Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

Causia Mathivannan

Functional studies of Group II silicanins in the diatom *Thalassiosira pseudonana* cell wall biomineralization

Master's thesis in Chemical Engineering and Biotechnology
Supervisor: Olav Vadstein & Tore Brembu
June 2022

NTNU
Norwegian University of
Science and Technology

Causia Mathivannan

Functional studies of Group II silicanins in the diatom *Thalassiosira pseudonana* cell wall biomineralization

Master's thesis in Chemical Engineering and Biotechnology
Supervisor: Olav Vadstein & Tore Brembu
June 2022

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

Preface

This master's thesis is a continuation of the course TBT4500 - Biotechnology, Specialization Project, conducted from January 2022 to June 2022 at the Department of Biotechnology and Food Science (IBT) at the Norwegian University of Science and Technology (NTNU). The thesis was accomplished as a part of the ongoing research project, DIASIL (2018-2022): *How to build a glass house: Revealing fundamental components of diatom cell wall biomineralization*, at the ACMS (Analysis and Control of Microbial Systems) research group. This master's thesis is part of completing the Master of Science (M.Sc) degree in Chemical Engineering and Biotechnology.

I want to thank my supervisors, Professor Olav Vadstein and research scientist Tore Brembu, for the insightful and informative guidance during the thesis. Especially, I want to thank Tore Brembu and Ph.D. student Marthe Caroline Grønbech Hafskjold for their help in the laboratory, both in answering questions and being present in the lab for guidance on experimental procedures. In addition, I would like to thank Ph.D. student Annika Messemer for the introduction and follow-up of the thesis during the semesters. I also want to thank the other members of the ACMS research group for their help and advice during the laboratory work. With this, I also want to thank Thea Sværen, my laboratory partner, for the scientific discussions and collaboration in the laboratory work.

Finally, I want to thank my family and friends for their love and support. I am deeply grateful to all of you who have been a part of five memorable years in Trondheim.

Abbreviations

bp	basepair
BSR	blasticidin S-resistance
Cas9	CRISPR associated protein 9
cDNA	complementary DNA
CDS	coding DNA sequences
Corr	correction collar
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
Ct	cycle threshold
Cy5	cyanine 5
DNA	deoxyribonucleic acid
DSB	double stranded break
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacteic acid
ER	endoplasmic reticulum
et al.	et alia (and others)
G - phase	growth phase
gDNA	genomic DNA
GFP	green fluorescent protein
HDR	homology-directed repair
HF	hight fidelity (for restriction enzyme)
HNH	an endonuclease domain named for characteristic histidine and asparagine residues
HRM	high resolution melting
INDEL	insertion and/or deletion
LB	Luria Bertani
LCPA	long-chain polyamine
LHCF9	light harvesting fucoxanthin chlorophyll protein
M - phase	mitosis phase
mNG	mNeonGreen
NAT/NOU	nourseothricin N-acetyl transferase resistance
NHEJ	non-homologous end joining
NRF (ptPDS1-M1)	norflurazon resistance (phaeodactylum tricornutum phytoene desaturase-1)
NTC	non template control
oriT	origin of transfer
PaGfp	photoactivatable green fluorescent protein
PAM	protospacer-adjacent motif
pBKS	plasmid Bluescript KS(+)
PCR	polymerase chain reaction

PMT	photomultiplier tubes
RNA	ribonucleic acid
RuvC	an endonuclease domain named for an E. coli protein involved in DNA repair
RT	reverse transcriptase
S - phase	synthesis phase
SAP	silicalemma associated protein
SEM	scanning electron microscopy
sgRNA	single guide RNA
Sh-ble	zeocin resistance
Sin	silicanin
SDV	silica deposition vesicle
SP	signal peptide
ssDNA	single stranded DNA
TAE	tris-acetate-EDTA
TEM	transmission electron microscope
TM	transmembrane
tracrRNA	transactivating crRNA
YFP	yellow fluorescent protein
WT	wildtype

Abstract

Diatoms are one of the major groups of unicellular eukaryotes in the phytoplankton ecosystem. It is estimated that there are around 100,000 species, which account for 40 % of the total primary production. In addition, they play a critical role in the biochemical cycle of carbon. Diatoms' cell wall, the frustule, is believed to be one of the reasons for their bioecological success. Diatoms synthesize their frustule in a process called biomineralization, by transforming soluble silicic acid ($SiOH_2$), into amorphous, hydrogenated silica (SiO_2), inside the silica deposition vesicle (SDV) organelle. It is a complex process involving many different genes and components.

In this master's thesis, the roles of group II silicanins, Tp24708 and Tp24711, in the biomineralization of the cell wall of the diatom *Thalassiosira pseudonana* (*T. Pseudonana*) are studied. Attempts were made to create knockout mutants of the single PAM site, Tp24711P2, and the double PAM site, Tp24708/24711P1, and screen for CRISPR/Cas9-induced mutations in high resolution melting analysis. In addition, the localization of pTpPUC3-mNeonGreen-Tp24711 fusion protein in *T. pseudonana* was studied using a flow cytometer, fluorescence microscope, and confocal microscope. Furthermore, to investigate whether electroporation or biolistic shooting is the appropriate method for genetic transformation of *T. pseudonana*, an analysis of gene expression was performed by quantitative real-time PCR (qRT-PCR). The expression of the native gene *LHCF9* and the genes *Cas9*, *sgRNA*, and *Nat* from the plasmids were analysed. In the past, plasmids suitable for conjugation have been used in selection marker studies. Therefore, an attempt was made to make pBKS_hCas9M plasmid for the different selection markers, blasticidin, nourseothricin, and norflurazone. In order to investigate whether the new plasmids had an effect, a new selection marker study was also carried out.

Knockout mutations were not achieved in the genes encoding Tp24708 and Tp24711, and hence no new information about the function of the proteins in *T. pseudonana*. In the localization study with the mNeonGreen tagged Tp24711, mNeonGreen was observed in the cell division plane, suggesting that Tp24711 may be involved in the synthesis of SDVs. A gene expression study to compare the gene transformation methods was also performed, where it was found that electroporation can be a better suitable method than biolistic shooting. However, it is only possible to suggest, as it was kinases that were electroporated and were compared to group II silicanins that were transformed with biolistic shooting. In previous selection marker studies, a plasmid suitable for conjugation has been used, but since electroporation appears to be the most appropriate method for genetic transformation of *T. pseudonana*, it was desired to assemble pBKS_hCas9 plasmids with selection. The plasmids were not assembled correctly and thus did not give the expected results in the selection marker study.

Sammendrag

Kiselalger er en av hovedgruppene av encellede eukaryoter i planteplanktonøkosystemet. Det er anslått at det er rundt 100 000 arter, som står for 40% av den totale primærproduksjonen. I tillegg spiller de en kritisk rolle i den biokjemiske syklusen til karbon. Kiselalger cellevegg, også kalt *frustule*, antas å være en av årsakene til deres bioøkologiske suksess. Kiselalger syntetiserer sin cellevegg i en prosess som kalles biomineralisering, ved å transformere oppløselig kiselsyre ($SiOH_2$), til amorf, hydrogenert silika (SiO_2), inne i silikaavsetningsvesikkelen (SDV)-organellen. Det er en kompleks prosess som involverer mange forskjellige gener og komponenter.

I denne masteroppgaven studeres rollene til gruppe II silikaniner, Tp24708 og Tp24711, i biomineraliseringen av celleveggen til kiselalgen *Thalassiosira pseudonana* (*T. Pseudonana*). Det ble gjennomført forsøk for å lage knockout-mutanter av det enkle PAM-setet, Tp24711P2, og det doble PAM-setet, Tp24708/24711P1, som videre ble undersøkt for CRISPR/Cas9-induserte mutasjoner i høyoppløselig smelteanalyse (HRM). I tillegg er lokaliseringen av pTpPUC3-mNeonGreen-Tp24711 fusjonsprotein i *T. pseudonana* studert ved bruk av et flowcytometer, fluorescensmikroskop og konfokalt mikroskop. For å undersøke om elektroporering eller biolistisk skyting er den passende metoden for genetisk transformasjon av *T. pseudonana*, ble en analyse av genuttrykk utført ved kvantitativ real-time-PCR (qRT-PCR). Ekspresjonen av genet *LHCF9* og genene *Cas9, sgRNA* og *Nat* fra plasmidene ble analysert. Tidligere har plasmider egnet for konjugering blitt brukt i seleksjonsmarkørstudier og derfor ble det forsøkt å lage pBKS_hCas9M plasmid for de forskjellige seleksjonsmarkørene blasticidin, nourseothricin og norflurazon. For å undersøke om de nye plasmidene hadde en effekt ble det også gjennomført en ny seleksjonsmarkørstudie.

Knockout-mutasjoner ble ikke oppnådd i genene som koder for Tp24708 og Tp24711, og derfor var det heller ingen ny informasjon om funksjonen til proteinene i *T. pseudonana*. I lokaliseringsstudien med mNeonGreen-merket Tp24711, ble mNeonGreen observert i celledelingsplanet, noe som tyder på at Tp24711 kan være involvert i syntesen av SDV-er. Det ble også utført en genekspresjonsstudie for å sammenligne gentransformasjonsmetodene, hvor det ble funnet at elektroporering kan være en bedre egnet metode enn biolistisk skyting. Det er imidlertid bare mulig å foreslå, da det var kinaser som ble elektroporert og ble sammenlignet med gruppe II silikaner som ble transformert med biolistisk skyting. I tidligere seleksjonsmarkørstudier har et plasmid egnet for konjugering blitt brukt, men siden elektroporering ser ut til å være den mest hensiktsmessige metoden for genetisk transformasjon av *T. pseudonana*, var det ønsket å sette sammen pBKS_hCas9-plasmider med seleksjon. Plasmidene ble ikke satt sammen riktig og ga dermed ikke de forventede resultatene i seleksjonsmarkørstudien.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

1.1.1 Diatoms

Phytoplankton, also known as microalgae, are unicellular organisms consisting of photosynthetic prokaryotes and eukaryotes (Bowler et al., 2010). One of the major groups of eukaryotes in the phytoplankton ecosystem are diatoms. The diatoms are a diverse group, with about 100 000 different species, that are critically important for marine and freshwater ecology (Alverson et al., 2011). The size of diatoms varies from 2 µm to 400 µm (Hasle and Syvertsen, 1996), and they grow as single cells or multiple cells linked together like a chain. It is estimated that diatoms are responsible for 40 % of the total primary production in the oceans, which means that they are responsible for approximately 20 % of the global photosynthetic activity. Diatoms thus generate about as much organic carbon as all the terrestrial rainforests combined (Armbrust, 2009; Benoiston et al., 2017). In addition to being an important part of the biogeochemical cycling of carbon, nitrogen, phosphorus, silicon, and iron (Moore et al., 2013; Nelson et al., 1995), diatoms are also considered to produce and represent the primary input of the aquatic food chain, as they provide the basis for both marine and freshwater microorganisms and animal larvae (Falciatore et al., 2019; Kroth, 2007).

It is believed that red algae, green algae, land plants, and glaucophytes originate from a single endosymbiosis event, called primary endosymbiosis. It is a process where a photosynthetic cyanobacterium was taken up by a heterotrophic eukaryotic host cell and was further transformed into a plastid (Kroth, 2007). The evolution of diatoms has taken another complex path by secondary endosymbiosis. In secondary endosymbiosis, a primary photosynthetic eukaryotic algae containing a plastid was taken up by a heterotrophic eukaryotic cell and then transformed into a secondary plastid (Prihoda et al., 2012; Kroth, 2007). Unlike primary endosymbiosis, plastids in organisms resulting from secondary endosymbiosis are surrounded by four membranes instead of two (Prihoda et al., 2012). Representation of the processes involved in primary and secondary endosymbiosis are given in Figure 1.1.

As a result of endosymbiosis gene transfer, diatoms contain a complex mix of genes, which means they have both plant-like and animal-like genes. Among other things, diatoms have a complete urea cycle, which has not been observed in plants (Prihoda et al., 2012). Metabolic pathways have also been found to break down fats that are further produced as energy, as in animals, and metabolic pathways for making chemical intermediates for other reactions, such as plants (Fattorini, 2022).

Another feature that makes diatoms fascinating is their highly ornate external cell walls, called the frustule. The complex process of frustule synthesis has often been studied with respect to biomineralization. In addition to the frustule, other features make diatoms ideal organisms for biotechnological and nanotechnological applications (Kroth, 2007). Among other things, diatoms produce highly unsaturated fatty acids, which can be used as an alter-

native biofuel or as a source of omega-3 fatty acids in aquafeed production (Zendejas et al., 2011; Reitan et al., 2021).

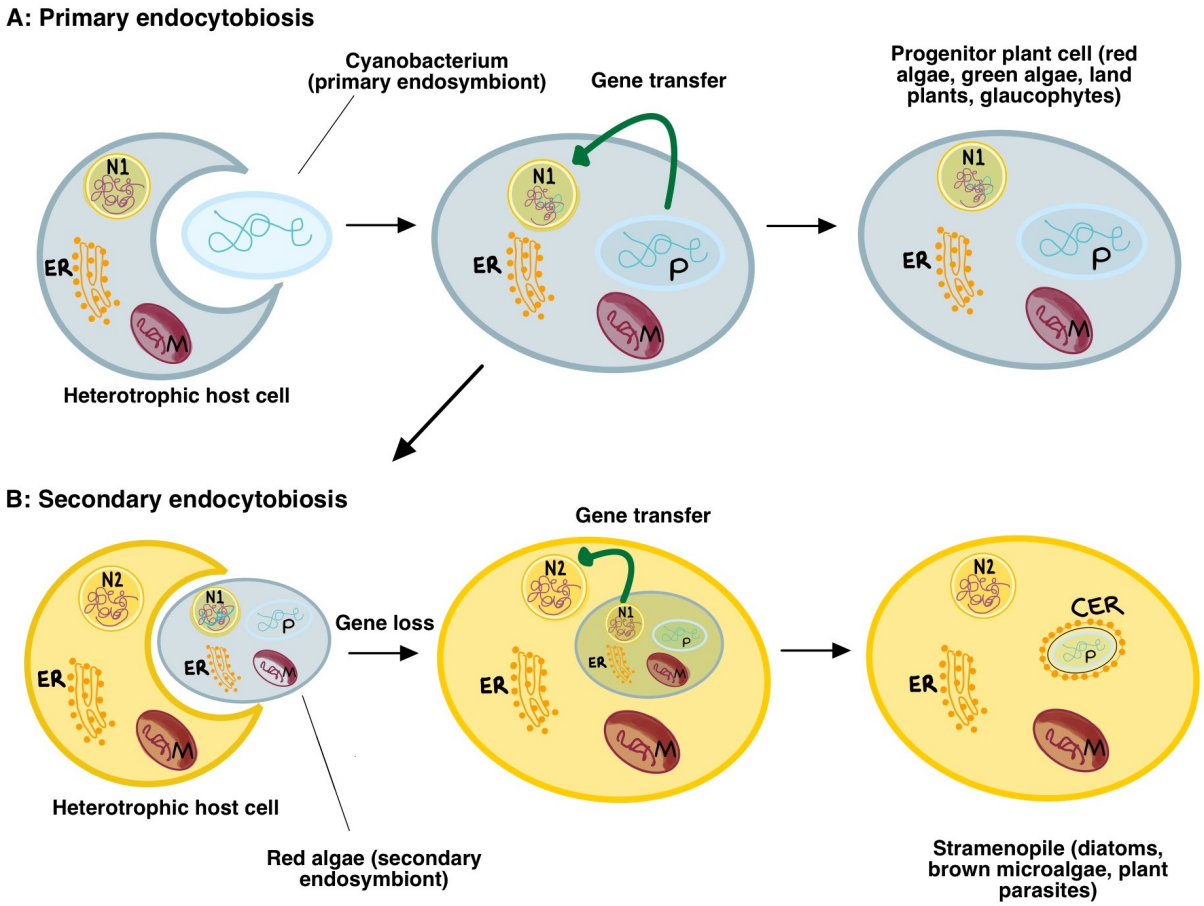


Figure 1.1: A representation of the processes involved in (a) primary and (b) secondary endosymbiosis. a: In a primary endosymbiosis, a heterotrophic eukaryote with a host cell nucleus (N1) will take up parts of the genome from a phototrophic cyanobacteria and will be transformed into a primary plastid surrounded by two membranes. The resulting progenitor plant cell contains parts of the original genes retained in the plastid genome and is further diverged to red algae, green algae, land plants, and glaucophytes. b: In a secondary endosymbiosis, a primary eukaryotic algae containing a chloroplast (P) will be taken up by a heterotrophic eukaryotic cell with a secondary host nucleus (N2) and will be transformed into a secondary plastid surrounded by four membranes, where the outer membrane in diatoms will contain bound ribosomes (chloroplast endoplasmic reticulum (CER)). The figure is recreated from (Kroth, 2007; Armbrust, 2009).

1.1.2 The diatom frustule

A defining feature that makes diatoms fascinating is their ability to control the precipitation of silica to form cell walls, called the frustule. The frustule is composed of silica (SiO_2), proteins and organic material, and is decorated primarily with hierarchically structured nano- and micro patterns of pores (Fattorini and Maier, 2021; Kröger and Poulsen, 2008). The diatoms absorb nutrients and get rid of waste through the pores that have a distinct pattern. Diatoms synthesize their frustule by transforming soluble silicic acid ($SiOH_2$) to amorphous, hydrogenated silica (SiO_2), inside the organelle called silica deposition vesicle (SDV) (Hildebrand et al., 2018). Frustule synthesis is thus an excellent example of biomineralization and

is seen as one of the reasons for the bioecological success of diatoms (Kröger and Poulsen, 2008).

The frustule has a rigid structure, which gives diatoms mechanical protection against predators of phytoplankton in the environment (Hildebrand and Lerch, 2015). However, they are not completely protected, as they are eaten by zooplankton (Liu et al., 2016). In addition, the porous structure allows diatoms to absorb nutrients. It is also believed that the physical structure of diatoms has an impact on how diatoms interact with surfaces and how they alter light that hits the cells. Therefore it is hypothesized whether it is possible that frustules can function as photoselective, photoprotective, or light-focusing structures (Romann et al., 2015).

The word *diatomas* originates from Greek and means "cut in half". This represents the silica frustule, which is composed of two overlapping halves (epitheca and hypotheca), that encase the protoplast. Since epitheca is slightly larger than hypotheca, they fit into each other like a petri dish (Falciatore et al., 2019). Each theca consists of a valve (top and bottom of the petri dish) and a cingulum (side of the petri dish). The valves are called epivalve and hypovalve, respectively, while the cingulums are called epicingulum and hypocingulum, respectively. The cingulums are formed by silicon-containing rings/strips called girdle bands (Kröger and Poulsen, 2008; Fattorini and Maier, 2021). The girdle bands have a simple structure and have a regular pattern with pores. On the other hand, the valves are made of several layers with structures and pore patterns, which vary in size and dimension based on species specificity (Babenko et al., 2022; Tommasi et al., 2017). The girdle bands consist of a series of overlapping circular structures, and the girdle band that overlaps with the valves is called the valvocopula, while the girdle band that overlaps the thecas is called the pleural band (Kröger and Poulsen, 2008).

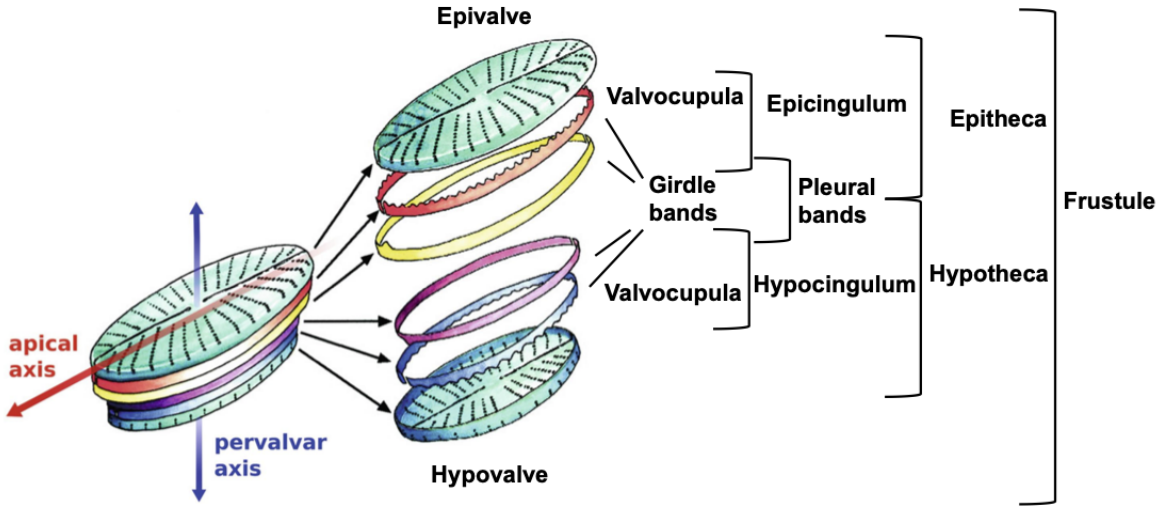


Figure 1.2: Composition of the frustule. The frustule is overlapped with epitheca and hypotheca. The thecas consist of the associated valve (epivalve and hypovalve) and cingulum (epicingulum and hypocingulum). The cingulums form the girdle bands, where the girdle band that overlaps the theca is called the pleural band, while the girdle band connected to the valves is called the valvocopula. The figure is adapted from (Zurzolo and Bowler, 2001)

Diatoms are divided into two main morphological categories, pennates, and centrics (Armbrust, 2009), according to the frustule symmetry with respect to the pervalvar axis (Figure 1.2). Centric diatoms have a radial symmetrical shape and are divided into two subgroups, radial and polar centric(Figure 1.3). Radial centrics have a radial symmetry around the pervalvar axis, while polar centrics are characterized by folded symmetry (Babenko et al., 2022). On the other hand, pennate diatoms have a bilaterally symmetrical shape, which is recognized by an elongated valve. Pennates are also divided into two subgroups, raphid pennates and araphid pennates (Figure 1.3). The subgroups depend on whether the valve contains of a raphe (raphid pennates) or not (araphid pennated). A raphe is a slit-like structure, found on the apical axis (Falciatore and Mock, 2022).

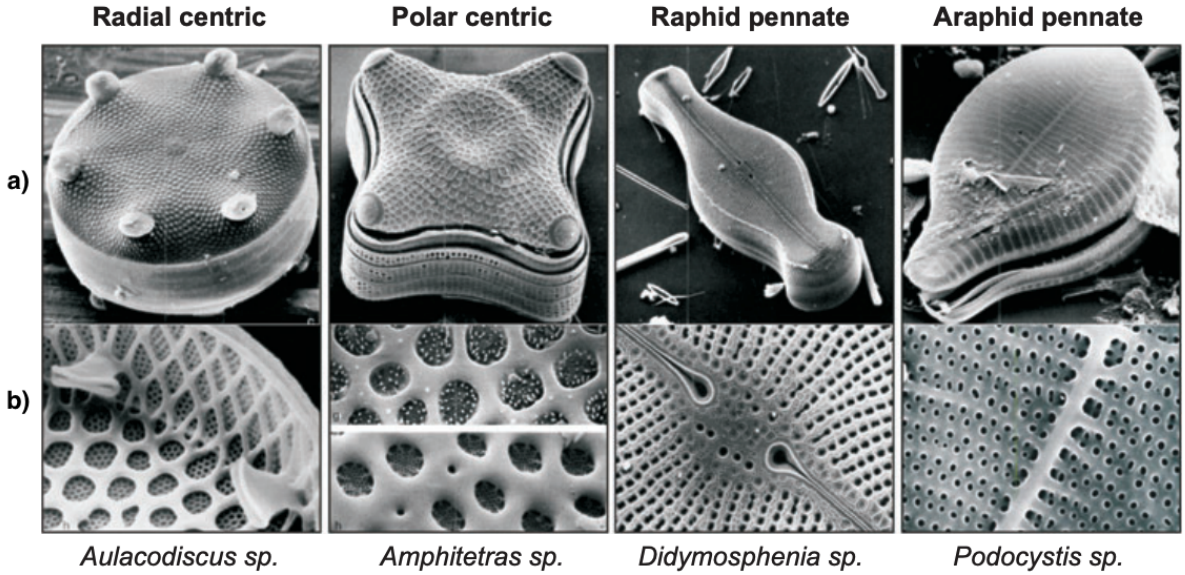


Figure 1.3: Scanning electron microscopy (SEM) images of (a) the morphology of the radial centric *Aulacodiscus sp.*, polar centric *Amphitetras sp.*, raphid pennate *Didymosphenia sp.* and the araphid pennate *Podocystis sp.*, and (b) their highly differentiated nano- and micropatterns. The figure is adapted from (Kröger and Poulsen, 2008).

1.1.3 Frustule synthesis in diatoms

Frustule synthesis depends on the cell cycle, as new frustules are synthesized when the cells divide. Diatoms reproduce asexually through mitotic division by vegetative growth, but some species divide by sexual reproduction to maintain cell size (Fattorini and Maier, 2021). The synthesis of new diatom frustule components takes place in the SDV organelle inside the protoplast and follows the four phases (G1, S, G2, and M) in the cell cycle of eukaryotic cells (Hildebrand, 2008). The valves and girdle bands are synthesized in SDV, which is surrounded by a lipid bilayer membrane called silicalemma. Before the valves and girdle bands exocytose, they are generated in the SDVs. SDVs are categorized into two types depending on what is synthesized; valve SDVs and girdle band SDVs. The valves are formed in valve SDVs and exocytosed to the cell surface. After the mother cell starts a new cell cycle, DNA replication occurs in the S phase, which is further followed by inter- and mitotic phases (G2 and M). During cytokinesis, the two daughter cells will remain in the mother frustule,

but will form a protoplast with silica. In each of the protoplasts a valve SDV will be formed, where synthesis of the new valve will take place. After the valve is formed, the cell will exocytose the contents of the SDVs, which cause the cells to expand when the silica is deposited, pushing the daughter thecas apart so the thecas divide. Cell A inherits the epitheca and will be larger than cell B, which instead inherits the hypotheca. At the same time as the girdle band SDV is synthesized, the girdle band synthesis begins in the intermediate phase (the blue circles given in 6a and 6b in Figure 1.4). In interphase (G1), silica is synthesized to expand the protoplast. The girdle bands are added in separate girdle band SDVs until full cell length is reached. At this point, cell division will be stopped by the formation of pleural bands and the cell cycle is completed, before a new cycle begins (Kröger and Poulsen, 2008; Fattorini and Maier, 2021; Falciatore and Mock, 2022).

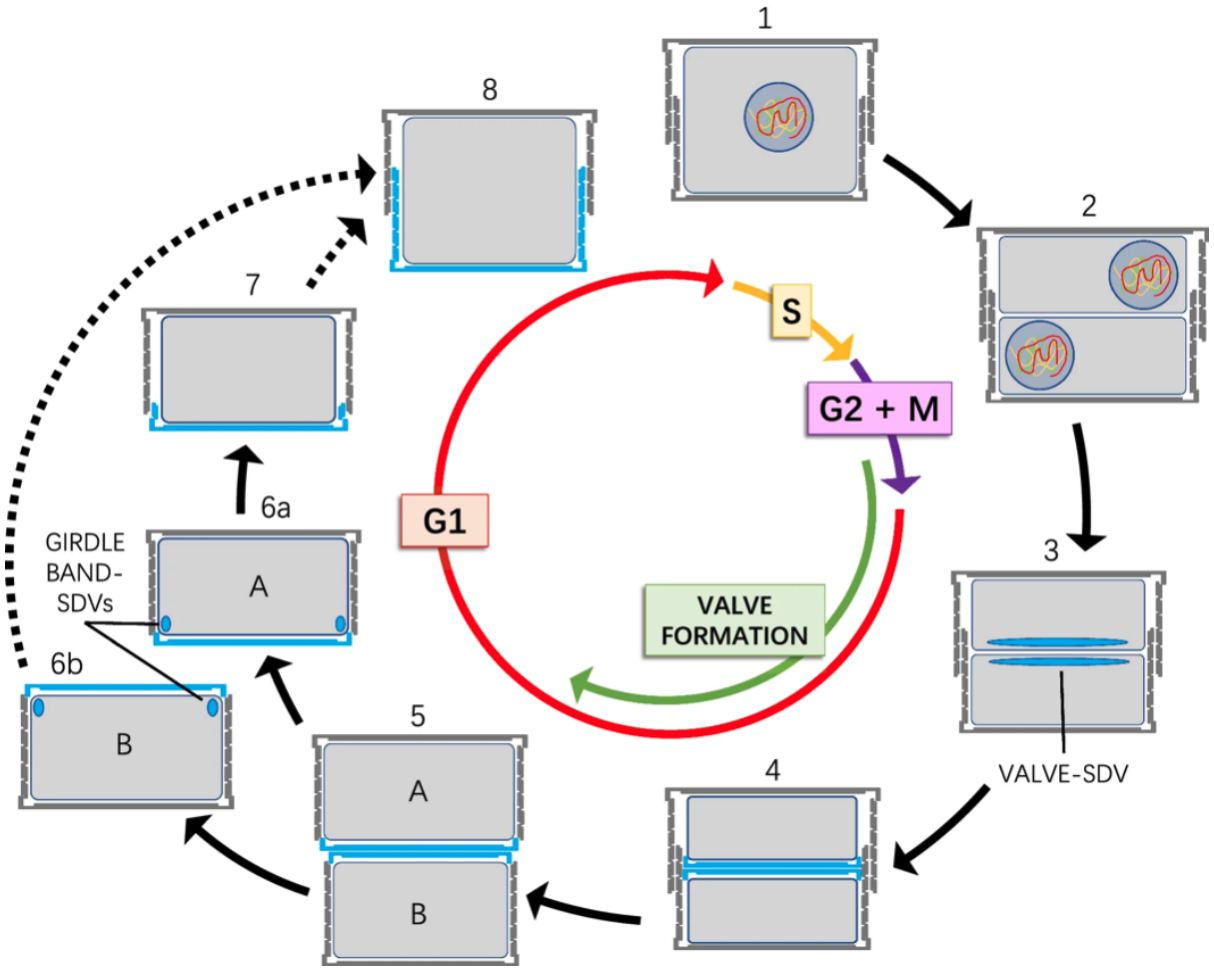


Figure 1.4: A representation of the frustule synthesis. In (1) a new cell cycle begins, followed by DNA replication in the S phase and the cells proceed to the G2 + M phase remaining in the mother frustule, where protoplasts will be formed. A SDV valve is made in each protoplast and the building elements of the hypotheca are synthesized in (3). In (4) the contents of the SDVs are exocytosed. In (5), the daughter cells have completed the separation and begin the development of the girdle bands. Since cell A inherited the mother epitheca, it will be slightly larger than cell B which inherits the hypotheca. Girdle band SDVs synthesize the building blocks for the formation of the girdle bands in (6), and they are exocytosed in (7). In (8), girdle length and cell volume increase as a result of girdle band synthesis. When the cell cycle is complete, a new one starts. The image is retrieved from (Fattorini and Maier, 2021).

1.1.4 Thalassiosira pseudonana

Thalassiosira pseudonana (T. pseudonana) is a centric diatom (Figure 1.5) and was the first diatom to have the genome sequenced, due to its short genome, 32.1 Mbp, (Saade and Bowler, 2009; Falciatore et al., 2019), which contains approximately 11,242 protein-encoding genes (Armbrust, 2009). *T. pseudonana* have circular valves with a diameter of 3.8 ± 0.4 µm and the average length of the cells are 6.5 ± 1.4 µm (Hildebrand et al., 2006). The valve of *T. pseudonana* has silica ribs that radiate from the center (Hildebrand et al., 2006). *T. pseudonana*, reproduces through mitotic division. Most diatoms have a so-called "cell size reduction – recovery cycle", where the cell size is gradually reduced by mitosis until cells undergo sexual reproduction and reach full cell size again. *T. pseudonana* (and *Phaeodactylum tricornutum (P. tricornutum)*) are among the diatoms that are able to bypass this process, possibly because the cell walls are flexible. In fact, sexual reproduction has not been observed in *T. pseudonana*. Therefore, they manage to maintain a constant size through the generations. The cell wall of *T. pseudonana* are decorated with silica ribs and pores, and structures called fultoportulae that extend from the thecas (Fattorini and Maier, 2021). From the time *T. pseudonana* was sequenced, it became a well-used model for diatoms and to study the generation, composition and biomineralization in diatoms by many researchers.

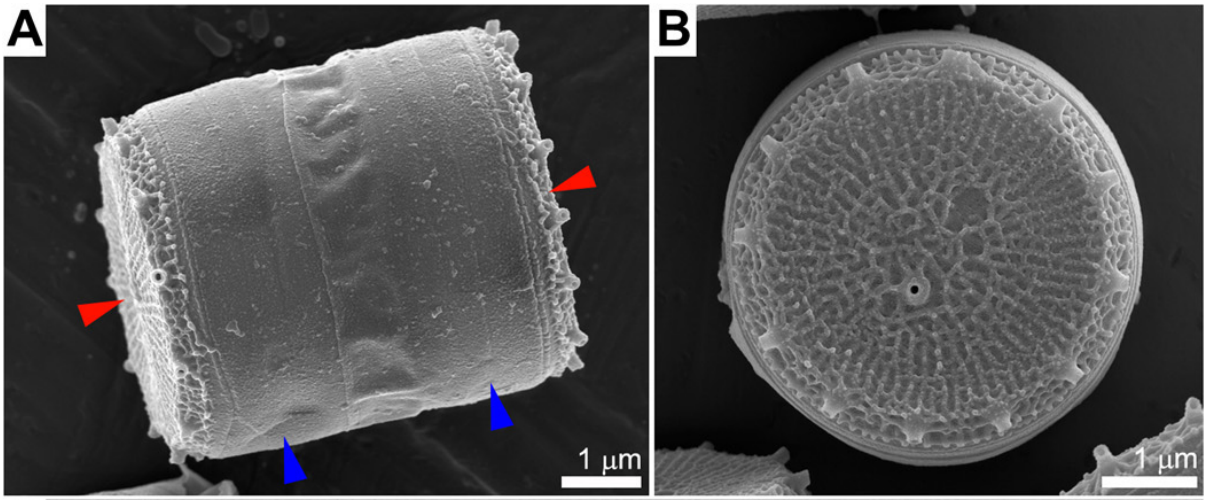


Figure 1.5: Scanning electron microscope (SEM) images of the Thalassiosira pseudonana dried cell walls. A) The side view of the cell, where red arrows indicate the valves, and blue arrows indicate the girdle bands. B) The view of the silica structures on the valve surface. Scale bar is 1 µm. The image is retrieved from (Kumar et al., 2020).

1.1.5 Silicalemma-associated proteins in frustule synthesis

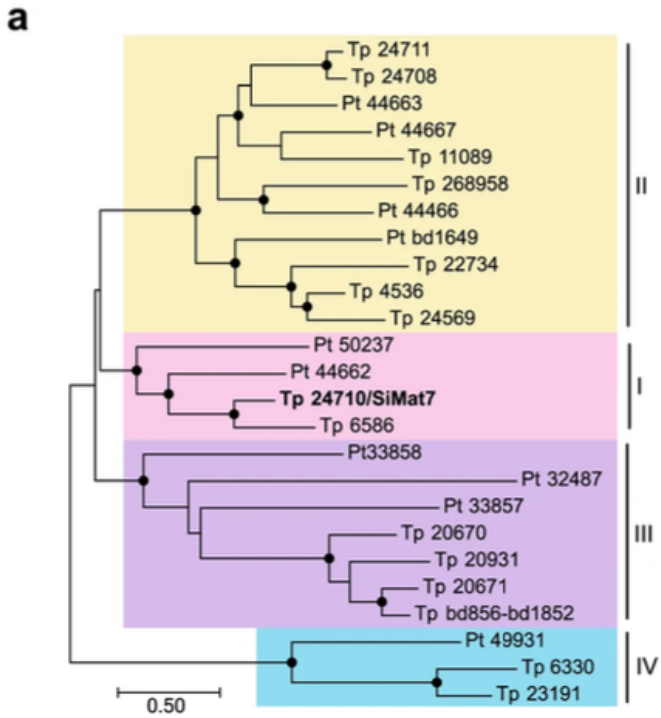
In order to gain an understanding of how the frustule is synthesized, several functional studies have been conducted on the components that are believed to be part of the process. Various components involved in biomineralization have been identified, including the organic compounds long-chain polyamines (LCPA) and chitin, and the proteins such as cingulines, silaffins and silacidines (Hildebrand et al., 2018). Studies on structure formation have also shown that silica formation in diatoms can be divided into three scales: nano-, micro- and

mesoscale (Hildebrand et al., 2006). The two known families of silicalemma-associated proteins are the silicalemma-associated proteins (SAPs) and the silicanins (Sins) (Hildebrand et al., 2018).

SAPs are characterized by a predicted endoplasmic reticulum (ER), signal peptide (SP), transmembrane (TM) domain, and sequence features characteristic of silica-associated proteins, including a short conserved cytosolic domain and a longer SDV luminal region rich in serine (Tesson et al., 2017). SAP1 was the first silica-associated protein identified and was further used to identify the other two silica-associated proteins SAP2 and SAP3. In diatoms, SAPs are involved in silica formation on the mesoscale. In *T. pseudonana*, the silicalemma-associated proteins SAP1 and SAP3 (TpSAP1 and TpSAP3) have been investigated. Green fluorescent protein (GFP) tagged TpSAP1 and TpSAP3 lines have shown localization patterns concentrated on the valve and girdle bands SDVs in *T. pseudonana*, indicating a role in frustule synthesis (Fattorini and Maier, 2021). The knockout lines for TpSAP1 and TpSAP3 showed altered frustule morphology (Tesson et al., 2017). In TpSAP1 knockout lines, the changes were observed in the valves and were characterized by misplaced pattern centers. In contrast, in TpSAP3 knockout lines, it was possible to observe that the height of silica ridges in the valves were reduced, which means that there is a reduction in silicification after the first formation of base layers (Hildebrand et al., 2018; Tesson et al., 2017). The observations thus confirmed the roles of TpSAP1 and TpSAP3 in frustule morphogenesis (Tesson et al., 2017).

Another family of highly conserved silicalemma-associated proteins are silicanins. Silicanin-1 (Sin1) was identified in *T. pseudonana* and consists as SAPs of a single TM domain, ER targeting, a short cytosolic domain and a longer SDV luminal domain (Figure 1.6b) (Tesson et al., 2017). Whereas SAPs are rich in serine, Sin1 is rich in asparagine and glutamine, and consists of 18 conserved cysteine (Cys) residues. Sin1 is also involved in silica morphogenesis and promotes silica polymerization by synergistic interaction with LCPAs (Kotzsch et al., 2017). Sin1 tagged with GFP has been observed in valve and girdle bands in *T. pseudonana*, and CRISPR/Cas9 knockout lines have shown normal cell growth, but reduced silicanin content in the frustule, abnormal silica deposition, and morphological changes. Modifications in mechanical properties can result in a reduction in cell wall strength and stiffness as a result of such alterations (Görlich et al., 2019).

A phylogenetic analysis of the Sin1-like proteins (Figure 1.6a) revealed that the diatom *T. pseudonana* have four subfamilies. Tp24710 in subgroup I is identical to Sin1. Two genes in subfamily II, Tp24711 and Tp24708, are the genes that are in focus of this study. All silicanins share the same general structure (Figure 1.6b). It consists of a SP for cotranslational import into the ER, an RXL protease cleavage site, eight conserved Cys, and a TM (Brembu et al., 2017).



(a) The Sin1 like subfamilies in *T. pseudonana*, where 4 subfamilies have been found in *T. pseudonana (Tp)* and *P. tricornutum (Pt)* (Brembu et al., 2017).

(b) Overall structure of Sin1 like proteins in *T. pseudonana*, consisting of signal peptide (SP), protease cleavage motif (RXL), and a transmembrane domain (TM) (Brembu et al., 2017)

Figure 1.6: The Sin1 like subfamilies in *T. pseudonana* and their overall structure. The image is adapted from (Brembu et al., 2017).

1.2 CRISPR/Cas9-based gene editing

Clustered regularly interspaced short palindromic repeats (CRISPR) with the CRISPR-associated protein 9 (Cas9) is a target gene editing technology. CRISPR/Cas9 genome editing is an adapted mechanism that has evolved from a naturally existing defense system against viruses in bacteria and archea (Hopes et al., 2016).

CRISPR/Cas9 allows a double-strand break (DSB) to be introduced at a specific target sequence, using the Cas9 nuclease and a guide RNA that directs the nuclease to the specific DNA sequence (Kroth et al., 2018). To generate DSB, it is necessary that the guide RNA fragment is complementary to the target site. The Cas9 nuclease carries two nuclease domains (HNH and RvuC), which cleave both DNA strands (Kroth et al., 2018). Single guide RNA (sgRNA) contains CRISPR RNA (crRNA) which consists of a complementary 20-nucleotide sequence to target DNA and transactivating crRNA (tracRNA) (Doudna and Charpentier, 2014). Together, crRNA and tracrRNA form a RNA duplex, and direct Cas9 to the cleavage site using standard RNA-DNA complementarity base pairing rules (Sander and Joung, 2014).

Cas9 uses its HNH domain to cleave the DNA strand at a site, 3-nucleotides away from the protospace-adjacent motif (PAM), which is complementary to the 20-nucleotide sequence, while RvuC cleaves the opposite of the complementary sequence (Doudna and Charpentier, 2014). Using the CRISPR/Cas9 genome editing technique, any DNA sequence can be cleavaged by modifying the 20 nucleotide sequence to complement the DNA target site close to a PAM. (Figure 1.7).

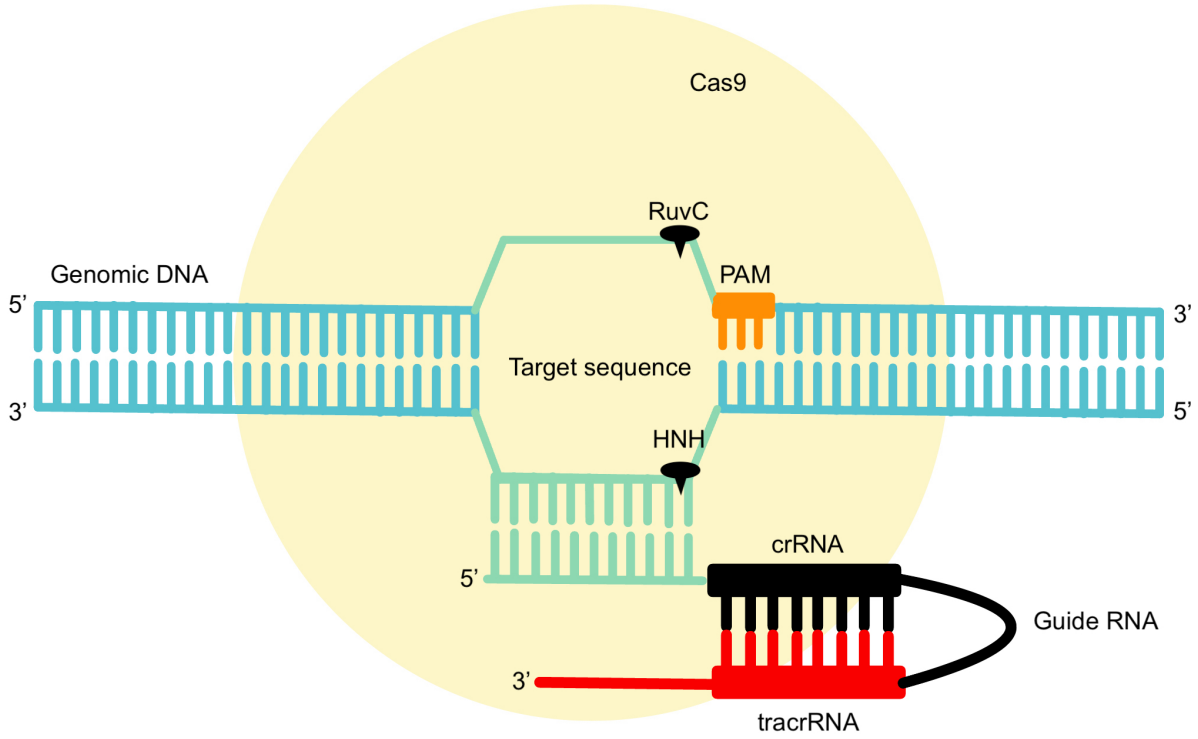


Figure 1.7: A representation of the CRISPR/Cas9 system. CRISPR/Cas9 induces double-strand break (DSB) in the DNA target sequence, using the Cas9 nuclease and a guide RNA. The sgRNA consisting of CRISPR RNA (crRNA) and transactivating crRNA (tracRNA) directs the Cas9 nuclease, consisting of the nuclease domains HNH and RvuC to the DNA target sequence. Cas9 uses its HNH domain to cleave the DNA strand at a site, 3-nucleotides away from the protospace-adjacent motif (PAM), which is complementary to the 20-nucleotide sequence, while RvuC cleaves the opposite of the complementary sequence. The image is adapted from (Doudna and Charpentier, 2014).

Following a DSB, repair of the DNA sequence can occur by two different cellular DNA repair pathways (Figure 1.8), where genomic DNA can be modified: non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Yang et al., 2020). The result of gene editing will thus depend on which of the repair paths takes place. Repair of the DNA sequence with NHEJ can lead to the introduction of insertions or deletions of different lengths. Such a repair will cause interference to the translation reading frame, leading to knockout mutations in the DNA target sequence (Sander and Joung, 2014). Repair of the DNA sequence with HDR can lead to editing of the target sequence by introducing specific point mutations or insertion of a sequence by recombination of a template DNA sequence (Sander and Joung, 2014).

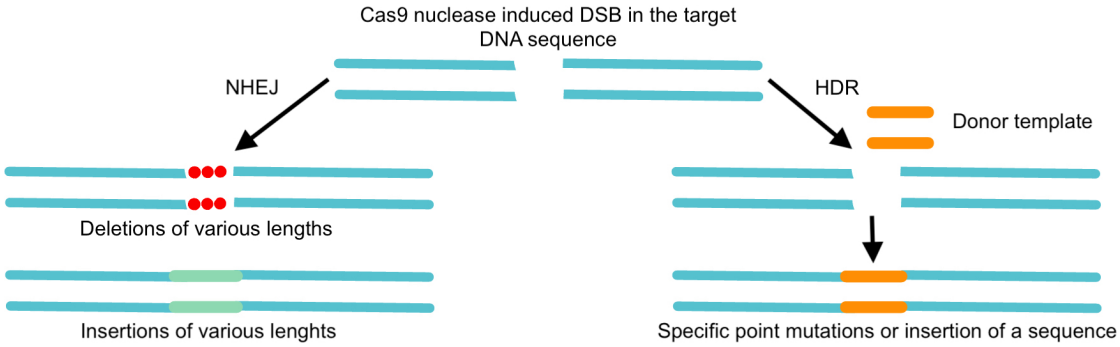


Figure 1.8: A representation of the two different cellular DNA repair pathways, non-homologous end joining (NHEJ) and homology-directed repair (HDR). A repair of the DNA sequence with NHEJ can lead to the introduction of insertions or deletions of different lengths, while a repair with HDR lead to specific point mutations or insertion of a template DNA sequence. The image is adapted from (Sander and Joung, 2014).

CRISPR/Cas9-based genome editing for marine algae, including diatoms, have been developed over the previous 6 years. The urease gene in T. pseudonana was successfully edited using CRISPR/Cas9 in 2016 (Hopes et al., 2016), and the first comprehensive approach to establishing a system for diatoms was published in 2016 (Nymark et al., 2016). Several genetic transformation methods can be used to introduce the CRISPR/Cas9 system into algal cells.

1.3 Genetic transformation methods

Genetic transformation is often used to characterize organisms and study the complexity of organisms and their biological functions and processes. Genetic transformation is the introduction of foreign DNA into a cell, and the first successful documented genetic transformation of a diatom was done by *T. Dunahay et al.* in 1995. Two species of diatoms were genetically transformed by introducing a plasmid vector by microprojectile bombardment, also known as biolistic shooting (Dunahay et al., 1995). Biolistic shooting was the first stable genetic transformation method established in the diatom *(P. tricornutum* (Falciatore et al., 1999). *(P. tricornutum)* is a pennate diatom and thus has properties other than *T. pseudonana.* Biolistic shooting remained the only method of genetic transformation of diatoms for many years, until effective protocols have been found using conjugation (Karas et al., 2015) and electroporation (Niu et al., 2012).

Biolistic shooting is a method used to shoot coated microcarriers directly into *T. pseudonana* cells using a gene gun. In biolistic shooting, tungsten particles coated with DNA are shot into the target cells with the force of pressurized helium, which causes some of the genetic material to enter the cells and transform them. Microcarriers with tungsten particles are the most commonly used in biolistic shooting, but gold particles can also be used. The disadvantage of biolistic shooting, is that it requires high amounts of plasmid, but cause low transformation rates, cell damage and integration of part of plasmid DNA into the host

genome that can affect random genetic elements (Sharma et al., 2018). However, high mutation efficiency has been observed when biolistic shooting is used in combination with CRISPR/Cas9. Since the DNA is randomly incorporated into the genome, insertions will vary for each mutant (Sharma et al., 2018). The complications of random incorporation have long been overlooked, even when CRISPR/Cas9 technology is being developed, but it has emerged in *George et al. (2020)* that other implications such as off-target effects from random integration of gene constructs can occur and cause concern. In this way, generating a precise knockout with biolisitc shooting is suboptimal (George et al., 2020).

Bacterial conjugation is another gene transformation method that has been established in recent times (Karas et al., 2015). Conjugation allows for transformation by direct cell-to-cell contact. In a conjugation, a bacterial vector, such as *Escherichia coli (E. coli)*, transfers the genetic material as an episome to the desired host cell. The epiosome can be constructed to contain the desired gene, promoters, terminator, selection markers, and can be used for the conjugation. In contrast to biolisitc shooting, the DNA in a conjugation will be an episome and will thus be independent of the host genome and can be removed (Sharma et al., 2018). In biolistic shooting, the inserted DNA can not easily be removed and thus it is an irreversible process. Since conjugation is a reversible process, it will be more predictable and easier to control (George et al., 2020).

Electroporation is a method used to transfect cells with DNA using an electric pulse (Carter and Shieh, 2015). When an electric pulse is applied to the cells, temporary pores in the cell membrane will let the cell take up DNA fragments (Carter and Shieh, 2015). Shortly after the pulsations, the pores will be closed and hence there will be no permanent damage to the cells. If suitable electric currents are not used, such as a high electric field, the pores will not close and result in cell death (Ozyigit, 2020). Like biolistic shooting, the foreign DNA will be incorporated into the host genome.

Methods of genetic transformation have both advantages and disadvantages. Conjugation has the advantage of delivering larger DNA fragments than biolistic shooting and electroporation, and it requires smaller amount of cells, as in electroporation and biolistic shooting there is random incorporation of the DNA into the genome, which means that more cells are needed (Gutiérrez and Lauersen, 2021). With biolistic shooting, the cells will also experience a lot of stress, and this is a more expensive method. Transformation parameters in biolistic shooting must also be optimized for each algae and target cell (Gutiérrez and Lauersen, 2021). Biolistic transformation can also result in the integration of a large number of copies of the plasmid (George et al., 2020). Electroporation is a cheaper method, but random incorporation is a disadvantage of this method as well. Conjugation is not dependent on the host genome, and thus the plasmid can be eliminated by easily removing the selection marker (Gutiérrez and Lauersen, 2021). Several attempts by other master's students as a part of the research project DIASIL at NTNU have concluded that there are also problems with conjugation, even though this has not been published (Bentzen, 2020; Harris, 2021; Sjevelås, 2021; Skisland, 2021).

1.4 Selection markers in diatoms

Selection markers are important in studies with genetic transformation, as they are a part of the screening. In diatoms, the availability of antibiotics and the number of selection markers are very low, which means that genetic modifications of diatoms are limited to two independent screening steps (Buck et al., 2018). It is mainly the antibiotics that were found decades ago, zeocin (Sh-ble) and Nourseothricin N-acetyl transferase (Nat) that are most used (Apt et al., 1996). The selection markers were found in studies with the diatom *P. tricornutum*, (Apt et al., 1996). Although several antibiotics and herbicides can be harmful to microalgae, it is often the large amount that causes mortality. Sh-ble has been shown to be harmful at low concentrations in *P. tricornutum*, while Nat can have higher concentrations without doing harm. In *T. pseudonana*, it has been shown that Nat is 100 times less toxic than Sh-ble. *T. pseudonana* is thus so sensitive to zeocin that it does not function as a selection marker and therefore Nat has remained the most common selection marker for *T. pseudonana* (Poulsen et al., 2006).

Blasticidin-S deaminase (Bsr) has recently been found to be a successful selection marker in *P. tricornutum* (Buck et al., 2018). In the study by *Buck et al. (2018)* it was also shown that Bsr only confers resistance to blasticidin-S and does not alter the effectiveness of Sh-ble and Nat (Buck et al., 2018). In a study by *Taparia et al. (2019)*, the first successful herbicide selection marker, Norfluarzon in *P. tricornutum* was found, by generating a point mutation in the phytoene desaturase-1 (PtPDS1) gene (Taparia et al., 2019). Norfluarzone is a bleaching herbicide, that inhibits one of the steps in the carotenoid biosynthetic pathway when it binds to PDS1. In the mutations, it was observed that norflurazone did not bleach the cells, which means that inhibition was avoided (Taparia et al., 2019). Although antibiotics and herbicides have been studied on other diatoms, it is believed that they may be useful in *T. pseudonana* as well.

1.5 Thesis objective

The purpose of the master's thesis is to use CRISPR/Cas9 gene editing in combination with different biovisualization techniques and biochemical characterization to investigate the roles of group II silicanins in biomineralization of the cell wall of the diatom *T. pseudonana*. The master's thesis is a continuation of the specialization project in the course TBT4500 ((Mathivannan, 2021)). In the specialization project, two members of the group II silicanins, Tp24708 and Tp24711 were studied. Two PAM sites, a single PAM site, Tp24711P2, and a double PAM site, Tp24708/24711P1, from the coding sequences, were gene transformed into *T. pseudonana* by biolistic shooting. pTpPUC3-mNeonGreen-Tp24711 fusion protein was transformed into *T. pseudonana* by electroporation. As a continuation of the specialization project, it is desired to see how Tp24708 and Tp24711 affect frustule synthesis. In addition, it is desirable to adapt the already existing plasmids for studies of selection marker by conjugation suitable for electroporation instead. Based on the purposes of the master's thesis, the given sub-goals were defined:

1. Creating knockout mutants of the single PAM site, Tp24711P2, and the double PAM site, Tp24708/24711P1, and screen for CRISPR/Cas9-induced mutations in high-resolution melting.
2. Study the localization of pTpPUC3-mNeonGreen-Tp24711 fusion protein in *T. pseudonana*, using flow cytometer, fluorescence microscope and confocal microscope.
3. Analyse the gene expression of *T. pseudonana* in biolistic transformation mutants compared to electroporation mutants.
4. Creating pBKS-hCas9M plasmid for the different selection markers, blasticidin, nourseothricin and norflurazone, followed by a selection marker study with the plasmids.

2 Materials and Methods

2.1 Materials

Tables with microorganism strains, kits, components, instruments, software and medium with associated supplier are provided in Appendix B and Appendix H, together with reaction mixes in Appendix C and primers in Appendix D. New England Biolabs (NEB) provided all restriction enzymes, whereas ThermoScientific, SigmaAldrich, and VWR provided other common compounds and chemicals. SigmaAldrich provided the primers, and Eurofins (Lightrun) did the Sanger sequencing.

2.2 Culture conditions

The *T. pseudonana* strain (CCMP 1335), ordered from The Culture Collection of Algae and Protozoa (CCAP), were grown as liquid axenic culture in an oxygen-rich environment using a cell culture flask with a filter cap (VWR), at 18 °C in constant light (175 µmol m^{-1} s^{-1}). The cultivation medium L1 was made from autoclaved (20 minutes, 121 °C) sterile filtered (0.2 µm) seawater (from Trondheimsfjorden, provided by NTNU Sealab), with sterile filtrated (2 µm) nutrients according to *Guillard and Hargraves (1993)*, as given in Table B.1 (Appendix B).

To keep the cells alive and in the exponential growth phase, they were diluted 10 times with L1 medium every two weeks and where diluted every three days. Transgenic lines of *T. pseudonana* were first grown on L1 agar plates with nourseothricin (nou) for 2-3 weeks before being picked and grown axenically with L1 medium with nou (100 µg mL^{-1}) at the same conditions. To confirm that the culture were alive, the cells were measured in Attune™ NxT Flow Cytometer and examined with a Axio Imager.Z2 Fluorescence microscope to ensure healthy and uncontaminated culture. In studies with different selection markers, transgenic lines were grown on L1 plates with different types of antibiotics and herbicides (Table 2.1) with different concentrations.

Table 2.1: A list of antibiotics and herbicides used during the thesis.

Purpose	**Antibiotics and herbicide**
Bacterial transformation	Kanamycin
Transgenic lines of *T. pseudonana* and study of selection markers	Nourseothiricin
Assembly of pBKS_hCas9M with selcetion	Ampicillin
Study of selection markers	Blasticidin
Study of selection markers	Norfluarzone

2.3 General protocols

2.3.1 Heat-shock transformation of *E. coli*

The heat-shock transformation method was used to insert a plasmid or ligation product into *E. coli*. DH5α competent *E. coli* (100 µL) from -80 °C freezer were thawed on ice. Thawed plasmid (0.5 µL) or ligation product (5 µL) was added to DH5α competent *E. coli*. The cells were heat shocked by incubation at 42 °C for 45 seconds in a heat block and incubated on ice for 2 minutes immediately after. Luria Bertani (LB) (Table B.2, Appendix B) medium (1 mL) was added to the transformation under sterile conditions and the transformation was grown at 37 °C in a heat block with shaking (220 rpm) for 1 hour. The transformation mix (100 µL) was spread on LB plates (37 °C) containing antibiotics (Table 2.1) and incubated over night at 37 °C. The plates were stored at 4 °C for further use.

2.3.2 Overnight cultures

One colony from the plate of the transformation was picked with a pipette tip and dropped with the tip in LB medium (10 mL) with antibiotics (100 µg mL^{-1}). The culture was then incubated at 37 °C overnight with shaking (220 rpm).

2.3.3 Electroporation of *T. pseudonana*

An electroporation protocol for the transformation of the diatom Chaetoceros muelleri, has been established by *Yin and Hu (2021)* and was further adjusted for electroporation of *T. pseudonana*.

100 mL *T. pseudonana* culture were grown as described in Subsection 2.2. The culture (2x10^6 cells/mL) were counted in Attune™ NxT Flow Cytometer and centrifuged at 3000 xg, 4 °C, for 10 minutes. The supernatant was decanted, and the *T. pseudonana* cells were resuspended in 1 mL ice-cold sorbitol (500 mM) and centrifuged at 1000 xg, 4 °C, for 8 minutes. This step was repeated twice, before resuspending the *T. pseudonana* cells in 100 µL ice cold sorbitol (500 mM). 5-7 µg linearized plasmid and 80 µg salmon sperm DNA (8 µL, 10 µg/µL) were added to the suspension and mixed with a pipette. The solution was then incubated on ice for 30 minutes and was then transferred to ice cold 2 mm electroporation cyvette.

The electroporation was performed with a Bio-Rad Gene Pulser Xcell Electroporation System in an exponential decay mode with an electrical pulse of 500 V, capacitance of 25 µF and a resistance of 400 Ω. Immediately after the electroporation, the solution was transferred to a conical Falcon tube (15 mL), with 10 mL L1 medium, and incubated over night at about 30 µmol photons m^{-2} s^{-1} at 18 °C.

The next day the cells were harvested by centrifuging at 1000 xg for 10 minutes. The supernatant was decanted and the cells were resuspended in 600 µL L1 medium, and 200 µL and 300 µL respectively, were spread on L1 plates (1/2 x L1 medium with 1/2x nutrients and 1x vitamins, 0,5% agar), containing antibiotics. The plates were incubated at 18 °C in constant

light, until colonies appeared (~ 14 days).

2.4 Creating knockout mutants of Tp24708 and Tp24711

The vector plasmid pTpPUC3 (Figure A.1) was modified to contain the Cas9 gene and sgRNA in the vector plasmid pTPPUC3 hCas9 (Mock) U6 complete (Figure A.2). This was further modified to include single PAM sites and double PAM sites for two of the group II silicanins, Tp24708 and Tp24711. From these silicanins, a single PAM site (Tp24711P2) and a double PAM site (Tp24708/24711P1) were chosen, and used to generate single and double knockout-lines in *T. pseudonana* (Table 2.2). The plasmids pTpPUC3 hCas9 Sg mut 24711P2 (Tp24711P2) and pTpPUC3 hCas9 Sg mut Tp24708/24711P1 (Tp24708/24711P1) were made and sequenced by Annika Messemer at the ACMS lab.

Table 2.2: The single PAM site, Tp24711P2, and the double PAM site, Tp24708/24711P1, used to generate single and double knockout-lines in *T. pseudonana*, with associated PAM sequence and orientation.

PAM site	PAM sequence	Orientation
Tp24708/24711PAM1	ACGGAACCTCACAAGACGCTGGG	Reverse
Tp24711PAM2	ATTGTTCTCCTCGACGCGACGGG	Reverse

The Tp24711P2 and Tp24708/24711P1 plasmids were amplified from heatshock transformation (Subsection 2.3.1) and overnight cultures (Subsection 2.3.2). To extract the plasmid, a mini-prep with E.Z.N.A® Plasmid DNA Mini Kit II (Omega Bio-tek), was performed according to the E.Z.N.A® Plasmid DNA Mini Kit II Spin Protocol (Omega Bio-tek, 2019). The concentration were measured after extraction on a NanoDrop One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Scientific). The size was confirmed by performing a restriction enzyme digestion with the dual cutter restriction enzyme KpnI-HF (New England Biolabs, Table C.1, Appendix C), followed by a gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1 h and 20 min). GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths, and 6x Loading Dye (Thermo Scientific) was used for staining the DNA. The lengths were compared to the expected fragment length in SnapGene.

The pTpPUC3 hCas9 Sg mut 24711P2 and pTpPUC3 hCas9 Sg mut 247081/24711P1 plasmids were transfected into *T. pseudonana* cells by electroporation as given in Subsection 2.3.3. Grown colonies were picked and cultivated in 24-well plates (VWR) with 1 mL L1 medium, containing nourseothricin (100 µg mL^{-1}).

2.4.1 Screening for CRISPR/Cas9 induced mutations by high resolution melting (HRM)

High resolution melting (HRM) analysis is a post-PCR method for characterisation of variations in nucleic acid sequences (Hoffmann et al., 2007; Kalthoff et al., 2013). To perform HRM, a larger region around the target DNA sequence must be amplified by PCR, with specific DNA saturation dyes, that only fluoresce in the presence of double stranded DNA (dsDNA). The PCR product will be amplified dsDNA. In the HRM process, the amplicon DNA

is gradually heated. When the melting temperature is reached, the amplicon will denature. Since the fluorescent saturating dye only fluoresces when it binds dsDNA, denaturation of DNA will lead to a reduction in fluorescence and makes it possible to detect the fluorescence. The temperature, where the fluorescence is detected, generates a melting curve. In the melting curve, the midpoint is described as the point where there is an equal amount of dsDNA and single stranded DNA (ssDNA) (Foroni et al., 2017). PCR amplicons with different sequence length, GC content and strand complementarity will give different melting curves. By analyzing the results from HRM, small differences in PCR melting curves between mutants and wildtype (WT) sample can be detected and makes it possible to find whether it is homozygous WT, homozygous mutant or heterozygous WT and mutant (Foroni et al., 2017). A representation of the screening procedure is presented in Figure 2.1.

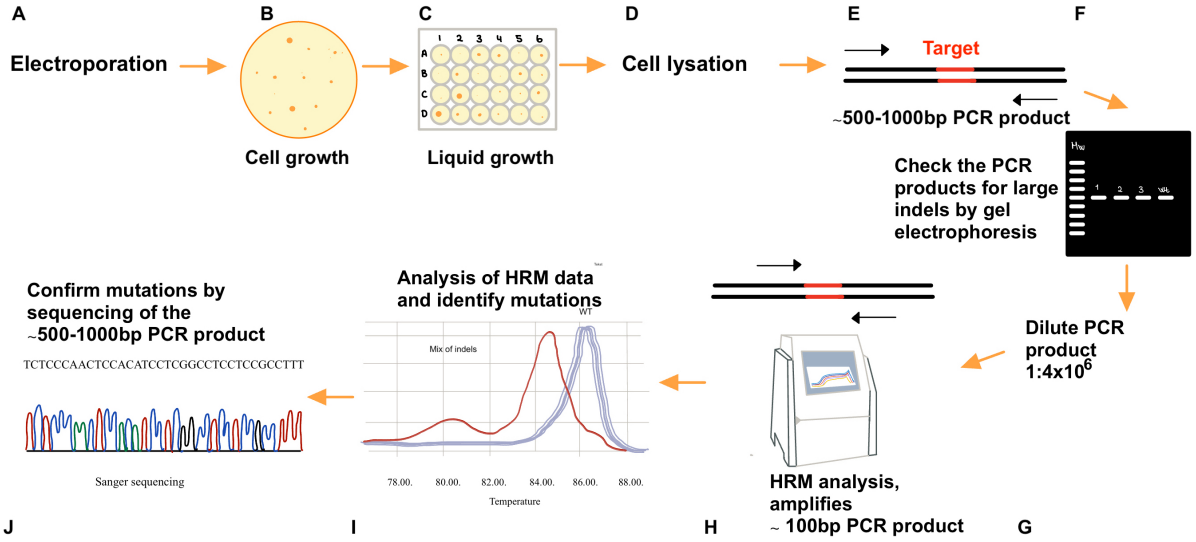


Figure 2.1: A representation of the screening procedure to identify CRISPR/Cas9-induced mutations by HRM in the target region the of Tp24708 and Tp24711 genes. A: Transfect the pTpPUC3 hCas9 Sg mut 24711P2 and pTpPUC3 hCas9 Sg mut 247081/24711P1 plasmids into *T. pseudonana* cells by electroporation and plate out on L1 agar containing nourseothricin (100 µg mL^{-1}) B: Let the colonies appear on the plates (~ 2-4 weeks). C: Cultivate the cells in 24-well plates with 1 mL L1 medium, containing nourseothricin (100 µg mL^{-1}) and let the cells grow (~ 1-2 weeks). D: Use a small amount of the cells to make lysate. E: Use the lysate as DNA template and amplify a region around the target site by PCR; F. Check the PCR products for large indels by gel electrophoresis. G: Dilute the PCR product 1:4x10^6. H: Perform HRM analysis to detect smaller indels around the target site. I: Analyse the HRM data. Study the difference in melting temperature between mutants and WT sample to identify samples with CRISPR/Cas9-induced mutations. J: Confirm mutations detected by HRM by submitting purified PCR product from (E) with HRM primers to Sanger sequencing. Figure is recreated from (Nymark et al., 2017).

Screening was performed as described in *Nymark et al. (2017)*, with some modification. HRM was performed on the colonies appeared from the electroporation plates of Tp24708/Tp24711P1 and Tp24711P2. 10 µL from each culture cultivated in 24-well plates (VWR) with 1 mL L1 medium, containing nourseothricin (100 µg mL^{-1}) were lysed as described in *Hopes et al. (2016)*. The cells were spun down (17000 xg for 1 min) and the supernatant was decanted. The cells were resuspended in 20 µL lysis buffer (10% Triton X-100, 20 mM Tris–HCl pH 8, 10 mM EDTA) and kept on ice for 15 min and incubated for 10 min at 95 °C.

In order to check for large indels by PCR, a region of ~ 560 bp for surrounding the target site (Tp24708/24711P1) in the Tp24708 genes and a region of ~ 750 bp for surrounding the target site (Tp24708/24711P1 and Tp24711P2) in the Tp24711 genes were amplified with the Phusion plus DNA polymerase kit (Thermo Scientific), where 1 µL of the lysate was used as DNA template. The reaction setup is given in Table 2.3 and thermocycle profile used is given in Table 2.4. The insert specific primers and the sequences used are given in Table D.1, Appendix D. Gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1 h) was performed to check the PCR products for potential indels. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths, and 6x Loading Dye (Thermo Scientific) was used for staining the DNA. The lengths were compared to the expected fragment length in a WT *T. pseudonana*.

Table 2.3: Reaction setup (50 µL) for PCR amplification with Phusion Plus DNA Polymerase (Thermo Scientific).

Components	**Volume [µL]**	**Final concentration**
5X Phusion Plus Buffer	10.00	1x
Forward primer	2.50	0,5 µM
Reverse primer	2.50	0,5 µM
10mM dNTPs	1.00	200 µM each
Template DNA	1.00	-
5X Phusion GC Enhancer	10.00	1X
Phusion Plus DNA Polymerase	0.50	-
Water, nuclease free	22.50	-
Total volume	50.00	-

Table 2.4: Thermocycling profile for PCR amplification with Phusion Plus DNA Polymerase (Thermo Scientific).

Program	**Temperature** (°C)	**Time**	**Cycles**
Initial denaturation	98	30 sec	1
Denaturation	98	10 sec	
Annealing	60	30 sec	34
Extension	72	15-30s/kb	
Final extension	72	10 min	1
	4	∞	

If the agarose gel showed clear and distinct bands expressed, the PCR products were further screened by HRM. A dilution series from 1:40 to 1:4x10^6, of the PCR products with bands

were made. 15 µL of the mastermix made with LightCycler® 480 High Resolution Melting Master (Table 2.5) with specific HRM primers (from stock solution of 100 µM primer: 2 µL of forward and reverse primer + 96 µL water, Table D.1, Appendix D) for single and double knockout mutant screening were added to 96 well plates (LightCycler® 480, Roche). 5 µL of the 1:4x10^6 dilution was used as a template to amplify a short indel, a region of ~ 150 bp for surrounding the target site in the Tp24708 and Tp24711 genes. The 96 well plates (LightCycler® 480, Roche) were sealed with sealing foil (LightCycler® 480, Roche) and spund for 2 min at 1500 xg. LightCycler® 96 instrument (Roche) was used to amplify the PCR product with the thermocycling profile given in Table 2.6. The the raw data was analysed with LightCycler® 96 software 1.1 (Roche).

Table 2.5: Reaction setup (15 µL) for HRM amplification with LightCycler® 480 High Resolution Melting Master (Roche).

Components	**Volume [µL]**	**Final concentration**
H_2O, PCR-grade	10.00	-
HRM master mix	0.60	1x
MgCl 25 mM	2.40	3 mM
Primer 200 nM	2.00	200 nM each
Total volume	15.00	-

Table 2.6: Thermocycling profile for HRM amplification with LightCycler® 480 High Resolution Melting Master (Roche).

Program	**Temperature** (°C)	**Time (s)**	**Ramp** (°C s^{-1})	**Cycles**
Pre incubation	95	600	4.4	1
3 step amplification	95	10	4.4	45
	63	10 (TD 63 °C, 1 cycle → 55 °C (-1 °C)	2.2	
	72	20	4.4	
High resolution melting	95	60	4.4	1
	40	60	2.2	
	65	1	2.2	
	95	1	0.7	
Cooling	37	30	1.0	1

If the melting curves deviated from the melting curve for WT, it was assumed that there could be single (Tp24711P2) or double (Tp24708/24711P1) knockout mutations. The PCR products for the samples with suspected mutations were purified with the Monarch PCR & DNA Cleanup Kit (5 µg) (New England BioLabs Inc.) and the concentrations were measured on a NanoDrop One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Scientific). Purified PCR products were submitted for Sanger sequencing with specific primers (Table D.1, Appendix D) performed by Eurofins (Lightrun).

2.5 Localization study of pTpPUC3-mNeonGreen-Tp24711 fusion protein in *T. pseudonana*

pTpPUC3-mNeonGreen-Tp24711 fusion protein plasmid was constructed by Marita Gresseth at the ACMS lab and was transformed into DH5α competent *E. coli* by heatshock transformation (Subsection 2.3.1) and overnight cultures (Subsection 2.3.2) (Gresseth, 2020). To extract the plasmid, a mini-prep with E.Z.N.A® Plasmid DNA Mini Kit II (Omega Bio-tek). The concentration were measured after extraction on a NanoDrop One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Scientific).

pTpPUC3-mNeonGreen-Tp24711 was introduced into *T. pseudonana* cells by electroporation as given in Subsection 2.3.3. Grown colonies were cultivated in 24-well plates (VWR) and were further grown as liquid axenic culture in an oxygen-rich environment using a cell culture flask with a filter cap (VWR). A localization study of pTpPUC3-mNeonGreen-Tp24711 transformants were performed, by studying the transformants in flow cytometer, fluorescence microscope, and confocal microscope. The reason why both fluorescence microscopy and confocal microscopy were used to study the transformants is that in a fluorescence microscope the whole sample can be studied simultaneously, as the sample is evenly flooded by light from a light source, while in a confocal microscope it is possible to focus on specific point of the sample, which is exposed to light from a light source. In a conventional fluorescence microscope, noise from parts of the object outside the plane being observed is also seen, which can make it difficult to study the transformants.

2.5.1 Flow cytometer

Attune™ NxT Flow Cytometer (Thermo Fisher Scientific) equipped with a blue (488 nm, 50 mW) and a yellow (561 nm, 50 mW) laser, was used to examine different aspects of the transformed cells compared to WT. The flow cytometer measured forward scatter (FSC), side scatter (SSC) and fluorescence for 100 µL sample for 1 minute. FSC is light that is scattered in the forward direction and is proportional to the size of the particle. FSC can therefore be used to quantify the particle size. SSC, on the other hand, is the light that is scattered at larger angles and can thus be used to measure the structural complexity of the cells. In order to measure the fluorescence of the pTpPUC3-mNeonGreen-Tp24711 transformants, BL1 (530/30 nm) and BL3 (695/40 nm) channelss were used. Since some of the events (cells) that were measured may be dead cells or cell debris, and dead events were removed by gating only the cells with high BL3 fluorescence. To compare which transformants expressed the most NeonGreen, overlays were made and the transformants that differed the most compared to the WT and expressed the highest fluorescence were further studied with a fluorescence microscope and confocal laser scanning microscope. All measurements were retrieved from Attune™ NxT Software.

2.5.2 Fluorescence microscope

Axio Imager.Z2 Fluorescence microscope with a Plan-Apochromat 40x /0.95 Ph 3 Corr objective, was used to study the size and shape of the pTpPUC3-mNeonGreen-Tp24711 transformants. In addition, the WT sample was also studied, to compare the difference. 10 µL of each sample were transferred to a 76x26 mm microscope slide (VWR) and a 20x20 mm cover glass (VWR) was placed on top of the sample to seal and the microscope slide was placed in the microscope. HXP-120 UV light source was used to measure mNeonGreen fluorescence and the autofluorescence, while the TL VIS-LED lamp was used as the light source to measure brightfield. The microscope had a 505-555 nm emission filter, called EGFP or PaGfp, to detect mNeonGreen fluorescence, and a 665-715 nm emission filter, called Cy5, to detect autofluorescence emitted from the chloroplasts. In addition, a brightfield channel was used. ZEISS ZEN Software was used for image processing and to measure the lengths of the cells.

2.5.3 Confocal microscope

Leica SP8 Confocal Microscope with HC PL APO CS 63x/1.20 water lens, was used to study specific points of the sample pTpPUC3-mNeonGreen-Tp24711 transformants and WT sample. 20 uL of each sample was added to a 76x26 mm microscope slide (VWR) with double-sided tape (Grace Bio-Labs SecureSeal ™ imaging spacer). A 20x20 mm cover glass (VWR) was placed on top of the tape to seal and the microscope slide was placed inverted in the microscope. Diode 638 (laser intensity 1 %) and OPSL 488 (laser intensity 5 %) lasers were used, together with PMT1 channel with optical filter 493-540 nm for emission, to detect mNeonGreen fluorescence and PMT 3 channel with optical filter 651-779 nm for emission, to detect autofluroscense emitted from chloroplast. In addition, PMT channel was used for bright field. The images were taken in 512x512 or in 1024x512 format with a speed of 100 Hz and a line average of 5. Leica LAS AF was used for image processing and to measure the lengths of the cells.

2.6 Analysis of gene expression in *T. pseudonana*

To investigate whether electroporation or biolistic shooting is the appropriate method for genetic transformation of *T. pseudonana*, an analysis of gene expression was performed by quantitative real-time PCR (qRT-PCR). The expression of the native gene *LHCF9* and the genes *Cas9*, *sgRNA* and *NAT* from the plasmids were analysed.

2.6.1 RNA isolation

30 mL liquid cell culture from the colonies picked after electroporation and biolistic shooting (from the project assignment) were grown as described in Subsection 2.2 and centrifuged at 3000 xg, 4 °C, for 10 minutes. The supernatant was decanted, and the cells were resuspended in left over medium. In order to snap-freeze the cells with liquid nitrogen, the cell culture were transferred to 2 mL Eppendorf tubes and centrifuged at 10,000 xg for 1 minute. The supernatant was removed and the tubes were placed in liquid nitrogen. Pre-cooled (-80 °C) 5 mm stainless steel bead (QIAGEN) was added to each tube and the tubes were placed with counterweight in pre-cooled (-80 °C) TissueLyser Adapter Sets (QIAGEN). Homogenization of the cells was performed using TissueLyser (QIAGEN) at 25 s^{-1} for 2 minutes. To disrupt the tissue, RLT lysis buffer (450 µL, 10 µL β-Mercaptoethanol per 1 ml Buffer RLT) was added and the samples were shaken in room tempered TissueLyser Adapter Sets (QIAGEN) at 25 s^{-1} for 2 minutes. To help the disruption of tissue, the samples were incubated for 3 minutes at 56 °C.

RNA isolation was performed with RNeasy Plant Mini Kit (QIAGEN). 96 % ethanol and buffer RW1 were added to the lysate and the lysate was washed with buffer RPE according to the Quick-Start Protocol RNeasy® Plant Mini Kit (QIAGEN, 2016*b*). The isolated RNA was eluted in 40 µL RNAse free water and the concentration was measured on a NanoDrop One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Scientific).

2.6.2 Synthesis of cDNA

Complementary DNA (cDNA) was prepared according to Quick-Start Protocol QuantiTect® Reverse Transcription Kit (QIAGEN, 2016*a*). The genomic DNA (gDNA) elimination reactions (21 µL) were prepared on ice as given in Table C.2 and were incubated at 42 °C for 2 minutes and stored on ice. 7 µL of the reaction was transferred to another tube as a -reverse transcriptase (-RT) control, without RT enzyme. -RT control was used to check for gDNA contamination. The reverse-transcription reactions were prepared on ice as given in Table 2.7 and the template RNA from the gDNA elimination reaction (Table C.2) was added to the transcription master mix to synthesize cDNA and -RT control. The reactions were incubated at 42 °C for 15 minutes, followed by deactivation of RT by incubating the reactions at 95 °C for 3 minutes. The cDNA was diluted 1:10 before being stored on ice.

Table 2.7: A general recipe for reverse-transcription reaction (QIAGEN, 2016*a*). Instead of Quantiscript Reverse Transcriptase, RNAse free water was used for the -RT control reactions. To synthesize cDNA and -RT control, template RNA from the gDNA elimination reaction (Table C.2) must be added to the reverse transcription master mix.

Components	**Volume [µL]**	
	for cDNA mix	for -RT mix
Reverse transctiption master mix		
Quantiscript Reverse Transcriptase	1	-
Quantiscript RT Buffer, 5x	4	2
RT Primer Mix	1	0.5
RNAse free water	-	0.5
Template RNA		
gDNA elimination reaction	14	7
Total volume	20	10

2.6.3 Quantitative real-time PCR (qRT-PCR)

Quantitative real-time PCR (qRT-PCR) is a method where RT is followed by PCR to detect and quantify gene expression. The fluorescence dye that binds to the cDNA measures the increasing intensity of the fluorescent signal that shows the amount of dsDNA amplicons. The fluorescence intensity increases with the amount of dsDNA that is synthesized with each PCR cycle, the amount of DNA produced (Kralik and Ricchi, 2017). The point where the fluorescence signal exceeds the background signal and is detectable, is called the cycle threshold (Ct). Ct value makes it possible to represent the relative gene expression. A higher Ct value means that there was a lower initial concentration of mRNA. Inversely, a lower Ct value means that the gene is higher expressed (Godbey, 2022). By studying gene expression, it is possible to compare transgene expression levels from different genetic transformation methods. Using target specific primers, the expression of the target sequence can be examined (Keer, 2008).

To ensure that the reaction components are functional and that the efficiency of the assay is acceptable, *LHCF9* target-specific primers were used as a positive control. In addition, target-specific primers for the *sgRNA*, *Cas9* and nourseothricin acetyltransferase (*NAT*) target sequences were used to analyse gene expression. In addition, non-template controls (NTCs) and -RT controls were used to detect gDNA contamination. 15 µL of the mastermix made with LightCycler® 480 SYBR Green I Master (Table 2.8) with target spesific primers (Table D.1, Appendix D) were added to 96 well plates (LightCycler® 480, Roche). The diluted cDNA samples (5 µL) were used as template. The 96 well plates (LightCycler® 480, Roche) were sealed with sealing foil (LightCycler® 480, Roche) and the qRT-PCR was performed by using LightCycler® 96 instrument (Roche) with the thermocycling profile given in Table 2.9. The raw data was analysed with LightCycler® 96 software 1.1 (Roche).

Table 2.8: Reaction setup (15 µL) for qRT-PCR with LightCycler® 480 SYBR Green I Master (Roche).

Components	**Volume [μL]**
H_2O, PCR-grade	3
PCR primers, 10 x concentration of each	2
LightCycler 480 Probes Master, 2x concentration	10
Total volume	15

Table 2.9: Thermocycling profile for qRT-PCR with with LightCycler® 480 SYBR Green I Master (Roche).

Program	**Temperature** (°C)	**Time (s)**	**Ramp** (°C s^{-1})	**Cycles**
Pre incubation	95	600	4.4	1
3 step amplification	95	10	4.4	45
	55	10	2.2	
	72	10	4.4	
Melting	95	5	4.4	1
	65	60	2.2	
	97	1	0.2	
Cooling	40	10	1.5	1

2.7 Assembly of pBKS-hCas9M-Bsr/Nat/ptPDS1-M1

Studies of alternative selection markers with plasmids suitable for conjugation have previously been performed. In order to make the plasmids smaller as well as better suited for electroporation, pBKS-hCas9M with selection markers bsr (Figure 2.2), nat (Figure 2.3) and ptPDS1-M1 (Figure 2.4) were assembled in a multi-step process.

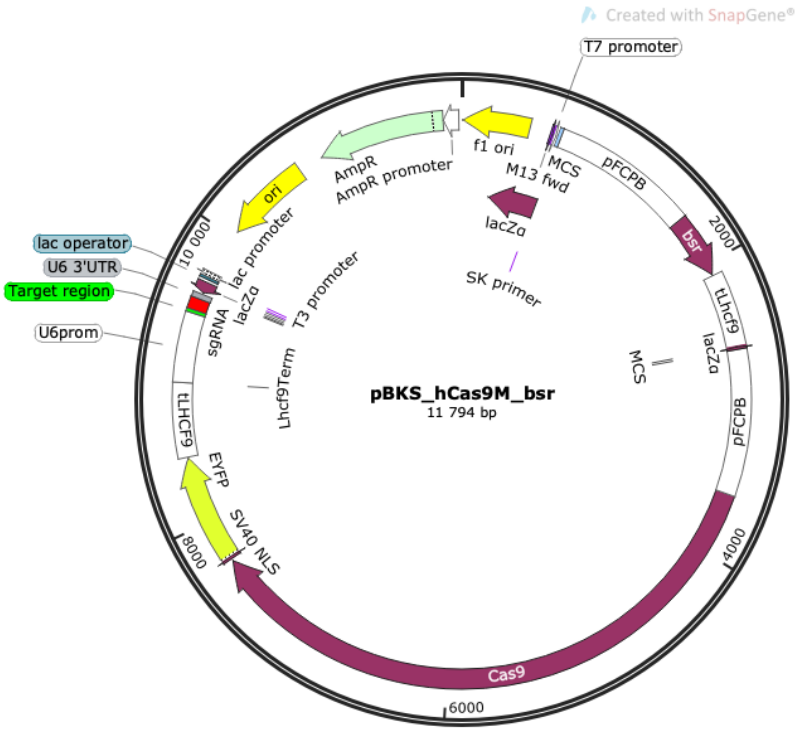


Figure 2.2: Map of vector plasmid pBKS_hCas9M_bsr. The pBKS plasmid is assembled to contain the blasticidine resistance (bsr) cassette for resistance in *T. pseudonana*, ampicillin resistance (AmpR) cassette for resistance in *E. coli*, the Cas9 fragment controlled by pFCPB, with SV40 NLS and a C terminal YFP tag, and a sgRNA cassette controlled by a U6 promoter. The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

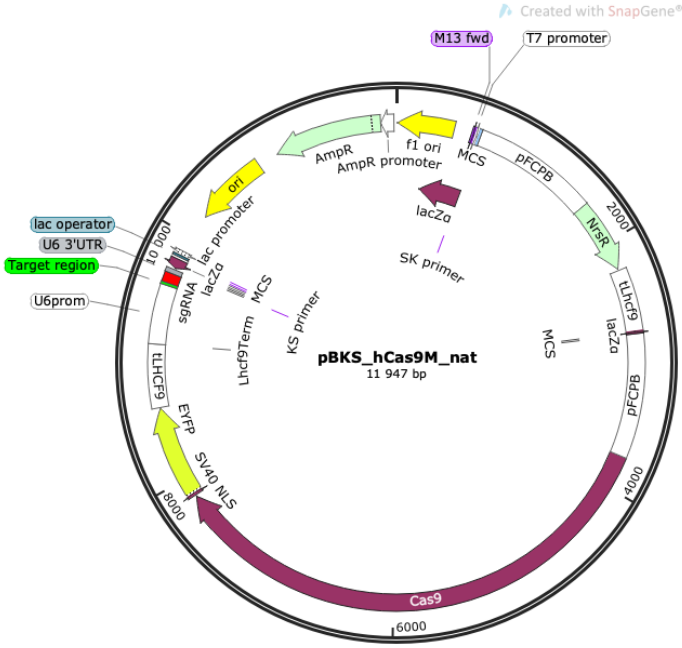


Figure 2.3: Map of vector plasmid pBKS_hCas9M_nat. The pBKS plasmid is assembled to contain the nourseothricin resistance (nat) cassette for resistance in *T. pseudonana*, ampicillin resistance (AmpR) cassette for resistance in *E. coli*, the Cas9 fragment controlled by pFCPB, with SV40 NLS and a C terminal YFP tag, and a sgRNA cassette controlled by a U6 promoter. The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

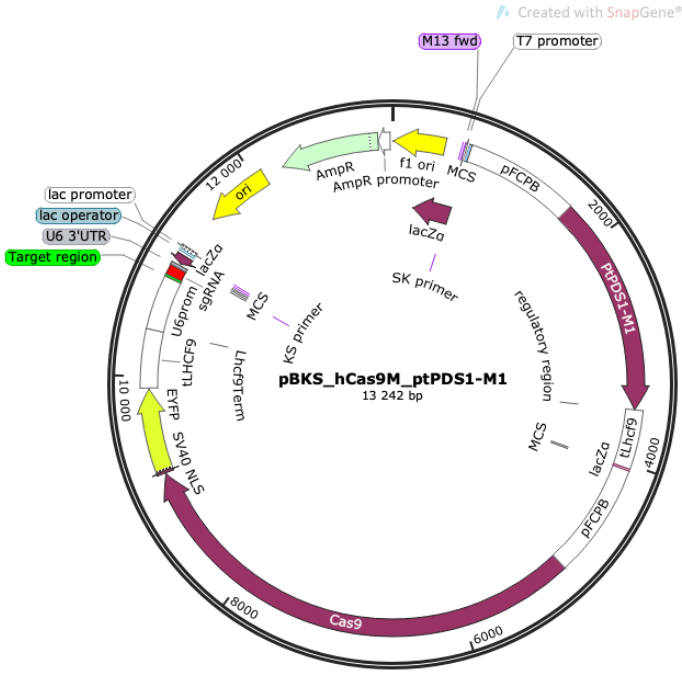


Figure 2.4: Map of vector plasmid pBKS_hCas9M_ptPDS1-M1. The pBKS plasmid is assembled to contain the norflurazon resistance (ptPDS1-M1) cassette for resistance in *T. pseudonana*, ampicillin resistance (AmpR) cassette for resistance in *E. coli*, the Cas9 fragment controlled by pFCPB, with SV40 NLS and a C terminal YFP tag, and a sgRNA cassette controlled by a U6 promoter. The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

All fragments used to assemble the plasmids were amplified by PCR using Phusion Plus DNA Polymerase (Thermo Scientific), as given in Table 2.3, with specific primers (Table D.1, Appendix D). Thermocycling profile for PCR amplification with Phusion Plus DNA Polymerase (Thermo Scientific) is given in Table 2.4, where the extension time varies depending on the fragment size (Table 2.10) to be amplified. To confirm the size of the PCR amplicons, gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1h) was performed. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths, and 6x Loading Dye (Thermo Scientific) was used for staining the DNA.

Table 2.10: PCR products for assembly of pBKS-hCas9M-Bsr/Nat/ptPDS1-M1

PCR product	**Fragment size [bp]**	**Template**
pBKS_BsaImut	2958	pBKS
Bsr fragment	1958	Blastisidine - pTpPUC-Bsr
Nat fragment	2105	Nourseothricin - pTpPUC3_hCas9MsgMut
ptPDS1-M1 fragment	3400	Norflurazone - pTpPUC3PtPDS1
hCas9M	7010	pTpPUC3_hCas9MsgMut

In order to assemble pBKS-hCas9M with selection markers, bsr (Figure 2.2), nat (Figure 2.3) and ptPDS1-M1 (Figure 2.4), it was necessary to carry out a site-directed mutagenesis of the pBKS plasmid. The plasmid was previously heat-shock transformed and extracted using E.Z.N.A® Plasmid DNA Mini Kit II (Omega Bio-tek) by Tore Brembu at the ACMS lab. A mutagenesis PCR was performed and the pBKS_BsaImut PCR product was cut overnight with the DpnI (New England Biolabs, Table C.1, Appendix C) restriction enzyme (37 °C). To transform pBKS_BsaImut into DH5α competent *E. coli* as given in Subsection 2.3.1, the PCR product was purified with the Monarch PCR & DNA Cleanup Kit (5 µg) (New England BioLabs Inc). The transformants were screened using BsaI-HFv2 and DraIII-HF (New England Biolabs, Table C.1, Appendix C) restriction digest (45 min at 37°C) and gel electrophoresis. If the gel looked correct, the samples were submitted for sequencing before the next fragments were inserted.

The selection markers bsr, nat and ptPDS1-M1 were inserted into pBKS_BsaImut using Gibson Assembly® Master Mix (New England Biolabs). pBKS_BsaImut was linearized with XbaI and PstI (New England Biolabs, Table C.1, Appendix C) for 1h at 37 °C and the selection marker fragments (Table 2.10) were amplified with specific primers by PCR. The products were purified with the Monarch PCR & DNA Cleanup Kit (5 µg) (New England BioLabs Inc). Since the linearized product was to be purified, heat inactivation was not necessary. However, complete digestion was assumed. Gibson Assembly reaction mixes (20 µL) were prepared for each of the selection markers with Gibson Assembly Master Mix (2X) (10 µL), 100 ng vector and fragment < 200 bp in 2:1 molar ratio. The reaction mix was incubated at 50 °C for 15 minutes and 2 µL was transformed into DH5α competent *E. coli.* To confirm that

the colonies picked for overnight culture were positive clones, a PCR colony screening was performed. PCR colony screening was done by swirling the picked colony in water, nuclease free (10 µL) before dropping it into LB medium. Water, nuclease free with the colony was used as a template for the PCR reaction given in Table 2.11, with specific primers (Table D.1, Appendix D). The thermocycle profile is given in Table 2.12. Gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1 h) was performed to check the PCR products for potential indels. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths, and 6x Loading Dye (Thermo Scientific) was used for staining the DNA. Since insert-specific primers are used, positive clones will give bands, while negative clones will not. The positive clones were extracted with a mini-prep with E.Z.N.A® Plasmid DNA Mini Kit II (Omega Bio-tek). The positive clones were submitted for Sanger sequencing with specific primers (Table D.1, Appendix D) performed by Eurofins (Lightrun). Testcutting was performed by restriction enzyme digestion with SacII + XhoI (New England Biolabs, Table C.1, Appendix C), followed by a gel electrophoresis.

Table 2.11: Reaction setup (25 µL) for PCR colony screening with VWR Red Taq DNA Polymerase Master Mix.

Components	**Volume [µL]**	**Final concentration**
Taq 2x Master Mix	12.5	1x
Forward primer	1	0.2 µM
Reverse primer	1	0.2 µM
Template	0.5	-
Water, nuclease free	10	-
Total volume	25	-

Table 2.12: Thermocycling profile for PCR colony screening with VWR Red Taq DNA Polymerase Master Mix.

Program	**Temperature** (°C)	**Time**	**Cycles**
Initial denaturation	98	5 min	1
Denaturation	95	30 sec	
Annealing	53	30 sec	34
Extension	72	10s	
Final extension	72	5 min	1
	4	∞	

The same procedure was used to insert the Cas9 fragment (Table 2.10). pBKS with bsr, nat and ptPDS1-M1 fragments were linearized with HindIII and SaI. The Cas9 fragment was PCR amplified with specific primers for each selection marker (Table D.1, Appendix D) and was inserted to pBKS_bsr, pBKS_nat and pBKS_ptPDS1-M using Gibson Assembly® Master Mix

(New England Biolabs). The reaction mix were transformed into DH5α competent *E. coli* and a PCR colony screening was performed to confirm the positive clones.

2.7.1 Study of pBKS with selection markers

To study selection markers for *T. pseudonana*, the plasmids with different resistance genes (Table 2.13) made in Subsection 2.7 were used.

Table 2.13: Plasmids with different resistance genes to study selection markers for *T. pseudonana*.

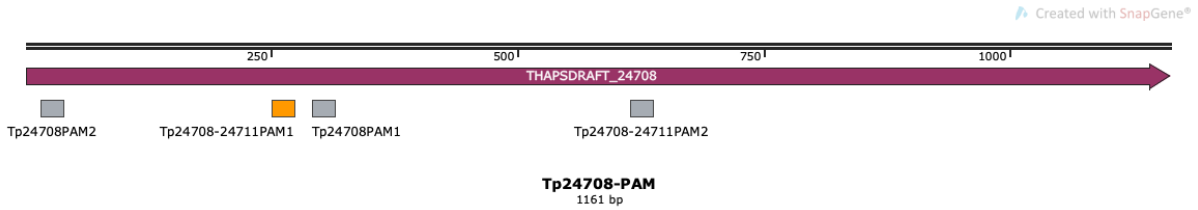
Plasmid	**Length [bp]**	**Resistence**	**Plasmid map**
pBKS_Bsr	4850	Blasticidin	Figure A.3
pBKS_Nat	4997	Noursethricin	Figure A.4
pBKS_PtPDS1-M1	6292	Norflurazon	Figure A.5

Electroporation (Subsection 2.3.3) was performed for each selection marker plasmid. The plasmids were linearized with XbaI (New England Biolabs, Table C.1, Appendix C) for 1h at 37 °C and heat inactivated at 65 °C for 15 minutes. The electroporated cells were plated out on selective plates with different concentrations (Bsr: 5, 7.5 and 10 µM, Nou: 100 and 150 µg/mL, and Nrf: 1.5, 3 and 5 µM).

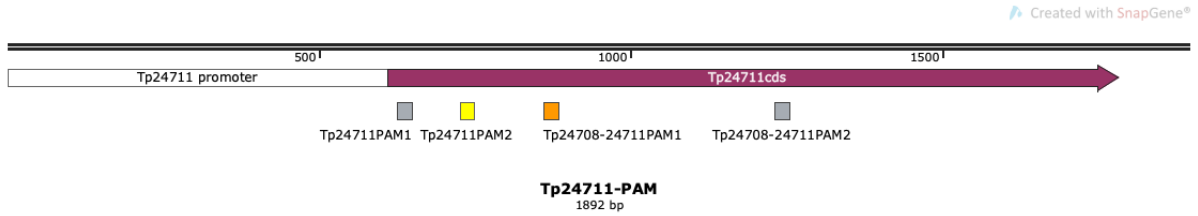
3 Results

3.1 Creating knockout mutants of Tp24708 and Tp24711

To create knockout mutants of Tp24708 and Tp24711, a vector containing Cas9 nuclease and sgRNA with specific sequences for targeting the genes was created and transferred to *T. pseudonana*. The pTpPUC3 hCas9 Sg mut 24711P2 (Tp24711P2) and pTpPUC3 hCas9 Sg mut Tp24708/24711P1 (Tp24708/24711P1) were constructed by Annika Messemer at the ACMS lab. The PAM sites in Tp24708 and Tp24711 are given as coding DNA sequences (cds) in Figure 3.1a and Figure 3.1b, respectively. Tp24711P2 is a single PAM site in the Tp24711 cds and is used to generate single knockout mutations, while Tp24708/24711P1 is a double PAM site, which is identical in both genes and allows knockout of both genes by targeting same region.



(a) The coding DNA sequence of Tp24708, where the location of the PAM sites (marked as boxes) in the coding region (marked in purple) is given. The Tp24708/24711PAM1 site is given as a orange box, and was used to generating double knockouts in *T. pseudonana*. Image is retrieved from SnapGene (Dotmatics, 2021).



(b) The promoter and coding DNA sequence of Tp24711, where the location of the PAM sites (marked as boxes) in the coding region (marked in purple) is given. The Tp24711PAM2 site is given as a orange box, and was used generating single knockout in *T. pseudonana*, while Tp24708/24711PAM1 site is given as a orange box, and was used to generating double knockouts in *T. pseudonana*. Image is retrieved from SnapGene (Dotmatics, 2021).

Figure 3.1: The coding DNA sequence of Tp24708 and Tp24711

3.1.1 Transformation of *T.pseudonana*

In the specialization project, a biolistic shooting of the already constructed plasmids was carried out (Mathivannan, 2021). In order to obtain sufficient yield of the plasmids, a bacterial transformation of the plasmids were performed, followed by a control restriction enzyme digestion and gel electrophoresis to confirm the size of the plasmids (Figure E.1 in Appendix E). The plasmids looked correct on the gel and had sufficient yield and thus linearized plasmids were transformed into *T. pseudonana* by electroporation, in an attempt to create stable knockout mutations.

The transformation was considered successful when colonies appeared on the plates. Another electroporation was also performed by Tore Brembu at ACMS lab, which was also included in the HRM screening process. Nourseothricin-resistant colonies picked from the bi-

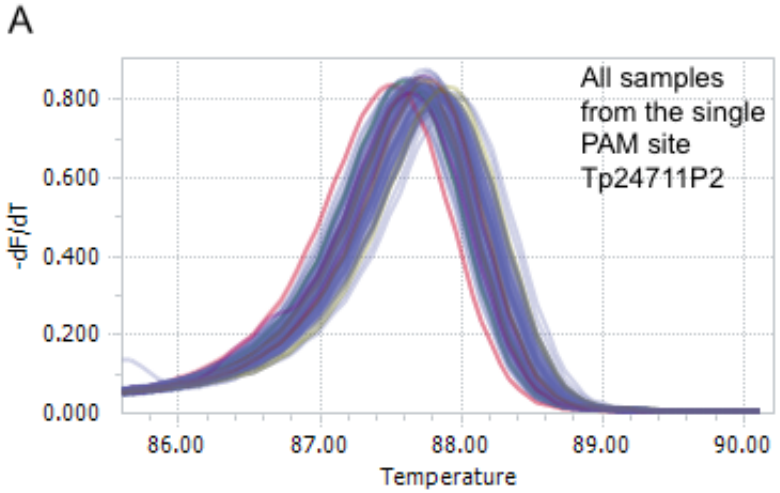
olistic shooting plates (from the specialization project) and from the electroporation plates, were grown in liquid medium with nourseothricin for selection. There was growth in a total of 31 clones of Tp24708/24711P1 and 49 clones of Tp24711P2. The clones were further lysed prior to HRM screening for CRISPR/Cas9-induced mutations.

3.1.2 Screening for CRISPR/Cas9 induced mutations by HRM

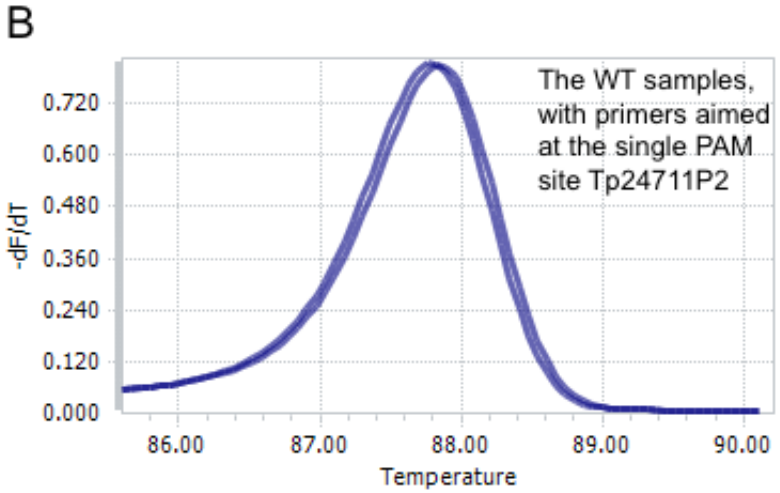
The lysed clones and WT were amplified for larger indels by PCR, followed by gel electrophoresis. The PCR products showed indels that corresponded to the WT, by gel electrophoresis (Figure E.2, Figure E.3 and Figure E.4 in Appendix E), but were still further diluted and used for HRM screening to potentially detect minor mutations in a small indel around the target region. In a gel electrophoresis, minor mutations will not appear, as the resolution is not good enough. Hence gel electrophoresis will only separate fragments with larger differences in length. With HRM, on the other hand, it is possible to detect smaller mutations, as there is a smaller indel around the target site that is amplified and analyzed.

An analysis of the melting curve for all samples from the single PAM site Tp24711P2 (Figure 3.2a) shows that the samples deviate little from the WT sample (Figure 3.2b). The samples that deviated the most were selected and submitted for Sanger sequencing for detection of any possible mutations along with the WT sample (Figure 3.2c). For the double PAM site Tp24708/24711P1, it was necessary to examine both the cds of Tp24708 and Tp24711. The melting curve of all the samples (Figure 3.3a), with primers directed towards Tp24708, shows slightly larger deviations from the WT sample (Figure 3.3b), but since some of the parallels deviate from each other as well, several samples (Figure 3.3c) were submitted for Sanger sequencing. The melting curve of all the samples (Figure 3.4a), with primers directed towards Tp24711, shows a slight deviation form the WT sample (Figure 3.4b). Since it was desirable that the mutation takes place in both the cds of Tp24708 and Tp24711, several of the same samples were submitted with both primers directed for the cds of Tp24708 and Tp24711, even if it was not possible to observe clear differences between WT and the same samples in the melting curve. The samples submitted to Sanger sequencing, where there are visible deviations from WT (marked in dark blue) for Tp24711P2 and Tp24708/24711P1 are given in Figure 3.2c, Figure 3.3c and Figure 3.4c.

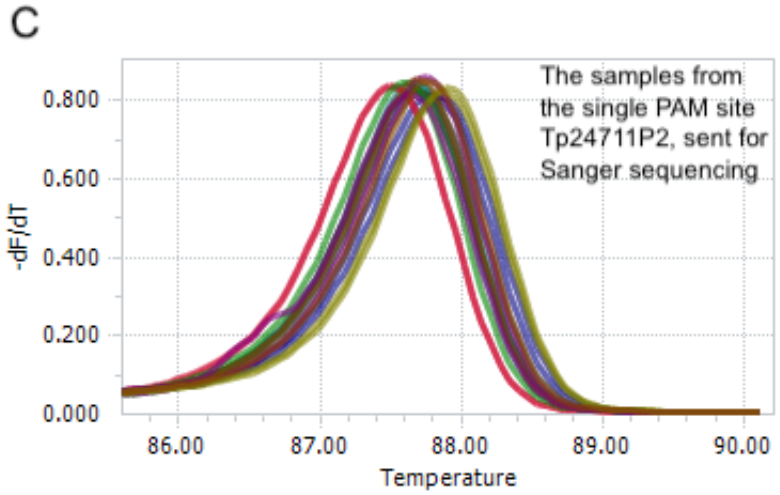
By analyzing the Sanger sequencing results in the software Snapgene (Dotmatics, 2021), followed by ICE (Synthego, 2022), it was confirmed that no mutations had occurred in any of the Tp24711P2 and Tp24708/24711P1 clones. Thus, no further analyses were performed.



(a) Melting curves from HRM screening of all the samples from the single PAM site Tp24711P2, where the WT samples are marked in dark blue.

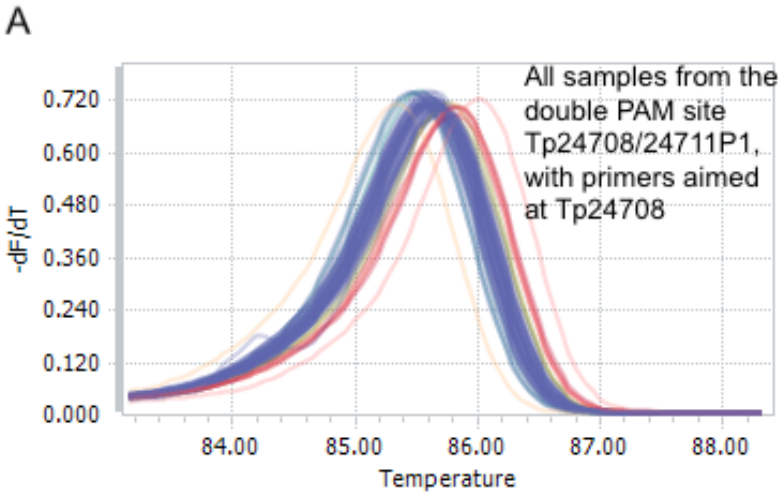


(b) Melting curves from HRM screening of the WT samples, with primers aimed at the single PAM site Tp24711P2.

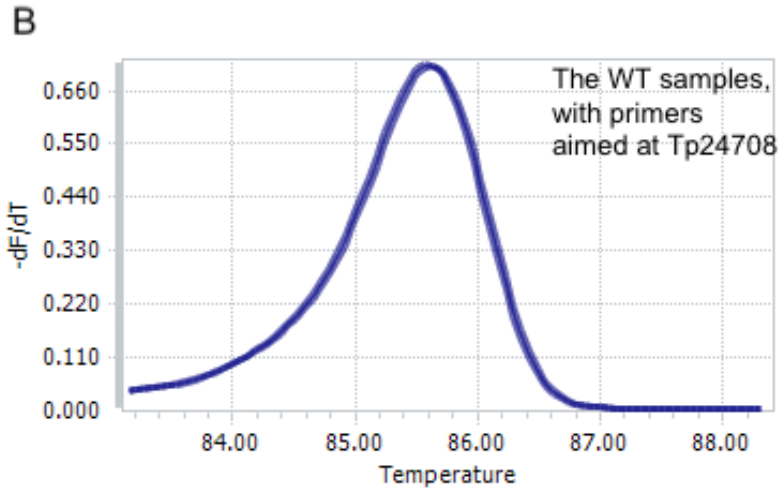


(c) Melting curves from HRM screening of the samples from the single PAM site Tp24711P2, submitted for Sanger sequencing. The WT samples are marked in dark blue. The samples that deviated the most from the WT samples were selected.

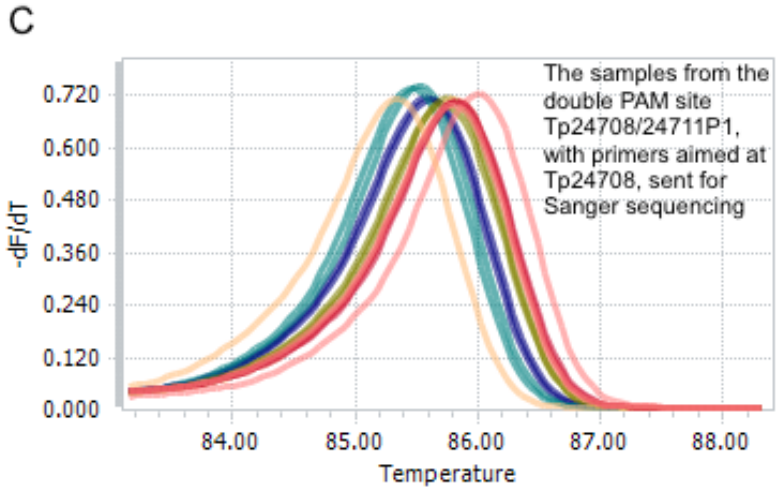
Figure 3.2: Melting curves of the single PAM site Tp24711P2, from HRM screening, where A) are the melting curves of all the Tp24711P2 samples, B) are the melting curves of the WT samples, and C) are the melting curves of the samples submitted for Sanger sequencing. Images are retrieved from LightCycler® 96 software 1.1 (Roche).



(a) Melting curves from HRM screening of all the samples from the double PAM site Tp24708/24711P2, with primers aimed at the cds of Tp24708. The WT samples are marked in dark blue.

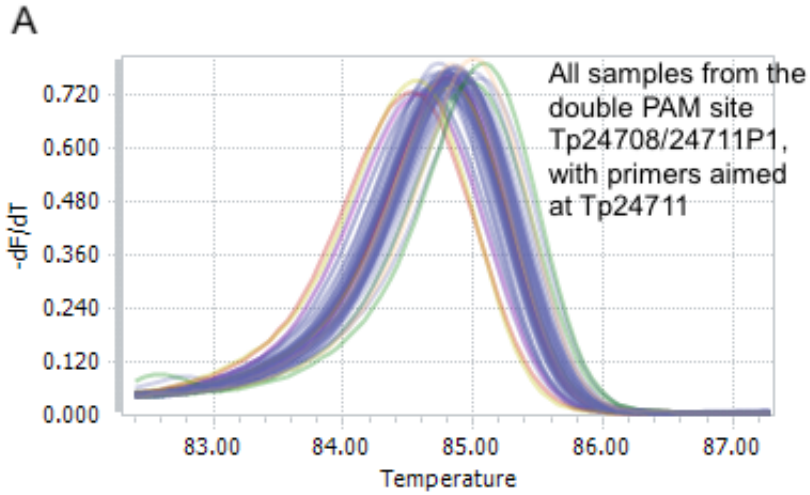


(b) Melting curves from HRM screening of the WT samples, with primers aimed at the cds of Tp24708 in the double PAM site Tp24708/24711P2.

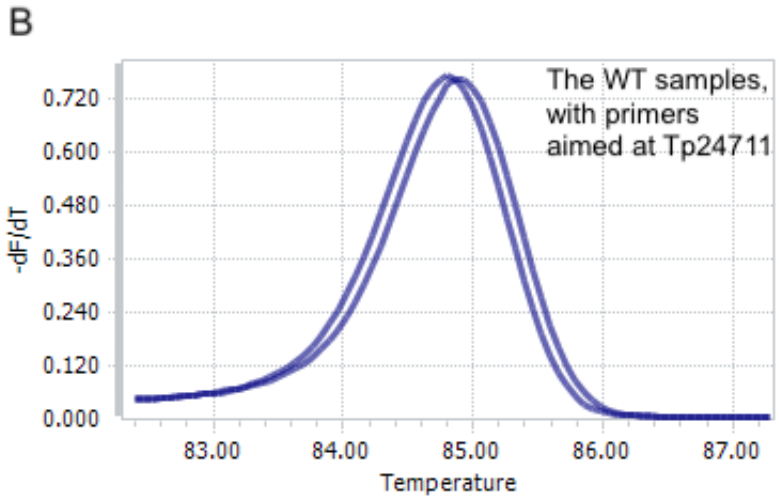


(c) Melting curves from HRM screening of the samples, with primers aimed at the cds of Tp24708 in the double PAM site Tp24708/24711P2, submitted for Sanger sequencing. The WT samples are marked in dark blue. The samples that deviated the most from the WT samples were selected.

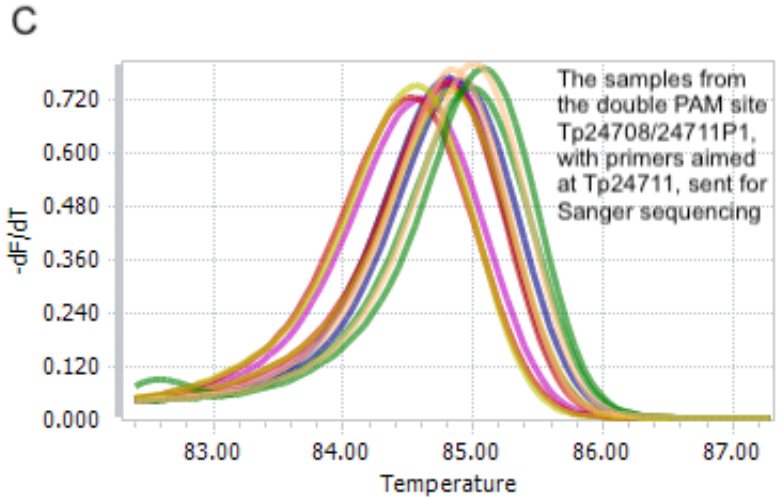
Figure 3.3: Melting curves of Tp24708/24711P1, with primers aimed for the cds of Tp24708, from HRM screening, where A) are the melting curves of all the Tp24708/24711P1 samples, with primers aimed for the cds of Tp24708, B) are the melting curves of the WT samples, and C) are the melting curves of the samples submitted for Sanger sequencing. Images are retrieved from LightCycler® 96 software 1.1 (Roche).



(a) Melting curves from HRM screening of all the samples from the double PAM site Tp24708/24711P2, with primers aimed at the cds of Tp24711. The WT samples are marked in dark blue.



(b) Melting curves from HRM screening of the WT samples, with primers aimed at the cds of Tp24711 in the double PAM site Tp24708/24711P2.



(c) Melting curves from HRM screening of the samples, with primers aimed at the cds of Tp24711 in the double PAM site Tp24708/24711P2, submitted for Sanger sequencing. The WT samples are marked in dark blue. The samples that deviated the most from the WT samples were selected.

Figure 3.4: Melting curves of Tp24708/24711P1, with primers aimed for the cds of Tp24711, from HRM screening, where A) are the melting curves of all the Tp24708/24711P1 samples, with primers aimed for the cds of Tp24711, B) are the melting curves of the WT samples, and C) are the melting curves of the samples submitted for Sanger sequencing. Images are retrieved from LightCycler® 96 software 1.1 (Roche).

3.2 Localization of pTpPUC3-mNeonGreen-Tp24711 in *T. pseudonana*

In order to study the location of mNeonGreen tagged Tp24711 in *T. pseudonana*, the previously constructed plasmid pTpPUC3-mNeonGreen-Tp24711 (Gresseth, 2020) was transformed into *T. pseudonana* by electroporation. Nourseothricin-resistant colonies picked from the electroporation plates, were grown in liquid medium with nourseothricin for selection. There was growth in a total of six clones, which were studied in flow cytometer, fluorescence microscope, and confocal microscope.

3.2.1 Flow cytometer

All transformed clones were studied in a Attune™ NxT Flow Cytometer to study which deviated the most from the WT sample, and expressed the highest fluorescence. The clones were then measured with BL1 and BL3 detectors, where it was found out which clones were the most fluorescent. BL3 detector detects fluorescence in a range that includes mNeonGreen. Living cells were gated with BL3 detector, which detects autofluorescence from chloroplasts, which were absent in dead cells. The clones and the mean fluorescence are given in the Figure 3.5, and as shown in the Table 3.1, there was a high fraction of living cells in all clones. Based on the highest mean fluorescence expressed (Figure 3.5, highlighted in orange) and the most deviating from the overlays in the Figure 3.6, mNeonGreen-Tp24711-s2 and mNeonGreen-Tp24711-s6 were used for further study of the localization of mNeonGreen-Tp24711 fusion protein in *T. pseudonana* using fluorescence microscope and confocal microscope.

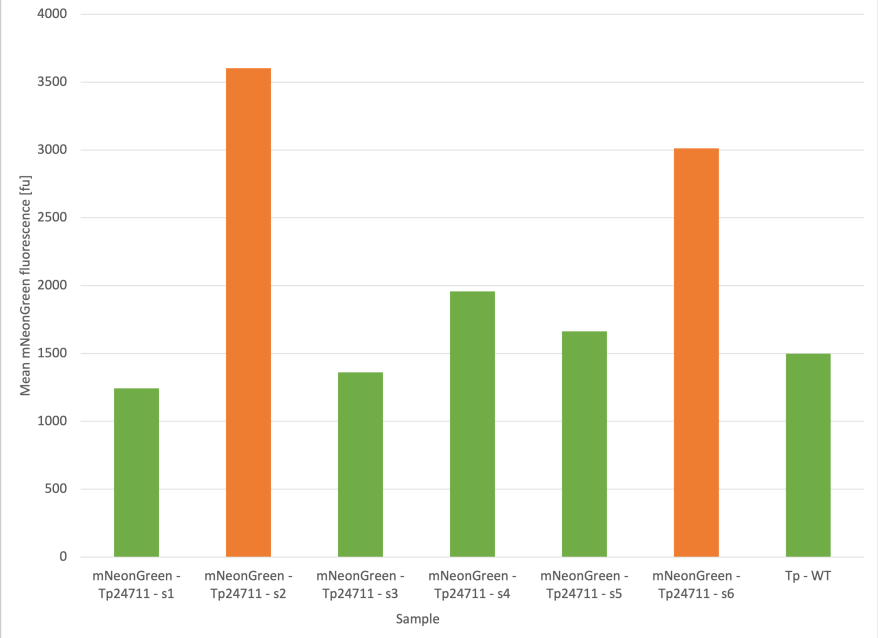


Figure 3.5: The mean fluorescence of mNeonGreen tagged Tp24711 in *T. pseudonana* clones measured with BL1 (530/30 nm) in a flow cytometer. A WT sample was also measured for comparison. mNeonGreen-Tp24711-s2 and mNeonGreen-Tp24711-s6 (highlighted in orange) expressed the the most mNeonGreen fluorescence and were used for further study of the ocalization of pTpPUC3-mNeonGreen-Tp24711 in *T. pseudonana* using fluorescence microscope, and confocal microscope. The mean mNeonGreen fluorescense [fu] was measured with Attune™ NxT Flow Cytometer.

Table 3.1: Cell density [cells/µL] and fraction of living [%] gated with BL3 (695/40 nm) for transformed mNeonGreen - Tp24711 clones and WT for comparison. Cell density and fraction are measured with Attune™ NxT Flow Cytometer.

Sample	Cell density [cells/µL]	Fraction of living cells [%]
mNeonGreen - Tp24711 - s1	1343.78	99
mNeonGreen-Tp24711-s2	682.05	96
mNeonGreen - Tp24711 - s3	638.66	96
mNeonGreen - Tp24711 - s4	800.96	95
mNeonGreen - Tp24711 - s5	622.59	97
mNeonGreen-Tp24711-s6	780.96	98
Tp - WT	323.44	93

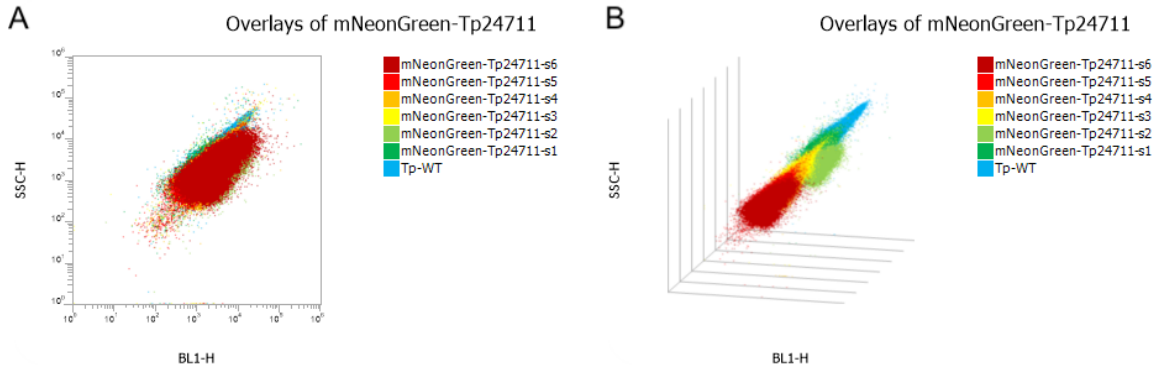


Figure 3.6: Overlay of all transformed mNeonGreen - Tp24711 clones and WT for comparison, where SSC used to measure the structural complexity of the cells is plotted against mNeonGreen fluorescence measured with BL1 ((530/30 nm)). The figures show how Tp - WT differs from the other samples and that mNeonGreen-Tp24711-s2 (marked with light green) and mNeonGreen-Tp24711-s6 (marked with dark red) differ most from Tp - WT (marked with blue). A) shows the overlays of mNeonGreen tagged Tp24711 in *T. pseudonana*, while B) is remade in 3D to get a better view. The figure was made with Attune™ NxT Software.

3.2.2 Fluorescence microscopy

mNeonGreen-Tp24711-s2 and mNeonGreen-Tp24711-s6, which showed highest fluorescence signal in the flow cytometry screen were further studied with Axio Imager.Z2 Fluorescence microscope. All cells were imaged with ZEISS ZEN Software, where the fluorescence from both red and green spectra was detected. The red spectrum fluorescence illustrates the autofluorescence emitted from the chloroplasts, while the green spectrum fluorescence illustrates fluorescence emitted from mNeongreen and autofluorescence. In addition, the cells were imaged with brightfield and all images were overlapped in order to compare the images. Additional images confirming the observations described are given in Appendix F.

The image obtained from fluorescence microscopy of mNeonGreen-Tp24711-s2 cells is given in Figure 3.7. In a fluorescence microscopy image, it will be possible to study a larger plane at the same time. In Figure 3.7, a closer look has been taken, where it is possible to

observe a small overlap (yellow). The mNeonGreen fluorescence perfectly overlaps with the autofluorescence in a small area, but also separated from the autofluorescence. There is some noise from the other planes that have not been studied, which makes the image a bit blurry. However, it is possible to observe green fluorescence dispersed throughout the cytosol, and it may appear that Tp24711 is present in SDVs, as there is higher mNeonGreen expressed between the autofluorescence. In Figure F.1 from the Appendix F, another image of mNeonGreen-Tp24711-s2 cells is given, where it can also look like there is little mNeonGreen that is expressed separately from the autofluorescence. In some places it is also possible to observe only green fluorescence in the cells, which means that there are also dead cells in the sample.

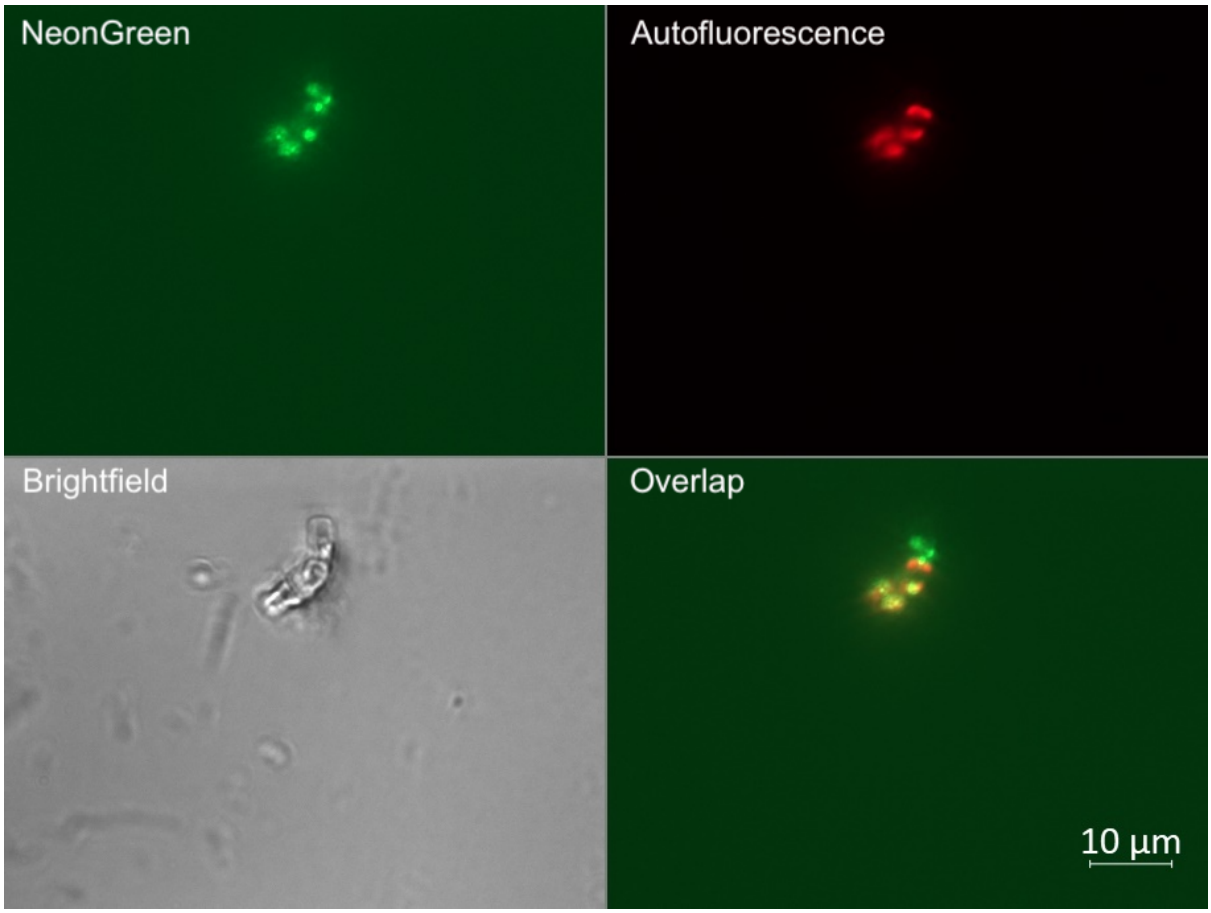


Figure 3.7: Fluorescence microscope image of mNeonGreen-Tp24711-s2 expressed in *T. pseudonana*. Upper left: PaGfp channel (505-555 nm), to detect mNeonGreen fluorescence (green), upper right: Cy5 channel (665-715 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a brightfield channel and lower right: overlap of both fluorescence channels. The scale bar is 10 µm. The image is retrieved from ZEISS ZEN Software.

In Figure 3.8 of mNeonGreen-Tp24711-s6 cells, it is possible to observe that mNeonGreen fluorescence is separated from the autofluorescence in some parts of the cells. As it appears that several of the cells are close to each other, it is difficult to say which part of the frustule mNeonGreen is expressed in. In Figure F.2, from Appendix F, another image of mNeonGreen-Tp24711-s6 cells is given, where it can be seen that mNeonGreen is expressed

as a ring in the middle of some of the cells. This ring with mNeonGreen can be a division plan when the cells divide. A weak division plane can also be observed in some cells in Figure 3.8. In addition, it may appear that mNeongreen is dispersed throughout the cytosol in most cells.

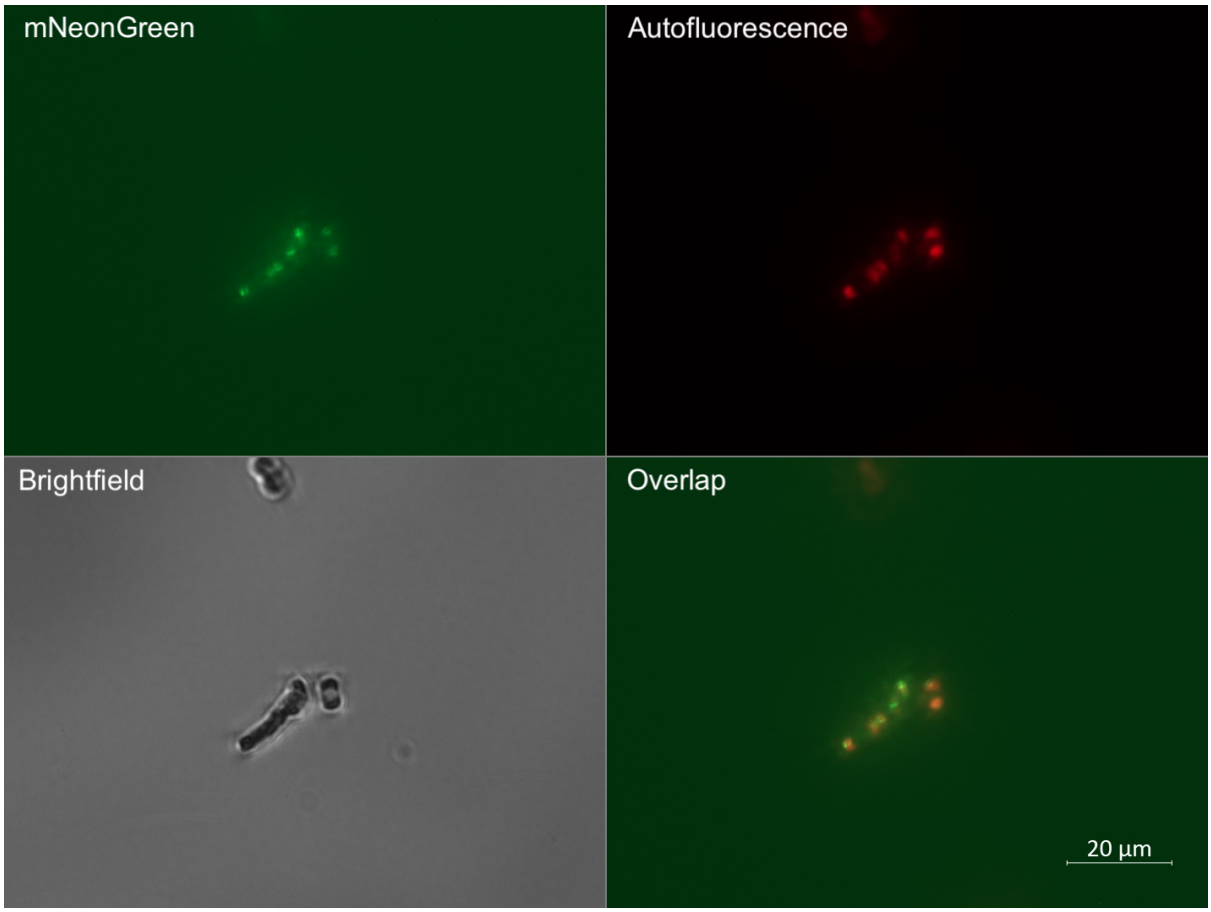


Figure 3.8: Fluorescence microscope image of mNeonGreen-Tp24711-s6 expressed in *T. pseudonana*. Upper left: PaGfp channel (505-555 nm), to detect mNeonGreen fluorescence (green), upper right: Cy5 channel (665-715 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a brightfield channel and lower right: overlap of both fluorescence channels. The scale bar is 20 µm. The image is retrieved from ZEISS ZEN Software.

3.2.3 Confocal microscopy

For a closer look at the cells in the mNeonGreen Tp24711-s2 and mNeonGreen-Tp24711-s6, they were studied in the Leica SP8 Confocal Microscope. All cells were imaged with Leica LAS AF, and as with the fluorescence microscope, both the red and green spectra were used. PMT1 channel was used to detect mNeonGreen (493-540 nm), PMT3 channel was used to detect autofluorescence (651-779 nm) emitted from chloroplast and PMT channel was used for brightfield. In addition, an image was created with an overlap of all the channels. Additional images confirming the observations described are given in Appendix F.

The image obtained from confocal microscopy of mNeonGreen-Tp24711-s2 cells is given in Figure 3.9. Confocal and fluorescence microscope images (Figure 3.7) show similar distribution, where mNeonGreen is expressed separately from the autofluorescence. The image is

composed of two cells close together, where mNeonGreen is expressed dispersed in the cytosol, but also in girdle band SDVs. Higher expression of mNeonGreen in girdle band SDVs can also be observed in Figure F.3 and Figure F.4.

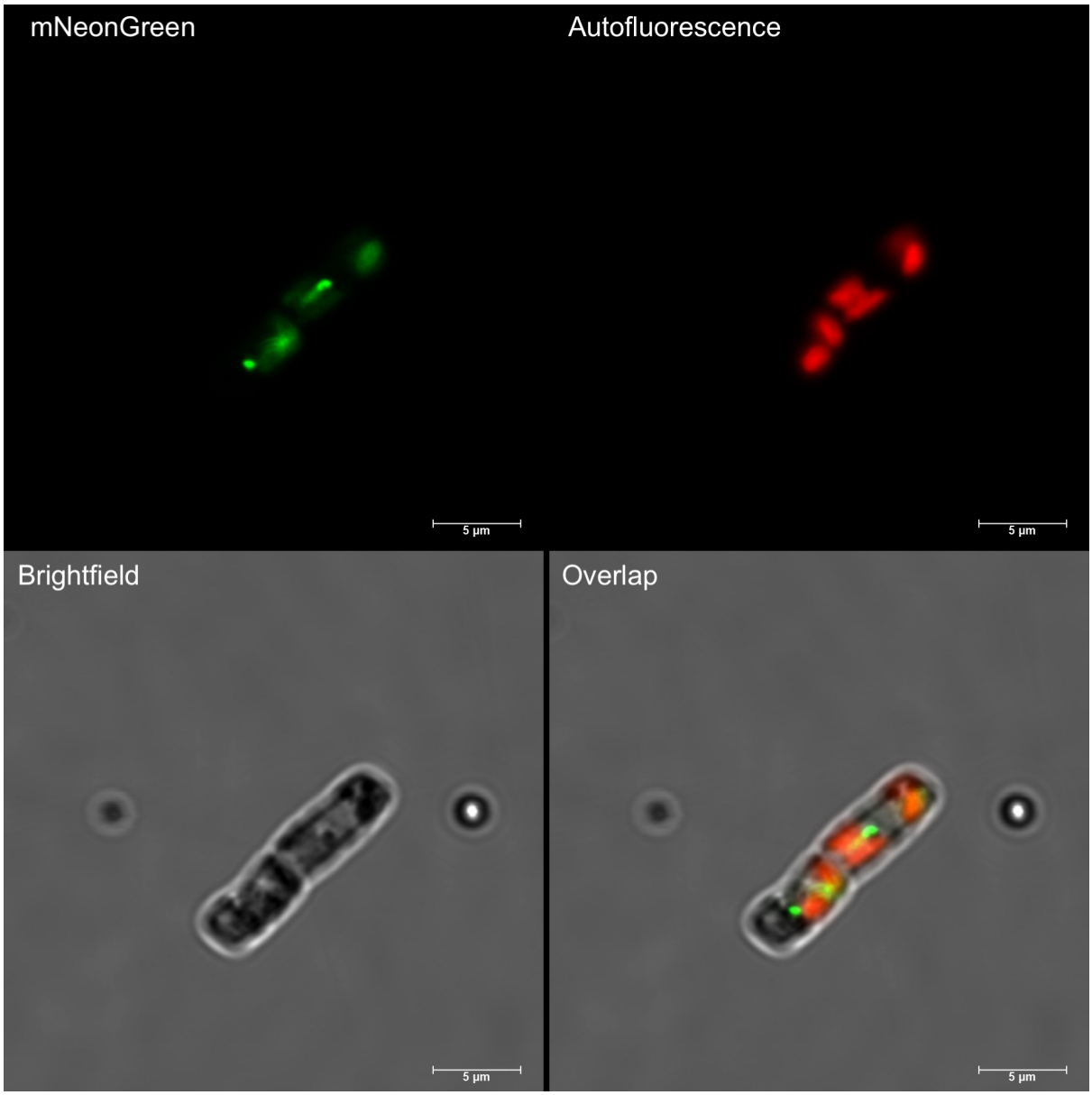


Figure 3.9: Confocal microscope image (512x512) of mNeonGreen-Tp24711-s2 expressed in *T. pseudonana*. Upper left: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), upper right: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a PMT brightfield channel and lower right: overlap of all channels. The scale bar is 5 µm. The image is retrieved from Leica LAS AF Software.

Figure 3.10 is another image of the transgenic clone mNeonGreen-Tp24711-s2, which expresses mNeonGreen-Tp24711 in *T. pseudonana* provided. It is possible to observe that there is a clear overlap between mNeonGreen and the autofluorescence, where mNeonGreen is highest expressed.

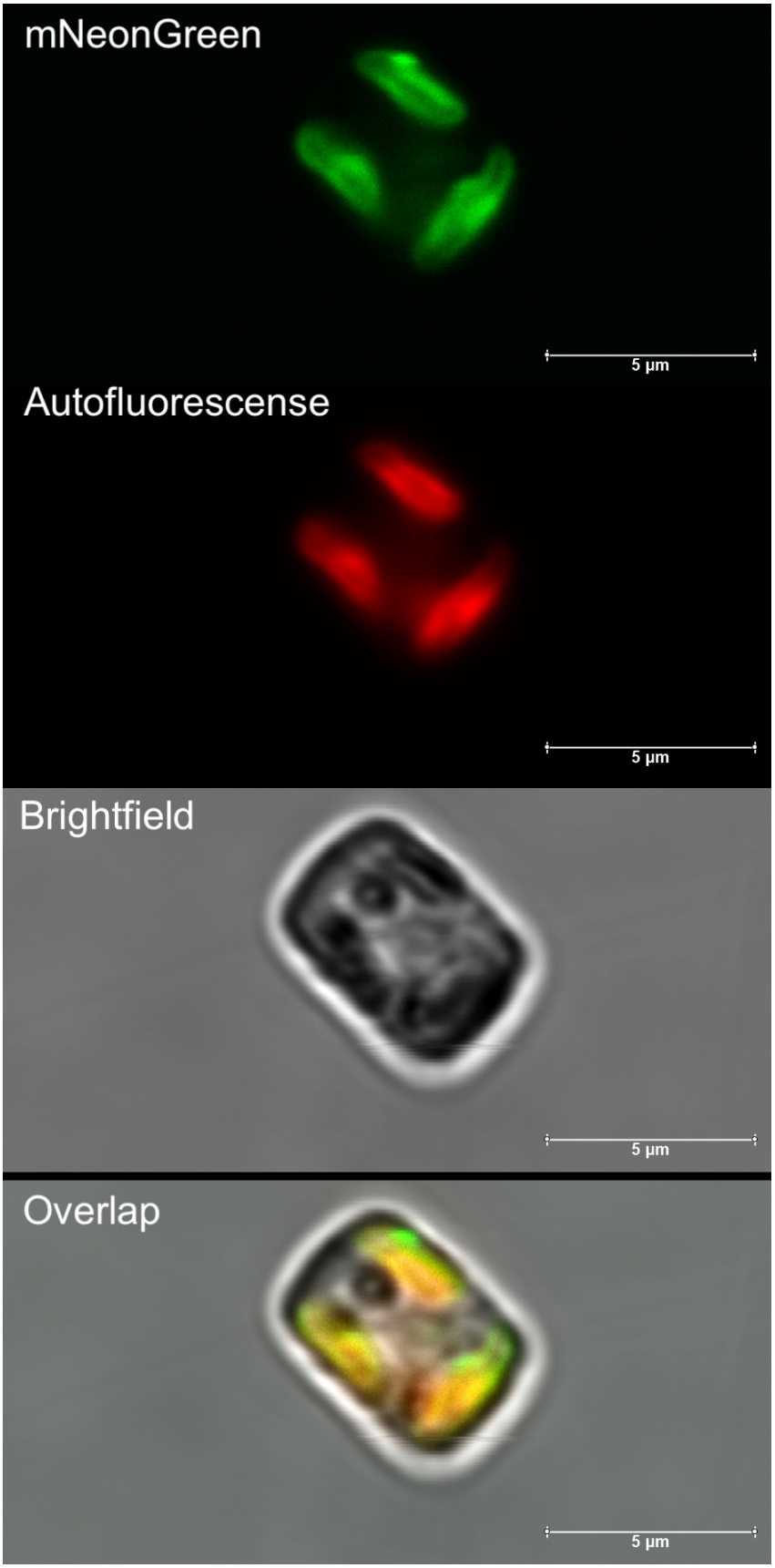


Figure 3.10: Confocal microscope image (1024x512) of mNeonGreen-Tp24711-s2 expressed in *T. pseudonana*. First part: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), second part: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, third part: a PMT brightfield channel and fourth part: overlap of all channels. The scale bar is 5 µm. The image is retrieved from Leica LAS AF Software.

The image obtained from the confocal microscopy of mNeonGreen-Tp24711-s6 cells, is given in Figure 3.11. Although there is noise in the background, it is clear that it is a single cell, with weak mNeonGreen dispersed throughout the cytosol, but also here strongly expressed mNeonGreen can be observed in the girdle band SDVs.

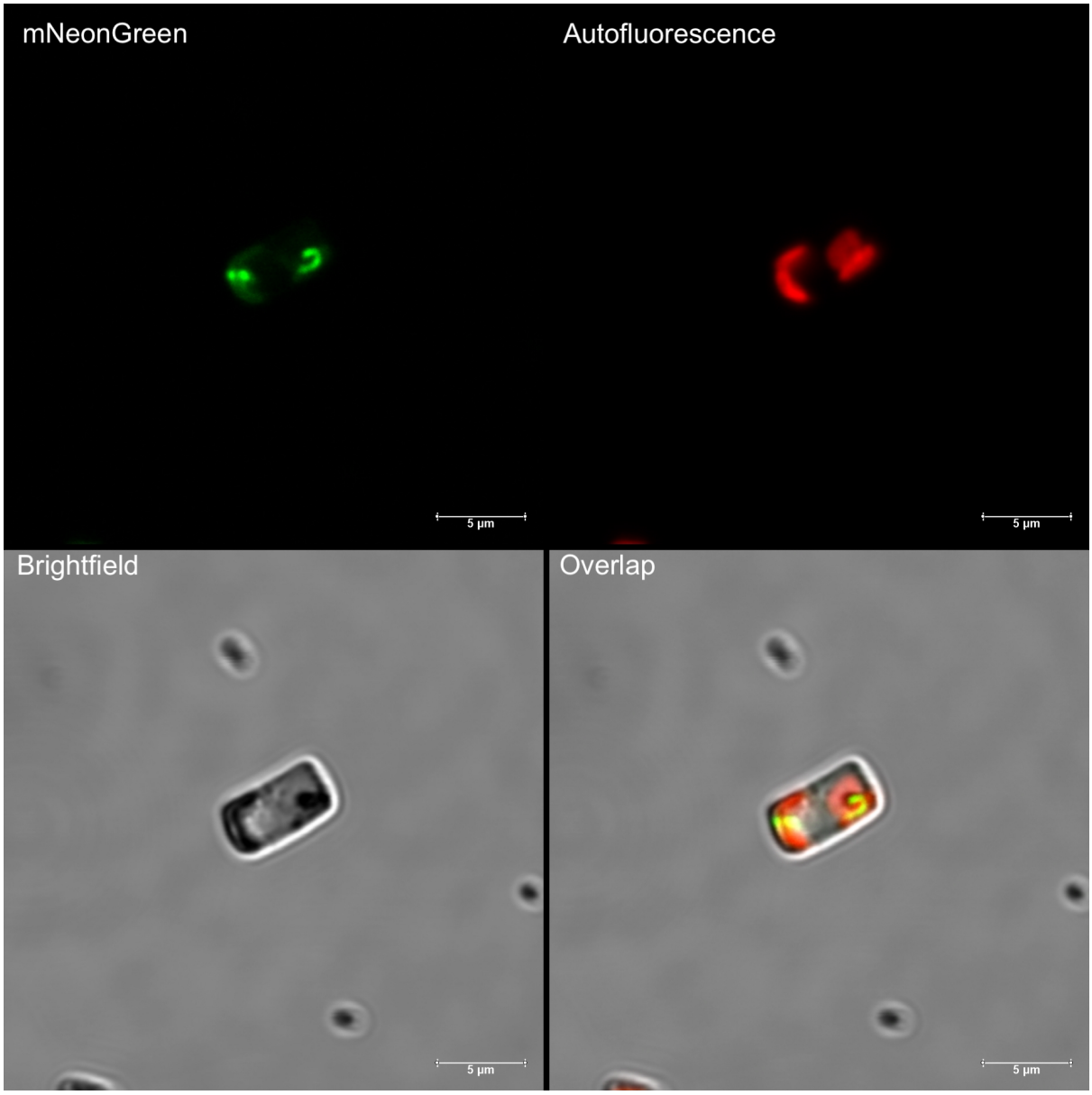


Figure 3.11: Confocal microscope image (512x512) of mNeonGreen-Tp24711-s6 expressed in *T. pseudonana*. Upper left: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), upper right: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a PMT brightfield channel and lower right: overlap of all channels. The scale bares are 5 µm. The image is retrieved from Leica LAS AF Software.

In Figure 3.12 it is clear that mNeonGreen is very weakly expressed through the cytosol, but there is a very clear difference in valve SDV, where it is very highly expressed. High expression in valve SDV can also mean that the cell is about to divide. In Figure F.5, on the other hand, mNeonGreen is evenly expressed, but also higher expressed in girdle band SDVs. In Figure F.6, it may also look like mNeonGreen is highest expressed in the thecas.

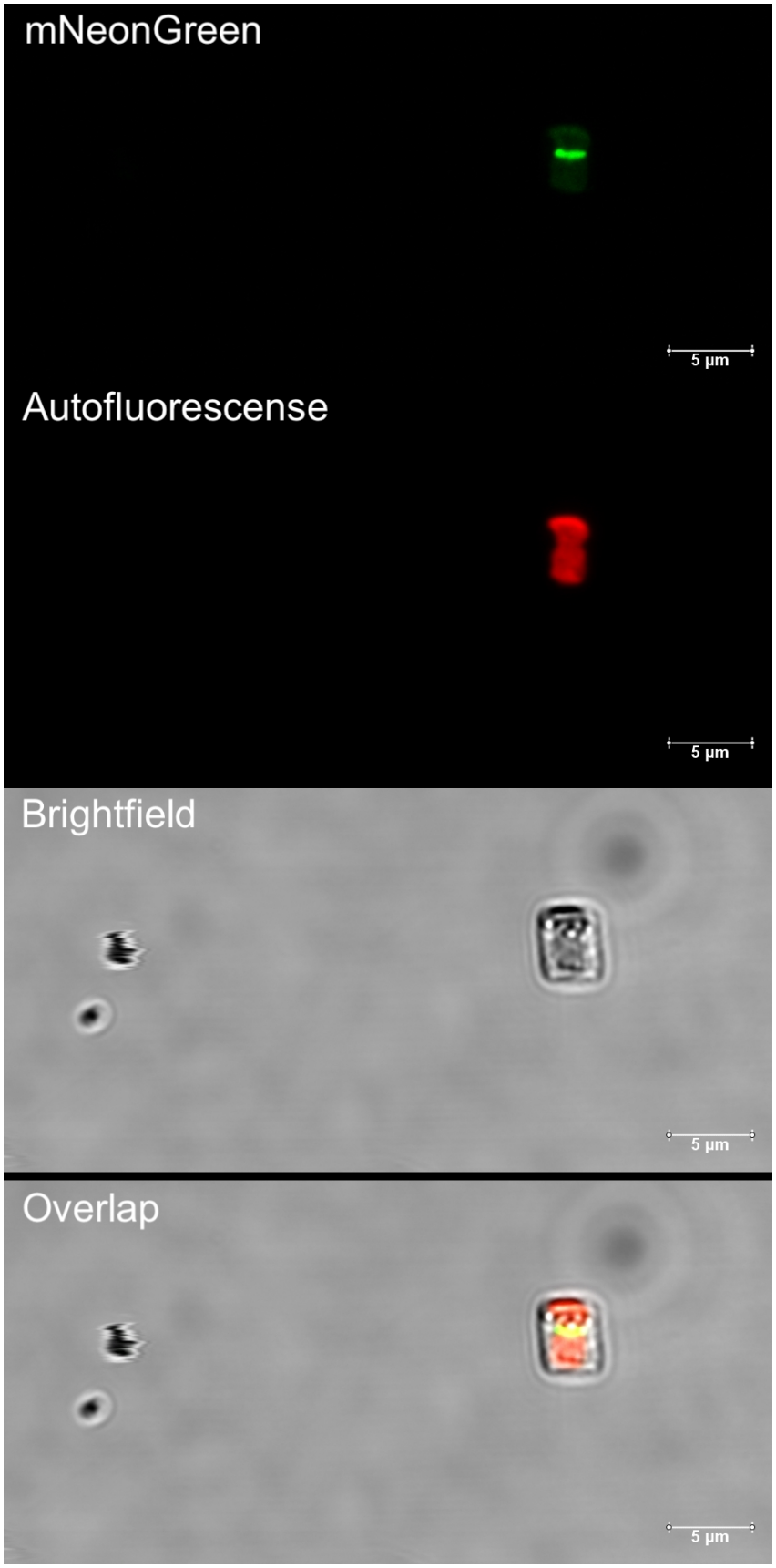


Figure 3.12: Confocal microscope image (1024x512) of mNeonGreen-Tp24711-s6 expressed in *T. pseudonana*. First part: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), second part: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, third part: a PMT brightfield channel and fourth part: overlap of all channels. The scale bares are 5 µm. The image is retrieved from Leica LAS AF Software.

3.3 Gene expression of *T. pseudonana*

Gene expression analysis with qRT-PCR was performed to determine if electroporation or biolistic shooting is the best approach for genetic transformation of *T. pseudonana*. By comparing the expression of *Cas9*, *LHCF9*, *NAT* and *sgRNA* for the different PAM sites, it is possible to see how the gene expression is expressed in relation to WT and which method gives the highest gene expression. Three silicon transporter-associated kinases from *Sværen (2021)*, were compared with two Tp24708/24711P1 clones and WT sample. Silicon transporter-associated kinases (Tp14322 r77) were transformed into *T. pseudonana* by electroporation, while Tp24708/24711P1 was transformed into *T. pseudonana* by electroporation. Figure 3.13 shows the average Ct value for the different gene expressions in the different clones.

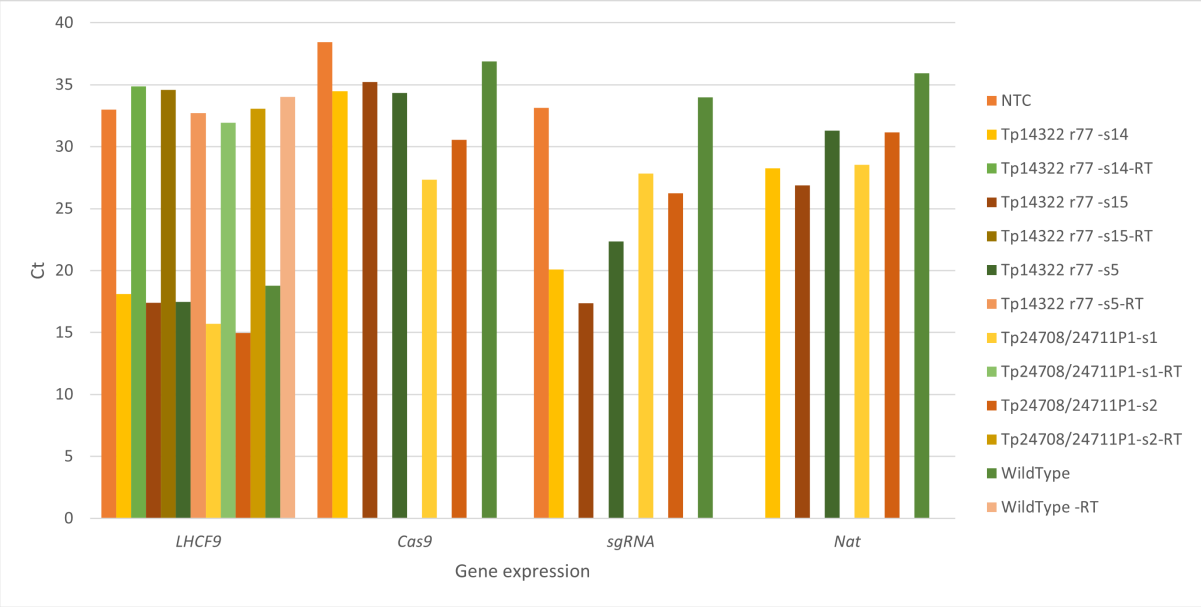


Figure 3.13: An analysis of the gene expression of *T. pseudonana* performed by qRT-PCR. The expression of the native gene *LHCF9* and the genes *Cas9*, *sgRNA* and *NAT* from the plasmids were analysed. The average Ct value for the relative gene expression for each of the clones, where three silicon transporter-associated kinases (Tp14322 r77) clones (transformed by electroporation), were compared to two Tp24708/24711P1 cloned (transformed by electroporation) is compared with WT sample.

Since high Ct value means that there is relatively low gene expression, it is possible to observe that there is low expression of *Cas9* in all clones. The NTC sample and the WT sample have the highest Ct value, which fits well with the fact that these should not contain *Cas9*. The WT sample is not transformed and will therefore not have a plasmid, and therefore not the associated genes either. *LHCF9* is used as a positive control and is a light harvesting fucoxanthin chlorophyll protein, which is highly expressed in *T. pseudonana*. The results from qRT-PCR are correct, as all the clones have low Ct for *LHCF9*, which means that they are highly expressed, while NTC and -RT samples, which are negative controls, have high Ct, which means that they are low expressed or do not have a template.

It is possible to observe that the Ct values for *sgRNA* for the different clones are expressed to

different degrees. The three silicon transporter-associated kinases (Tp14322 r77) clones that were transformed into *T. pseudonana* by electroporation, has lower Ct values than the two Tp24708/24711P1 samples transformed in *T. pseudonana* by biolistic shooting. The results from the *sgRNA* expression in the clones make it possible to suggest that electroporation is a better method than biolistic shooting. There was also a high Ct value for *sgRNA* in the WT sample, which is also true to the fact that WT sample should not express *sgRNA*.

Since all the clones grow in cultures with L1 and Nou (100 µm/µL), it was expected that there would be low Ct in all samples with *NAT*. However, this is not the case as all samples show to have relatively low gene expression with high Ct value. Based on the low gene expression of *NAT*, it is possible to assume the clones have a different mechanism for resisting antibiotics. In contrast to the clones, the WT sample is expected to have a high Ct, as it is a negative control for *NAT* expression, just like *Cas9* and *sgRNA* as well.

3.4 Assembly of pBKS-hCas9M with selection

Alternative selection markers plasmids suitable for conjugation have been previously made. In Kristin Lillebo Bentzen's master thesis, it was found that conjugation is not a suitable method for genetic transformation (Bentzen, 2020). Since it has also been found that electroporation is at least as suited as biolistic shooting in Subsection 3.3, it was desirable to assemble plasmids of pBKS-hCas9 with selection that is better suited for electroporation.

In site-directed mutagenesis, the specifically designed mutagenesis primers were used to perform a site-specific mutagenesis to remove a BsaI site from the pre-existing pBKS plasmid, so that the PCR product generated was pBKS_BsaImut. The purified PCR product was transformed into higly competent DH5α competent *E. coli* cells, and the colonies were screened by restriction digestion and the result is given in Figure 3.14. The results from the gel electrophoresis and Sanger sequencing show that the composition of pBKS_BsaImut is correct, and was confirmed by the cds from SnapGene (Figure G.1).

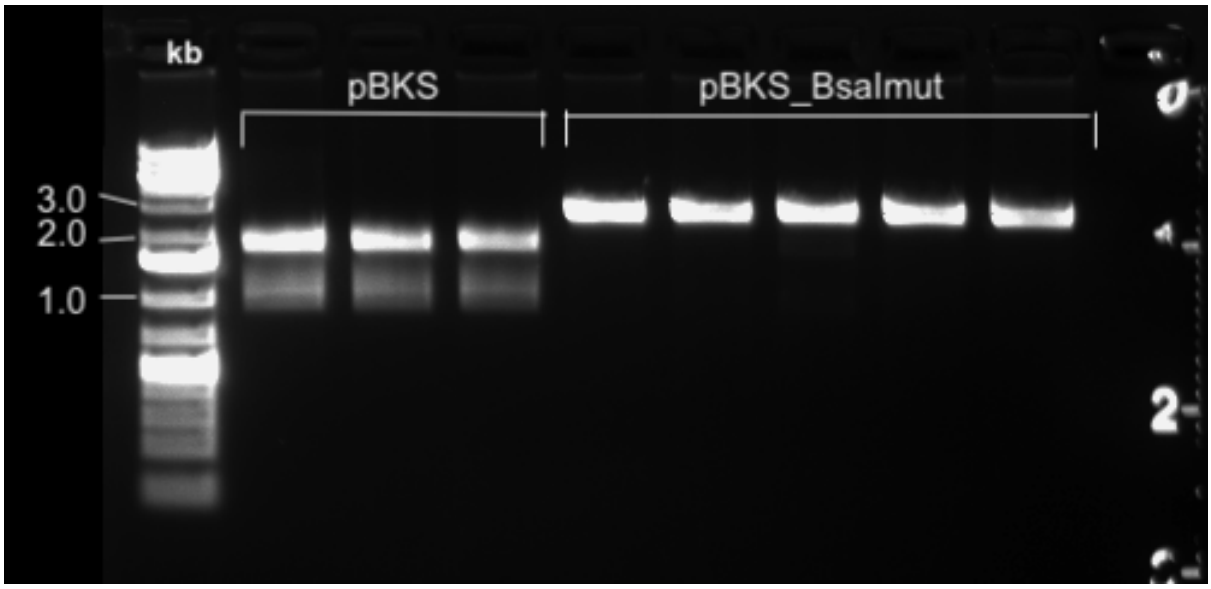


Figure 3.14: Gel electrophoresis (1.0 %) analysis of pBKS and pBKS_BsaImut after restriction digestion with BsaI-HFv2 and DraIII-HF. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation.

Furthermore, selection marker fragments were amplified from various plasmids and assembled with linearized pBKS_BsaImut by Gibson Assembly. The samples that looked best on the colony screening were extracted and analyzed by gel electrophoresis after control restriction digestion with SacIII and XhoI (Figure 3.15 og Figure 3.16). All pBKS_Bsr and pBKS_Nou samples appeared to be correct on the gel, and were submitted to Sanger sequencing. pBKS_Nrf, on the other hand, did not look as good on the gel, but all the samples were still submitted to Sanger sequencing.

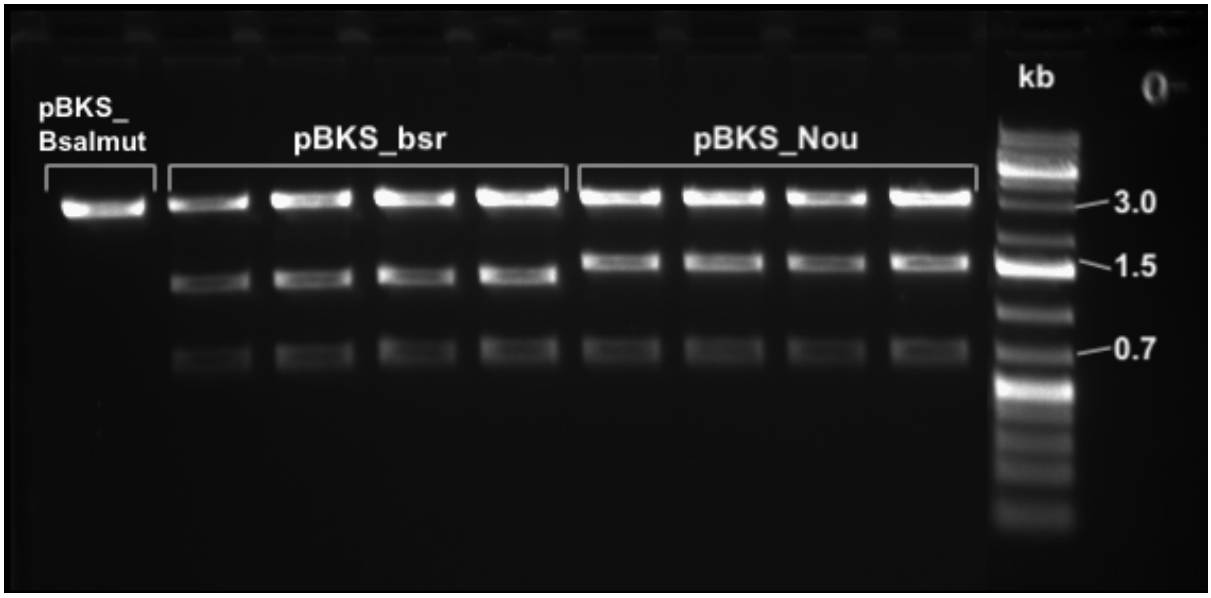


Figure 3.15: Gel electrophoresis (1.0 %) analysis of pBKS_Bsr and pBKS_Nou after restriction digestion with SacII and XhoI. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation.

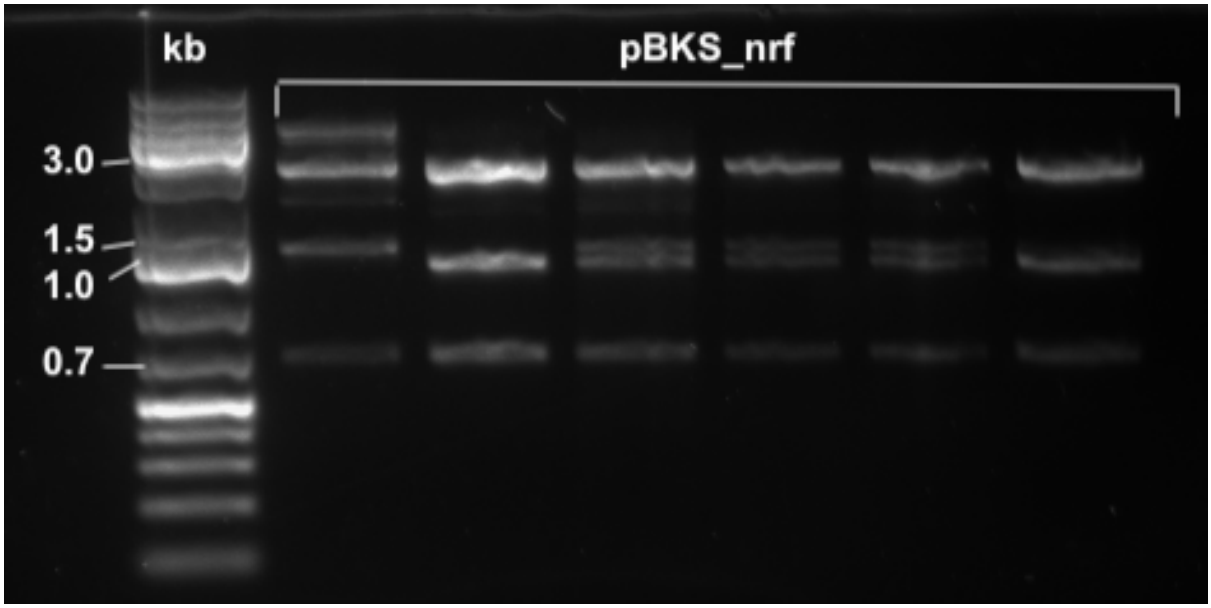


Figure 3.16: Gel electrophoresis (1.0 %) analysis of pBKS_Nrf after restriction digestion with SacII and XhoI. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation.

In the Sanger sequencing results from pBKS_Bsr, which were compared with the cds in SnapGene (Figure G.2), mutations occured in the overlap between the primers. It was also possible to see that, for example the GAT codon had mutated to GAC. Since both GAT and GAC condon encode the aspartate aminoacid, this has no effect on the plasmid. The problem with compostion of pBKS_Bsr is that the stop codon for Bsr has not been included. Without a stop codon, it will not be possible to produce specific proteins. For Sanger sequencing results of pBKS_Nou however, there are several mutations. The biggest change is

in the codon GCC, which encodes alanine that has mutated to aspartate (GAC). Alanine is a hydrophobic amino acid, which is non polar, while aspartate is negatively charged, which will cause the entire structure to be affected. Such changes can lead to changes in protein folding or function, which can potentially lead to changes in phenotype. Even if some of the samples of pBKS_Nrf look correct on the gel (Figure 3.16), it is possible to see in the Sanger sequencing results that were compared to cds in SanpGene (Figure G.4) that this is not the case. It is possible to see in the figure that the Nrf selection marker has not been inserted in any of the samples. Thus, assembly of pBKS with selection markers was not successful.

hCas9 fragment was also attempted to be inserted into the plasmids. Although the Cas9 fragment looked good on the gel (not included), the plasmid would not be assembled, as the plasmids themselves were assembled incorrectly. An electroporation was performed on the assembled pBKS plasmids with selection. After electroporation, they were plated at different concentrations, but there were no growth on the plates after two weeks.

4 Discussion

The main purpose of the master's thesis was to use CRISPR/Cas9 gene editing in combination with different biovisualization techniques and biochemical characterization to investigate the roles of group II silicanins in the biomineralization of the cell wall of the diatom *T. pseudonana*. In order to investigate if the two members of group II silicanins, TP24708 and Tp24711, are involved in the frustule synthesis, three main approaches were used: (1) create and characterize knockout mutants by CRISPR/Cas9-based gene editing; and (2) study the intracellular localization by flow cytometer, fluorescence microscope, and confocal microscope; and (3) analyse the gene expression of *T. pseudonana* in biolistic transformation mutants compared to electroporation mutants. In addition, it was desired to create a pBKS_hCas9 plasmid with different selection markers, which are better suited for electroporation.

Creating knockout mutants of the single PAM site, Tp24711P2, and the double PAM site, Tp24708/24711P1, by CRISPR/Cas9-based gene editing was not achieved. Vectors containing the Cas9 nuclease and specific sgRNA for targeting Tp24711P2 and Tp24708/24711P1 were earlier created and sequenced by Annika Messemer at the ACMS lab. A bacterial transformation of the plasmids were made to obtain sufficient yield and were confirmed with control restriction digest, before being transferred into *T. pseudonana* by electroporation. PCR screening of the Tp24711P2 clones and the Tp24708 / 24711P1 clones, showed that the PCR products from the CRISPR clones were the same size as WT, which means that none of the clones had larger indels. Therefore, the PCR products were further used to investigate for minor mutations around the target site. There were no significant differences between normalized melting curves for the transgenic lines expressing Cas9 together with the specific guide RNAs, and the WT samples from HRM screening of *T. pseudonana*, targeting Tp24708/24711P1 and Tp24711P2. Sanger's sequencing results from the samples that deviated the most confirmed that there were no mutations in any of the samples.

The first study on stable CRISPR/Cas9-based gene editing in algae is by *Nymark et al. (2016)*, using a vector encoding both a codon-optimized synthetic Cas9 and sgRNA to target the gene encoding *P. tricornutum* chloroplast signal recognition particle 54. The study found biallelic modifications, and almost complete absence of the WT allele (Nymark et al., 2016). In the study by *Hopes et al. (2016)*, knockout lines in *T. pseudonana* were created using a CRISPR system with two sgRNAs (Hopes et al., 2016). However, creating knockout mutations in *T. pseudonana* has not been achieved in several of the master projects performed as a part of the research project DIASIL at NTNU (Bentzen, 2020; Harris, 2021; Sjevelås, 2021; Skisland, 2021). Since knockout mutations were not created, there was a small possibility that the targeted genes were vital, and creating knockout mutations would be lethal. In a genome-wide knockout analysis in yeasts, it has been found that up to a quarter of the genes in a genome can be essential (Du, 2020). Therefore, it may also be possible to assume that the group II silicanins, Tp24708 and Tp24711, may be crucial, although this has not been confirmed. In the study of *Görlich et al. (2019)*, it was found that *Sin1* is important for

the biogenesis of diatom cell walls. Although the mutants appeared to grow normally, the study showed that function and structure were affected (Görlich et al., 2019). If Tp24708 and Tp24711 are essential proteins in *T. pseudonana*, a mutation generated by the Cas9 nuclease might create a lethal phenotype. This would, however, imply that even a heterozygous mutation would result in the lethal phenotype, making this a doubtful explanation.

Another reason why no mutations were obtained may be inactive sgRNA. In the study of *Görlich et al. (2019)*, *Sin-1* knockout mutations of *T. pseudonana* were generated. Two DNA sequences encoding *Sin-1*-specific guide RNAs (gRNA-1 and gRNA-2) were tested. Sequencing of *Sin-1*-specific genomic PCR products revealed sequence ambiguities at the predicted Cas9-induced double-strand break site targeted by gRNA-1. However, it was found that no *Sin1* mutations were obtained in the region targeted by gRNA-2. It was therefore assumed that sgRNA was inactive (Görlich et al., 2019). sgRNA can also be ineffective in making DSBs in the target sites of DNA. The inefficiency may be due to poor Cas9 or target site detection. Verification of the presence of sgRNA was not performed to confirm this assumption in this thesis either. However, there was growth on the plates with *NAT* after electroporation, which means that the *NAT* gene must have been successfully integrated. On the other hand, gene expression study (Figure 3.13) also showed low expression of *NAT* for both electroporation and biolistic shooting, which means that there may not have been enough pressure from the antibiotic Nou to preserve the plasmids. In that case, there must be another mechanism for resisting antibiotics. The presence of Cas9 was also confirmed on some of the clones screened (Figure 1.7, with primers from Table D.1 in Appendix D). It may also be that Cas9 is low expressed, which results in insufficient Cas9 nuclease for efficient cleavage in the PAM site (Yuen et al., 2017). Thus, there may be normal growth after electroporation, although some of the genes are poorly expressed and will not create knockout mutations. *sgRNA* was highly expressed in electroporation compared to biolistic shooting (Figure 3.13). This contradicts the fact that sgRNA is inactive. However, it is difficult to confirm this as the electroporation results used in the gene expression study are from kinases.

The inefficiency of plasmids may also be one of the reasons why knockout mutations were not created. The ineffectiveness of the plasmids may be due to instability in the vector. In *Sharma et al. (2018)*, it appears that episomic retrieval of the original pPtPuc3 vector from *P. tricornutum* cells shows that only 50 % of the episomes retain their original size, indicating that rearrangements of the diatom episome may occur during propagation in *E. coli* (Sharma et al., 2018). It is known that electroporation requires large amounts of cells, as there is random incorporation of DNA into the genome (Gutiérrez and Lauersen, 2021). When it is uncertain whether the plasmids maintain the correct size without reorganization during propagation in *E. coli*, a higher amount of plasmids is required. Since the plasmids themselves are made suitable for conjugation, the size and amount will lead to the instability of the plasmids when they are electroporated.

It is also possible that successful repair mechanisms initiated by DSB have repaired the mutations. HDR provides greater repair security than NHEJ, although NHEJ is the main route

for repair in DSBs in all cell cycle phases (Reynolds et al., 2012). The study by *Reynolds et al. (2012)* found that the repair of DSBs by NHEJ is strongly regulated with route selection and repair kinetics depending on the chemical complexity of DSB. In *Guo et al. (2018)*, it has also been observed that NHEJ is inherently accurate in repairing Cas9-induced DNA double-strand breaks. With an accurate NHEJ, there may thus also be cases where no mutations occur (Guo et al., 2018).

According to *Guo et al. (2018)*, instead of depending on the chance of NHEJ being error-prone, a purposeful indel of a specified size should be introduced to enhance the success rate of creating knockout lines (Guo et al., 2018). It may have been easier to create mutations in the group II silicanins by generating a specific mutation. Because there is no certainty that a desirable mutation will occur, the strategy utilized in this thesis is too unpredictable. Using several sgRNA targeting sites with a particular distance to cut off the exact required indel might therefore be a safer approach to generating the intended knockout. This method has also been observed in *Hopes et al. (2016)*, when two sgRNAs were used to introduce a precise 37 nt deletion early in the coding region of the urease gene, which led to biallelic mutations (Hopes et al., 2016).

It is even more challenging to get a mutation in a double PAM site, as the mutation must be present in both genes. The study by *Sharma et al. (2021)* also described how several genes within the same family could be knocked out to generate a phenotype. Off-target mutations were generated in up to five genes using the previously established CRISPR/Cas9 system used in *P. tricornutum*, where the off-target effects of Cas9 were exploited (Sharma et al., 2021). Tp24708/ 24711P1 is identical in both genes, which means that the probability of generating mutations should be higher, as off-target effects limit the use of CRISPR/Cas9. However, there were no clear visible shifts in Figure 3.3a and Figure 3.4a, as all samples had normalized melting curves corresponding to the WT samples.

The vector, pTpPUC3-hCas9, is large and has many redundant properties that are not required for electroporation. Therefore, a smaller vector will be a better alternative for electroporation. An assay was attempted to assemble a pBKS_hCas9 plasmid with selection to make plasmids suitable for electroporation. pBlueScript (pBKS) phagemids (plasmids of phage origin) are cloning vectors designed to simplify cloning and sequencing procedures (Alting-Mees and Short, 1989). With pBKS it is possible to construct nested deletions for DNA sequencing, generate RNA transcripts in vitro and site-specific mutagenesis and gene mapping (Alting-Mees and Short, 1989). Thus, it is very well suited for assembling a plasmid as desired, with, for example, hCas9 and selection. A smaller vector, with only the essential properties, will increase the diffusion rate through the temporary pores during an electrical pulse. The assembly of the pBKS_hCas9 plasmids with selection was unsuccessful, as there were several deficiencies in the plasmid. Defects and mutations can cause both the structure and function of the plasmids to be affected. Without the stop codon for Bsr in pBKS_Bsr (Figure A.3), it will not be possible to signal a halt to the protein synthesis in the cell, and it will not be possible to produce Bsr. Without the stop codon, the Bsr gene will grow until

it stops working. This is also the reason why no colonies grew on the plates with Bsr after the electroporation. However, in the reading frames of the pBKS_Bsr plasmid (Figure G.2), it is possible to observe that there is a stop codon only ten codons after Bsr. In other words, this means that the protein in the current vector will only be ten amino acids longer. Thus, the plasmid may be functional, but it should be re-cloned anyway, as it is uncertain how the function is affected. In pBKS_Nat, there were several mutations, including a mutation changing the amino acid alanine to aspartate. A mutation in which the codon encodes another amino acid will affect the structure of the protein. In addition, aspartic acid (acidic amino acid) is very different from alanine (aliphatic), which increases the chance of a non-functional protein. For pBKS_Nrf, the selection marker was not assembled into the vector. As the pBKS plasmid with selection itself could not be correctly assembled, it was impossible to insert hCas9 into the plasmids. To generate the desired plasmids, all the vectors and fragments must be reassembled with new specific primers. In the same way as assembling plasmids for selection marker study, it can be suggested to construct new plasmids targeting PAM sites, which are better suited to create knockout mutations in *T. pseudonana* by electroporation.

The localization of pTpPUC3-mNeonGreen-Tp24711 fusion protein in *T. pseudonana* was studied by flow cytometry, fluorescence microscopy, and confocal microscopy. The clones were analysed in a flow cytometer to see which clones expressed the highest mNeonGreen fluorescence. Mean mNeonGreen fluorescence was measured in *T. pseudonana* samples. The two clones, mNeonGreen Tp24711-s2 and mNeonGreen Tp24711-s6 deviated the most and were selected to study the localization further. Both samples had twice the mean mNeonGreen fluorescence. The discrepancy could also be observed in the plot, where SSC used to measure the structural complexity of the cells was plotted against mNeonGreen fluorescence measured with BL1. It was clear that mNeonGreen Tp24711-s2 and mNeonGreen Tp24711-s6 had a more complex structure than the other samples. The plots of mNeonGreen Tp24711-s2 and mNeonGreen Tp24711-s6 clones (Figure 3.6) have an extra top that forms a shoulder on the right side of the plot. This can be interpreted as a population of cells with higher fluorescence than the rest. This agrees with the fluorescence microscopy images, where almost all cells appear to have distinct mNeonGreen fluorescence.

The images from fluorescence microscopy and confocal microscopy showed both fluorescence and autofluorescence. Many of the images, especially in the overlap, showed that the green fluorescence was similar to the autofluorescence. However, there were some places where mNeonGreen was expressed and did not overlap with the autofluorescence. The images indicate that mNeonGreen was expressed in the SDV structures. In Figure 3.12, among others, it may look like mNeonGreen is expressed in the division plane for cell division. The division plane corresponds to point 3 in Figure 1.4, where the SDV valves are synthesized. mNeonGreen is also often observed in cytosolic dispersion throughout the cells, meaning that the proteins are cytosolic. It can thus be assumed that Tp24711 acts as cytosolic linkers between the cytoskeleton and proteins in silicalemma. On the other hand, silicanins have a transmembrane domain, which means that they are permanently linked to a membrane.

The cytosolic fluorescence may be the background.

It was desirable to examine the samples during the exponential phase, as it was assumed that mNeonGreen would be highest expressed in this phase. However, this was not the case, as it could be observed that there were few cells that divided with mNeonGreen located in the division plane in valve SDVs. Since most of the cells that underwent division had an accumulation of the protein in the division plane, it is likely that it is involved in the valve formation. Evidence also showed that mNeonGreen was expressed in girdle band SDVs, corresponding to point 6 in Figure 1.4. It can thus also be assumed that mNeonGreen is also involved in girdle band formation. In *Kotzsch et al. (2017)*, GFP fluorescence was also observed in living cells' valve and girdle band regions. This was also the case in the GFP label SAP1 and SAP3 (Tesson et al., 2017). Thus, it is possible to assume that Tp24711 is involved in the frustule synthesis. However, this localization could be misinterpreted, as several processes associated with the cell division are closely related to frustule synthesis. Expression of mNeonGreen in one of those processes can thus be misinterpreted and it may be possible to believe that the localization is associated with the frustule synthesis. Furthermore, because the N-terminal mNeonGreen tagged proteins may affect protein sorting, the observed localization of the mNeonGreen tagged proteins may not reflect the actual localization of the native proteins in the cell. The plasmid map of pTpPUC3-mNeonGreen-Tp24711 (Figure A.6) shows that mNeonGreen is not located entirely on the N-terminus, but after approximately 40 amino acids of Tp24711. mNeonGreen is therefore not cut off together with the signal peptide. Studies of the subcellular localization of a C-terminal fusion protein will establish that the observed localization is not the consequence of incorrect protein sorting. According to *Kotzsch et al. (2017)*, the localization is quite similar for N- and C-terminal fusion proteins, but the C-terminal green fluorescent protein is obviously cut off during biomineralization. N-terminal GFP is incorporated into the biosilica (Kotzsch et al., 2017).

Some cells did not have any fluorescent. Cells without autofluorescence are dead or either responding to stressful conditions. Silicanins are probably only expressed in a limited part of the cell cycle. Therefore, only some of the cells are expected to have clear fluorescence. The samples were studied in different periods, which means that they were not always in the exponential phase. If the samples go into a stationary phase, they can go into a stressful phase, and mNeonGreen can become a silent gene. In order to take better images, attempts were made to set up new cultures and dilute the culture regularly. Diatoms are known to address epigenetic change when they experience stress (Tirichine and Bowler, 2011). In *Gresseth's master's thesis (2020)*, however, it has been observed that the transformants give expression to having regained the mNeonGreen fluorescence. Hence, it is possible to assume that the silencing is reversible. If the cells are able to regain the expression of mNeonGreen, it means that Tp24711 is not vital. In stressful situations, non-essential genes will not be prioritized (Teng et al., 2013).

In order to investigate if electroporation or biolistic shooting is the best approach for the genetic transformation of *T. pseudonana*, a gene expression study with qRT-PCR was per-

formed. However, there is no reasonable basis for concluding anything from this analysis. Still, it is possible to suggest a proposal on which method is best suited for creating knockout mutations, as there are both similarities and differences between the plasmids. The plasmids used are identical except for the short sgRNA fragment. The transformants of the kinases constructed by CRISPR/Cas9 were generated by electroporation, while the silicanin transformants were generated by biolistic transformation. Analysis of the qRT-PCR results (Figure 3.13), shows that parts of the results are unexpected. The WT samples are as expected. *Cas9*, *sgRNA* and *NAT* are low expressed in WT, while *LHCF9* is highly expressed. Although the samples are not directly comparable, as some of the samples are kinases, it is possible to assume that electroporation is a better method. There should also have been differences in the Ct values for the expression of *Cas9* and *NAT* in the different clones, but this is not the case. The NTC samples and -RT samples have high Ct values, which means that these are low expressed. Low to almost no expression implies that there has been little to no contamination. *Cas9* is low expressed in all clones, relatively similar to NTC. For biolistic shooting, this may be true, as no mutations have been found. In the kinase, on the other hand, this is uncertain.

LHCF9, on the other hand, is as expected. -RT samples are low expressed, while all other samples are highly expressed. The *LHCF9* gene is located in the genome, while all other genes studied are located in the vector, pTpPUC3-Cas9. In the WT sample, it will be the only gene that is expressed. *NAT* was also low expressed in all samples, which was not expected as there was growth on all plates after transformation. The clones may have used a different mechanism for resisting Nou. *sgRNA* is highly expressed in the kinases, but not the Tp24708/24711P1 transformants. This may be correct, as there were no mutations in the biolistic shooting. The expression of *sgRNA* also makes it possible to say that electroporation is a better suited method, but this cannot be determined until more analyses are performed. Low expression of *sgRNA* and *Cas9* by biolistic shooting implies that it will be challenging to create knockout mutations, similar to the results found earlier. In addition, it should also be investigated whether the expression of *sgRNA* and *Cas9* can be increased by modifying the promoters used.

5 Future work

There is a lot that can be done differently for future work. Among other things, new plasmids should be fitted or methods changed to be able to make knockout mutations with group II silicanins, Tp24708 and Tp24711, in *T. pseudonana*. A flow cytometer should be used more frequently to look for clones. Several selection marker studies should also be done to determine if other selection markers can be used instead of Nou. It may also be interesting to study the frustule morphology in transmission electron microscopy (TEM) and/or scanning electron microscopy (SEM) to gain a better understanding of *T. pseudonana*. The knockout mutations may say something about how the frustule changes during the synthesis.

Several experiments with mNeonGreen will also provide a better understanding of the function of *T. pseudonana*. For example, several confocal microscopy examinations can be performed to verify observed locations. This can be done by examining C-terminal fusion protein. Studies can also be done where the expression of mNeonGreen is examined over a period in confocal microscopy. This will give a clearer picture of how mNeonGreen is expressed throughout the cell cycle. By making specific mutations in a known target site, it may also be possible to see how mutations play their role in diatoms which can be further investigated to learn more about its exact functions.

In a future gene expression study, to investigate whether electroporation or biolistic shooting is the appropriate method for genetic transformation of *T. pseudonana*, it may be desirable to use the same plasmids that are transformed into *T. pseudonana* in different ways. Most importantly, more replicates should be made, as it will give a better result of which method is better suited.

In order to construct a more suitable plasmid for electroporation, an attempt should also be made to reassemble pBKS_hCas9 plasmids with selection. It should also be considered whether new primers should be made. If the plasmids are constructed correctly, it may also be a suggestion to carry out a new selection marker study.

6 Conclusions

The aim of this thesis was to use CRISPR/Cas9 gene editing in combination with different biovisualization techniques and biochemical characterization to investigate the roles of group II silicanins in the biomineralization of the cell wall of the diatom *T. pseudonana*. The purpose was studied in various ways that gave a good understanding of the frustule of *T. pseudonana*. However, it is difficult to draw a conclusion based on the results.

Creating knockout mutants of the single PAM site, Tp24711P2, and the double PAM site, Tp24708/24711P1, and screen for CRISPR/Cas9-induced mutations in high-resolution melting were not achieved. However, it should be attempted again in a future work, so that it is possible to investigate the functions of the proteins.

mNeonGreen-tagged Tp24711 was investigated in confocal microscopy and were clearly observed in the region where valve SDVs and newly synthesized valves are located during cell division. mNeonGreen fluorescence was also observed in the girdle band area. Hence, it is possible to assume that Tp24711 may be involved in the synthesis of valves and girdle bands.

A gene expression assay to investigate whether electroporation or biolistic shooting is the appropriate method for genetic transformation of *T. pseudonana*, was also performed by qRT-PCR, which showed low expression of the transgenes, *Cas9*, *Nou* and *sgRNA* in the transgenic clones, while native *LHCF9* was higher expressed. *sgRNA* was higher expressed in the clones generated by electroporation compared to biolistic shooting. Therefore electroporation is suggested to be a better suited method for genetic transformation of *T. pseudonana*. Such studies can also be performed in future attempts to create knockout mutations, as it will provide an understanding of which transformation method is best suited, and also whether *Cas9* and *sgRNA* are sufficiently expressed to create knockout mutations. It was also found that *Nat* expression was low in all samples, which may indicate that there are other mechanisms for resisting antibiotics.

The plasmids used in the previous studies both to create knockout motions and in selection markers are constructed for conjugation. It may be desirable to construct a plasmid that is better suited for electroporation. To construct a more suitable plasmid for electroporation, attempts were made to assemble pBKS_hCas9 plasmids with selection. It was also intended to conduct a selection marker study with the new plasmids. The plasmids were not assembled correctly and thus did not give the expected results in the selection marker study.

References

Alting-Mees, M. and Short, J. (1989), 'pBluescript II: gene mapping vectors', *Nucleic Acids Research* **17**(22), pp. 9494–9494. doi: 10.1093/nar/17.22.9494.

Alverson, A. J., Beszteri, B., Julius, M. L. and Theriot, E. C. (2011), 'The model marine diatom thalassiosira pseudonana likely descended from a freshwater ancestor in the genus cyclotella', *BMC Evolutionary Biology* **11**(125). doi: 10.1186/1471-2148-11-125.

Apt, K. E., Grossman, A. R. and Kroth-Pancic, P. G. (1996), 'Stable nuclear transformation of the diatomPhaeodactylum tricornutum', *Molecular and General Genetics MGG* **252**(5), pp. 572–579. doi: 10.1007/bf02172403.

Armbrust, E. V. (2009), 'The life of diatoms in the world's oceans', *Nature* **459**(7244), pp. 185–192. doi: 10.1038/nature08057.

Babenko, I., Friedrich, B. M. and Kröger, N. (2022), Structure and morphogenesis of the frustule, *in* 'The Molecular Life of Diatoms', Springer International Publishing, p. 287–312. doi: 10.1007/978-3-030-92499-7_11.

Benoiston, A.-S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S. and Bowler, C. (2017), 'The evolution of diatoms and their biogeochemical functions', *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**(1728) (20160397). doi: 10.1098/rstb.2016.0397.

Bentzen, K. L. L. (2020), Functional studies of two genes encoding closely related group ii silicanins in the diatom *Thalassiosira pseudonana*, Master's thesis, NTNU, Trondheim.

Bertani, G. (1951), 'Studies on lysogenesis I: The mode of phage liberation by lysogenic escherichia coli', *Journal of Bacteriology* **62**(3), pp. 293–300. doi: 10.1128/jb.62.3.293-300.1951.

Bowler, C., Vardi, A. and Allen, A. E. (2010), 'Oceanographic and biogeochemical insights from diatom genomes', *Annual Review of Marine Science* **2**(1), pp. 333–365. doi: 10.1146/annurev-marine-120308-081051.

Brembu, T., Chauton, M. S., Winge, P., Bones, A. M. and Vadstein, O. (2017), 'Dynamic responses to silicon in thalasiossira pseudonana - identification, characterisation and classification of signature genes and their corresponding protein motifs', *Scientific Reports* **7**(4865). doi: 10.1038/s41598-017-04921-0.

Buck, J. M., Bártulos, C. R., Gruber, A. and Kroth, P. G. (2018), 'Blasticidin-s deaminase, a new selection marker for genetic transformation of the diatom *Phaeodactylum tricornutum*', *PeerJ* **6**(e5884). doi: 10.7717/peerj.5884.

Carter, M. and Shieh, J. (2015), Chapter 11 - gene delivery strategies, *in* 'Guide to Research Techniques in Neuroscience', Elsevier, p. 239–252. doi: 10.1016/b978-0-12-800511-8.00011-3.

Dotmatics (2021), 'Snapgene®', https://www.snapgene.com. Version: 6.0.3.

Doudna, J. A. and Charpentier, E. (2014), 'The new frontier of genome engineering with CRISPR-cas9', *Science* **346**(6213) (1258096). doi: 10.1126/science.1258096.

Du, L.-L. (2020), 'Resurrection from lethal knockouts: Bypass of gene essentiality', *Biochemical and Biophysical Research Communications* **528**(3), pp. 405–412. doi: 10.1016/j.bbrc.2020.05.207.

Dunahay, T. G., Jarvis, E. E. and Roessler, P. G. (1995), 'Genetic transformation of the diatoms cyclotella cryptica and navicula saprophila 1', *Journal of Phycology* **31**(6), pp. 1004–1012. doi: 10.1111/j.0022-3646.1995.01004.x.

Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C. and Bowler, C. (1999), 'Transformation of nonselectable reporter genes in marine diatoms', *Marine Biotechnology* **1**(3), pp. 239–251. doi: 10.1007/pl00011773.

Falciatore, A., Jaubert, M., Bouly, J.-P., Bailleul, B. and Mock, T. (2019), 'Diatom molecular research comes of age: Model species for studying phytoplankton biology and diversity', *The Plant Cell* **32**(3), pp. 547–572. doi: 10.1105/tpc.19.00158.

Falciatore, A. and Mock, T., eds (2022), *The Molecular Life of Diatoms*, Springer International Publishing. doi: 10.1007/978-3-030-92499-7.

Fattorini, N. (2022), 'Studying the targeting of frustule-associated proteins in the centric diatom thalassiosira pseudonana'. PhD thesis. Philipps-Universität Marburg. doi: 10.17192/Z2022.0105.

Fattorini, N. and Maier, U. G. (2021), 'Targeting of proteins to the cell wall of the diatom thalassiosira pseudonana', *Discover Materials* **1**(5). doi: 10.1007/s43939-021-00005-z.

Foroni, L., Reid, A. G., Gerrard, G., Toma, S. and Hing, S. (2017), Molecular and cytogenetic analysis, *in* 'Dacie and Lewis Practical Haematology', Elsevier, p. 126–164. doi: 10.1016/b978-0-7020-6696-2.00008-4.

George, J., Kahlke, T., Abbriano, R. M., Kuzhiumparambil, U., Ralph, P. J. and Fabris, M. (2020), 'Metabolic engineering strategies in diatoms reveal unique phenotypes and genetic configurations with implications for algal genetics and synthetic biology', *Frontiers in Bioengineering and Biotechnology* **8**(513). doi: 10.3389/fbioe.2020.00513.

Godbey, W. (2022), The polymerase chain reaction (PCR), *in* 'Biotechnology and its Applications', Elsevier, p. 219–246. doi: 10.1016/b978-0-12-817726-6.00010-1.

Gresseth, M. (2020), Investigating fluorescence in transformants expressing a silicanin-mneongreen fusion protein, and the effect of cytoskeleton inhibitors on frustule biosynthesis in the diatom *Thalassiosira pseudonana*, Master's thesis, NTNU, Trondheim.

Guillard, R. R. L. and Hargraves, P. E. (1993), 'Stichochrysis immobilis is a diatom, not a chrysophyte', *Phycologia* **32**(3), pp. 234–236. doi: 10.2216/i0031-8884-32-3-234.1.

Guo, T., Feng, Y.-L., Xiao, J.-J., Liu, Q., Sun, X.-N., Xiang, J.-F., Kong, N., Liu, S.-C., Chen, G.-Q., Wang, Y., Dong, M.-M., Cai, Z., Lin, H., Cai, X.-J. and Xie, A.-Y. (2018), 'Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/cas9-mediated genome editing', *Genome Biology* **19**(170). doi: 10.1186/s13059-018-1518-x.

Gutiérrez, S. and Lauersen, K. J. (2021), 'Gene delivery technologies with applications in microalgal genetic engineering', *Biology* **10**(4) (265). doi: 10.3390/biology10040265.

Görlich, S., Pawolski, D., Zlotnikov, I. and Kröger, N. (2019), 'Control of biosilica morphology and mechanical performance by the conserved diatom gene silicanin-1', *Communications Biology* **2**(245). doi: 10.1038/s42003-019-0436-0.

Harris, I. E. (2021), Functional studies of a family of ankyrin repeat-containing proteins in the diatom *Thalassiosira pseudonana* hypothesized to be involved in frustule morphogenesis, Master's thesis, NTNU, Trondheim.

Hasle, G. R. and Syvertsen, E. E. (1996), Chapter 2 - marine diatoms, *in* C. R. Tomas, ed., 'Identifying Marine Diatoms and Dinoflagellates', Academic Press, San Diego, p. 5–385. doi: 10.1016/b978-012693015-3/50005-x.

Hildebrand, M. (2008), 'Diatoms, biomineralization processes, and genomics', *Chemical Reviews* **108**(11), pp. 4855–4874. doi: 10.1021/cr078253z.

Hildebrand, M. and Lerch, S. J. (2015), 'Diatom silica biomineralization: Parallel development of approaches and understanding', *Seminars in cell & developmental biology* **46**, pp. 27–35. doi: 10.1016/j.semcdb.2015.06.007.

Hildebrand, M., Lerch, S. J. L. and Shrestha, R. P. (2018), 'Understanding diatom cell wall silicification—moving forward', *Frontiers in Marine Science* **5**(125). doi: 10.3389/fmars.2018.00125.

Hildebrand, M., York, E., Kelz, J. I., Davis, A. K., Frigeri, L. G., Allison, D. P. and Doktycz, M. J. (2006), 'Nanoscale control of silica morphology and three-dimensional structure during diatom cell wall formation', *Journal of Materials Research* **21**(10), pp. 2689–2698. doi: 10.1557/jmr.2006.0333.

Hoffmann, M., Hurlebaus, J. and Weilke, C. (2007), 'High-resolution melting curve analysis on the lightcycler® 480 pcr system', *Nat. Methods* **2**(82377).

Hopes, A., Nekrasov, V., Kamoun, S. and Mock, T. (2016), 'Editing of the urease gene by CRISPR-cas in the diatom thalassiosira pseudonana', *Plant Methods* **12**(49). doi: 10.1186/s13007-016-0148-0.

Kalthoff, D., Beer, M. and Hoffmann, B. (2013), 'High resolution melting analysis: rapid and precise characterisation of recombinant influenza a genomes', *Virology Journal* **10**(284). doi: 10.1186/1743-422x-10-284.

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., Gillard, J. T., Beeri, K., Ellisman, M. H., Glass, J. I., III, C. A. H., Smith, H. O., Venter, J. C., Allen, A. E., Dupont, C. L. and Weyman, P. D. (2015), 'Designer diatom episomes delivered by bacterial conjugation', *Nature Communications* **6**(6925). doi: 10.1038/ncomms7925.

Keer, J. T. (2008), Chapter 7. quantitative real-time PCR analysis, *in* 'Essentials of Nucleic Acid Analysis', Royal Society of Chemistry, p. 132–166. doi: 10.1039/9781847558213-00132.

Kotzsch, A., Gröger, P., Pawolski, D., Bomans, P. H. H., Sommerdijk, N. A. J. M., Schlierf, M. and Kröger, N. (2017), 'Silicanin-1 is a conserved diatom membrane protein involved in silica biomineralization', *BMC Biology* **15**(65). doi: 10.1186/s12915-017-0400-8.

Kralik, P. and Ricchi, M. (2017), 'A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything', *Frontiers in Microbiology* **8**(108). doi: 10.3389/fmicb.2017.00108.

Kroth, P. (2007), *Molecular Biology and the Biotechnological Potential of Diatoms*, Springer, New York, p. 23–33. doi: 10.1007/978-0-387-75532-8_3.

Kroth, P. G., Bones, A. M., Daboussi, F., Ferrante, M. I., Jaubert, M., Kolot, M., Nymark, M., Bártulos, C. R., Ritter, A., Russo, M. T., Serif, M., Winge, P. and Falciatore, A. (2018), 'Genome editing in diatoms: achievements and goals', *Plant Cell Reports* **37**(10), pp. 1401–1408. doi: 10.1007/s00299-018-2334-1.

Kröger, N. and Poulsen, N. (2008), 'Diatoms—from cell wall biogenesis to nanotechnology', *Annual Review of Genetics* **42**(1), pp. 83–107. doi: 10.1146/annurev.genet.41.110306.130109.

Kumar, S., Rechav, K., Kaplan-Ashiri, I. and Gal, A. (2020), 'Imaging and quantifying homeostatic levels of intracellular silicon in diatoms', *Science Advances* **6**(42). doi: 10.1126/sciadv.aaz7554.

Liu, H., Chen, M., Zhu, F. and Harrison, P. J. (2016), 'Effect of diatom silica content on copepod grazing, growth and reproduction', *Frontiers in Marine Science* **3**(89). doi: 10.3389/fmars.2016.00089.

Mathivannan, C. (2021), Functional studies of group II silicanins in diatom cell wall biomineralization. TBT4500 - Biotechnology, specialization project.

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., Galbraith, E. D., Geider, R. J., Guieu, C., Jaccard, S. L., Jickells, T. D., Roche, J. L., Lenton, T. M., Mahowald, N. M., Marañón, E., Marinov, I., Moore, J. K., Nakatsuka, T., Oschlies, A., Saito,

M. A., Thingstad, T. F., Tsuda, A. and Ulloa, O. (2013), 'Processes and patterns of oceanic nutrient limitation', *Nature Geoscience* **6**(9), pp. 701–710. doi: 10.1038/ngeo1765.

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. and Quéguiner, B. (1995), 'Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation', *Global Biogeochemical Cycles* **9**(3), pp. 359–372. doi: 10.1029/95gb01070.

New England Biolabs (2022), 'Nebcloner® - restriction enzyme single/double digestion', https://nebcloner.neb.com/#!/redigest. Version: 1.12.0.

Niu, Y.-F., Yang, Z.-K., Zhang, M.-H., Zhu, C.-C., Yang, W.-D., Liu, J.-S. and Li, H.-Y. (2012), 'Transformation of diatom *Phaeodactylum tricornutum* by electroporation and establishment of inducible selection marker', *BioTechniques* **52**(6), pp. 1–3. doi: 10.2144/000113881.

Nymark, M., Sharma, A. K., Hafskjold, M. C. G., Sparstad, T., Bones, A. M. and Winge, P. (2017), 'CRISPR/Cas9 gene editing in the marine diatom phaeodactylum tricornutum', *Bio Protocol* **7**(15) (e2442). doi: 10.21769/bioprotoc.2442.

Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M. and Winge, P. (2016), 'A CRISPR/cas9 system adapted for gene editing in marine algae', *Scientific Reports* **6**(24951). doi: 10.1038/srep24951.

Omega Bio-tek (2019), *E.Z.N.A® Plasmid DNA Mini Kit II*, Omega Bio-tek, 400 Pinnacle Way, Suite 450 Norcross, GA 30071. Revision Number: v5.2.

Ozyigit, I. I. (2020), 'Gene transfer to plants by electroporation: methods and applications', *Molecular Biology Reports* **47**(4), pp. 3195–3210. doi: 10.1007/s11033-020-05343-4.

Poulsen, N., Chesley, P. M. and Kröger, N. (2006), 'MOLECULAR GENETIC MANIPULATION OF THE DIATOM THALASSIOSIRA PSEUDONANA (BACILLARIOPHYCEAE)', *Journal of Phycology* **42**(5), pp. 1059–1065. doi: 10.1111/j.1529-8817.2006.00269.x.

Prihoda, J., Tanaka, A., de Paula, W. B. M., Allen, J. F., Tirichine, L. and Bowler, C. (2012), 'Chloroplast-mitochondria cross-talk in diatoms', *Journal of Experimental Botany* **63**(4), pp. 1543–1557. doi: 10.1093/jxb/err441.

QIAGEN (2016*a*), *Quick-Start Protocol QuantiTect® Reverse Transcription Kit*, QIAGEN. Cat. No. 205311.

QIAGEN (2016*b*), *Quick-Start Protocol RNeasy® Plant Mini Kit*, QIAGEN. Cat. No. 74904.

Reitan, K. I., Øie, G., Jørgensen, H. and Wang, X. (2021), 'Chemical composition of selected marine microalgae, with emphasis on lipid and carbohydrate production for potential use as feed resources', *Journal of Applied Phycology* **33: 6**, pp. 3831–3842. doi: 10.1007/s10811-021-02586-x.

Reynolds, P., Anderson, J. A., Harper, J. V., Hill, M. A., Botchway, S. W., Parker, A. W. and O'Neill, P. (2012), 'The dynamics of ku70/80 and DNA-PKcs at DSBs induced by ionizing radiation is dependent on the complexity of damage', *Nucleic Acids Research* **40**(21), pp. 10821–10831. doi: 10.1093/nar/gks879.

Romann, J., Valmalette, J.-C., Chauton, M. S., Tranell, G., Einarsrud, M.-A. and Vadstein, O. (2015), 'Wavelength and orientation dependent capture of light by diatom frustule nanostructures', *Scientific Reports* **5**(17403). doi: 10.1038/srep17403.

Saade, A. and Bowler, C. (2009), 'Molecular tools for discovering the secrets of diatoms', *BioScience* **59**(9), pp. 757–765. doi: 10.1525/bio.2009.59.9.7.

Sander, J. D. and Joung, J. K. (2014), 'CRISPR-cas systems for editing, regulating and targeting genomes', *Nature Biotechnology* **32**(4), pp. 347–355. doi: 10.1038/nbt.2842.

Sharma, A. K., Nymark, M., Flo, S., Sparstad, T., Bones, A. M. and Winge, P. (2021), 'Simultaneous knockout of multiple *LHCF* genes using single sgRNAs and engineering of a high-fidelity cas9 for precise genome editing in marine algae', *Plant Biotechnology Journal* **19**(8), pp. 1658–1669. doi: 10.1111/pbi.13582.

Sharma, A. K., Nymark, M., Sparstad, T., Bones, A. M. and Winge, P. (2018), 'Transgene-free genome editing in marine algae by bacterial conjugation – comparison with biolistic CRISPR/cas9 transformation', *Scientific Reports* **8**(14401). doi: 10.1038/s41598-018-32342-0.

Sjevelås, C. V. (2021), Morphological changes in the frustule of *Thalassiosira pseudonana* by genetic manipulation and cultivation conditions, Master's thesis, NTNU, Trondheim.

Skisland, M. M. (2021), Functional studies of two silicon transporter-associated kinases in the diatom *Thalassiosira pseudonana*, Master's thesis, NTNU, Trondheim.

Synthego (2022), 'Ice analysis', https://ice.synthego.com/#/#about. V3.0.

Taparia, Y., Zarka, A., Leu, S., Zarivach, R., Boussiba, S. and Khozin-Goldberg, I. (2019), 'A novel endogenous selection marker for the diatom phaeodactylum tricornutum based on a unique mutation in phytoene desaturase 1', *Scientific Reports* **9**(8217). doi: 10.1038/s41598-019-44710-5.

Teng, X., Dayhoff-Brannigan, M., Cheng, W.-C., Gilbert, C. E., Sing, C. N., Diny, N. L., Wheelan, S. J., Dunham, M. J., Boeke, J. D., Pineda, F. J. and Hardwick, J. M. (2013), 'Genome-wide consequences of deleting any single gene', *Molecular Cell* **52**(4), pp. 485–494. doi: 10.1016/j.molcel.2013.09.026.

Tesson, B., Lerch, S. J. L. and Hildebrand, M. (2017), 'Characterization of a new protein family associated with the silica deposition vesicle membrane enables genetic manipulation of diatom silica', *Scientific Reports* **7**(13457). doi: 10.1038/s41598-017-13613-8.

Tirichine, L. and Bowler, C. (2011), 'Decoding algal genomes: tracing back the history of photosynthetic life on earth', *The Plant Journal* **66**(1), pp. 45–57. doi: 10.1111/j.1365-313x.2011.04540.x.

Tommasi, E. D., Gielis, J. and Rogato, A. (2017), 'Diatom frustule morphogenesis and function: a multidisciplinary survey', *Marine Genomics* **35**, pp. 1–18. doi: 10.1016/j.margen.2017.07.001.

Yang, H., Ren, S., Yu, S., Pan, H., Li, T., Ge, S., Zhang, J. and Xia, N. (2020), 'Methods favoring homology-directed repair choice in response to CRISPR/cas9 induced-double strand breaks', *International Journal of Molecular Sciences* **21**(18) (6461). doi: 10.3390/ijms21186461.

Yuen, G., Khan, F. J., Gao, S., Stommel, J. M., Batchelor, E., Wu, X. and Luo, J. (2017), 'CRISPR/cas9-mediated gene knockout is insensitive to target copy number but is dependent on guide RNA potency and cas9/sgRNA threshold expression level', *Nucleic Acids Research* **45**(20), pp. 12039–12053. doi: 10.1093/nar/gkx843.

Zendejas, F. J., Benke, P. I., Lane, P. D., Simmons, B. A. and Lane, T. W. (2011), 'Characterization of the acylglycerols and resulting biodiesel derived from vegetable oil and microalgae (thalassiosira pseudonana and phaeodactylum tricornutum)', *Biotechnology and Bioengineering* **109**(5), pp. 1146–1154. doi: 10.1002/bit.24395.

Zurzolo, C. and Bowler, C. (2001), 'Exploring bioinorganic pattern formation in diatoms. a story of polarized trafficking', *Plant Physiology* **127**(4), pp. 1339–1345. doi: 10.1104/pp.010709.

Appendices

A SnapGene maps

pTpPUC3:

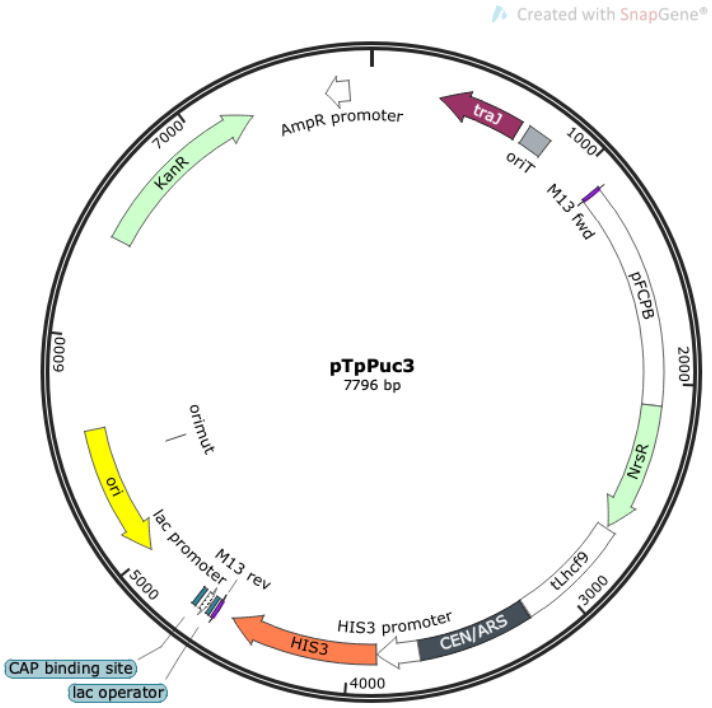


Figure A.1: Map of vector plasmid pTpPUC3. The pTpPUC3 plasmid contain the nourseothricin resistance cassette (NrsR) for resistance in *T. pseudonana*, kanamycin resistance cassette (KanR) for resistance in *E. coli*, the yeast derived sequence CEN/ARS-HIS3, which is responsible for replication and segregation in diatoms, and an origin of transfer (oriT). The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

pTPPUC3 hCas9(Mock) U6 complete:

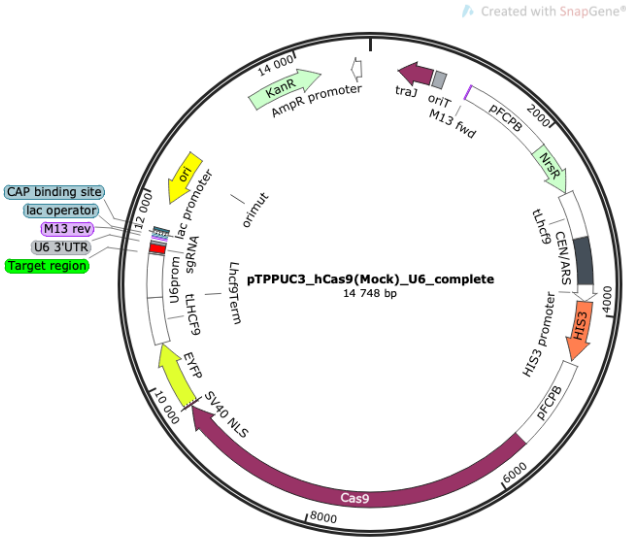


Figure A.2: Map of vector plasmid pTPPUC3 hCas9(Mock) U6 complete. The pTPPUC3 hCas9(Mock) U6 complete plasmid contain the nourseothricin resistance cassette (NrsR) for resistance in *T. pseudonana*, kanamycin resistance cassette (KanR) for resistance in *E. coli*, the yeast derived sequence CEN/ARS-HIS3, which is responsible for replication and segregation in diatoms, an origin of transfer (oriT), the Cas9 fragment controlled by pFCPB, with SV40 NLS and a C terminal YFP tag, and a sgRNA cassette controlled by a U6 promoter. The plasmid map is retrieved from SnapGene (Dotmatics, 2021)..

pBKS_bsr:

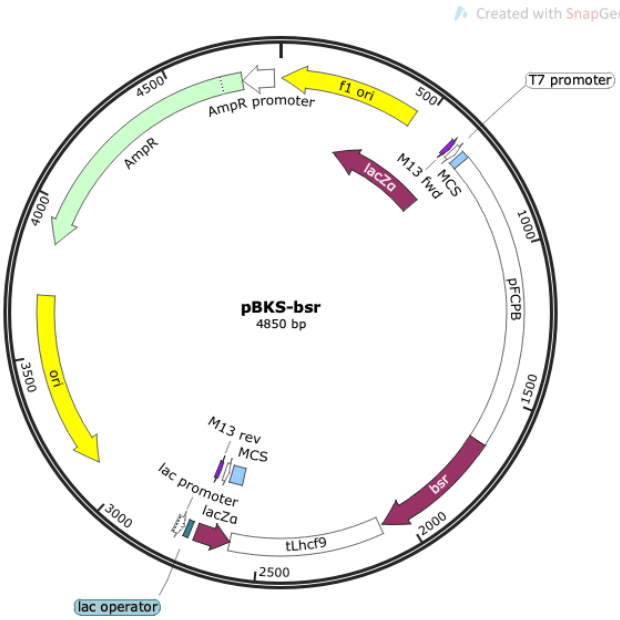


Figure A.3: Map of vector plasmid pBKS_bsr. The pBKS plasmid is assembled to contain the blasticidine resistance (bsr) cassette for resistance in *T. pseudonana* and ampicillin resistance (AmpR) cassette for resistance in *E. coli*. The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

pBKS_nat:

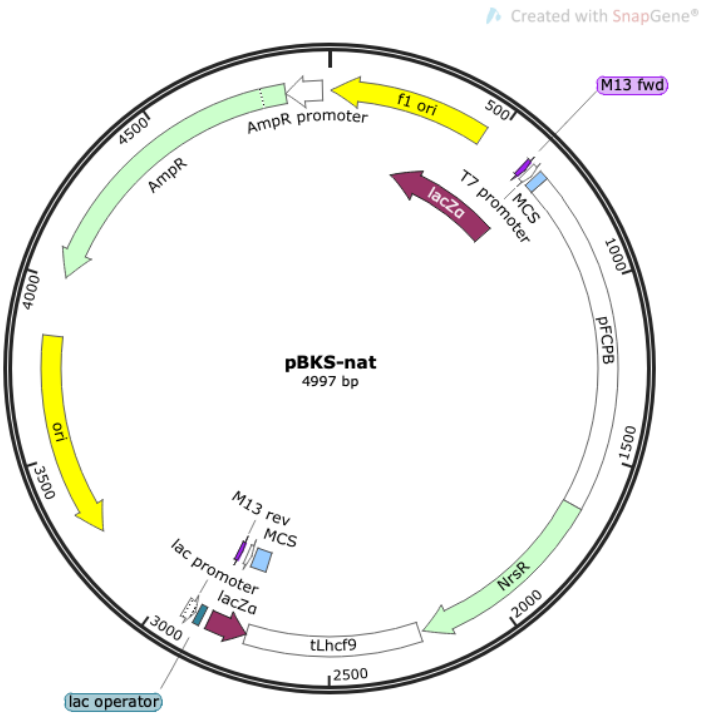


Figure A.4: Map of vector plasmid pBKS_nat. The pBKS plasmid is assembled to contain the nourseothricin resistance (nat) cassette for resistance in *T. pseudonana* and ampicillin resistance (AmpR) cassette for resistance in *E. coli.* The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

pBKS_ptPDS1-M1:

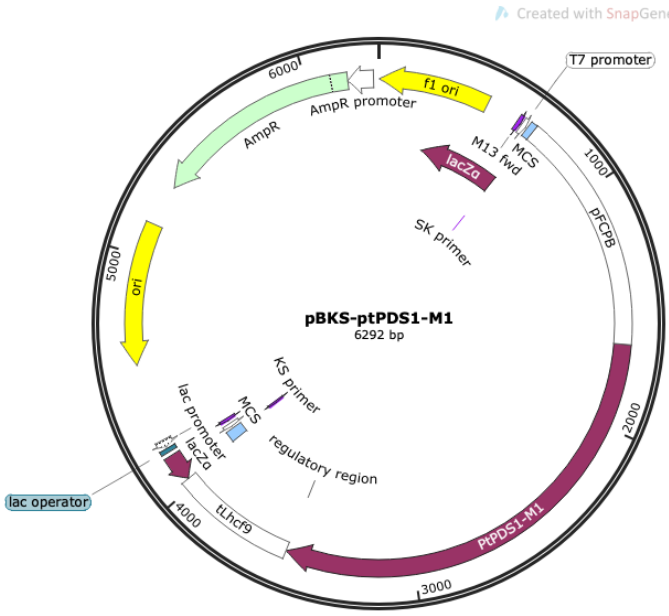


Figure A.5: Map of vector plasmid pBKS_ptPDS1-M1. The pBKS plasmid is assembled to contain the norflurazon resistance (ptPDS1-M1) cassette for resistance in *T. pseudonana* and ampicillin resistance (AmpR) cassette for resistance in *E. coli.* The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

pTpPUC3-mNeonGreen-Tp24711:

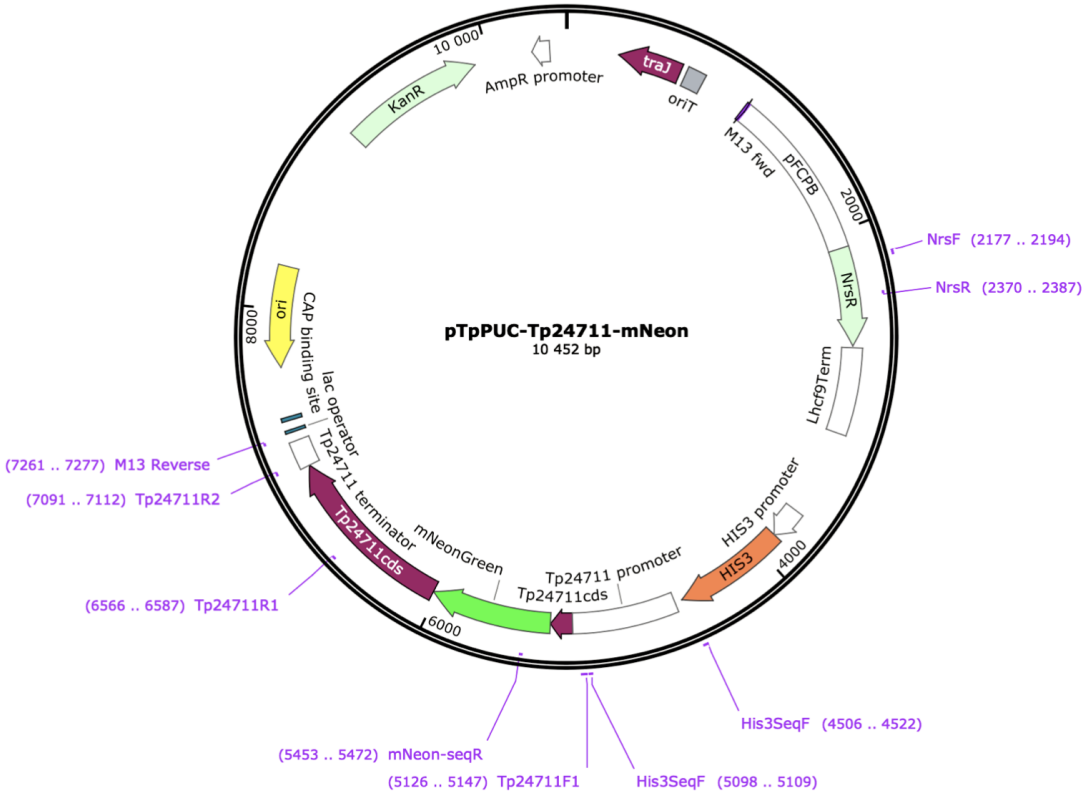


Figure A.6: Map of vector plasmid pTpPUC3-mNeonGreen-Tp24711. The pTpPUC3 plasmid contain the nourseothricin resistance cassette (NrsR) for resistance in *T. pseudonana*, kanamycin resistance cassette (KanR) for resistance in *E. coli*, the HIS3, which is responsible for replication and segregation in diatoms, and an origin of transfer (oriT). The mNeonGreen fusion gene is inserted before the coding DNA sequence of Tp24711. The plasmid map is retrieved from SnapGene (Dotmatics, 2021).

B Culture media and solutions

Table B.1: Recipe for L1 medium and agar plates (Guillard and Hargraves, 1993). The stock solutions of nutrients were pre-made. For medium, the nutrients were added by sterile filtration (2 μm) to cooled down 1 L autoclaved (20 minutes, 121 °C) sterile filtered (0.2 μm) seawater (from Trondheimsfjorden, provided by NTNU Sealab). For agar plates, filtered seawater was mixed with 1: 2 with distilled water and autoclaved for 20 minutes at 121 °C. Nutrients were added with sterile filtration after the autoclaved solution was cooled down to approximately 55 °C. Antibiotics were added by pipetting to medium and the agar plate solutions if needed.

Stocks	**Components**	**Amount per liter**
(1) L1 trace elements	$Na_2EDTA \cdot 2H_2O$	4.36 g
	$FeCl_3 \cdot 6H_2O$	3.15 g
	$MnCl_2 \cdot 4H_2O$	1 mL
	$ZnSO_4 \cdot 7H_2O$	1 mL
	$CoCl_2 \cdot 6H_2O$	1 mL
	$CuSO_45 \cdot H_2O$	1 mL
	$Na_2Mo_4 \cdot 2H_2O$	1 mL
	H_2SeO_3	1 mL
	$NiSO_4 \cdot 6H_2O$	1 mL
	Na_3VO_4	1 mL
	K_2CrO_4	1 mL
(2) Vitamin mix	Cyanocobalamin (Vitamin B_{12})	1 mL
	Thiamine HCL (Vitamin B_1)	200 mg
	Biotin (Vitamin H)	10 mL
(3) Sodium metasilicate	$Na_2SiO_3 \cdot 9H_2O$	30.0 g
Medium	**Components**	**Amount per liter**
	$NaNO_3$	1 mL
	$NaH_2PO_4 \cdot H_2O$	1 mL
	(1) L1 trace elements	1 mL
	(2) Vitamin mix	0.5 mL
	(3) Sodium metasilicate	1 mL
	Agar for microbiology (Only for agar plates)	10 g

Table B.2: Recipe for Luria Bertani (LB) medium and agar plates (Bertani, 1951). The components from the given suppliers, were added to 1 L of distilled water and autoclaved for 20 minutes at 121 °C. For agar plates with antibiotic, the antibiotic was added after the autoclaved solution was cooled down to approximately 55 °C.

Components	**Amount [g/L ddH_2O]**	**Supplier**
Tryptone	10	VWR Chemicals BDH®
Yeast extract	5	Oxoid
NaCl	5	VWR Chemicals BDH®
Agar for bacteriology (Only for agar plates)	15	VWR Chemicals BDH®

Table B.3: Recipe for Lysis buffer for *T. pseudonana*. The components were added to distilled water by sterile filtration (0.2 μm).

Components	**Concentration**
Triton X-100	10%
Tris–HCl pH 8	20 mM
EDTA	10 mM

Table B.4: Recipe for 50x TAE buffer. The components were added to a 1 L bottle and filled with distilled water. 1x TAE buffer was preppared from this.

Components	**Amount**	**Final oncentration**
Tris base	242 grams	2 M
Glacial acetic acid	57.1 mL	1 M
0.5 M EDTA, pH 8.0	100 mL	0.05 M

C Reaction mixes

Table C.1: A general recipe for restriction enzyme digestion (New England Biolabs, 2022). The reaction mixture was adjusted to the desired volume. If two or more restriction enzymes were used, 1 μL of each was added and the amount of Nuclease-free water was adjusted accordingly.

Components	**Amount (50 μL reaction**
DNA	1 μg
10X rCutSmart Buffer	5 μL
Restriction enzyme	1.0 μL
Nuclease-free water	to 50 μL

Table C.2: A general recipe for genomic DNA (gDNA) elimination reaction (QIAGEN, 2016*a*).

Components	**Volume [µL]**	**Final concentration**
gDNA Wipeout Buffer, 7x	3 µL	1x
Template RNA	1 µg	-
RNAse free water	up to 21 µL	-
Total volume	21 µL	-

D Primers

Table D.1: Primers used for PCR amplification, HRM and sequencing

	Primer set	Sequence (5´-3´)
PCR primers		
Tp24708/24711	Tp24708F1	GCCTTCAAAACGACAATGTGGA
Targeting for Tp24708	Tp24708PAM2qSR	TACCTCCGTTCTCGGAACAGTA
Tp24708/24711	Tp24711F1	GAGACGACGAATCCTTTGCCAC
Targeting for Tp24711	Tp24711R1	TCCTCCGTTCTCGGAACAGTAG
Tp24711P2	Tp24711F1	GAGACGACGAATCCTTTGCCAC
	Tp24711R1	TCCTCCGTTCTCGGAACAGTAG
Tp24708/24711	Tp24708PAM1qSF	ACATGTCCCTCAAGTTCCAG
Targeting for Tp24708	Tp24708PAM1qSR	CATGGAGCATGAATCCGTTG
HRM primers		
Tp24708/24711	Tp24711PAM1qSF	ACATGTCCCTCAAGTTTCAG
Targeting for Tp24711	Tp24711PAM1qSR	CTCCATAGAACAAGAAGTCG
Tp24711P2	Tp24711SPAM2qSF	GCCATGGAAGCACCCAGCA
	Tp24711SPAM2qSR	GGAGTGGTAGCATCCCTGA
qRT-PCR primers for clones and controls		
LHCF9	qTpLhcf9F	CCATGATGGGAATTCTTGGACT
	qTpLhcf9R	AGCCGAATGTAACCATTGTGCT
sgRNA	Tp24708/24711PAM1F	ATTGtACGGAACCTCACAAGACGCTA
	qsgRNA	TCAAGTTGATAACGACTAGCC
Cas9	qhCas9F	GGCATAAGCCCGAGAATATC
	qhCas9R	TCCTCTTCATCCTTTCCCTAC
Nat	qNatF	GCCATCGAGGCACTGGATGGGT
	qNatR	CGTCGGGGAACACCTTGGTCAG
Assembly of pBKS_hCas9M_Bls/Nou/Nrf		
	pBKS_BsaImutF	GATACCGCGAGAGCCACGCTCACCGGCTCCAGATTTATCAGC
	pBKS_BsaImutR	GAGCCGGTGAGCGTGGCTCTCGCGGTATCATTGCAGCACTGGGGC
	Gib_SelMF	GGTGGCGGCCGCTAGCTTGCGCTTTTTCCGAGA
	Gib_SelMR	ATCGATAAGCTTGATATCGAATTCCTGCAGGGAGAACTGGAGCAGCTACTAC
	M13fwd	TGTAAAACGACGGCCAGT
	M13rev	CAGGAAACAGCTATGAC
	Gib_hCas9M-pKS-Nrf/Nou F	GTTCTCCCTGCAGGAATTCGATATCAAGCTTGCGCTTTTTCCGAGAAC
	Gib_hCas9M-pKS-Nrf/Nou R	GGCCCCCCCCTCGAGGGCATGCCTGCAGGTCAATAGAAT
	Gib_hCas9M-pKS-Bls R	GTACCGGGCCCCCCCCGCATGCCTGCAGGTCAATAGAAT
	CasM_F2	GATCTCGACAATCTGCTGGC
	CasM_R2	GTGCGCTGTTTGCGC
Control of integrated Cas9		
	Cas9M_F1	GGCTCGATATCGGCACAAAC
	Cas9M_R3	CGACCTTAACGGTCTGC

E Images from gel electrophoresis

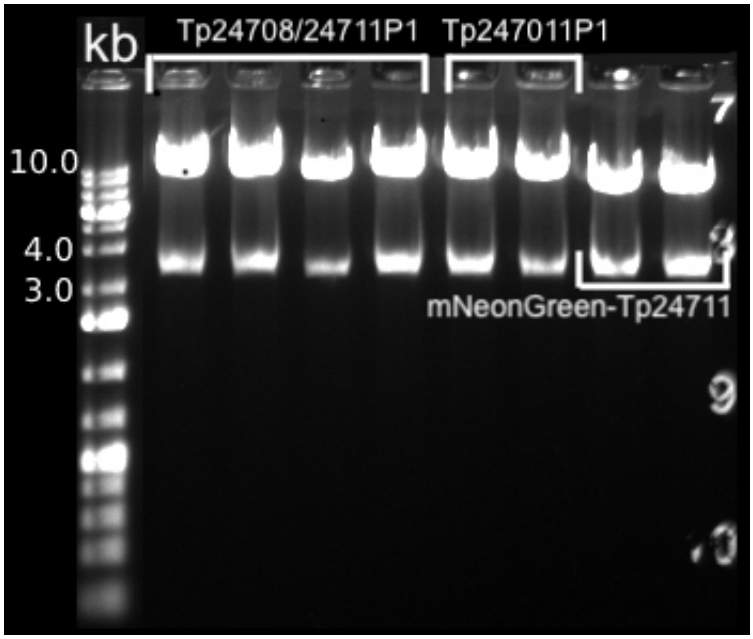


Figure E.1: Control restriction enzyme digestion performed on the three plasmids, Tp24708/24711P1, Tp24711P1 and mNeongreen-Tp24711, to verify the size of the plasmid, performed with the dual cutter restriction enzyme KpnI-HF followed by a gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1h). GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths.

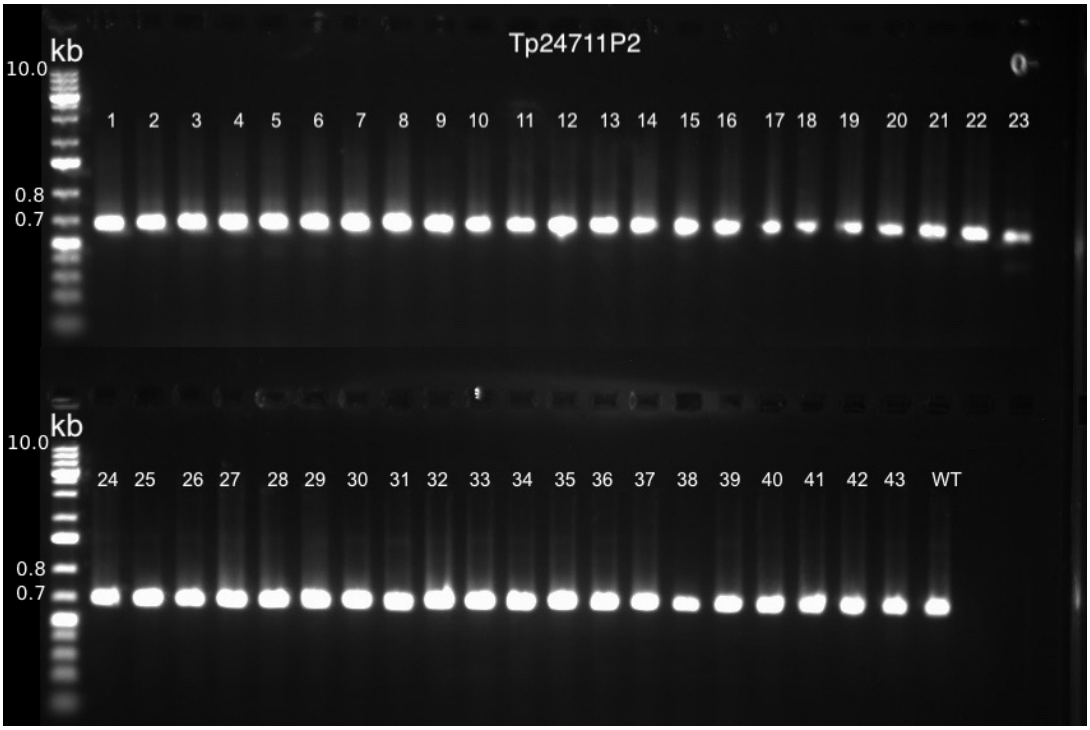


Figure E.2: To confirm the size of 43 Tp24711P2 clones amplified for larger indels by PCR, a gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1h) was performed. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths. The 43 clones were compared to the WT.

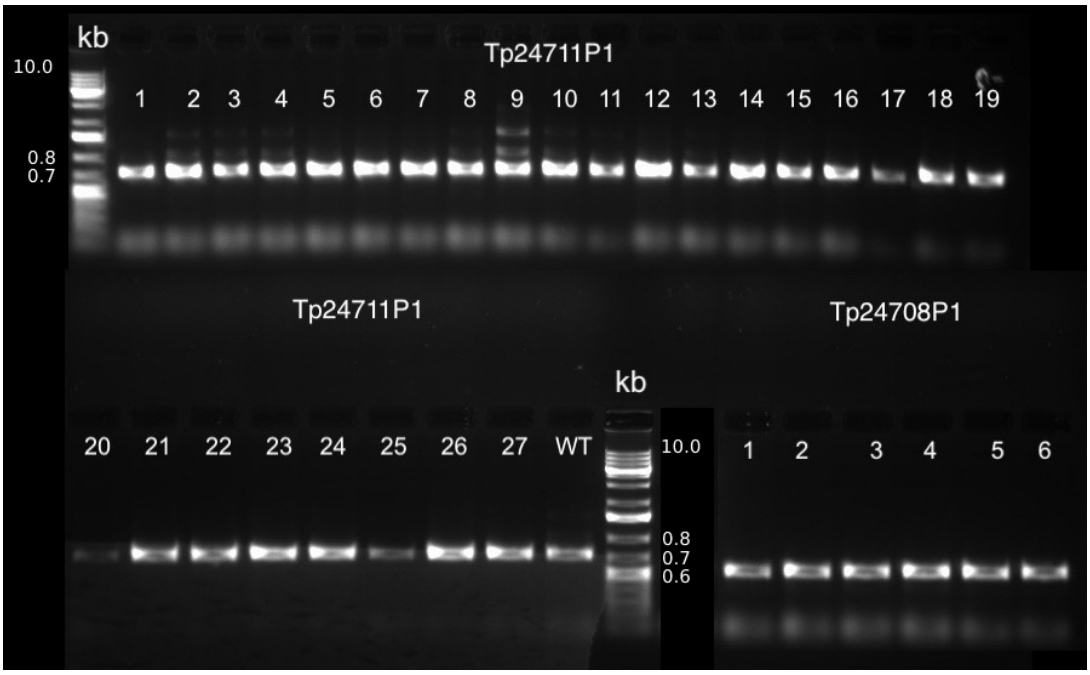


Figure E.3: To confirm the size of 27 Tp24708/24711P1 clones targeted for Tp24711 and 6 clones targeted for Tp24708, which where amplified for larger indels by PCR, a gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1h) was performed. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths. The clones were compared to the WT.

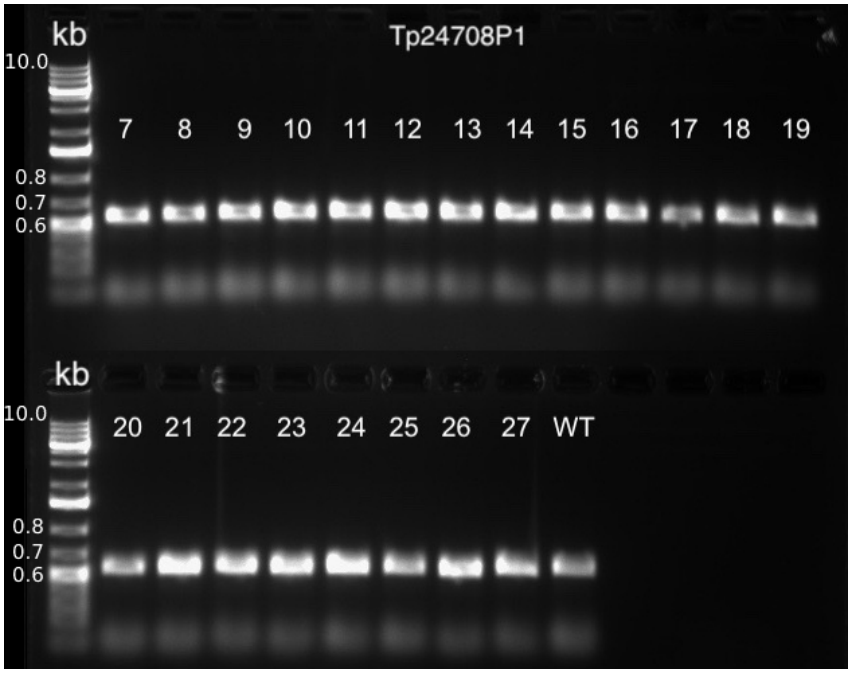


Figure E.4: To confirm the size of Tp24708/24711P1 clones targeted for Tp24708, which where amplified for larger indels by PCR, a gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1h) was performed. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths. The clones were compared to the WT.

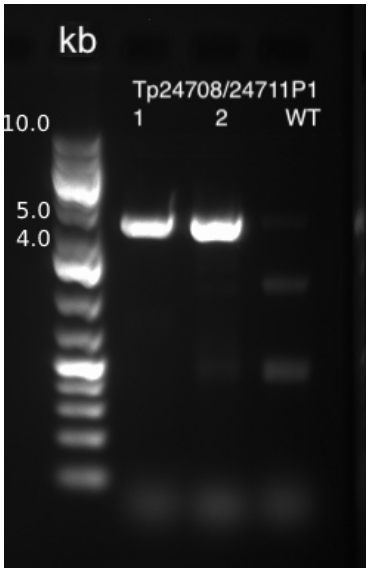


Figure E.5: An example of verification of integrated Cas in transformants by PCR, where two Tp24708/24711P1 clones were compared to WT. A gel electrophoresis (1% agarose in Tris-acetate-EDTA buffer with GelRed® Nucleic Acid Gel Stain (0.05 M, Biotium), 120V for 1h) was performed. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as a reference for sizing and approximation of the fragment lengths.

F Images of pTpPUC3-mNeonGreen-Tp24711 in *T. pseudonana*

F.1 Fluorescence microscopy

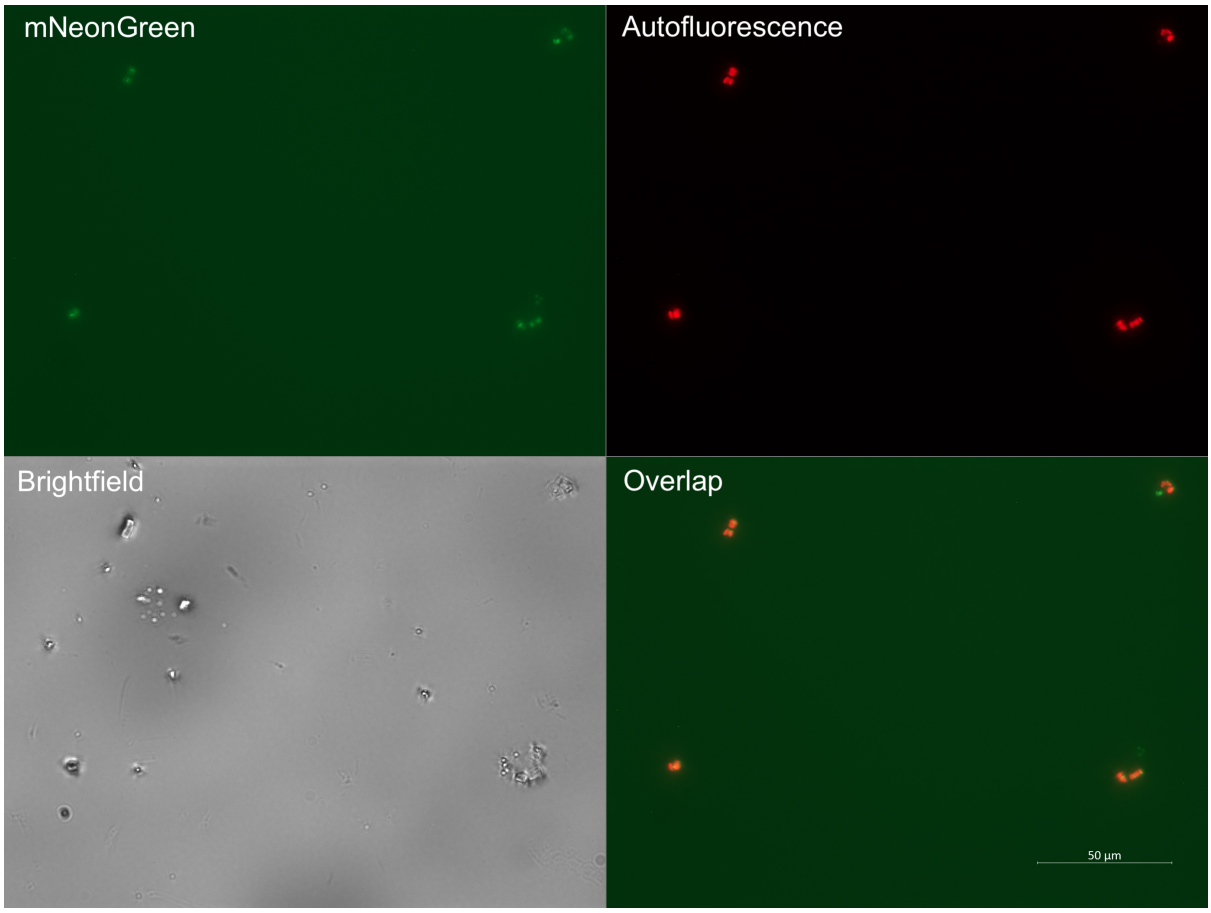


Figure F.1: Fluorescence microscope image of mNeonGreen - Tp24711 - s2 expressed in *T. pseudonana*. Upper left: PaGfp channel (505-555 nm), to detect mNeonGreen fluorescence (green), upper right: Cy5 channel (665-715 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a brightfield channel and lower right: overlap of both fluorescence channels. The scale bares are 50 µm. The image is retrieved from ZEISS ZEN Software.

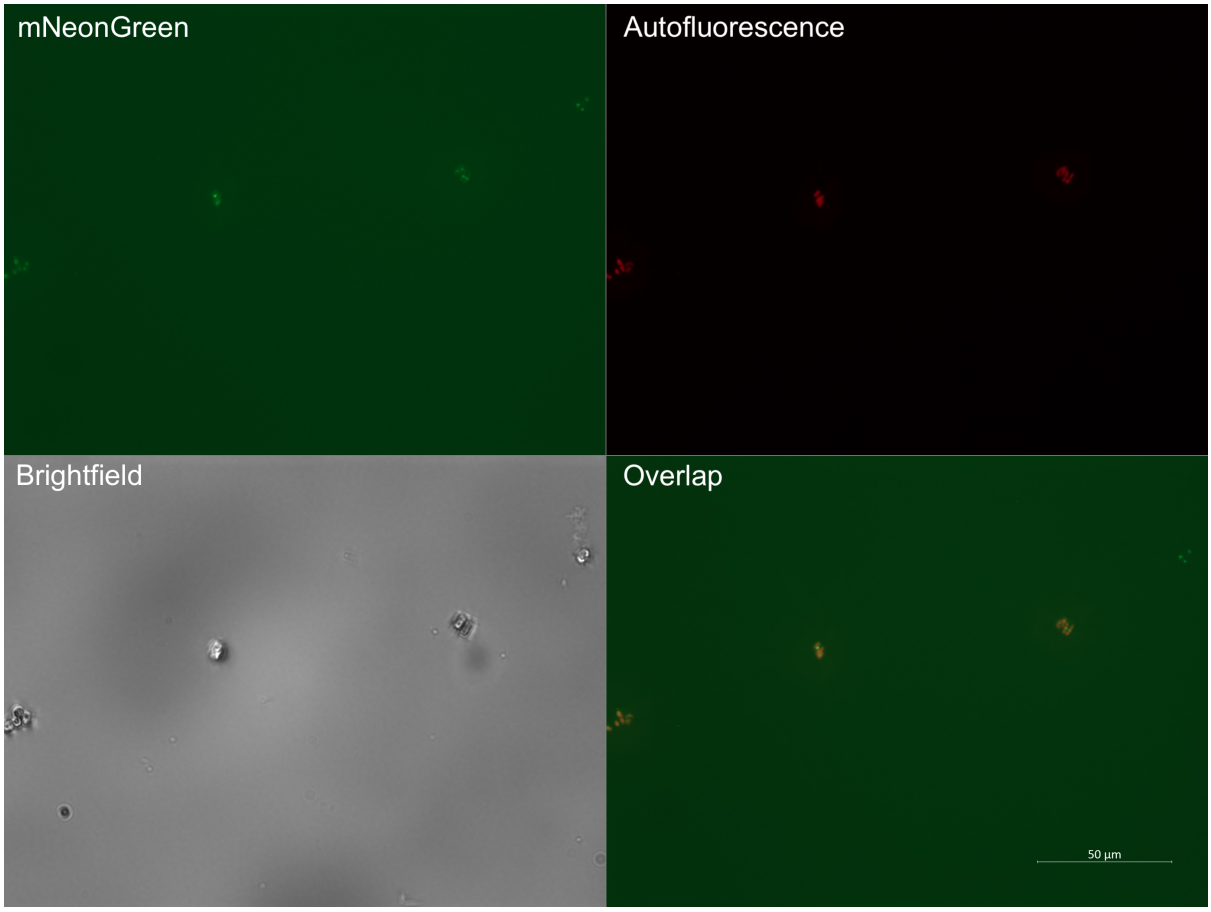


Figure F.2: Fluorescence microscope image of mNeonGreen - Tp24711 - s6 expressed in *T. pseudonana*. Upper left: PaGfp channel (505-555 nm), to detect mNeonGreen fluorescence (green), upper right: Cy5 channel (665-715 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a brightfield channel and lower right: overlap of both fluorescence channels. The scale bares are 50 µm. The image is retrieved from ZEISS ZEN Software.

F.2 Confocal microscopy

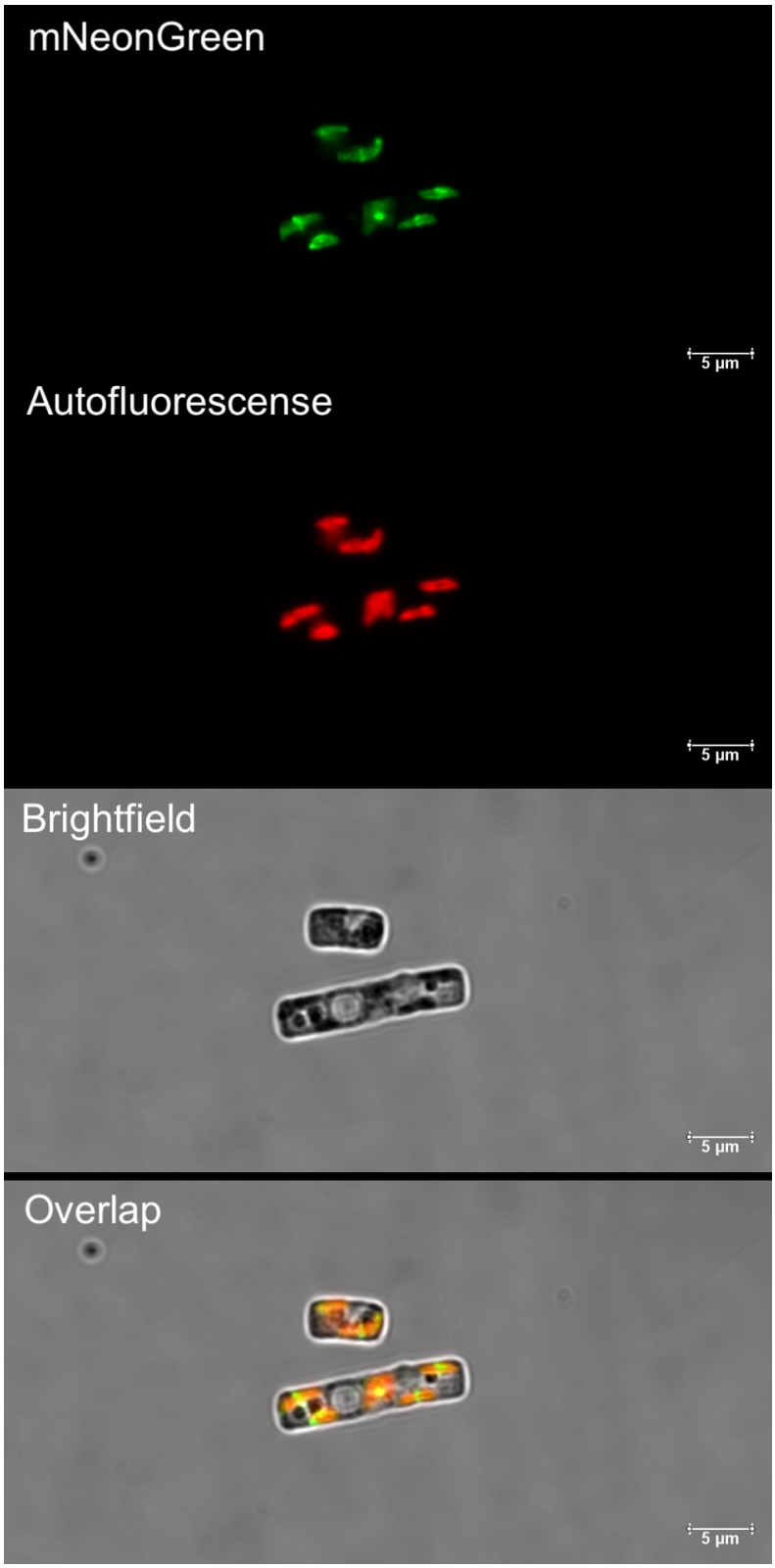


Figure F.3: Confocal microscope image (1024x512) of mNeonGreen - Tp24711 - s2 expressed in *T. pseudonana*. First part: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), second part: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, third part: a PMT brightfield channel and fourth part: overlap of all channels. The scale bares are 5 µm. The image is retrieved from Leica LAS AF Software.

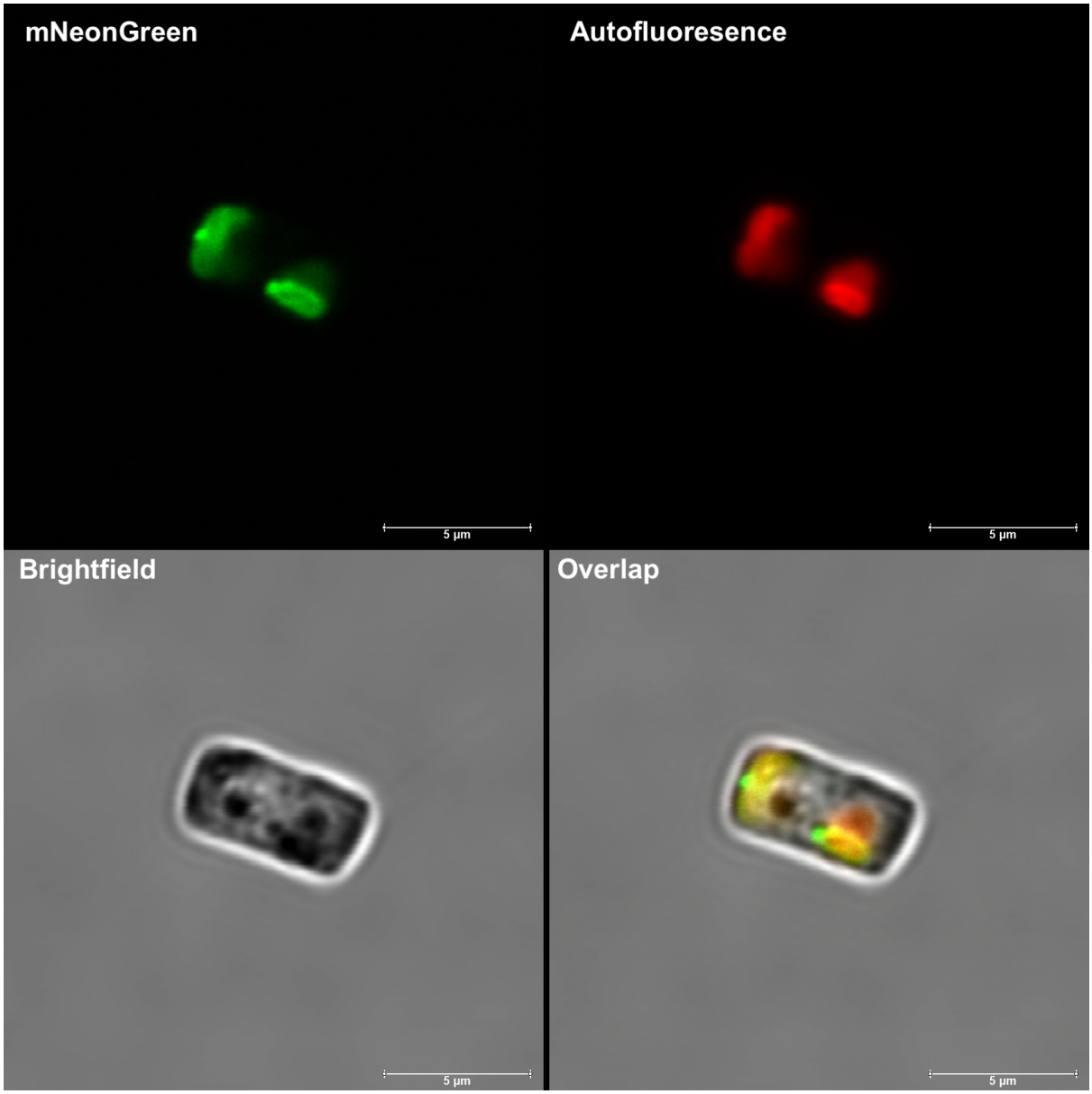


Figure F.4: Confocal microscope image (512x512) of mNeonGreen - Tp24711 - s2 expressed in *T. pseudonana*. Upper left: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), upper right: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a PMT brightfield channel and lower right: overlap of all channels. The scale bares are 5 µm. The image is retrieved from Leica LAS AF Software.

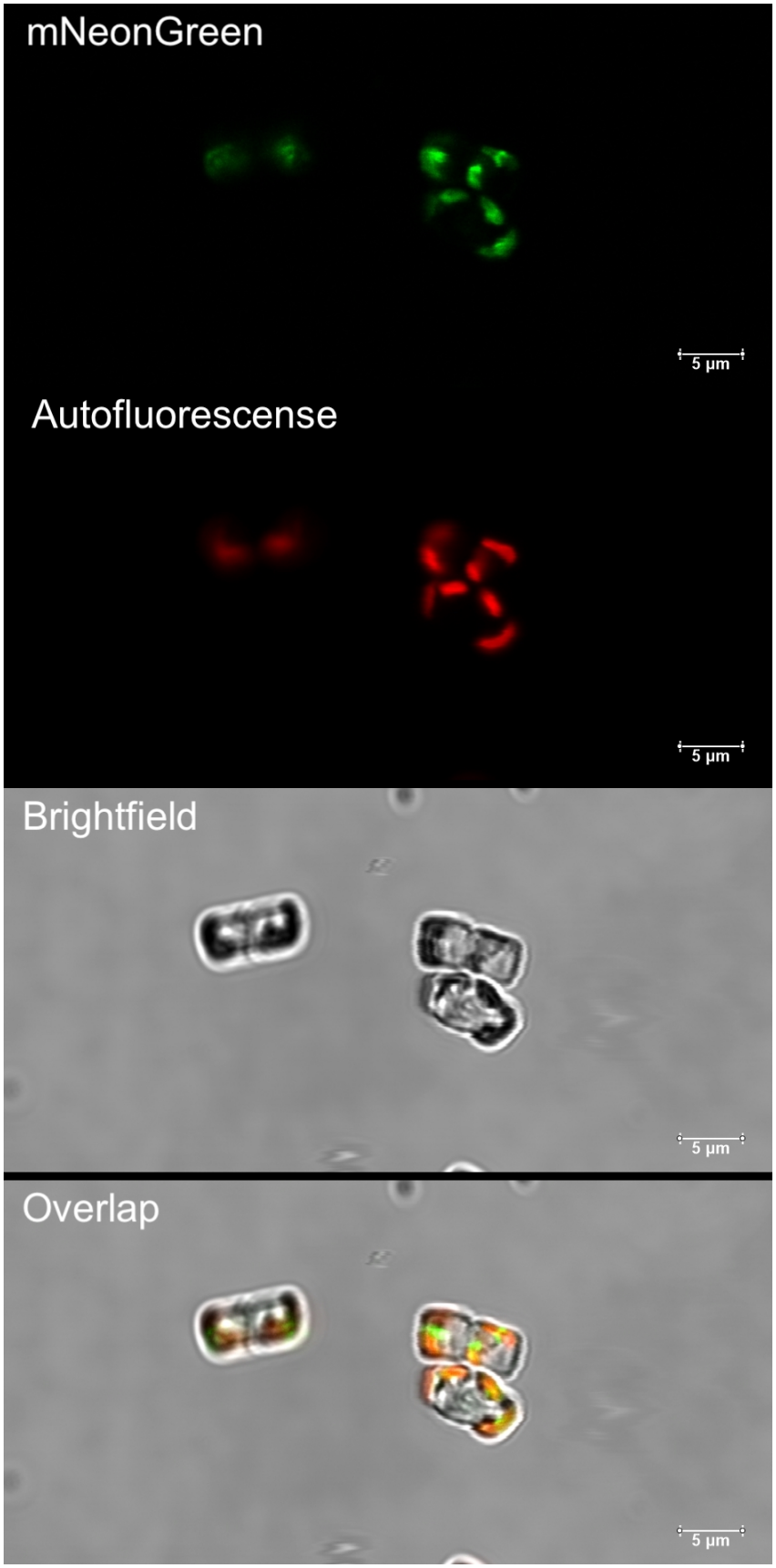


Figure F.5: Confocal microscope image (1024x512) of mNeonGreen - Tp24711 - s6 expressed in *T. pseudonana*. First part: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), second part: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, third part: a PMT brightfield channel and fourth part: overlap of all channels. The scale bares are 5 µm. The image is retrieved from Leica LAS AF Software.

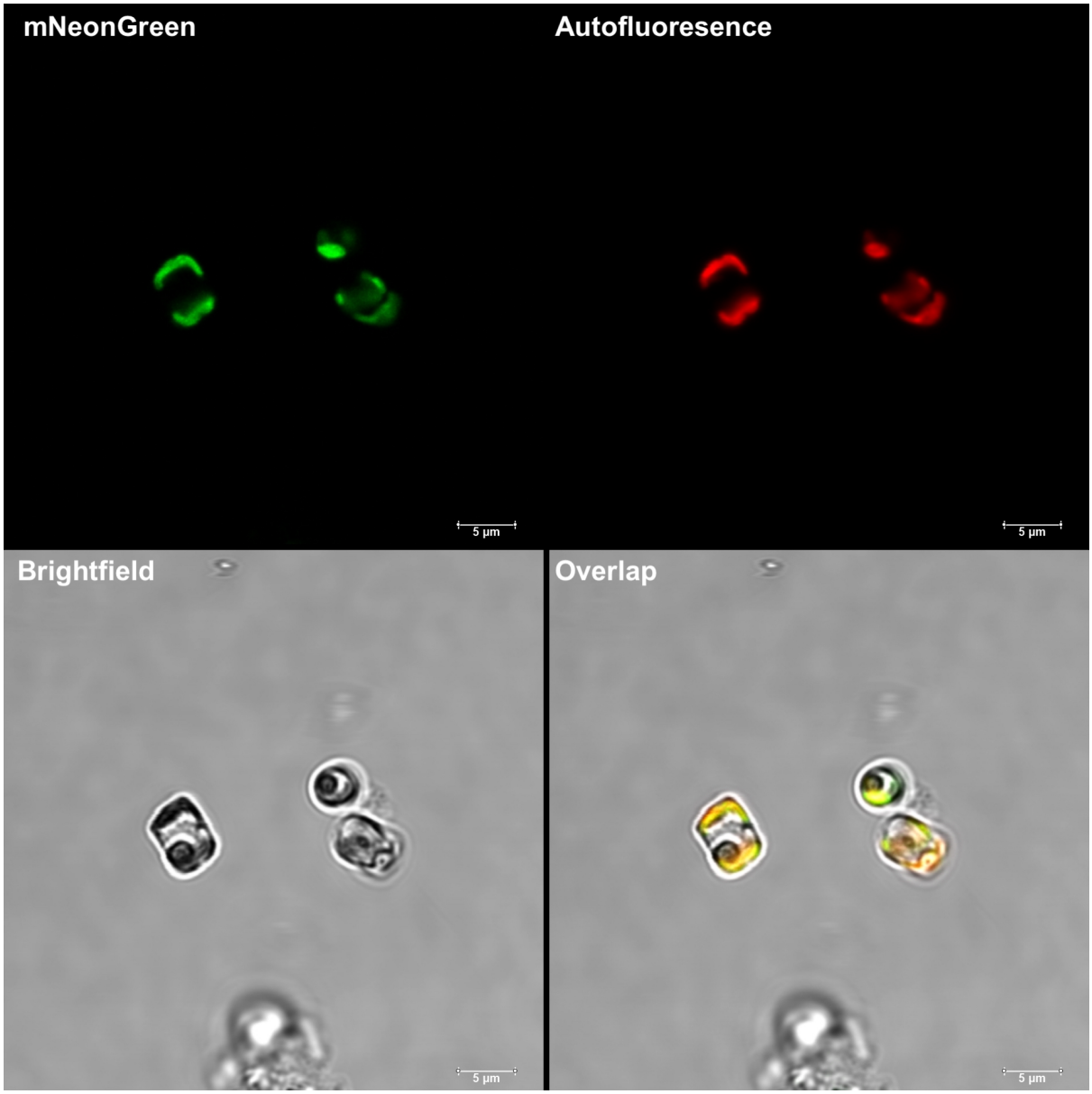


Figure F.6: Confocal microscope image (512x512) of mNeonGreen - Tp24711 - s6 expressed in *T. pseudonana*. Upper left: PMT1 channel (493-540 nm), to detect mNeonGreen fluorescence (green), upper right: PMT 3 channel (651-779 nm) to detect autofluorescence (red) emitted from the chloroplasts, lower left: a PMT brightfield channel and lower right: overlap of all channels. The scale bares are 5 µm. The image is retrieved from Leica LAS AF Software.

G Sequencing results of the new pBKS plasmid assembled with selection

pBKS_BsaImut:

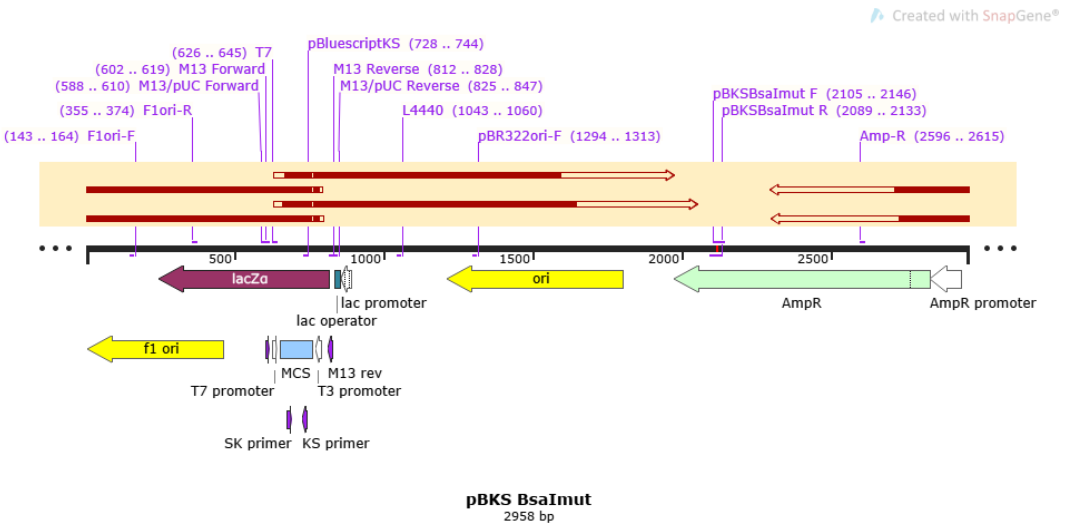


Figure G.1: The coding sequence of pBKS_BsaImut, with two of the most matched Sanger sequencing results aligned. The primers used in the samples submitted for Sanger sequencing are pBKS_BsaImut_F and pBKS_BsaImut_R. There are mutations in the overlap of the primers, which do not affect the plasmid. The worst ends of aligned sequence traces are hidden. The coding sequence is retrieved from SnapGene (Dotmatics, 2021).

pBKS_bsr:

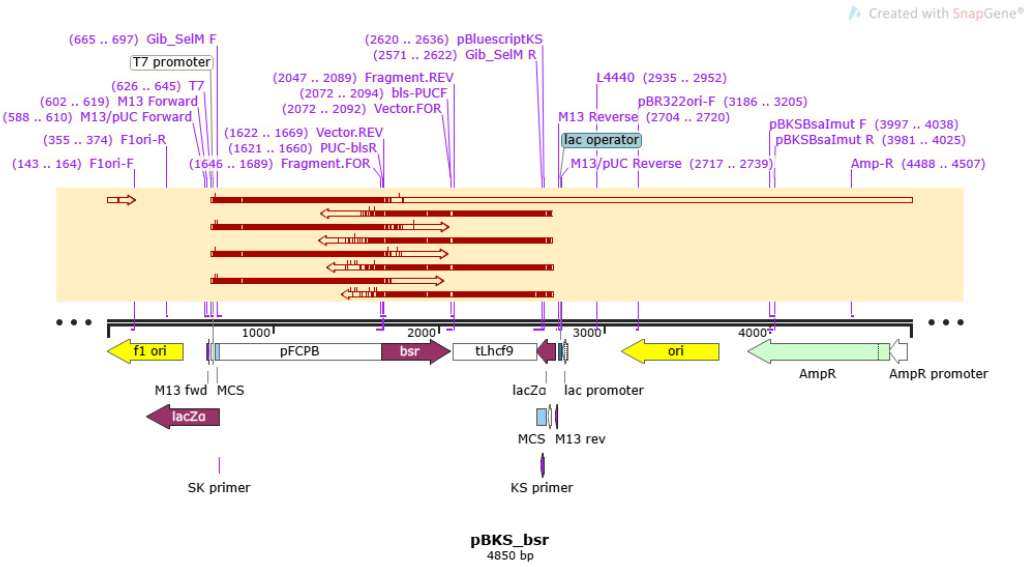


Figure G.2: The coding sequence of pBKS_bsr, with all the Sanger sequencing results aligned. The primers used in the samples submitted for Sanger sequencing are M13fwd and M13rev. There are mutations in the overlap of the primers, which do not affect the plasmid, but there are other mutations that may have an effect on the plasmid. The worst ends of aligned sequence traces are hidden. The coding sequence is retrieved from SnapGene (Dotmatics, 2021).

pBKS_nat:

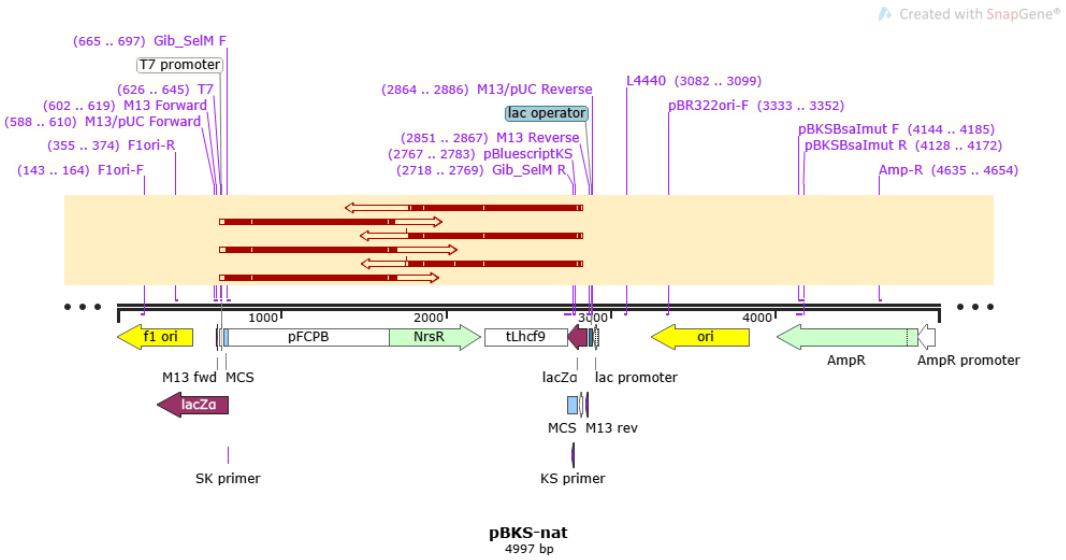


Figure G.3: The coding sequence of pBKS_nat, with all the Sanger sequencing results aligned. The primers used in the samples submitted for Sanger sequencing are M13fwd and M13rev. There are mutations in the overlap of the primers, which do not affect the plasmid, but there are other mutations that may have an effect on the plasmid. The worst ends of aligned sequence traces are hidden. The coding sequence is retrieved from SnapGene (Dotmatics, 2021).

pBKS_ptPDS1-M1:

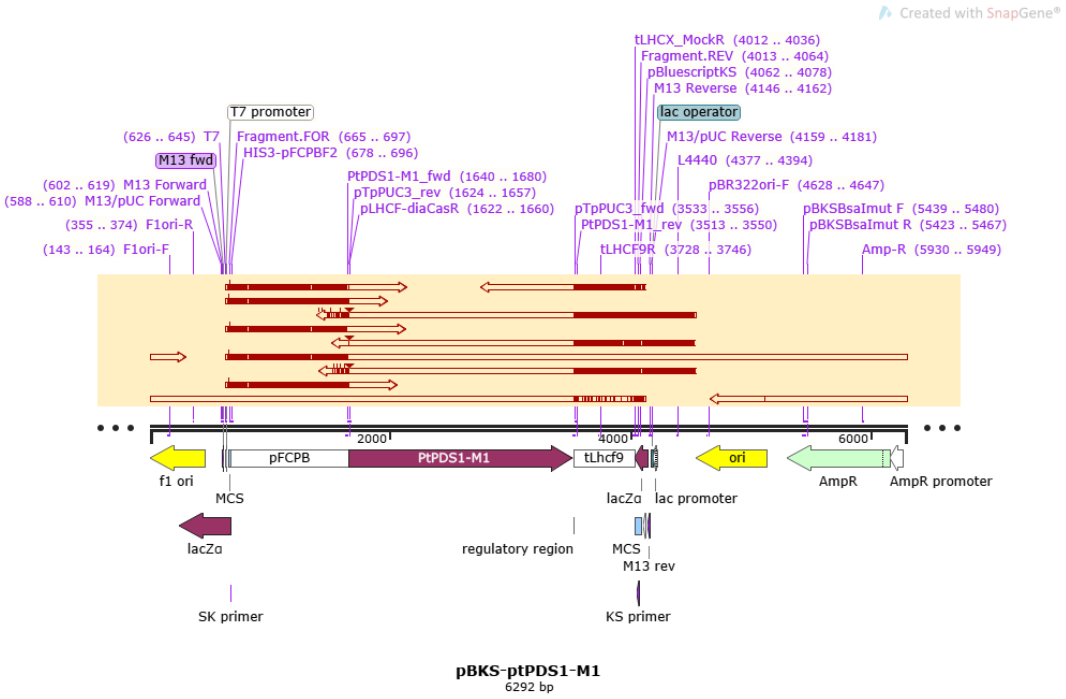


Figure G.4: The coding sequence of pBKS_ptPDS1_M1, with all the Sanger sequencing results aligned. The primers used in the samples submitted for Sanger sequencing are M13fwd and M13rev. There are mutations in the overlap of the primers, which do not affect the plasmid, but there are other mutations that may have an effect on the plasmid. The worst ends of aligned sequence traces are hidden. The coding sequence is retrieved from SnapGene (Dotmatics, 2021).

H Compounds and instruments

Table H.1: A list of strains, with associated microorganism, used during the thesis.

Microorganism	Strain
E.coli	DH5α
T. pseudonana	CCMP 1335

Table H.2: A list of kits with associated supplier, used during the thesis.

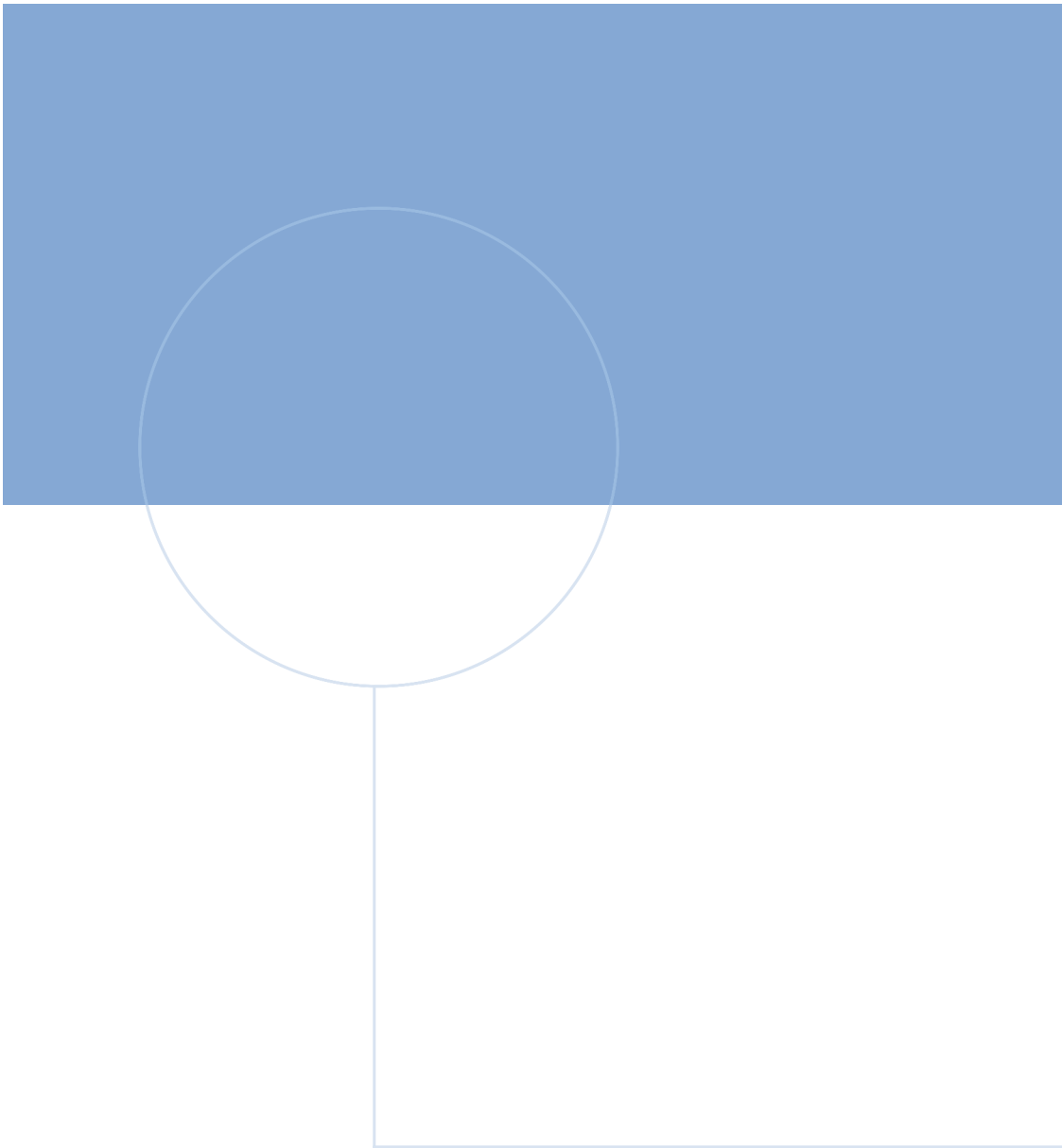
Kit	Supplier
E.Z.N.A® Plasmid DNA Mini Kit II	Omega Bio-tek
Gibson Assembly® Master Mix kit	New England Biolabs
LightCycler® 480 High Resolution Melting Master	Roche
LightCycler® 480 SYBR Green I Master kit	Roche
Monarch® PCR & DNA Cleanup Kit (5 µg)	New England Biolabs
Phusion plus DNA polymerase kit	Thermo Scientific
QuantiTect® Reverse Transcription kit	QIAGEN
Quick-Start Protocol QuantiTect® Reverse Transcription Kit	QIAGEN
RNeasy Plant Mini Kit	QIAGEN

Table H.3: A list ofall software with associated developer, used during the thesis.

Name	Developer
Attune™ NxT	Thermo Fisher Scientific
ExCel	Microsoft
Gel documentation system	Syngene
ICE	Synthego
Leica LAS A	Leica microsystems
LightCycler96	Roche
SnapGene	Dotmatics
ZEISS ZEN	Zeiss

Table H.4: A list of equipment with associated supplier, used during the master thesis.

Equipment	**Supplier**
Adhesive PCR Sealing Foil Sheets	Thermo scientific
Attune™ NxT Flow-cytometer	Thermo Fisher Scientific
Autoclave	Astell
Axio Imager.Z2 Fluorescence microscope	Zeiss
Centrifuge	Eppendorf™ and Thermo scientific
Cover glass	VWR
Culture flask with filter cap	VWR
Double-sided tape	race Bio-Labs SecureSeal™ imaging spacer
Gel electrophoresis	Consort
Gel tray	Thermo Scientific
Gene Pulser Xcell Electroporation System	Bio-Rad
Laboratory incubator	Termaks AS
Leica SP8 Confocal Microscope	Leica microsystems
LightCycler 96	Roche
LinRegPCR	LinRegPCR
Microcentrifuge	VWR
Micro tube 1.5ml	Sarstedt
Microscope slide	VWR
LightCycler® 480 Multiwell plate 96	Roche
NanoDrop One/OneC Microvolume UV-Vis Spectrophotometer	Thermo Scientific
PCR	Bio-Rad, T100 Thermal Cycler
LightCycler® 480 Sealing foil	Roche
Stackable incubator shaker	Infors HT
Stainless steel bead	QIAGEN
Stereo microscope M3C	Wild Heerbrugg
Thermo–Shaker	Grant-bio
TissueLyser Adapter Sets	QIAGEN
UV-visible spectrophotometer	Auxilab S.L
Vortex	VWR and Heidolph
VWR Reagent Reservoirs	VWR



NTNU

Norwegian University of
Science and Technology