2O	12.375

Supplementary Table F.2: PCR thermocycler program used for colony PCR.

Cycles	Process	Temp. (°C)	Duration
1	Initial denaturing	94	5 min
35	Denaturing	94	30 sec
	Annealing	55	30 sec
	Extension	72	1 min
1	Final extension	72	10 min

F.2 PCR setup for amplification of target area

Supplementary Table F.3: PCR reaction mixture used for amplification of target area.

Reagent	Volume (µL)
Template (cell lysate)	2.5
Buffer 10X	2.5
dNTP (2 mM)	2.5
Forward primer (10 µM)	1.25
Reverse primer (10 µM)	1.25
DreamTaq DNA Polymerase (5 U/µL)	0.125
H_2O	14.875

Supplementary Table F.4: PCR thermocycler program used for amplification of target area.

Cycles	Process	Temp. (°C)	Duration
1	Initial denaturing	94	5 min
34	Denaturing	94	30 sec
	Annealing	59	30 sec
	Extension	72	30 sec
1	Final extension	72	5 min

F.3 PCR setup for HRM analysis

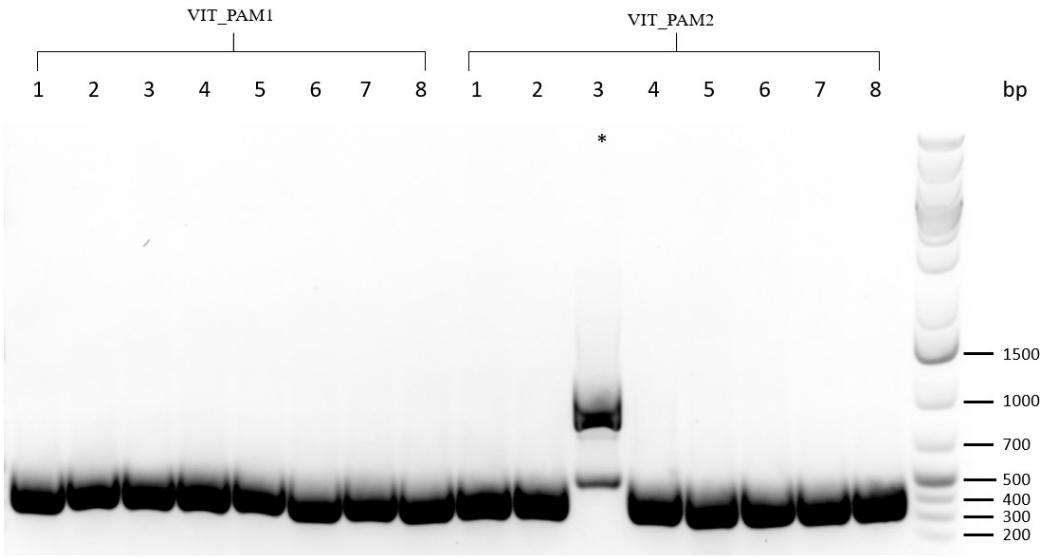
Supplementary Table F.5: PCR reaction mixture used for HRM screening for edited lines of *P. tricornutum*. A combined F + R primer solution was prepared with 2 µM forward and 2 µM reverse primer.

Reagent	Volume (µL)
HRM master mix	10
MgCl (25 mM)	2.4
F + R Primer	2
PCR-grade water	0.6

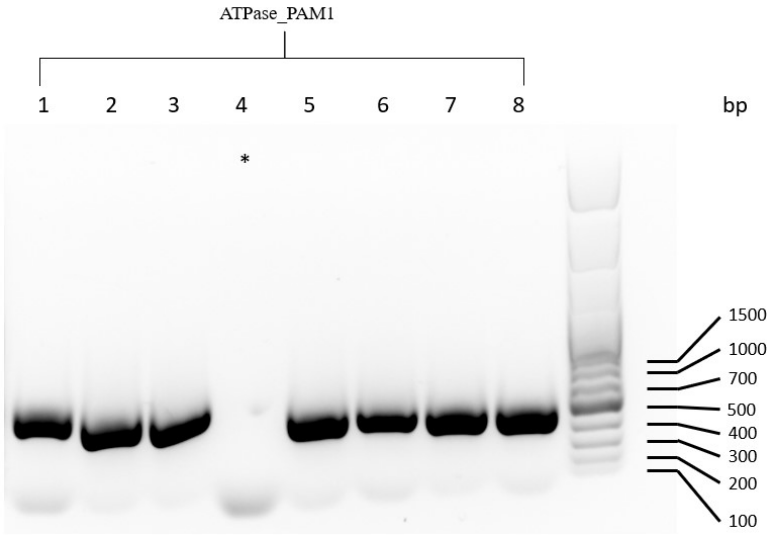
Supplementary Table F.6: HRM program used to screen for edited lines of *P. tricornutum.*

Cycles	Process	Temp. (°C)	Duration (sec)
1	Preincubation	95	600
40	Amplification	95	10
		TD 64 → 55 (-1)	
		72	15
1	High resolution melting	95	60
		40	60
		65 → 95 (2.2 per sec.)	
1	Cooling	37	30

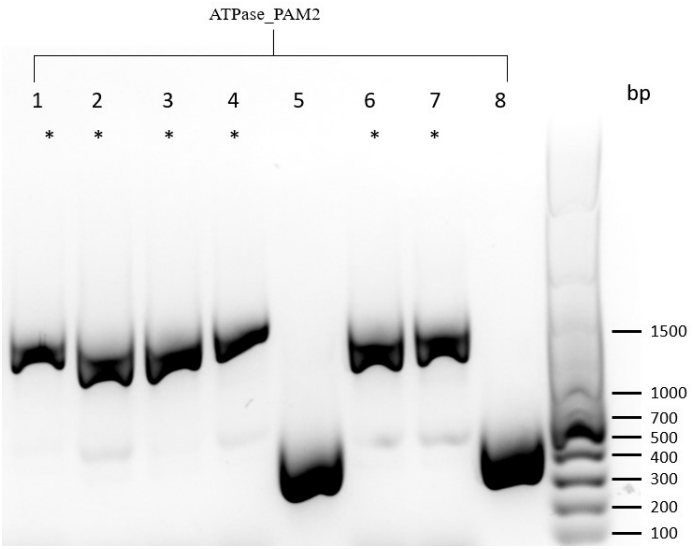
Appendix G: Gel images from PCR colony screening



Supplementary Figure G.1: Gel pictures from the colony PCR screening of the VIT_PAM1 and VIT_PAM2 plasmids. Agarose gel electrophoresis was carried out with the PCR products (5 µL; predicted amplicon size 484 bp), using GeneRuler 1 kb plus DNA ladder (5 µL; Appendix B.5) in the last well as reference. PCR products indicating unsuccessful adapter insertion are marked with an asterisk (*).



Supplementary Figure G.2: Gel pictures from the colony PCR screening of the ATPase_PAM1 plasmid. Agarose gel electrophoresis was carried out with the PCR products (5 µL; predicted amplicon size 484 bp), using GeneRuler 1 kb plus DNA ladder (5 µL; Appendix B.5) in the last well as reference. PCR products indicating unsuccessful adapter insertion are marked with an asterisk (*).



Supplementary Figure G.3: Gel pictures from the colony PCR screening of the ATPase_PAM2 plasmid. Agarose gel electrophoresis was carried out with the PCR products (5 µL; predicted amplicon size 484 bp), using GeneRuler 1 kb plus DNA ladder (5 µL; Appendix B.5) in the last well as reference. PCR products indicating unsuccessful adapter insertion are marked with an asterisk (*).

Appendix H: PAM protocol

Supplementary Table H.1: The protocol for variable chlorophyll a fluorescence measurement carried out with Multi-Colour PAM, including the measuring steps, their incremental irradiance levels (Photosynthetically active radiation; PAR) and duration of each exposure interval.

Step	PAR (µmol photons m^{-2} s^{-1})	Time (s)
Incubation	0 (Darkness)	180
0	0	30
1	2	30
2	4	30
3	7	30
4	31	30
5	57	30
6	123	30
7	177	30
8	244	30
9	350	30
10	469	30
11	599	30
12	753	30
13	905	30
14	1313	30

Appendix I: Calculated parameters from characterization of mutant lines

Supplementary Table I.1: The calculated means and standard deviations (n = 3) for the experimental parameters maximum specific growth rate (K'_{max}), maximum quantum yield of PSII ($\phi_{PSIImax}$), minimum saturating irradiance (E_k) and maximum relative electron transport rate ($rETR_{max}$) for the V1, V2, A1 and A2 mutant lines and wild type (WT) of *P. tricornutum* grown in f/2 medium (control; CT), f/2 medium containing zinc (Zn^{2+}) and f/2 medium containing cadmium (Cd^{2+}).

Line (Treatment)	Mean K'_{max}	SD K'_{max}	Mean $\phi_{PSIImax}$	SD $\phi_{PSIImax}$	Mean E_k	SD E_k	Mean $rETR_{max}$	SD $rETR_{max}$
WT (CT)	1.594	0.036	0.640	0.009	121.667	5.400	37.767	1.528
WT (Zn^{2+})	1.477	0.042	0.641	0.006	121.667	4.477	37.133	0.987
WT (Cd^{2+})	1.524	0.023	0.616	0.047	119.967	2.548	37.800	0.624
V1 (CT)	1.223	0.059	0.628	0.006	110.667	4.302	33.200	1.308
V1 (Zn^{2+})	1.116	0.034	0.614	0.012	112.833	4.388	31.933	0.404
V1 (Cd^{2+})	1.169	0.030	0.622	0.008	117.267	5.431	34.433	0.814
V2 (CT)	1.130	0.085	0.626	0.007	110.733	6.331	33.633	1.457
V2 (Zn^{2+})	1.024	0.103	0.628	0.005	116.967	2.658	34.767	0.551
V2 (Cd^{2+})	1.053	0.069	0.626	0.005	112.200	4.782	34.733	1.258
A1 (CT)	1.277	0.048	0.568	0.003	119.600	1.947	31.567	0.058
A1 (Zn^{2+})	1.121	0.059	0.586	0.005	110.533	4.822	29.900	1.114
A1 (Cd^{2+})	1.295	0.044	0.574	0.004	111.167	2.857	30.500	0.265
A2 (CT)	1.145	0.021	0.546	0.002	105.567	3.301	26.367	0.252
A2 (Zn^{2+})	0.931	0.035	0.557	0.007	102.867	3.482	26.667	0.513
A2 (Cd^{2+})	1.077	0.026	0.556	0.006	99.933	7.881	25.567	1.270

Appendix J: Statistical analysis

J.1 Analysis of parameters from *P. tricornutum* lines grown in f/2 control medium

Supplementary Table J.1: Statistical analysis of the calculated experimental parameters (Appendix I.) for the V1, V2, A1 and A2 mutant lines and wild type (WT) of *P. tricornutum* grown in f/2 medium without heavy metals. One-way ANOVA with Tukey's post hoc test was conducted to identify significant differences in parameter means between cell lines. Statistically significant differences are marked green, and statistically insignificant differences are marked red. Statistical significance level $P < 0.05$.

Treatments compared	Parameter (P-value)			
	K'_{max}	$\phi_{PSIImax}$	$rETR_{max}$	E_k
WT – V1	<.001	.998	<.001	.231
WT – V2	<.001	.992	<.001	.239
WT – A1	<.001	<.001	<.001	1.000
WT – A2	<.001	<.001	<.001	.011
V1 – V2	.676	1.000	1.000	1.000
V1 – A1	.994	<.001	.720	.537
V1 – A2	.868	<.001	<.001	.984
V2 – A1	.091	.001	.369	.548
V2 – A2	1.000	<.001	<.001	.983
A1 – A2	.185	.809	<.001	.042

J.2 Analysis of parameters from *P. tricornutum* lines grown in f/2 medium supplemented with zinc

Supplementary Table J.2: Statistical analysis of the calculated experimental parameters (Appendix I.) for the V1, V2, A1 and A2 mutant lines and wild type (WT) of *P. tricornutum* grown in f/2 medium containing zinc. One-way ANOVA with Tukey's post hoc test was conducted to identify significant differences in parameter means between cell lines and between individual cell lines growth in the presence and absence of zinc. Statistically significant differences are marked green, and statistically insignificant differences are marked red. Statistical significance level $P < 0.05$.

Treatments compared	Parameter (P-value)			
	K'_{max}	$\phi_{PSIImax}$	$rETR_{max}$	E_k
WT – V1	<.001	.553	<.001	.554
WT – V2	<.001	.997	.190	.993
WT – A1	<.001	.002	<.001	.217
WT – A2	<.001	<.001	<.001	.002
WT – WT (CT)	.347	1.000	1.000	1.000
V1 – V2	.703	.994	.053	.998
V1 – A1	1.000	.459	.393	1.000
V1 – A2	.012	.001	<.001	.367
V1 – V1 (CT)	.471	.995	.936	1.000
V2 – A1	.631	.044	<.001	.909
V2 – A2	.694	<.001	<.001	.040
V2 – V2 (CT)	.499	1.000	.973	.926
A1 – A2	.009	.400	.015	.752
A1 – A1 (CT)	1.000	.942	.693	.514
A2 – A2 (CT)	.002	1.000	1.000	1.000

J.3 Analysis of parameters from *P. tricornutum* lines grown in f/2 medium supplemented with cadmium

Supplementary Table J.3: Statistical analysis of the mean calculated experimental parameters (Appendix I.) for the V1, V2, A1 and A2 mutant lines and wild type (WT) of *P. tricornutum* grown in f/2 medium containing cadmium. One-way ANOVA with Tukey's post hoc test was conducted to identify significant differences in parameter means between cell lines and between individual cell lines growth in the presence and absence of cadmium. Statistically significant differences are marked green, and statistically insignificant differences are marked red. Statistical significance level $P < 0.05$.

Treatments compared	Parameter (P-value)			
	K'_{max}	$\phi_{PSIImax}$	$rETR_{max}$	E_k
WT – V1	<.001	1.000	.010	1.000
WT – V2	<.001	1.000	.026	.736
WT – A1	<.001	.049	<.001	.560
WT – A2	<.001	<.001	<.001	<.001
WT – WT (CT)	.939	.686	1.000	1.000
V1 – V2	.356	1.000	1.000	.985
V1 – A1	.241	.012	.001	.937
V1 – A2	.705	<.001	<.001	.005
V1 – V1 (CT)	.992	1.000	.947	.893
V2 – A1	<.001	.006	<.001	1.000
V2 – A2	1.000	<.001	<.001	.120
V2 – V2 (CT)	.890	1.000	.979	1.000
A1 – A2	.002	.945	<.001	.206
A1 – A1 (CT)	1.000	1.000	.983	.624
A2 – A2 (CT)	.953	1.000	.999	.965

J.4 Analysis of cell counts from *P. tricornutum* lines grown with and without heavy metals at day 17 of growth experiment

Supplementary Table J.4: Statistical analysis of the mean cell counts at day 17 of the growth experiment (Appendix I.) for the V1, V2, A1 and A2 mutant lines and wild type (WT) of *P. tricornutum* grown in f/2 medium (Control; CT), f/2 medium containing zinc (Zn^{2+}) and f/2 medium containing cadmium (Cd^{2+}). One-way ANOVA with Tukey's post hoc test was conducted to identify significant differences in relevant cell count means. Statistically significant differences are marked green, and statistically insignificant differences are marked red. Statistical significance level $P < 0.05$.

Treatments compared		P-value
WT (CT)	WT (Zn^{2+})	.015
	WT (Cd^{2+})	<.001
	V1 (CT)	<.001
	V2 (CT)	<.001
	A1 (CT)	<.001
	A2 (CT)	<.001
WT (Zn^{2+})	V1 (Zn^{2+})	<.001
	V2 (Zn^{2+})	<.001
	A1 (Zn^{2+})	<.001
	A2 (Zn^{2+})	<.001
WT (Cd^{2+})	V1 (Cd^{2+})	<.001
	V2 (Cd^{2+})	.001
	A1 (Cd^{2+})	<.001
	A2 (Cd^{2+})	<.001
V1 (CT)	V1 (Zn^{2+})	1.000
	V1 (Cd^{2+})	<.001
V1 (Zn^{2+})	V1 (Cd^{2+})	<.001
V2 (CT)	V2 (Zn^{2+})	.975
	V2 (Cd^{2+})	.001
V2 (Zn^{2+})	V2 (Cd^{2+})	<.001
A1 (CT)	A1 (Zn^{2+})	.992
	A1 (Cd^{2+})	<.001
A1 (Zn^{2+})	A1 (Cd^{2+})	<.001
A2 (CT)	A2 (Zn^{2+})	1.000
	A2 (Cd^{2+})	<.001
A2 (Zn^{2+})	A2 (Cd^{2+})	<.001

Andreas Holm Loriente

Investigating heavy metal tolerance in the diatom *Phaeodactylum tricornutum*

NTNU
Kunnskap for en bedre verden